

## The detrimental effects of a suggestion of sexism in an instruction situation <sup>☆</sup>

Glenn Adams <sup>a,\*</sup>, Donna M. Garcia <sup>b</sup>, Valerie Purdie-Vaughns <sup>c</sup>, Claude M. Steele <sup>d</sup>

<sup>a</sup> *Department of Psychology, University of Toronto, 100 St. George Street, Toronto, ON, Canada M5S 3G3*

<sup>b</sup> *Department of Psychology, University of Kansas, 1415 Jayhawk Blvd., Lawrence, KS 66045-7556, USA*

<sup>c</sup> *Department of Psychology, Yale University, 2 Hillhouse Ave. (Box 208205), New Haven, CT 06520-8205, USA*

<sup>d</sup> *Department of Psychology, Stanford University, Jordan Hall, Bldg 420, Stanford, CA 94305-2130, USA*

Received 18 March 2004; revised 25 August 2005

Available online 15 December 2005

### Abstract

This research investigates the hypothesis that the mere suggestion of sexism can harm women's experience of an instruction situation. Across three experiments, women exposed to the suggestion about the sexism of a male instructor reported a less positive experience, performed worse on a logic test, and rated the instructor as less competent than did women who were not exposed to the suggestion. The same harmful consequences did not befall men, even when they were potential targets of the alleged sexism. To interpret results, the authors emphasize the concept of social identity threat: the concern that one will be the target not only of stereotypes about inferiority, but also a more general hostility based on a salient social identity. Results suggest the need to expand conceptions of discrimination to include systemic forms of identity threat that can be sufficient to produce harm, even in situations where differential treatment is initially absent.

© 2005 Elsevier Inc. All rights reserved.

*Keywords:* Discrimination; Social identity threat; Stereotype lift; Stereotype threat; Systemic devaluation

Consider the situation of a male instructor conducting a tutorial for a female student. Now imagine that this situation becomes tainted with the suggestion of sexism. How does the suggestion of sexism change the instruction experience? Does it transform the situation into one that demands greater vigilance and causes distraction? Does it spoil the interpersonal dynamics of the situation, increasing the woman's discomfort and damaging her rapport with the instructor? In short, is the mere suggestion of sexism—

apart from the issue of actual discrimination—sufficient to cause harm to potential targets of sexism? Such questions are hard to answer in everyday life because suggestions of sexism and actual discrimination often co-occur. In the present research, we use a laboratory paradigm to investigate the consequences of a suggestion of sexism while holding constant the potential for discriminatory treatment.

### Direct and indirect consequences of systemic devaluation

Women in contemporary American settings face systemic devaluation associated with an environment of sexism (U.S. Department of Commerce, 2000). This environment of sexism impacts women's lives through direct acts of differential treatment. With respect to education, women face barriers to success in domains like math and science. These barriers include gendered socialization practices in which girls learn to devalue these domains (e.g.,

<sup>☆</sup> This research was partially supported by a Stanford Graduate Research Opportunity Grant to the first author. We thank Monica Biernat, Paul Davies, and Benoit Monin for their comments on earlier drafts. We also thank Duane Buscher, Maile Cabral, Annice Chen, Trinity Donovan, Stacy Fambro, Mike Fannen, Relief Jones, Fahad Khan, Erin Kidwell, Sarah Kurien, Eddie Sandoval, Michael Sessoms, James Silver and Nicole West for their assistance in collecting data.

\* Corresponding author.

E-mail address: [adams@psych.utoronto.ca](mailto:adams@psych.utoronto.ca) (G. Adams).

Eccles, Jacobs, & Harold, 1990). They also include blatant discrimination from classroom authorities, as when teachers treat girls in ways that suggest low expectations (e.g., Constantinople, Cornelius, & Gray, 1988).

As serious as these direct forms of discrimination are, we propose that the harmful consequences of an environment of sexism are not limited to such acts of differential treatment. In addition, the realistic threat of discrimination, rendered plausible by a cultural legacy of sexism, can be sufficient to poison interaction and elicit harmful outcomes—even in circumscribed situations that would otherwise appear to be free from differential treatment. One can consider these outcomes to be consequences of systemic devaluation to the degree that they reflect and require a cultural environment of sexism. These outcomes constitute *indirect* consequences of systemic devaluation to the extent that they occur in the absence of direct, differential treatment.

### Detrimental effects of a suggestion of sexism

To investigate the hypothesis that the threat of discrimination can be sufficient to produce discrimination-like outcomes, we conducted a series of three experiments in which we manipulated the threat of discrimination by having a confederate either mention or not mention that an instructor seemed sexist. Although ideological prescriptions and anecdotal evidence propose that this suggestion of sexism might increase women's motivation to do well, previous research indicates that the mere suggestion of sexism may have multiple negative consequences.

#### Comfort

First, regardless of whether an instructor is actually sexist, the mere suggestion of sexism may constitute a “threat in the air” (Steele, 1997) or contribute to a “chilly climate” (Constantinople et al., 1988; Hall & Sandler, 1982; Hyde & Kling, 2001) that not only renders the instruction situation less comfortable for potential targets of the alleged sexism, but also fosters short-term disengagement from the instruction situation (Major, Spencer, Schmader, Wolfe, & Crocker, 1998) and long-term disidentification with the instruction domain (Major & Schmader, 1998; Major et al., 1998; Schmader, Major, & Gramzow, 2001). One source of discomfort triggered by the suggestion of sexism is the localized threat of unfair treatment inside the circumscribed instruction setting. Learning that an instructor may be hostile toward one's gender could cause anyone—man or woman—to experience discomfort about the instruction situation. An additional source of discomfort that may be specific to women is a more systemic threat associated with broader devaluation outside the instruction setting. For women, the suggestion of sexism in a circumscribed instruction setting may activate associations to the larger environment of sexism, triggering a more general sense of hostility that men do not share.

#### Learning and performance

Second, regardless of whether an instructor is actually sexist, the mere suggestion of sexism may be sufficient to interfere with learning and performance. The suggestion of sexism may interfere with learning by distracting attention and resources away from the content of instruction to the task of monitoring the situation for signs of sexism (Frable, Blackstone, & Scherbaum, 1990; Lord & Saenz, 1985). The person need not believe the suggestion for such distraction to occur; instead, the suggestion may trigger an *automatic vigilance* (Pratto & John, 1991) for threat-related information that has a similarly distracting effect.

Likewise, the suggestion of sexism may interfere with performance not only by increasing distraction, but also through mechanisms like anxiety, disengagement, or evaluation apprehension (e.g., Geen, 1991; Spencer, Steele, & Quinn, 1999; Steele, Spencer, & Aronson, 2002; Wine, 1971). Again, targets of the alleged sexism need not believe the suggestion for such harms to occur. Instead, the suggestion of sexism may be sufficient to trigger a diffuse arousal that interferes with learning or test performance, even if the target concludes that sexism is not relevant (Inzlicht & Ben-Zeev, 2000; O'Brien & Crandall, 2003).

#### Evaluations of the instructor

Third, regardless of whether an instructor is actually sexist, the mere suggestion of sexism may be sufficient to harm evaluations of the instructor. People exposed to the suggestion may perceive the instructor to be less fair, less warm, or less effective than people who are not exposed to the suggestion (see Brown & Dobbins, 2004). The importance of this outcome lies in its potential to interfere with the instructor–student relationship. The suggestion of sexism may be sufficient to taint evaluation of the instructor, decrease interpersonal rapport, and ultimately produce a less positive interaction.

#### Varieties of social identity threat

An overarching framework from which to consider these detrimental consequences is the concept of *social identity threat*: the broad set of concerns that arise when some aspect of the environment signals the danger that a person might be evaluated on the basis of a threatened social identity (Steele et al., 2002; see also Branscombe, 1998). This concept is an extension of the more specific predicament of *stereotype threat*: the concern that one's performance will confirm or be interpreted in light of cultural stereotypes about inferiority in a specific performance domain (Steele et al., 2002). Social identity threat refers to the broader concern that one will be the target not only of specific stereotypes about inferiority, but also a more general hostility or prejudice directed against a salient social identity.

To clarify the distinction between the general phenomenon of social identity threat and the more specific instance

of stereotype threat, consider the case of a man who overhears a suggestion that his female math instructor dislikes men. This suggestion is unlikely to trigger stereotype threat—concern about being the target of stereotypes about inferiority—because math is a domain in which men benefit from positive stereotypes. However, the suggestion of gender bias may nevertheless promote social identity threat to the extent that it triggers the man's concern that he will be the target of hostility or unfair treatment (rather than stereotypes about inferiority).

Besides the distinction between stereotype threat and social identity threat, another important distinction is the difference between local and systemic forms of identity threat. To illustrate this distinction, compare the parallel cases of the woman described in the opening paragraph and the man described in the preceding paragraph. Both the woman and the man may experience a localized form of identity threat associated with the potential for bias in the immediate instruction setting. This localized form of identity threat may be sufficient to cause harm to the potential targets, regardless of whether they are devalued (woman) or relatively advantaged (man) in society at large, because both people may be threatened by the possibility that their respective instructors are hostile towards them. If the detrimental consequences of a suggestion of sexism were limited to such local forms of identity threat, then one would expect men exposed to the suggestion about anti-male sexism from a female instructor to suffer the same harm as women exposed to the suggestion about anti-female sexism from a male instructor.

Without denying the possibility of harm to male targets, we propose that the consequences of a suggestion of sexism will be greater for female targets. This is because women exposed to the suggestion about sexism of a male instructor suffer an additional threat of systemic oppression (e.g., stereotypes about inferiority or pervasive sexual objectification) that men exposed to the suggestion about sexism of a female instructor do not share. This broader threat of systemic oppression gives the same suggestion of bias a more sinister meaning for women than for men (Schmitt, Branscombe, Kobrynowicz, & Owen, 2002). Rather than a stray cloud in an otherwise clear sky, it implies darker clouds looming on the horizon. Applied to the present research, the distinction between local and systemic forms of threat suggests the following hypothesis: Although the suggestion of sexism may be sufficient to cause harm for potential male targets, the consequences will be more harmful for women exposed to the suggestion about sexism of a male instructor than for men exposed to the suggestion about sexism of a female instructor.

### The present research

To investigate whether the mere suggestion of sexism, holding constant the potential for actual discrimination, is sufficient to produce multiple negative outcomes, we conducted three experiments in which we exposed tutorial stu-

dents to the suggestion that their instructor was sexist. Study 1 investigated whether women who were exposed to the suggestion that a male instructor seemed sexist would report a less positive experience of the instruction situation, perform worse on a standardized test, and rate the instructor less competent than would women who were not exposed to this suggestion. Study 2 investigated whether the detrimental effects of the suggestion of sexism are the general consequence of exposure to the suggestion or are instead limited to potential targets of the alleged sexism. Study 3 investigated whether the detrimental effects of the suggestion of sexism derive from the circumscribed predicament that we have referred to as localized identity threat, or instead reflect a more systemic form of identity threat.

### Study 1

We have hypothesized that the mere suggestion of sexism will have harmful consequences for women in an instruction situation. In contrast, previous theory and research have indicated that the suggestion of sexism can sometimes have the relatively beneficial effect of buffering potential targets—and, in particular, their self-esteem—from the sting of potentially prejudiced, negative feedback (Crocker, Voelkl, Testa, & Major, 1991; see also Major, Quinton, & McCoy, 2002; Schmitt & Branscombe, 2002). The present study extends this paradigm beyond self-esteem to consider other important outcomes. Regardless of whether the suggestion of sexism protects self-esteem, we hypothesize that it will have less beneficial consequences for experience of the instruction situation, performance, and evaluations of instructor competence.

### Method

#### Design

We manipulated two factors in a design taken from previous research on the buffering hypothesis (Crocker et al., 1991, Study 1). The first factor was the presence or absence of a suggestion that the instructor might be sexist. The second factor involved a feedback manipulation that framed performance on a difficult task in either a positive or negative light.

#### Participants

A total of 49 women from introductory psychology courses at Stanford University were recruited by telephone to participate in a study about logic tests. The primary criterion for inclusion was that the participant had taken neither the Graduate Record Examination (GRE) and similar tests nor a preparatory course for such tests.

#### Procedure

Each participant reported to an office in the psychology department where a research coordinator asked her to wait, ostensibly because the study was running late. After a few minutes, the coordinator directed the participant to the

testing room and told her to announce her presence to the instructor even though he might be occupied with someone else. When the participant arrived at the appointed room, she found it occupied by a male instructor and a female student who was a confederate of the experimenter. The instructor invited the participant to sit and quickly excused himself, leaving the participant alone with the female confederate.

*Suggestion manipulation.* While the instructor was gone, the confederate had 1 min to enact the suggestion-of-sexism manipulation. In the course of an otherwise unscripted conversation, the confederate informed half of the participants (assigned at random to the suggestion condition) that the study was not so bad, even if the instructor did seem sexist. For the other half of the participants (who comprised the no-suggestion condition) the confederate omitted the "sexist" phrase. It is important to note that the instructor did not hear these comments; consequently, he was unaware of the condition to which a given participant was assigned.

*Feedback manipulation.* The instructor returned to the testing room, dismissed the confederate, and explained to the participant that she would experience one of several types of tutorial, after which she would be asked to attempt a standardized logic test. Before the tutorial, the instructor asked the participant to go to a chalkboard and solve a "diagnostic" item similar to those that appeared on the later test. This chalkboard problem set the stage for the feedback manipulation. After allowing the woman to work for 5 min, the instructor interrupted with his assessment of her performance. Half of the participants were assigned at random to receive standard, relatively positive feedback. The instructor informed these women that, although they could still benefit from the upcoming tutorial, their "pattern of reasoning" was better than most other participants he had seen. The remaining participants received relatively negative feedback. The instructor told these women that they apparently did not have much experience with tests like this and that their pattern of reasoning could benefit from the upcoming tutorial.

*Pretest questionnaire.* As the instructor prepared the tutorial, each participant completed a pretest questionnaire with the first set of dependent measures. These measures included a series of 7-point scale items assessing participants' responses to the instructor and his feedback. First, participants responded to five items—the extent to which the feedback was (a) fair, (b) due to the participant's gender, and (c) due to the instructor's prejudice, and the extent to which participants believed that the instructor was (d) prejudiced and (e) fair—that assessed their perceptions about the relevance of prejudice to the feedback they received. Two additional items instructed participants to rate the appropriateness of the instructor's standard of difficulty and the degree to which the adjective *competent* characterized the instructor. The purpose of the former item was to provide information about the effectiveness of

the feedback manipulation. Participants in the positive feedback condition who received praise for their performance were likely to perceive that the instructor held more lenient standards than did participants in the negative feedback condition who receive criticism for their similar performance. As with other measures, participants received assurances that their responses would be confidential and anonymous.

The pretest questionnaire also included a subset of 16, 5-point scale items from the "Current Thoughts Scale" (CTS; [Heatherton & Polivy, 1991](#)), a 20-item instrument designed to measure state self-esteem. The four deleted items concerned physical appearance and during pretesting tended to arouse undue suspicion about possible deception.

*Tutorial.* Next came the tutorial, designed to be tedious but also a potential boost to performance. It consisted of two example problems which mirrored the first two items of the ensuing logic test and were taken directly from past versions of the Graduate Management Admission Test (GMAT). First was a sequencing problem with items about the relative position of six brands of ant poison along dimensions of sweetness and deadliness. Second was a grouping problem with items about possible arrangements of seven players to form a team of four.

The tutorial proceeded according to a well rehearsed script that provided little opportunity for interaction. The instructor used an overhead projector to demonstrate the solution of the example problems while providing running commentary of the solution process. He was trained to minimize extraneous conversation with the participant and to respond in a businesslike manner to any questions or comments. The length of the tutorial was approximately 10 min.

*Logic test.* The logic test followed immediately after the tutorial. It consisted of four scenarios with six items each, adapted from the analytical sections of past GMAT exams. The first scenario included six items about the relative position of six panda bears along dimensions of fatness and wetness. The second scenario included six items about the possible arrangements of eight players into two teams of four. The third scenario included six items about possible arrangements of 10 marbles of five different colors into a series of five cans. The fourth scenario included six items about possible arrangements of four statues with different characteristics (gender, profession, and potential to glow in the dark) given various constraints on position.

The instructor informed the participant that she had 25 min to complete the 24 items. A normal administration of such a test allows 35 min for completion; consequently, participants were under considerable time pressure about which they were not explicitly warned.

*Posttest questionnaire.* After 25 min, the instructor returned to the testing room, directed the participant to stop work, and again administered a measure instructing the participant to rate the instructor's competence on

a 7-point scale. Because the participants completed a similar item in the pretest questionnaire, this measure permitted an assessment of the possibility that evaluations of competence diverge over time with increasing instructor–student interaction.

Upon completion of the posttest measure, the instructor directed the participant to return to the coordinator's office. The research coordinator interviewed each participant at length regarding her experience in the study and focused especially on any suspicion about the suggestion manipulation. Following this extensive interview, the coordinator debriefed participants, emphasized the rationale for deception, summarized relevant research, shared hypotheses, and addressed any remaining concerns a participant might have about her experience in the study.

### Results

Of the 27 women exposed to the suggestion manipulation, four expressed suspicion about the confederate or the true purpose of the study. In addition, one of the 22 women assigned to the no-suggestion condition had already taken the GRE. We excluded data from these five participants, resulting in a final number of 44 participants. Unless otherwise noted, we analyzed all results with a  $2 \times 2$  (Suggestion  $\times$  Feedback) analysis of variance (ANOVA). Differing degrees of freedom between analyses reflect instances when participants neglected to complete an item.

#### Experience of the instruction situation

Immediately after the feedback manipulation, participants completed five items that measured the extent to which they thought prejudice was relevant to the instruction situation. We computed the mean of each participant's responses for these items to create a composite measure of prejudice perception ( $\alpha = .80$ ). Analysis of this composite item revealed the hypothesized main effect of the suggestion,  $F(1, 38) = 11.64$ ,  $p = .002$ ,  $\eta^2 = .24$ . Women in the suggestion condition perceived greater prejudice ( $M = 3.0$ ,  $SD = 1.08$  on a 7-point scale) than did women in the no-suggestion condition ( $M = 2.04$ ,  $SD = .81$ ). There was no effect of the feedback factor, either as a main effect or in interaction with the suggestion factor (see Table 1).

Besides perceptions of prejudice, participants also indicated perceptions of the instructor's standard of difficulty. Consistent with the formulation of this outcome as a check of the feedback manipulation, analyses revealed only a main effect of feedback,  $F(1, 40) = 6.11$ ,  $p = .02$ ,  $\eta^2 = .13$ . Women who received negative feedback considered the standard of difficulty to be higher ( $M = 4.6$ ,  $SD = .73$  on a 7-point scale) than did women who received positive feedback ( $M = 4.0$ ,  $SD = 1.05$ ). No other main effects or interactions reached conventional levels of significance.

Together, these results suggest that women were sensitive to the feedback manipulation, but this sensitivity was not reflected in perceptions of prejudice. Instead, only the suggestion manipulation influenced women's perceptions of prejudice.

#### Self-esteem

Previous research has focused on the consequences of a suggestion of sexism for self-esteem. To assess this outcome, we averaged across the 16 items from the CTS and created a self-esteem score for each participant ( $\alpha = .90$ ). The  $2 \times 2$  ANOVA for these scores revealed only a main effect of feedback,  $F(1, 40) = 7.36$ ,  $p = .01$ ,  $\eta^2 = .16$ . Regardless of suggestion condition, self-esteem of women who received negative feedback was significantly lower ( $M = 3.39$ ,  $SD = .58$  on a 5-point scale) than that of women who received positive feedback ( $M = 3.82$ ,  $SD = .47$ ). Consistent with previous research, results provided no evidence that the suggestion of sexism harmed self-esteem. However, there was also no evidence that the suggestion of sexism buffered self-esteem from the threatening implications of negative feedback (see Table 1 for condition means).

#### Performance on the logic test

We hypothesized that the suggestion of sexism would have a detrimental effect on test performance. Results supported this hypothesis, revealing only a main effect of the suggestion factor,  $F(1, 40) = 9.2$ ,  $p = .004$ ,  $\eta^2 = .19$ . Regardless of feedback, women in the suggestion condition performed significantly worse on the logic test ( $M = 9.1$ ,  $SD = 2.46$  out of 24 items) than did women in the no-suggestion condition ( $M = 11.6$ ,  $SD = 2.84$ ). Parallel to results

Table 1  
Mean responses for outcome measures in Study 1

Measures	No suggestion		Suggestion	
	Negative feedback	Positive feedback	Negative feedback	Positive feedback
Prejudice in setting	2.10 (0.87)	1.98 (.78)	3.42 (1.17)	2.65 (0.90)
Instructor's standard	4.73 (0.79)	3.90 (1.29)	4.55 (0.69)	4.00 (0.85)
Self-esteem	3.53 (0.61)	3.85 (0.60)	3.24 (0.54)	3.79 (0.37)
Performance	11.27 (2.94)	11.90 (2.85)	8.82 (2.40)	9.42 (2.58)
Instructor competence (pre-test)	5.20 (0.79)	5.90 (1.10)	4.90 (1.20)	5.00 (0.85)
Instructor competence (post-test)	5.90 (0.54)	6.30 (0.95)	5.18 (0.98)	5.42 (1.08)

Note.  $N = 44$  (10–12 per cell). Responses for self-esteem are on a 5-point scale. All other responses are on 7-point scales. Standard deviations appear in parentheses.

for self-esteem, there was no evidence that the suggestion of sexism buffered performance from the threatening implications of negative feedback.

#### *Evaluations of instructor competence*

The final outcome of interest was participants' evaluations of instructor competence. Although our primary hypothesis was that the suggestion would harm evaluations of instructor competence, we measured this outcome at the beginning and end of the procedure to explore the possibility that evaluations of instructor competence would diverge over time. To assess this possibility, we added time of measure (pretest or posttest) as a within-subjects factor to the basic, Feedback  $\times$  Suggestion design and conducted a  $2 \times 2 \times 2$  mixed-design ANOVA on ratings of instructor competence. Results of this analysis provided no indication that evaluations diverged over time (i.e., the Suggestion  $\times$  Time of Measure interaction was not significant,  $F(1,38) = .37$ ,  $p = .55$ .) Instead, results revealed two main effects. One was a main effect of the within-subject variable (Time of Measure),  $F(1,38) = 8.28$ ,  $p = .007$ ,  $\eta^2 = .18$ , indicating that participants rated the instructor as more competent at the end of the study ( $M = 5.69$ ,  $SD = 1.00$ ) than they did following initial feedback ( $M = 5.23$ ,  $SD = 1.03$ ). More important was a main effect of the suggestion factor,  $F(1,38) = 7.51$ ,  $p = .009$ ,  $\eta^2 = .17$ . As hypothesized, women in the suggestion condition ( $M = 5.13$ ,  $SD = .79$ ) rated the instructor as less competent than did women in the no-suggestion condition ( $M = 5.83$ ,  $SD = .86$ ), regardless of the type of feedback they received or when they completed the competence measure. No other effects reached conventional levels of statistical significance ( $ps > .17$ ).

#### *Discussion*

Our hypothesis was that the suggestion of sexism would trigger harmful consequences for women in an instruction situation. We found strong support for this hypothesis. Women exposed to the suggestion reported a less positive experience of the instruction situation (i.e., perceived greater prejudice), performed worse on a standardized test, and rated the instructor as less competent than did women who were not exposed to the suggestion.

The harmful effects of the suggestion of sexism did not extend to self-esteem. This pattern is broadly consistent with previous work investigating why people from stigmatized groups do not internalize cultural representations that portray them in negative light (Crocker & Major, 1989). Results of the present study imply that people can maintain a sense of self-worth, even when the threat of sexist devaluation has detrimental consequences for other important outcomes.

Although results provided no evidence that the suggestion of sexism harmed self-esteem, they also provided no evidence that it buffered self-esteem (or any other outcome) from the impact of negative feedback. Instead, women who

received negative feedback reported equally low self-esteem, regardless of whether or not they were exposed to the suggestion of sexism.<sup>1</sup> Recent research has suggested that, to buffer self-esteem from negative feedback, alleged discrimination should be blatant or unambiguous, with no possibility that the negative feedback might be justified by lack of competence (Major, Kaiser, & McCoy, 2003; Major, Quinton, & Schmader, 2003). It is unlikely that the suggestion of sexism used in the present study met those conditions. Because most participants performed poorly at the chalkboard task, they may have found their negative feedback to be somewhat justified and difficult to discount. In any case, the purpose of the present study was not to provide a definitive test of the buffering hypothesis, but instead to contrast any effects that the suggestion of sexism might have on self-esteem with its harmful consequences for other important outcomes.<sup>2</sup>

#### **Study 2**

We have proposed that the harmful consequences of exposure to a passing suggestion of sexism are the product of women's status as potential targets of the alleged sexism. An alternative possibility is that the consequences have little to do with being a potential target, but instead are more general effects of the suggestion (e.g., negative labeling or distraction-producing strangeness) that would happen to *anyone* who was exposed to it. We investigated this issue in Study 2 by comparing the effects of the identical suggestion of sexism among women and a sample of participants who were not potential targets of the alleged sexism—men. If the detrimental effects of the suggestion of sexism are a product of being a potential target, then one should observe these effects among women but not among men.

#### *Method*

To test this hypothesis, we retained the suggestion manipulation and introduced sex of participant as a factor. Because the feedback manipulation of Study 1 had no effect

<sup>1</sup> The present study may have lacked sufficient power to detect evidence for the buffering hypothesis if the pattern of means had supported the hypothesis. It is therefore important to emphasize that lack of support for the buffering hypothesis is not a question of statistical significance. Instead, the pattern of results actually runs counter to the buffering hypothesis—specifically, exposure to the suggestion in the context of negative feedback resulted in the most harmful outcomes of any condition—a situation that additional power would not change.

<sup>2</sup> It was with this purpose in mind that we based the design of the present study on previous investigations of the buffering hypothesis. However, as an anonymous reviewer noted, the design afforded a plausible, alternative variety of buffering hypothesis: namely, that positive feedback might act as an affirmation that buffered participants from the detrimental effects of the suggestion of sexism (see Schimel, Arndt, Banko, & Cook, 2004). Results of the present study failed to support this variety of buffering hypothesis, too. The detrimental effects of a mere suggestion of sexism were sufficiently strong that they overwhelmed any buffering effect of positive feedback.

on the primary outcomes of interest, we abandoned it in Study 2.

### Participants

We recruited women ( $n = 30$ ) and men ( $n = 29$ ) from two different sources at Stanford University with the same criterion for inclusion as Study 1. One source was a series of introductory psychology courses, and these people received course credit in return for their participation. Another source was a pool of students who received \$7.00 for participation.

### Procedure

The procedure for Study 2 was nearly identical to that of Study 1. Participants, now both men and women, reported to the office of a coordinator who directed them to a testing room where they found a male instructor and a woman who appeared to be the previous participant. As before, the woman was a confederate of the experimenter whose job it was to enact the suggestion of sexism manipulation by informing half of the participants that the instructor seemed sexist.<sup>3</sup> For the remaining participants (the no-suggestion condition), the confederate omitted the “sexist” phrase. As before, the instructor was unaware of the condition to which participants were assigned.

Because we abandoned the feedback manipulation of Study 1, it was no longer necessary to stage a chalkboard problem once the confederate was dismissed. Instead, the instructor began by briefly explaining the study and then conducted the same tutorial as in Study 1.

After the tutorial, the instructor administered a pretest questionnaire. The questionnaire included the same items from the CTS (Heatherton & Polivy, 1991) as the pretest questionnaire in Study 1. Otherwise, changes to the procedure of Study 2 required changes to the remaining pretest measures. First, elimination of the feedback manipulation rendered unnecessary the item about the instructor’s standard of difficulty. Second, the addition of men rendered problematic the items about perceived prejudice. Men expressed undue suspicion about an initial version of the procedure that included questions about the male instructor’s possible prejudice, so we discarded these items from Study 2 and instead asked participants to rate the degree to which the adjectives *comfortable* and *friendly* characterized the instruction situation.

When the participant had completed this first set of measures, the instructor began the same logic test as in Study 1. After 25 min, the instructor asked the participant to stop work and administered the same posttest rating of instructor competence as in Study 1. (Because results of Study 1 provided no evidence that the effects of the suggestion diverged over time, we discarded the pretest rating of instructor competence.) In addition, participants completed

eight items asking them to recall features of the tutorial (e.g., *What color of ink did the instructor use for the second tutorial example?*). These items assessed the extent to which participants attended to the tutorial and thereby provided a criterion for inclusion in data analyses. One man in the suggestion condition failed to recall correctly any of the eight features, so we excluded his data from subsequent analyses. When participants completed the tutorial recall items, the coordinator entered, conducted a thorough debriefing, and carefully probed participants for suspicion.

### Results

The first hypothesis was that, as in Study 1, women exposed to the suggestion of sexism would suffer worse outcomes than women who were not exposed to the suggestion. To test this hypothesis, we conducted an analysis for each outcome (with  $\alpha = .05$ ) that assessed the simple main effect of the suggestion manipulation within women. The second hypothesis was that the harmful effects of the suggestion would not extend to men, who are not potential targets of the alleged sexism. To test this hypothesis, we assessed the interaction in a Sex  $\times$  Suggestion ANOVA. Despite this conceptual order, we follow convention by reporting results of ANOVA first, followed by the test assessing the simple main effect of the suggestion within women. For sake of comparison (rather than a test of hypotheses), we also report the parallel analysis assessing the simple main effect of the suggestion within men. With the exception of the 5-point self-esteem measures, all measures were based on 7-point Likert-type scales.

#### Experience of the instruction situation

Immediately after the tutorial, participants rated the friendliness and comfort of the instruction situation. Because these two items were correlated,  $r(58) = .59$ ,  $p < .0001$ , we combined them to form a composite measure. The ANOVA for this measure revealed a main effect of sex,  $F(1, 55) = 10.512$ ,  $p = .002$ ,  $\eta^2 = .16$ , such that women ( $M = 4.15$ ,  $SD = 1.23$ ) reported a less positive experience than did men ( $M = 5.03$ ,  $SD = .92$ ). This main effect was qualified by the hypothesized Sex  $\times$  Suggestion interaction,  $F(1, 55) = 5.17$ ,  $p = .03$ ,  $\eta^2 = .09$ , indicating that the identical suggestion of sexism had different effects among women and men. Analyses of simple main effects indicated that women in the suggestion condition reported a less positive experience than women in the no-suggestion condition,  $F(1, 55) = 3.95$ ,  $p = .05$ ,  $\eta^2 = .07$ . In contrast, men reported a similar experience across conditions  $F(1, 55) = 1.34$ ,  $p = .25$ ,  $\eta^2 = .03$  (see Table 2 for means).

#### Self-esteem

As in Study 1, there was no evidence that the suggestion of sexism harmed self-esteem. The ANOVA for the self-esteem measure revealed no reliable effects,  $ps > .13$  (see Table 2 for means).

<sup>3</sup> Participants’ remarks during the debriefing interview indicated that men understood the confederate’s statement to mean that the instructor was sexist against women.

Table 2  
Mean responses for outcome measures in Study 2

Measures	Women		Men	
	No suggestion	Suggestion	No suggestion	Suggestion
Experience of setting	4.53 (1.04)	3.77 (1.32)	4.80 (0.88)	5.33 (0.90)
Self-esteem	3.90 (0.64)	3.87 (0.54)	4.11 (0.41)	4.06 (0.59)
Performance	11.80 (2.88)	9.53 (2.90)	10.40 (3.44)	12.93 (4.56)
Instructor competence	6.07 (0.80)	5.00 (1.07)	5.87 (0.99)	5.80 (0.86)

Note.  $N = 59$  (14–15 per cell). Responses for self-esteem are on a 5-point scale. All other responses are on 7-point scales. Standard deviations appear in parentheses.

### Performance on the logic test

The ANOVA for performance revealed only the hypothesized Sex  $\times$  Suggestion interaction,  $F(1, 55) = 7.20$ ,  $p = .01$ ,  $\eta^2 = .12$ , indicating that the identical suggestion of sexism had different effects within sex. Analyses of simple effects indicated a marginal tendency for women in the suggestion condition to perform worse than women in the no-suggestion condition,  $F(1, 55) = 3.19$ ,  $p = .08$ ,  $\eta^2 = .05$ . In contrast, men in the suggestion condition performed better than men in the no-suggestion condition,  $F(1, 55) = 4.03$ ,  $p = .05$ ,  $\eta^2 = .07$  (see Table 2 for means).

### Evaluations of instructor competence

The ANOVA for posttest ratings of instructor competence revealed a significant main effect of suggestion,  $F(1, 55) = 4.85$ ,  $p = .03$ ,  $\eta^2 = .08$ . This effect was qualified by the hypothesized Sex  $\times$  Suggestion interaction,  $F(1, 55) = 4.68$ ,  $p = .04$ ,  $\eta^2 = .08$ , indicating that the same suggestion of sexism had different effects among women and men. Analyses of simple main effects indicated that women in the suggestion condition rated the instructor to be less competent than did women in the no-suggestion condition,  $F(1, 55) = 9.70$ ,  $p = .003$ ,  $\eta^2 = .15$ . In contrast, men across conditions did not differ in ratings of instructor competence,  $F(1, 55) = 0$  (see Table 2 for means).

### Discussion

Study 2 replicated results of Study 1, showing again that the suggestion of sexism was sufficient to produce harmful consequences for women in an instruction situation. Women exposed to the suggestion reported a less positive experience, tended to perform worse on the logic test, and rated the instructor as less competent than did women who were not exposed to this suggestion. In addition, Study 2 extended results of Study 1 by showing that the identical suggestion of sexism did not produce similarly harmful consequences for men. Framed in general terms, the harmful effects of the suggestion were limited to potential targets of the alleged sexism and did not extend to participants who were not potential targets.

Although we hypothesized that the harmful effects of the suggestion of sexism on women's outcomes would not extend to men, we did not anticipate that it would promote positive outcomes. Yet, men who were exposed to the sug-

gestion of sexism actually performed *better* than men who were not exposed to the suggestion. This pattern is consistent with the phenomenon of *stereotype lift* (Walton & Cohen, 2003), in which the same laboratory conditions that harm performance of people from disadvantaged or negatively stereotyped groups boost performance of people from advantaged or positively stereotyped groups. We return to this pattern after presenting results of Study 3.

### Study 3

Results of Study 2 indicate that the suggestion of sexism did not cause harm for men when they were not potential targets of the alleged sexism. However, the question remains: What happens when men *are* potential targets? Here, the distinction between local and systemic forms of identity threat becomes important. This distinction starts with the observation that anyone—man or woman—may experience a localized form of gender-based hostility in a given instruction setting. If the harmful effects of the suggestion that we observed in previous studies are solely the result of this localized identity threat, then one should observe the similarly harmful consequences for men who are exposed to the suggestion about anti-male sexism of a female instructor.

In contrast, we propose that the harmful effects of the suggestion observed in the previous studies are partly the consequence of a more systemic form of identity threat related to concerns about hostility and acceptance that extend beyond the immediate testing situation. This additional set of concerns gives the parallel suggestion of sexism a more sinister meaning for women than for men (see also Schmitt & Branscombe, 2002; Schmitt, Branscombe, & Postmes, 2003). Because women must contend with an additional set of systemic threats that men do not share, we hypothesize that exposure to parallel suggestion of sexism will have more harmful consequences for potential female targets than for potential male targets. We tested this hypothesis in Study 3.

Besides enabling a direct comparison of the suggestion's effects on female and male targets, the case of male targets also provided a unique opportunity to compare the costs of local disadvantage with the benefits of systemic advantage. To the extent that men are potential targets of alleged bias, they may suffer from local forms of identity threat that are



Table 3  
Means responses for outcome measures in Study 3

Measures	Women		Men	
	No suggestion	Suggestion (sexist male)	No suggestion	Suggestion (sexist female)
Experience of setting	4.50 (1.04)	3.64 (1.39)	5.31 (0.96)	4.56 (1.06)
Self-esteem	3.58 (0.90)	4.02 (0.80)	4.20 (0.45)	4.10 (0.63)
Performance	9.56 (2.99)	7.45 (3.25)	6.94 (2.61)	8.00 (2.91)
Instructor competence	5.33 (0.97)	4.83 (1.79)	5.17 (1.15)	5.50 (0.86)

Note.  $N = 72$  (18 per cell). Responses for self-esteem are on a 5-point scale. All other responses are on 7-point scales. Standard deviations appear in parentheses.

sufficient to produce harm. Simultaneously, the suggestion of sexism may also trigger broader associations to systems of meaning in which men are advantaged relative to women. These associations may provide symbolic resources (e.g., positive stereotypes) that not only mitigate harmful effects of local disadvantage, but also promote positive outcomes.

### Method

#### Participants

Women ( $n = 39$ ) and men ( $n = 38$ ) from an introductory psychology course at the University of Kansas participated in the study in exchange for course credit. Participants had taken neither a GRE-type test nor a preparatory course for such a test.

#### Procedure

The procedure was identical to Study 2, with one exception. Rather than a female confederate who suggested that a male instructor was sexist against women, male participants now met a male confederate who suggested (or did not suggest) that a female instructor was sexist against men.

### Results

As in previous studies, the first hypothesis was that women exposed to the suggestion of sexism would suffer worse outcomes than women who were not exposed to the suggestion. To test this hypothesis, we conducted analyses for each outcome (with  $\alpha = .05$ ) that assessed the simple main effect of the suggestion within women. In addition, the design of Study 3 permitted an investigation of two different hypotheses regarding the source of the suggestion's harm. If the harmful consequences of the suggestion reflect experience of local identity threat, then one would expect parallel suggestions of sexism to have similar effects for potential male and female targets (i.e., a main effect of the suggestion manipulation in a Sex  $\times$  Suggestion ANOVA). In contrast, if the harmful consequences of the suggestion reflect experience of systemic identity threat, then one would expect parallel suggestions of sexism to have different effects for potential male and female targets (i.e., the interaction effect in a Sex  $\times$  Suggestion ANOVA).

As in Study 2, we conform to convention by reporting results of ANOVA first, followed by the planned test assessing the simple main effect of the suggestion within

women. For sake of comparison, we again report the parallel analysis assessing the simple main effect of the suggestion within men. With the exception of the 5-point self-esteem measures, all measures were based on 7-point Likert-type scales. As before, we retained only participants who correctly answered at least one of the eight tutorial-recall questions. We excluded five participants (three women in the no-suggestion condition and one man in each condition) who failed this criterion.

#### Experience of the instruction situation

As in Study 2, we again computed the mean of each participant's ratings of *friendly* and *comfortable* to create a single, composite measure of comfort with the instruction situation,  $r(72) = .41$ ,  $p < .001$ . The ANOVA for this measure again revealed a main effect of sex,  $F(1, 68) = 10.56$ ,  $p = .002$ ,  $\eta^2 = .13$  and an additional, main effect of the suggestion factor,  $F(1, 68) = 9.25$ ,  $p = .003$ ,  $\eta^2 = .12$ . The first effect indicates again that, whether or not the suggestion was present, women ( $M = 4.07$ ,  $SD = 1.29$ ) rated the situation less positively than did men ( $M = 4.93$ ,  $SD = 1.06$ ). The second effect indicates that, regardless of sex, participants in the suggestion condition ( $M = 4.10$ ,  $SD = 1.07$ ) rated the situation less positively than did those in the no-suggestion condition ( $M = 4.90$ ,  $SD = 1.30$ ). The Sex  $\times$  Suggestion interaction was not significant,  $F(1, 68) = 0.04$ ,  $p = .84$ . Instead, the combination of main effects meant that women exposed to the suggestion suffered a particularly negative experience (See Table 3 for means).

#### Self-esteem

As in previous studies, there was no evidence that the suggestion of sexism harmed self-esteem. The ANOVA revealed only a main effect of sex, indicating that men scored higher on the measure than did women,  $F(1, 68) = 3.87$ ,  $p = .05$ ,  $\eta^2 = .05$ . (see Table 3 for means).<sup>4</sup>

<sup>4</sup> In this case, the Sex  $\times$  Suggestion interaction approached conventional levels of statistical significance,  $F(1, 68) = 2.43$ ,  $p = .12$ ,  $\eta^2 = .03$ . This pattern constitutes the only evidence from the present studies in support of the buffering hypothesis (Crocker & Major, 1989); specifically, the simple effect of the suggestion within women indicated a marginal tendency for women in the suggestion condition to report *greater* self-esteem than women in the no-suggestion condition,  $F(1, 68) = 3.19$ ,  $p = .08$ ,  $\eta^2 = .04$ . However, the pattern also illustrates the larger theme of the present research. Even when the suggestion of sexism has the relatively beneficial consequence of buffering women's self-esteem, it has harmful consequences for other important outcomes.

### Performance on the logic test

The ANOVA for test performance revealed only the hypothesized Sex  $\times$  Suggestion interaction,  $F(1,68)=4.49$ ,  $p=.04$ ,  $\eta^2=.06$ , indicating that parallel suggestions of sexism had different effects among women and men. Analyses of simple main effects indicated that women in the suggestion condition performed significantly worse than did women in the no-suggestion condition,  $F(1,68)=4.21$ ,  $p=.04$ ,  $\eta^2=.06$ . In contrast, men's performance did not differ significantly as a function of suggestion condition,  $F(1,68)=.89$ ,  $p=.35$  (see Table 3 for means).

### Evaluations of instructor competence

Mean scores for the last outcome of interest—posttest ratings of instructor competence—showed a similar pattern as in Study 2 (see Table 3). Women in the suggestion condition tended to rate the instructor as less competent than did participants in the other three conditions. However, the corresponding, Sex  $\times$  Suggestion interaction for this measure was not significant,  $F(1,68)=2.01$ ,  $p=.16$  (nor did either main effect approach conventional levels of statistical significance). Likewise, the planned test of the simple effect of the suggestion within women indicated that the tendency for women in the suggestion condition to evaluate the instructor less positively was not significant,  $F(1,68)=1.45$ ,  $p=.23$ ,  $\eta^2=.02$ . The parallel analysis indicated no tendency for men in the suggestion condition to evaluate the instructor any differently than did men in the no-suggestion condition,  $F(1,68)=0.64$ ,  $p=.43$ ,  $\eta^2=.01$ .

### Discussion

Results of Study 3 replicate the pattern of previous studies in an independent sample of women at a different university. Women exposed to the suggestion of sexism again reported a less positive experience and performed worse on a logic test than did women who were not exposed to this suggestion. Results for the third outcome of interest—evaluations of the instructor—revealed a similar pattern, although the difference in this case was not statistically significant.

In contrast to its uniformly harmful consequences for potential female targets, exposure to the parallel suggestion of sexism did not have uniformly harmful consequences for potential male targets. There was some evidence that men exposed to the suggestion had some experience of threat; specifically, they rated their experience of the tutorial setting to be less positive than did men who were not exposed to the suggestion. However, there was no evidence that the suggestion or its implication of threat harmed men's performance or evaluations of the instructor.

### Analyses across studies

The preceding studies provide consistent evidence that the mere suggestion of sexism can be sufficient to harm women's experience of an instruction situation. To evaluate

the consistency of this evidence in a statistically rigorous fashion, we examined the detrimental effects of the suggestion on women's outcomes meta-analytically across studies.<sup>5</sup> Preliminary analyses gave no indication that effect sizes varied across studies, so we proceeded with the meta-analyses. Results provide strong evidence that the suggestion of sexism had the hypothesized negative effects. Across all three studies, women exposed to the suggestion reported a less positive experience (weighted mean  $d=1.00$ ,  $z=3.95$ ,  $p<.0001$ ), performed worse on the logic test (weighted mean  $d=0.92$ ,  $z=3.95$ ,  $p<.0001$ ), and rated the instructor as less competent (weighted mean  $d=0.41$ ,  $z=2.88$ ,  $p=.002$ ) than did women who were not exposed to the suggestion.

Meta-analysis was inappropriate for examining effects of the suggestion on men because the meaning of the manipulation differed across the two studies in which men participated. However, this between-study difference afforded an alternative,  $2 \times 2$  (Study  $\times$  Suggestion) ANOVA that helps to clarify key concepts. Along one dimension, the study factor constitutes a manipulation of local identity threat: men's concern that they will be the target of hostility or bias in a discrete situation. Men in Study 2 overheard a suggestion about gender bias of a male instructor for which they were not potential targets (and therefore were *not* exposed to local identity threat), but men in Study 3 overheard a suggestion about gender bias of a female instructor for which they were potential targets (and therefore *were* exposed to local identity threat). Along the other dimension, the suggestion factor constitutes a manipulation of men's systemic advantage, which was more salient in the suggestion condition than in the no-suggestion condition. Of course, one must interpret this analysis with caution because other factors besides threat—for example, university setting and instructor gender—also varied across studies. Accordingly, we present the analysis not as a formal test of hypotheses, but instead as a way to organize discussion of results.

Beginning with experience of the instruction situation, the ANOVA revealed a significant Study  $\times$  Suggestion interaction,  $F(1,62)=8.13$ ,  $p=.006$ ,  $\eta^2=.12$ . Tests of simple main effects indicate that, in Study 2, men exposed to the suggestion about sexism of a male instructor (for which they were *not* potential targets) displayed a non-significant tendency to report a *more* positive experience of the instruction situation than did men in the no-suggestion

<sup>5</sup> The justification for the meta-analysis is enhanced by the fact that, with one exception, measures of the primary outcomes were identical across studies. The exception is experience of the instruction situation, which we measured via prejudice composite in Study 1, but via a combination of *comfortable* and *friendly* in Studies 2 and 3. Results reported here include prejudice-composite data from Study 1, reverse scored so that higher numbers signify a more positive experience (i.e., perception of less prejudice). However, conclusions regarding experience of the instruction setting do not change when one conducts the meta-analysis for this outcome using only data from Studies 2 and 3 (weighted mean  $d=1.00$ ,  $z=2.63$ ,  $p<.01$ ).

condition,  $F(1,62) = 3.07$ ,  $p = .085$ ,  $\eta^2 = .05$  (see Table 2). In Study 3, men exposed to the suggestion about sexism of a female instructor (for which they *were* potential targets) reported a *less* positive experience of the instruction situation than did men in the no-suggestion condition,  $F(1,62) = 5.34$ ,  $p = .024$ ,  $\eta^2 = .08$  (see Table 3). This pattern implies that the experience of localized threat is sufficient to produce discomfort among men, even when they benefit from systemic advantage.

Turning to the outcome of performance, the ANOVA revealed orthogonal main effects. A main effect of study indicated that, regardless of suggestion condition, Stanford men in Study 2 ( $M = 11.67$ ,  $SD = 4.17$ ) performed better than Kansas men in Study 3 ( $M = 7.53$ ,  $SD = 2.78$ ),  $F(1,62) = 24.16$ ,  $p = .001$ ,  $\eta^2 = .28$ . More important, a main effect of the suggestion manipulation indicated that, regardless of study, men who were exposed to the suggestion ( $M = 10.24$ ,  $SD = 4.45$ ) performed better than men who were not exposed to the suggestion ( $M = 8.58$ ,  $SD = 3.43$ ),  $F(1,62) = 4.26$ ,  $p = .04$ ,  $\eta^2 = .06$ . The Study  $\times$  Suggestion interaction was not significant,  $F(1,62) = .89$ ,  $p = .35$ . This pattern implies that the salience of systemic advantage provides a boost to men's performance in a way that is not moderated by the presence of localized threat.<sup>6</sup>

This pattern is consistent with the phenomenon of stereotype lift (Walton & Cohen, 2003) and with research showing that men, unlike women, are not harmed by knowledge that they have been selected for a task on the basis of gender rather than merit (Heilman, Lucas, & Kaplow, 1990; Heilman, Simon, & Repper, 1987; Major, Feinstein, & Crocker, 1994). The present work extends previous research by pitting the potential benefits of systemic advantage (e.g., stereotype lift) against the potential costs of local disadvantage (e.g., local identity threat). Results suggest that—at least for test performance in our laboratory paradigm—the beneficial effects of men's systemic advantage can be enough to offset harmful effects of local disadvantage.

## General discussion

The suggestion of sexism functions on the conceptual level as a manipulation of social identity threat: the concern that one might be a target not only of specific stereotypes about inferiority, but also a more general hostility based (in this case) on one's gender identity. Results of the present research support the hypothesis that this threat is sufficient to trigger harmful consequences for women in an instruction situation, even in the absence of any predisposition

toward discrimination on the part of the instructor. Women who were exposed to the suggestion of sexism reported less positive experience, performed worse on a standardized test, and rated the instructor as less competent than did women who were not exposed to the suggestion.

Results also support the hypothesis that the suggestion of sexism does not have identically harmful consequences for men. Our account of these differential consequences emphasizes a distinction between local and systemic varieties of identity threat.

On one hand, the suggestion of sexism constitutes a local form of identity threat that can be sufficient to harm experience of anyone who is a potential target, whether relatively devalued (i.e., women) or advantaged (men) in the surrounding environment. For example, exposure to the suggestion about sexism of a female instructor was sufficient to harm experience of the instruction setting for men in Study 3, even though they are typically not the targets of either stereotypes about inferiority or broad devaluation in the university settings in which we conducted this research.

On the other hand, the suggestion of sexism has associations to broader systems of devaluation and advantage that give the parallel instance more sinister implications for women than for men. For men, the threat associated with the suggestion is limited to the immediate situation and has few implications for life outside the instruction setting. Indeed, the suggestion may even trigger associations to men's relatively advantaged status and promote the phenomenon of stereotype lift. In contrast, the threat associated with the suggestion of sexism is not limited to the immediate situation for women, but implies a more chronic set of concerns that pervade professional and academic life. These broader concerns constitute a more systemic form of identity threat that men do not share. This threat weighs heavily on experience of the instruction setting, triggering the differential consequences for men and women that we observed in the present research.

## Limitations and future directions

Although the present results are generally consistent with this account, there may be other differences between women and men—for example, gender-based dispositions concerning sensitivity to threat—that might also underlie the observed differences. Rather than deny these possibilities, we simply note that they are not incompatible with the present account. A lifetime of engagement with gendered academic environments may shape women to acquire a dispositional vigilance for the possibility of discrimination (e.g., stigma consciousness; see Pinel, 1999). Although this dispositional vigilance is an adaptive strategy for life in a general environment of discrimination, it can have harmful consequences when applied to specific situations from which discrimination is initially absent.

In any case—whether temporary responses to a instance of identity threat or dispositional tendencies shaped by chronic experience of identity threat—our preferred

<sup>6</sup> The ANOVA for instructor competence revealed only a main effect of Study,  $F(1,62) = 4.29$ ,  $p = .04$ ,  $\eta^2 = .06$ . Regardless of suggestion condition, Stanford men in Study 2 ( $M = 5.83$ ,  $SD = .91$ ) rated the male instructor as more competent than Kansas men in Study 3 ( $M = 5.33$ ,  $SD = 1.01$ ) rated the female instructor. Whether this pattern reflects derogation of the female instructor or another factor that varies across studies is a question that the present analysis cannot answer.

account of observed results invokes the situationist worldview of social psychology (Ross & Nisbett, 1991). Rather than blame the victim and attribute observed harm to something inherent in the nature of women, this situationist worldview locates the source of observed results in the different worlds that men and women typically inhabit. Conclusive evidence for this account awaits additional research.

#### *Stereotype threat or social identity threat*

A second issue that remains from these studies is whether the harmful effects on women's experience are reducible to a specific concern about stereotypical beliefs (i.e., stereotype threat) or instead reflect a more general concern about hostility and prejudicial treatment (i.e., social identity threat). We designed the present studies to investigate the latter rather than the former. Studies of stereotype threat typically use manipulations that focus on the salience of social categories or beliefs about inferiority that are attached to these categories. In contrast, we designed the present manipulation (i.e., the suggestion of sexism) to invoke the broader experience of oppression in which the potential for hostility and prejudicial treatment were at least as relevant as stereotypical beliefs about inferiority. Likewise, studies of stereotype threat typically focus on the outcome of performance in relatively sanitized testing situations that are insulated from interpersonal interaction. In contrast, the present studies focus on a broader range of outcomes in the context of a prolonged interaction that renders the threat of hostility or prejudicial treatment a more relevant concern.

The distinction between social identity threat and the more specific phenomenon of stereotype threat is perhaps most clear in Study 3, where men who were exposed to the suggestion about the anti-male sexism of a female instructor reported a less positive experience than did men who were not exposed to this suggestion. It seems unlikely that this result is the product of stereotype threat, because the relevant performance domain (logic tests) is one for which men benefit from positive stereotypes. Indeed, as a possible reflection of these positive stereotypes, the analyses across studies indicate that men who were exposed to a suggestion of sexism generally performed better than men who were not exposed to the suggestion, even when (as in Study 3) they were potential targets of the alleged sexism. Instead, a more plausible source of the suggestion's detrimental effects is the general concern about hostility and prejudicial treatment that we have referred to as social identity threat. This concern was apparently sufficient to produce a less positive experience, despite the positively stereotyped nature of the performance domain.

Although the design and results of the present studies are consistent with a distinction between a broad category of social identity threat and the specific phenomenon of stereotype threat, a conclusive demonstration of this distinction awaits further research. One way to distinguish between these phenomena is to investigate the effects of the

suggestion of sexism in instruction domains for which women do not suffer from cultural stereotypes about inferiority (e.g., verbal skill). If the suggestion of sexism produces similar detrimental consequences in a non-stereotyped domain, then one can attribute these consequences to a more general phenomenon of social identity threat rather than the specific phenomenon of stereotype threat.

#### *Mediators of harm to performance*

A third issue that remains from the present research concerns potential mediators of the suggestion's harmful consequences, especially for test performance. Although the present studies did not explore this issue, they do provide relevant data. For example, one might hypothesize that the suggestion harms test performance by damaging women's self-esteem. Indeed, the fact that previous research has focused on the outcome of self-esteem may reflect the prevalent assumption in North American settings about the importance of self-esteem for a variety of outcomes, including test performance. However, we observed no evidence that self-esteem mediated the effect of the suggestion manipulation on performance, because the suggestion had no effects on self-esteem.<sup>7</sup>

One contribution of the present paradigm is to expand empirical attention beyond the moment of test performance to investigate the effects of the suggestion on the broader instruction process. Besides potential mediators of performance decrements that operate at the moment of performance, the detrimental effects of the suggestion of sexism might also be mediated by processes—like distraction or disengagement—that operate during moments of learning and instruction. By including the instruction phase in the research design, the present paradigm provides opportunities for future research to assess these potential mediators.

#### *Suggestion of sexism as chilly climate: Implications for motivation*

The present research emphasizes the consequences of the suggestion for experience of the instruction situation, not as a potential mediator of performance, but instead as an important outcome in its own right. Results indicate that the suggestion of sexism is sufficient to transform the situation into a "chilly climate" (Hall & Sandler, 1982) that women experience as less comfortable than when the suggestion is absent. An important direction for future research is to consider the consequences of this chilly climate for both short-term disengagement from the instruction setting and long-term disidentification from the

<sup>7</sup> We continue to address the question of mediation in our ongoing research. To date, this research provides no evidence that the suggestion of sexism has any effect on such plausible mediators of women's performance decrements as decreased confidence, lowered performance expectations, or other self-beliefs. The research has provided some evidence that the suggestion of sexism does increase such plausible mediators as test anxiety, evaluation apprehension, and self-reports of effort withdrawal. However, it has provided no evidence that increases in these potential mediators are reliably associated with decreases in women's test performance.

instruction domain. One can hypothesize that the negative consequences of the suggestion will differ for men and women as a function of local and systemic forms of identity threat. For women, the pervasive reach of systemic threat may promote disengagement, disidentification, or something akin to learned helplessness in response to the suggestion of sexism. The same harmful consequences may not extend to men, for whom the parallel suggestion about sexism of a female instructor may trigger not only a more limited threat, but also symbolic resources (associated with systemic advantage) that afford men greater confidence, agency, or motivation in response to this threat.

#### *Evaluations of instructor competence: Implications for the student–instructor relationship*

Likewise, the present research extends the focus of investigation beyond performance to impressions of instructor competence. The meta-analysis indicated that, across studies, women who were exposed to the suggestion of sexism rated the instructor to be less competent than did women who were not exposed to the suggestion. Although this pattern may be a result of biased perception (i.e., perceiving equally competent performance as less competent), it may also reflect a self-fulfilling prophecy whereby the suggestion of sexism, mediated through women's responses, actually led the instructor to perform less competently.

The possibility has important implications for interpretation of results. First, this possibility qualifies claims about the absence of discrimination from the present paradigm. Because the suggestion of sexism may trigger a self-fulfilling prophecy that elicits substandard treatment from the instructor, it is inappropriate to claim that the negative consequences of the suggestion happen in the absence of discrimination. Instead, we emphasize that any differential treatment that does result is itself a product of the suggestion that happens regardless of any predisposition for discrimination on the part of the instructor. Second, this possibility suggests yet another potential mediator of women's performance decrements. Women who are exposed to the suggestion of sexism may perform worse on the logic test, not because the suggestion damages their test-taking psyche, but instead because they receive less competent instruction from the male instructor.

#### *Implications of the target's perspective for conceptions of discrimination*

One of the most important insights to emerge from a growing body of research “from the target's perspective” (e.g., Oyserman & Swim, 2001) is that the detrimental consequences of systemic devaluation are not limited to direct acts of differential treatment by biased perpetrators (see Brown et al., 2003). Instead, the historical legacy of systemic devaluation constitutes a form of social identity threat that can be sufficient to cause harm, even in the absence of differential treatment. To distinguish these harmful outcomes from prevailing conceptions of discrimi-

nation (as differential treatment), we refer to them as “indirect” consequences of systemic oppression.

To illuminate the experience of potential targets, we conducted extensive debriefing interviews with participants. Although one must interpret these interviews with caution, we discuss them briefly to illustrate the indirect consequences of systemic oppression. When asked if there was anything unusual about the study, many women in the suggestion condition spontaneously mentioned the other student's remark that the instructor seemed sexist. However, almost all women emphasized that they did not allow themselves to be swayed by this remark; instead, they claimed to examine the instructor's behavior and to conclude on this basis that the accusation was unwarranted. Similarly, almost all women denied that this remark had any harmful effect on their outcomes. In fact, many women added that, if they had suspected sexism, they would have tried harder in an attempt to prove the instructor wrong. These comments assume an ironic quality when compared with measures taken during the experiment, which show that these women's experience, test performance, and impressions of the instructor would have been better if the suggestion were absent.

In the absence of differential treatment as a ready explanation, it is likely that women (and other observers) interpret their outcomes as a reflection of their dispositions—perhaps low ability, disinterest, or cold demeanor—rather than consequences of a situation tainted by the suggestion of sexism. In contrast, the idea of social identity threat reflects an attempt to extend conceptions of discrimination to incorporate these harmful consequences and correctly locate their ultimate source in systemic oppression. By emphasizing relatively indirect manifestations of systematic oppression, we do not mean to downplay the continued relevance of the direct, discriminatory treatment that women face on an everyday basis in situations outside the laboratory. Instead, we take this background differential treatment as a baseline and investigate the additional negative outcomes—beyond any structural disadvantage or instructor prejudices—that accrue to women as the result of a passing suggestion of sexism.

## References

- Branscombe, N. R. (1998). Thinking about one's gender group's privileges or disadvantages: Consequences for well-being in women and men. *British Journal of Social Psychology*, 37, 167–184.
- Brown, L. M., & Dobbins, H. (2004). Students' of color and European American students' stigma-relevant perceptions of university instructors. *Journal of Social Issues*, 60, 157–174.
- Brown, M. K., Carnoy, M., Currie, E., Duster, T., Oppenheimer, D. B., Shultz, M. M., et al. (2003). *Whitewashing race: The myth of a color-blind society*. Berkeley, CA: University of California Press.
- Constantinople, A., Cornelius, R., & Gray, J. (1988). The chilly climate: Fact or artifact? *Journal of Higher Education*, 59, 527–550.
- Crocker, J., & Major, B. (1989). Social stigma and self-esteem: The self-protective properties of stigma. *Psychological Review*, 96, 608–630.
- Crocker, J., Voelkl, K., Testa, M., & Major, B. (1991). Social stigma: The affective consequences of attributional ambiguity. *Journal of Personality and Social Psychology*, 60, 218–228.

- Eccles, J. S., Jacobs, J. E., & Harold, R. E. (1990). Gender role stereotypes, expectancy effects, and parents' socialization of gender differences. *Journal of Social Issues, 46*, 183–201.
- Frable, D., Blackstone, T., & Scherbaum, C. (1990). Marginal and mindful: Deviants in social interaction. *Journal of Personality and Social Psychology, 59*, 140–149.
- Geen, R. G. (1991). Social motivation. *Annual Review of Psychology, 42*, 377–399.
- Hall, R. M., & Sandler, B. R. (1982). *The campus climate: A chilly one for women? (Report of the Project on the Status and Education of Women)*. Washington, DC: Association of American Colleges.
- Heatherton, T. F., & Polivy, J. (1991). Development and validation of a scale for measuring state self-esteem. *Journal of Personality and Social Psychology, 60*, 895–910.
- Heilman, M. E., Lucas, J. A., & Kaplow, S. R. (1990). Self-derogating consequences of sex-based preferential selection: The moderating role of initial self-confidence. *Organizational Behavior and Human Decision Processes, 46*, 202–216.
- Heilman, M. E., Simon, M. C., & Repper, D. P. (1987). Intentionally favored, unintentionally harmed? Impact of sex-based preferential selection on self-perceptions and self-evaluations. *Journal of Applied Psychology, 72*, 62–68.
- Hyde, J. S., & Kling, K. C. (2001). Women, motivation, and achievement. *Psychology of Women Quarterly, 25*, 364–378.
- 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*, 365–370.
- Lord, C. G., & Saenz, D. S. (1985). Memory deficits and memory surfeits: Differential cognitive consequences of tokenism for tokens and observers. *Journal of Personality and Social Psychology, 49*, 918–926.
- Major, B., Feinstein, J., & Crocker, J. (1994). Attributional ambiguity of affirmative action. *Basic and Applied Social Psychology, 15*, 113–141.
- Major, B., Kaiser, C. R., & McCoy, S. K. (2003). It's not my fault: When and why attributions to prejudice protect self-esteem. *Personality and Social Psychology Bulletin, 29*, 772–781.
- Major, B., Quinton, W. J., & McCoy, S. K. (2002). Antecedents and consequences of attributions to discrimination: Theoretical and empirical advances. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 34, pp. 251–329). New York: Academic Press.
- Major, B., Quinton, W. J., & Schmader, T. (2003). Attributions to discrimination and self-esteem: Impact of group identification and situational ambiguity. *Journal of Experimental Social Psychology, 39*, 213–220.
- Major, B., & Schmader, T. (1998). Coping with stigma through psychological disengagement. In J. K. Swim & C. Stangor (Eds.), *Prejudice: The target's perspective* (pp. 219–241). San Diego, CA: Academic Press.
- 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*, 34–50.
- O'Brien, L. T., & Crandall, C. S. (2003). Stereotype threat and arousal: Effects on women's math performance. *Personality and Social Psychology Bulletin, 29*, 782–789.
- Oyserman, D., & Swim, J. K. (Eds.). (2001). Stigma: An insider's perspective. *Journal of Social Issues, 57*.
- Pinel, E. C. (1999). Stigma consciousness: The psychological legacy of social stereotypes. *Journal of Personality and Social Psychology, 76*, 114–128.
- Pratto, F., & John, O. P. (1991). Automatic vigilance: The attention-grabbing power of negative social information. *Journal of Personality and Social Psychology, 61*, 380–391.
- Ross, L., & Nisbett, R. E. (1991). *The person and the situation: Perspectives of social psychology*. New York, NY: McGraw-Hill Book Company.
- Schimmel, J., Arndt, J., Banko, K. M., & Cook, A. (2004). Not all self-affirmations were created equal: The cognitive and social benefit of affirming the intrinsic (vs extrinsic) self. *Social Cognition, 22*, 75–99.
- Schmader, T., Major, B., & Gramzow, R. H. (2001). Coping with ethnic stereotypes in the academic domain: Perceived injustice and psychological disengagement. *Journal of Social Issues, 57*, 93–111.
- Schmitt, M. T., & Branscombe, N. R. (2002). The meaning and consequences of perceived discrimination in disadvantaged and privileged social groups. *European Review of Social Psychology, 12*, 167–199.
- Schmitt, M. T., Branscombe, N. R., Kobrynowicz, D., & Owen, S. (2002). Perceiving discrimination against one's gender group has different implications for well-being in women and men. *Personality and Social Psychology Bulletin, 28*, 197–210.
- Schmitt, M. T., Branscombe, N. R., & Postmes, T. (2003). Women's emotional responses to the pervasiveness of gender discrimination. *European Journal of Social Psychology, 33*, 297–312.
- 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.
- Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist, 52*, 613–629.
- Steele, C. M., Spencer, S. J., & Aronson, J. (2002). Contending with group image: The psychology of stereotype and social identity threat. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 34, pp. 379–439). New York: Academic Press.
- U.S. Department of Commerce: Economics and Statistics Administration, U.S. Census Bureau 2000, September. Money income in the United States 1999: Current population reports: Consumer income. Retrieved February 4, 2003, <<http://www.census.gov/prod/2000pubs/p60-209.pdf>>.
- Walton, G. M., & Cohen, G. L. (2003). Stereotype lift. *Journal of Experimental Social Psychology, 39*, 456–467.
- Wine, J. (1971). Test anxiety and direction of attention. *Psychological Bulletin, 76*, 92–104.