Journal of Personality and Social Psychology
2007, Vol. 92, No. 4, 698–716

0022-3514/07/$12.00 DOI: 10.1037/0022-3514.92.4.698

Threatened by the Unexpected: Physiological Responses During Social Interactions With Expectancy-Violating Partners

Wendy Berry Mendes
Harvard University

Jim Blascovich
University of California, Santa Barbara

Sarah B. Hunter
Rand Corporation

Brian Lickel
University of Southern California

John T. Jost
New York University

Individuals who violate expectations increase uncertainty during social interactions. Three experiments explored whether expectancy-violating partners engender "threat" responses in perceivers. Participants interacted with confederates who violated or confirmed expectations while multiple measures were assessed, including cardiovascular reactivity, task performance, appraisals, and behavior. In Experiments 1 and 2, participants interacted with White or Latino confederates who described their family backgrounds as either high or low socioeconomic status. In Experiment 3, participants interacted with Asian or White confederates who spoke with expected accents or southern accents. Participants interacting with expectancy-violating partners (e.g., Asians with southern accents) exhibited cardiovascular responses consistent with threat, poorer task performance, and manifested negative and defeat-related behavior. Implications for decreasing prejudicial responses via uncertainty reduction are discussed.

Keywords: uncertainty, expectancy violations, cardiovascular reactivity, challenge and threat

Interacting with strangers can be stressful. During such interactions, individuals are simultaneously attending to their own thoughts, feelings, and behaviors, including self-presentational, self-regulatory, and impression management concerns, as well as trying to accurately perceive the characteristics of the person to establish conversational fluency, expectations, and social norms (DePaulo & Friedman, 1998; Gilbert, 1998; Gudykunst, 1984). To the extent that the stranger meets expectations, the interaction is likely to be routinized and predictable. However, when the stranger is unexpected or surprising in some way, routinized responses are likely to be disrupted and anxiety heightened (Hebl, Tickle, & Heatherton, 2000). In the present study, we tested the notion that expectancy violations engendered by interaction partners would create more situational uncertainty and increase psychological threat as marked cardiovascularly, affectively, and behaviorally.

We have previously observed that, in general, African Americans relative to European Americans engendered threat responses among White participants during a pleasant social interaction. More important, the degree of past intergroup contact with African Americans reported by White participants moderated responses, such that the greater the contact with African Americans the lower the threat responses. This finding underscores the importance of contact during social interactions. But what about situations in which one could not have experienced contact because the stranger is surprising or unexpected in some way, that is, a stranger who violates expectancies?

Expectancies function to help people predict the future on the basis of past experiences and knowledge (Olson, Roese, & Zanna, 1996). Hence, expectancy violations disrupt one's predictive ability and can create uncertainty. This uncertainty is likely to increase the demands of situations, diverting attentional resources away from the social interaction and toward internal demands associated with emotional regulation and stress management. In this case, we argue that expectancy-violating partners will be threatening for similar reasons that lack of contact is threatening, namely that the situation is novel and that uncertainty is higher, leading to overall threat evaluations and concomitant physiological responses. To examine how interactions with persons who violate expectations affect stress, motivation, and physiology, one needs an organizing framework that offers specific predictions regarding the relationships between psychological states and physiological, behavioral,

Wendy Berry Mendes, Department of Psychology, Harvard University; Jim Blascovich, Department of Psychology, University of California, Santa Barbara; Sarah B. Hunter, Rand Corporation; Brian Lickel, Department of Psychology, University of Southern California; John T. Jost, Department of Psychology, New York University.

The research was supported in part by National Research Service Award MH12013 awarded to Wendy Berry Mendes. We are grateful to our dedicated lab assistants at the University of California, Santa Barbara, and Harvard University for their assistance in conducting the experiments, scoring the physiological data, and behavioral coding.

Correspondence concerning this article should be addressed to Wendy Berry Mendes, Department of Psychology, Harvard University, 33 Kirkland, WJH 1420, Cambridge, MA 02138. E-mail: wbm@wjh.harvard.edu

and cognitive responses. We used the biopsychosocial model of challenge and threat as our organizing framework.

Antecedents of Challenge and Threat

Research on the challenge and threat model (Blascovich & Mendes, 2000; Blascovich & Tomaka, 1996) has identified motivational states associated with distinct patterns of cardiovascular (CV) reactivity and hormonal responses (Blascovich & Mendes, 2000; Blascovich & Tomaka, 1996; Epel, McEwen, Ickovics, 1998; Mendes, Ayduk, Gyurak, & Akinola, 2006). According to this model, individuals evaluating task demands as exceeding their personal resources are *threatened*; individuals evaluating resources as exceeding demands are instead *challenged* (e.g., Tomaka, Blascovich, Kelsey, & Leitten, 1993). Blascovich and Mendes (2000) specified components of demands to include *danger*, *uncertainty*, and *required effort*, and resources include *knowledge and abilities*, *dispositions*, and *external support*. We predicted that individuals' interactions with strangers who violate expectations will lead to threat responses primarily via two evaluation components identified by expectancy violation theory: uncertainty and required effort (Bettencourt, Dill, Greathouse, Charlton, & Mulholland, 1997; Jackson, Sullivan, & Hodge, 1993; Jussim, Coleman, & Lerch, 1987; see Olson et al., 1996, for a review).

Uncertainty

Uncertainty is generally viewed as aversive and has been associated with anxiety, in-group bias, and difficulty in adapting to novel situations (Gao & Gudykunst, 1990; Grieve & Hogg, 1999; van den Bos, 2001). When confronted with category-based expectancy violations, the initial automatic response is believed to be surprise and uncertainty (Bettencourt et al., 1997). To the extent that interaction partners violate expectancies, they increase one's uncertainty during social interactions, thereby increasing demand evaluations, and hence, threat.

Required Effort

Given the unpleasantness of uncertainty, individuals may strive to make sense of surprising or unexpected information. For example, individuals remember more expectancy-violating behavior than expectancy-consistent behavior (Srull & Wyer, 1989), implicating increased working memory and more elaboration and effortful processing (Bartholow, Fabiana, Gratton, & Bettencourt, 2001; Srull, Lichtenstein, & Rothbart, 1985; Stangor & Duan, 1991). Physiological research demonstrates that expectancy-violating information elicits larger late-positive potentials (after 300 ms), a component of the electroencephalogram (EEG) waveform associated with working memory (Bartholow et al., 2001; Bartholow, Pearson, Gratton, & Fabiani, 2003). Because of increases in uncertainty and effort associated with expectancy-violating information, we argue that social interactions with expectancy-violating strangers will result in threat states, whereas interactions with expectancy-confirming partners, which are more "familiar" or routinized, will not relatively increase demands and will bring about challenge states.

Outcomes of Challenge and Threat States

Since Cannon (1929) distinguished fight–flight motivation, researchers have sought to identify distinct physiological responses that represent different psychological states. Several theories have focused on activation of the sympathetic adrenal medullary (SAM) and the hypothalamic pituitary adrenal (HPA) cortical axes to differentiate psychological states (e.g., Dienstbier, 1989; Folkow, 1993; Henry, 1986; Lundberg & Frankenhaeuser, 1980; Rotenberg, Sirota, Elizur, 1996). Challenge and threat relies on relative activation of these axes and is based on an integration of Dienstbier's (1989) physiological "toughness" model and Folkman and Lazarus' (1985) appraisal theory. Challenge and threat theory argues that challenge states (high resource relative to demands) are dominated by SAM activation and threat states (high demands relative to resources) by HPA activation. It specifies that in motivated performance situations—active, self-, or goal-relevant situations that require cognitive processing—psychological states of challenge and threat produce distinct coordinated constellations of responses, which include physiological, affective, and behavioral markers. It is important to note that challenge and threat are psychological states that can be indexed by physiology, appraisal, and behavior and that the physiological response is not "threat" (or "challenge") per se, but rather a manifestation of the psychological state.

Physiological

Most physiological evidence differentiating challenge and threat has focused on CV differences (Blascovich & Mendes, 2000; Blascovich & Tomaka, 1996), though neuroendocrine evidence is accumulating (Epel et al., 1998; Mendes et al., 2006). Physiological responses mediated via SAM activation are characterized by enhanced cardiac performance, particularly left ventricular contractility (VC, which are changes in preejection period multiplied by -1) and cardiac output (CO), and, additionally, increases in epinephrine result in vasodilation or decreased systemic vascular resistance (total peripheral resistance; TPR). In contrast, threat is associated not only with SAM activation, again increasing VC, but also with HPA activation, which is associated with increases in cortisol. Additionally, increases in norepinephrine in threat states inhibit vasodilation and often produce vasoconstriction (i.e., increases in TPR).

Affective and Motivational

Challenge and threat states are generally, but not necessarily, tied to positive and negative affect as well as to approach and avoidance orientations, respectively (Blascovich, Mendes, & Seery, 2002; Tomaka et al., 1993; Tomaka & Palacios-Esquivel, 1997). Challenge states are often characterized by increased general positive affect and, more notably, decreased negative affect (Blascovich et al., 2002). Emotional responses measured using facial electromyography (EMG) suggests that threat states are associated with greater corrugator relative to zygomaticus regional activity, implicating greater negative relative to positive emotion (Tomaka, 1994).

Nonverbal Behavior

Behavioral responses associated with these motivational states have been less studied, though several lines of research inform our predictions. Because threat states are consistent with defeat re-

sponses, predictions of motor behaviors follow. First, general body orientation should reveal an avoidance or protective stance, one characterized by closed body posture and general orientation away from the "threatening" stimulus. Second, consistent with the orienting response, less general somatic activity should occur (see Stern, Ray, & Quigley, 2001). For example, freezing is an adaptive behavior when confronted with a novel stimulus in the environment or a potential threat. Freezing is characterized by immobility in humans and other species so that alerted individuals are motionless while monitoring a source that is either uncertain or possibly dangerous (Blanchard, Flannelly, & Blanchard, 1986; Marks, 1987). Consistent with this notion, the cognitive process model (Scherer, Zentner, & Stern, 2004) predicts that uncertainty appraisals will interrupt ongoing activity and result in motoric freezing allowing an organism time to reorient. Some recent evidence supports these predictions among infants. Infants exposed to expectancy violations, such as exposure to the unexpected voice qualities of an experimenter, exhibited more motoric freezing compared with those who were exposed to normal voices (Scherer et al., 2004).

Background for the Present Research

As we described above, past research has revealed that White participants interacting with African American same-sex confederates exhibited CV responses consistent with threat states and performed worse on a subsequent cognitive task (Blascovich, Mendes, Hunter, Lickel, & Kowai-Bell, 2001; Mendes, Blascovich, Lickel, & Hunter, 2002; see also Richeson & Shelton, 2003). In an attempt to identify moderators of threat responses, we assessed positive personal contact with African Americans using questions developed by Islam and Hewstone (1993), which assessed qualitative and quantitative out-group contact. We also assessed explicit racism (Modern Racism Scale [MRS]; McConahay, 1986), and the extent to which one was motivated to appear nonprejudiced (Motivations to Control Prejudice Reactions Scale; Dunton & Fazio, 1997). Only personal contact reliably moderated threat responses such that the greater the contact with African Americans. the lower the CV threat responses (specifically vascular resistance decreased and CO increased; Blascovich et al., 2001, Study 3). We argued that quality contact with out-group members can reduce threat during social interactions with them.

The aforementioned studies were conducted at a relatively affluent, nonurban university (University of California, Santa Barbara [UCSB]), in which the student population is predominantly White and upper-middle class (less than 3% African American students at the time). Although self-reported attitudes toward African Americans did not suggest strong bias (MRS: $M = -1.2$, $SD = 0.9$, on a $-4 - +4$ scale), the actual amount of personal contact with African Americans was low ($M = 2.8$, $SD = 1.4$, on a 1–7 scale). Consequently, participants exhibited threat responses during interactions with African Americans. We argued that participants perceived interactions with African Americans as more demanding for reasons including greater intergroup anxiety (Stephan & Stephan, 2000), increased uncertainty and ambivalence (Jones et al., 1984), and increased effort via stereotype suppression, guilt, or compensation (Devine, Monteith, & Zuwerink, 1991; Wyer, Sherman, & Stroessner, 2000).

Given that greater intergroup contact reduced threat responses (and because racial attitudes were not predictive of threat responses), we believe that the weight of the evidence points toward greater uncertainty, via unfamiliarity, with the interaction partner as a critical factor in engendering threat states with different race partners. Other evidence supports this idea. In two neuroimaging studies (Phelps et al., 2000), White participants' racial biases were measured using the Implicit Association Test (Greenwald, McGhee, & Schwartz, 1998) and a startle-probe paradigm following a functional magnetic resonance imaging (fMRI) scan during which they viewed unfamiliar White and Black faces (see Experiment 1) or familiar and positively regarded White and Black faces (see Experiment 2). Racial bias was positively related to amygdalar activation when participants were presented with unfamiliar Black faces but not when presented with familiar Black faces. Thus, familiarity in this case eliminated the relationship between racial bias and amygdalar activation. An important point here is that although amygdalar activation has been reliably linked to fear in rats (LeDoux, 2000), recent evidence suggests that the amygdala responds to novel stimuli as well (Wright et al., 2003).

We reasoned that if participants' threat responses during interactions with African Americans were because of lack of familiarity, then an a priori test of our hypothesis would be to assess participants' responses during social interactions with expectancy-violating (i.e., counterstereotypical or atypical) persons, who, by definition, would be unfamiliar. We predicted that social interactions with such partners would engender uncertainty and increased cognitive effort leading to increased threat relative to typical partners. We also sought to demonstrate that interactions with out-group members would not be threatening if participants had sufficient contact with them.

To develop our manipulations of expectancy violation, we analyzed demographic data from the University and conducted pilot studies to determine participants' expectancies about different ethnic groups. Demographic data revealed that all racial/ethnic groups had approximately equivalent parental incomes with the exception of Latinos, who as a group had the lowest parental incomes. Consistent with this, research has shown that Latinos are typically viewed by Whites as poor (Goodwin & Fiske, 1996) and low in social status (Jost, Pelham, & Carvallo, 2002). Therefore, we reasoned that White and Latino students would be expected by participants to differ socioeconomically. Specifically, Latino students would be expected to come from low-socioeconomic status (SES) families (based on stereotypes and demographic information), White students would be expected to come from high-SES families, and the reverse would be unexpected.

Pilot Studies

To confirm the expectancy violation manipulation and to provide a comparison across our past and present research, we conducted two pilot studies. The first pilot study ($N = 35$) consisted of administration of an in-depth questionnaire to assess numbers, race, ethnicity, SES, and closeness of participants' friends and acquaintances. SES was rated on a 9-point scale, anchored with the following numbers and descriptors: -4 (*poor*), 0 (*middle class*), and $+4$ (*wealthy*). Participants reported having almost twice as many Latino as African American friends and acquaintances (Ms = 7.8 vs. 4.1), $t(34) = -4.75$, $p < .0001$. Among their friends

and acquaintances, Latinos were rated as lower in SES than African Americans (Ms = −0.6 vs. 1.0), $t(31) = 2.52$, $p < .02$, and Latinos were rated as closer friends than African Americans (Ms = 0.8 vs. −0.1), $t(25) = -2.97$, $p < .01$. If close contact is an important moderator of threat responses, then we reasoned that Latinos in general may be less likely to engender threat responses than African Americans because of the large differences in our samples' reported history of interaction with Latinos relative to African Americans. This should be particularly true for Latinos who would confirm Whites' expectancies, specifically those Latinos who were perceived as lower in SES.

In Pilot Study 2, we tested perceptions of the ethnicity and SES pairings that we wished to manipulate. We created four one-page questionnaires that described each of the four combinations of ethnicity and status: 2 (ethnicity: Latino, White) × 2 (SES: high, low). Each description included a picture of either a Latino or a White male. The pictures were chosen on the basis of similarity in attractiveness, age, photographic quality, and facial expression (smiling). Descriptive information included name (either Tomás Garcia or Thomas Green), which was consistent with the ethnicity of the depicted person, and major (business-economics). For the high-SES condition, the description that followed the picture was:

> This is _____. He is a business-econ major at UCSB. Originally, he is from Palo Alto [CA]. His father owns a law firm, and his mother is a professor at Stanford. He has a little sister that goes to UCLA. In his spare time, he enjoys waterskiing, jet skiing, and other water-type sports. He also enjoys snowboarding. This past summer, he worked in his Dad's law office and then went to Europe for a couple of months.

For the low-SES condition, the description was as follows:

> This is _____ . He is a business-econ major at UCSB. Originally, he is from San Jose [CA]. He hasn't talked to his father in a few years, and his mother was just laid off from her job at a factory. He has four brothers and sisters. His oldest brother drives a cab, and his three younger siblings live at home with his mother. He has a couple of part-time jobs, so he doesn't have a lot of free time, but sometimes he'll hang out with his buddies and play basketball. This summer, he worked to save up some money.

Participants ($N = 205$) then reported level of agreement with two statements using 9-point response scales ranging from −4 (*strongly disagree*) to + 4 (*strongly agree*). The first was "(Tomás/Thomas) is a typical undergraduate at UCSB"; the second was "I have known several people similar to (Tomás/Thomas)." The third question asked participants to estimate the percentage of people in California who were similar to (Tomás/Thomas).

To test our predictions, we conducted analyses of variance (ANOVAs), using the three responses as the dependent variables and ethnicity and SES of the target person as the independent variables. Though there were main effects for SES—participants rated high SES as more typical for UCSB students and estimated low SES as more typical of people in California—all three variables yielded a significant interaction. Undergraduates reported that White-high SES described the most typical students at UCSB ($M = 0.6$), which differed from the other three conditions (Latino-low SES: $M = -0.3$, Latino-high SES: $M = -0.3$, White-low SES: $M = -0.9$), $F(1, 203) = 6.22$, $p < .01$. Though this effect was not entirely consistent with the predictions—White-high SES differed from Latino-low SES—this effect does show the propensity of participants to perceive "typical" students as White and of high SES.

More important, the interaction with the contact variable—had the participant ever known someone similar—did yield the expected directional interaction, $F(1, 203) = 8.60$, $p < .01$. Participants were more likely to report knowing someone similar when the person was described as White-high SES ($M = 1.3$) or Latino-low SES ($M = 1.1$) than participants who read the Latino-high SES ($M = -0.3$) or White-low SES ($M = -0.05$) descriptions. Finally, participants' percentage estimates of people in California who were similar to the target person yielded the expected interaction, $F(1, 199) = 10.72$, $p < .001$. Participants rated the Latino-low SES combination as the most common (44%), followed by White-high SES (40%), White-low SES (34%), and Latino-high SES (23%). All post hoc comparisons yielded differences between groups except for Latino-low SES and White-high SES. These data reveal that undergraduates report more contact and estimate more individuals similar to the stereotypical pairings of White-high SES and Latino-low SES than the counterstereotypical pairings of White-low SES and Latino-high SES.

Overview of Experiments and Hypotheses

We tested the prediction that interaction partners who violate expectations are threatening relative to those who do not in three experiments. In the first two experiments (Experiment 1 sampling women and Experiment 2 sampling men), we combined the characteristics established in the pilot studies to examine the effect of participants' responses to "stereotypical" versus "counterstereotypical" partners to study the effects of expectancy violations on threat responses. The third experiment required participants to interact with atypical partners whose characteristics would be less influenced by ideological beliefs. For all experiments, we hypothesized that relative to participants interacting with expected or typical partners, participants interacting with unexpected or atypical partners would exhibit CV responses consistent with threat, perform worse on a cognitive task, and attribute more negative traits to their partner. In Experiment 3, we coded the nonverbal and verbal behavior of participants from the videotaped interactions with partners and hypothesized that unexpected/atypical partners would engender threat and defeat responses characterized by less somatic activity, negative treatment, and general avoidance orientation than interactions with typical partners.

We designed experimental procedures with two goals in mind. First, the social interaction had to include tasks that would constitute a motivated performance situation—specifically, tasks that were active, goal relevant, and required instrumental cognitive responses. Second, we wished to create an interaction that would mimic (in some respects) the unfolding of a social interaction between strangers. To this end, the experimental design included a participant and confederate meeting and exchanging limited but informative background information, sharing personal information about their positive qualities, and then cooperating on an interdependent task. This was accomplished with three phases in the experiment: (a) information exchange—during which participants and partners (confederates) met face-to-face and exchanged background information; (b) speech delivery—during which participants prepared and delivered speeches on "working together"; and

(c) a word-finding task—during which participants and confederates engaged in cooperative word-finding tasks.

Experiment 1: Women—High and Low SES and Ethnicity

In Experiment 1, we recruited female participants for a laboratory study in which they interacted with another female participant (a confederate). Confederates were either White or Latina, and they described their background as either wealthy (high SES) or poor (low SES). After exchanging background information with each other, the dyad also engaged in a speech task and cooperative game. Physiological responses were recorded during the latter two tasks.

Method

Setting and Participants

A social psychophysiology laboratory containing separate control, participant preparation, and recording rooms as well as physiological recording, audiovisual, and computer equipment was the setting. We recruited female participants ($N = 47$; 78% White, 20% Asian, 2% unknown; mean age = 20.6, range = 19–24 years) who were prescreened for doctor-diagnosed heart murmurs, pacemakers, or pregnancy. We did not limit our analyses only to White participants because expectancy violations, as we have operationalized them, should occur within groups as well as across groups (see Kernahan, Bartholow, & Bettencourt, 2000). Participants were informed that they would receive either $10 or course credit for the experiment. Using the same SES scale from the pilot tests, participants rated themselves slightly above middle class ($M = 0.9$) on a 9-point scale anchored at −4 (*poor*), 0 (*middle class*) and +4 (*wealthy*).

Measures

Physiological measures. Cardiac and hemodynamic measures were recorded noninvasively using equipment meeting commercial and hospital safety standards and following guidelines established by the Society for Psychophysiological Research (e.g., Sherwood et al., 1990). A Minnesota Impedance Cardiograph (Model 304B, Chapel Hill, NC), a Cortronics (Model 7000) continuously inflated blood pressure monitor, and a Coulbourn electrocardiographic (ECG) amplifier/coupler (Model S75-11, Allentown, PA) provided physiological signals. The impedance signals were conditioned using Coulbourn amplifiers (Model S79-02, Allentown, PA).

Impedance cardiographic (ZKG) and ECG recordings provided continuous measures of cardiac performance. A tetrapolar aluminum/mylar tape electrode system is used in ZKG to provide basal transthoracic impedance (Z0) and the first derivative of basal impedance (dZ/dt). Two pairs of ZKG tape encircle the participant at the neck and the torso and are secured with electrodes. A 4mA AC 100-kHz current is passed through the two outer electrodes and measures basal impedance from the two inner electrodes. The ECG recordings were obtained using a Standard Lead II configuration (right arm, left leg, and a right-leg ground). A Cortronics blood pressure monitor provided continuous noninvasive recordings of blood pressure. An interactive software program (Kelsey & Guethlein, 1990) was used to record and score the cardiac and hemodynamic data.

We differentiated challenge and threat on the basis of CV reactivity (i.e., changes from baseline) focusing on VC, CO, and TPR, the latter derived from blood pressure and CO using the formula (mean arterial pressure/CO) × 80 (Sherwood et al., 1990).

Task performance. The word-finding task was similar to the game of Boggle. The task stimulus consisted of an 8 × 8 letter matrix presented on a computer monitor. The objective was to form as many words as possible in 4 min by linking adjacent letters to form words. Participants and confederates were instructed to take turns finding words, one at a time, and saying them aloud. We tracked the number and accuracy of their responses.

Self-Report Ratings

Each participant completed two posttask questionnaires. The first questionnaire followed the speech delivery task and included three questions regarding perceptions of stress, effort, and quality of performance. The second followed the word-finding task and assessed the same concepts as the postspeech and included manipulation checks. The first question required the participant to rate how poor or wealthy she thought her partner was on the same SES scale previously described.[1] This question was followed by an open-ended question asking her the ethnicity of her partner. We then asked the participant to rate her own SES background on the same scale as above and state her ethnicity.

Procedure

We randomly assigned participants to either a White or Latina partner (confederate) who would describe her background as low or high SES to create a 2 (ethnicity: White, Latina) × 2 (SES: low or high) between-subjects design. We used two Latina and two White female confederates who completed an extensive training program that focused on creating as much similarity in responses and reactions as possible. All confederates dressed in a similar neutral fashion (blue jeans and white t-shirt), allowing for ambiguity with regard to perceived SES on the basis of sartorial cues. Confederates were not aware of any of the study's hypotheses.

Initial interaction. Participants and confederates arrived and waited in front of separate doorways 10 m apart outside of the laboratory. The confederate ensured that no interaction took place in the hallway by avoiding eye contact and studying. Two experimenters greeted them and explained that the study involved "interpersonal styles and working together." They confirmed that the participant and the confederate did not know each other and explained that they would go to separate rooms to fill out forms but would see each other later.

One experimenter escorted the confederate to a preparation room, and the other escorted the participant to a separate room. The experimenter requested that the participant complete a consent form and background information sheet. The background informa-

[1] Though SES was manipulated with the three main constructs that constitute SES in mind—occupational prestige, education, and income—we only measured SES perceptions using one category of SES: income. A more nuanced measurement, which considered all aspects of SES, might reveal greater differences.

tion sheet queried the participant about her age, hometown, college major, parents' occupations, siblings, hobbies, sports, and extracurricular activities. A photograph was then taken of the participant that we attached to her completed background sheet. Concurrently, the confederate completed her background sheet on the basis of random assignment to SES, and we took a photograph of her as well.

Information exchange. The experimenter then escorted the participant to the confederate's experiment room and instructed the participant and the confederate to describe their backgrounds to each other (using the background information sheet as a guide) and left the room. The confederate spoke first and introduced herself as Alicia, a 20-year-old psychology major. Following this, the information varied, depending on SES condition, which was developed from the same pilot testing described earlier, but that yielded slightly different profiles for high- and low-SES males compared with females. The "advantaged" Alicia said she was from Montecito, California (a wealthy community), her father was an international lawyer in private practice, her mother was involved in lots of charities, and her older brother was a Harvard medical student. She said she liked to go shopping with her friends and play volleyball and described her summer plans to include waitressing at her parent's country club and volunteering at an animal shelter. The "disadvantaged" Alicia said she was from Goleta, California, her father was not currently working, and her mother cleaned houses. She had three younger brothers and sisters and one older brother who worked in an auto body shop. In her spare time, she enjoyed watching TV with her friends and playing softball. Her summer plans included waitressing at the International House of Pancakes and babysitting her siblings. Following this, the participant described her background.

Following this exchange, the experimenters reentered and asked the participant and confederate to trade their background information sheets with photos attached. The participant and confederate were then separated and escorted to different preparation rooms for the remainder of the study.

Speech delivery. We then applied physiological sensors to the participant. She was seated in an upholstered chair with a small lap tray. She was given a computer mouse as well as the confederate's background information sheet and photo (face down). The experimenter instructed the participant to sit quietly and relax for several minutes. A 5-min baseline period began once the signal was properly filtered and amplified. CV responses collected during this period served as baseline levels of physiological responses.

After the baseline period, the participant heard audiotaped instructions to turn over and review the confederate's background information sheet for 1 min. Afterwards, she received instructions for delivering a 3-min speech about working together that would be videotaped and shown to the other participant later on during the experiment. The participant was instructed to look into the camera and discuss "how well you work with other people . . . how well do you believe the other subject works with other people . . . and how well will the two of you work together." We presented the speech topic and instructions on a large television monitor in front of the participant for her reference. The participant was cued via intercom to begin the 1-min preparation period, to deliver, and then to end the speech. The participant received prompts via the intercom to elaborate on the speech themes if she stopped talking before the 3-min period expired. Following the speech, the participant completed the postspeech questionnaire, and then a 5-min recovery/rest period was recorded.

Word-finding task. Afterwards, the experimenter explained that the intercoms would be connected so that the two could hear each other for the next part of the experiment. After confirming that the participant and confederate could hear each other, they received instructions that they would be working together on a word-finding task similar to the game of Boggle that required them to take turns finding words and saying them out loud. They were informed that each would receive a $5 bonus if, as a team, they could find 26 words in 4 min (the participant always received the bonus money at the end of the experiment).

After the instructions, an ostensibly randomly generated matrix of letters appeared on the monitor, the participant and the confederate were told to begin, and the participant was instructed to find the first word. The confederate's responses came from a list of over 60 valid words in the matrix and were guided by timed prompts. The timing was designed on the basis of pretesting to represent typical performance ability. After 4 min, the experimenter informed the dyad that the task was completed, disconnected the audio system, and entered the recording room with the postword task questionnaire. After the participant completed the questionnaire, the experimenter removed the sensors and probed for suspicion. The experimenter then debriefed, paid, and thanked the participant.

Results

Participant Attrition

One participant was excluded because she expressed suspicion regarding the confederate, and 2 participants' physiological data were lost because of equipment malfunction. This attrition resulted in 46 participants with self-report and performance data, and 44 with physiological data.

Manipulation Checks

Participants' response to their partner's socioeconomic background yielded a main effect for SES, $F(1, 42) = 50.27$, $p < .0001$, confirming the SES manipulation. Participants paired with high-SES partners rated her background as wealthier ($M = 2.7$) than those paired with low-SES partners ($M = -0.8$). Both status manipulations significantly differed from zero (middle class) in the expected direction: high SES, $t(21) = 7.83$, $p < .0001$; low SES, $t(21) = -2.30$, $p < .04$. Participants' open-ended response regarding the ethnicity of her partner indicated that the ethnicity of White confederates was never misidentified. However, Latina confederates were correctly identified 64% of the time. The most common response for nonidentifications was "don't know" (80%), and the remaining incorrect responses consisted of participants believing their partners were Middle Eastern. Though participants were more likely to make identification errors when their partner was Latina and high SES (47%) than Latina and low SES (23%), this was not significant, $\chi^2(1, N = 28) = 1.7$, *ns*. We reran all subsequent analyses without the participants who made incorrect identification cases. For all critical tests of the hypothesis, the effect sizes were virtually identical to those using the full sample.

CV Measures

Scoring and analytic strategy. Mean VC, CO, and TPR values were calculated for each minute within each rest and task period. Our analytic strategy included several steps. First, we tested for any individual confederate effects by looking at responses within each ethnicity cell category (in all experiments reported here, we never observed any significant differences between confederates). We then tested baseline CV differences between conditions. Third, we confirmed that the tasks constituted motivated performance situations by testing whether sympathetic nervous system activation increased significantly from baseline by examining changes in VC and heart rate (HR) reactivity. To test our primary predictions, we performed multivariate analyses of variance (MANOVAs) to examine the effect of partner's ethnicity and SES on relative patterns of the CV reactivity markers of challenge and threat (VC, CO, and TPR) for each task. Last, we examined each CV measure via univariate analyses to confirm directional hypotheses.

We typically focused on the first minute of the tasks because of differences in task habituation between challenge and threat. Challenge reactivity, like the physiologically tough pattern, is characterized by quick CV habituation and recovery. Therefore, unless new elements are introduced within a task, participants are expected to habituate quickly during the tasks evaluated as challenging. In contrast, threat reactivity is predicted to recover slower than challenge reactivity (Dienstbier, 1989). Furthermore, increases in vascular reactivity recover slower than increases in cardiac reactivity (Kelsey et al., 1999).

Baseline differences. A MANOVA using the three variables—VC, TPR, and CO—from the last minute of baseline and the confederate's ethnicity and SES as the independent variables revealed no significant main effects or interaction (all multivariate Fs < 1.4). As is typical when baseline responses do not differ among levels of between-subjects factors, reactivity scores (differences from baseline) were used as the primary dependent variables (Kamarck et al., 1992). Reactivity scores were calculated for each measure by subtracting the last minute of baseline from each minute of each task.

Goal relevance. Separate univariate tests contrasted VC and HR reactivity scores to zero to verify goal relevance during the speech and word-finding tasks by experimental condition. Significant increases in VC and HR reactivity occurred for all tasks and conditions (speech delivery: all ps $< .0001$; word-finding task: all ps $< .001$) warranting challenge and threat analyses.

Challenge and threat: Speech delivery. Participants' CV responses from the speech task were analyzed by ethnicity and SES. Neither main effects nor the interaction were significant.

Challenge and threat: Word-finding task. We then analyzed physiological responses from the cooperative word-finding task. The prediction was that counterstereotypical partners would engender CV threat responses relative to stereotypical partners. A MANOVA testing the effects of partner's ethnicity and SES on CV responses from the first minute of the word task yielded a significant interaction (Wilks's $\Lambda = .82$), $F(3, 38) = 2.82$, $p < .05$ (see Figure 1), and significant univariate interactions for all CV variables: VC, $F(1, 42) = 4.35$, $p < .05$; CO, $F(1, 42) = 4.58$, $p < .05$; and TPR, $F(1, 42) = 8.70$, $p < .01$. As predicted, being paired with counterstereotypical partners (White-low SES and Latina-high SES partners) resulted in relatively less VC, lower CO, and increased TPR compared with participants interacting with partners whose background was stereotypical (Latina-low SES and White-high SES partners). In summary, interactions with stereotypical partners resulted in challenge responses, and interactions with counterstereotypical partners resulted in threat responses.

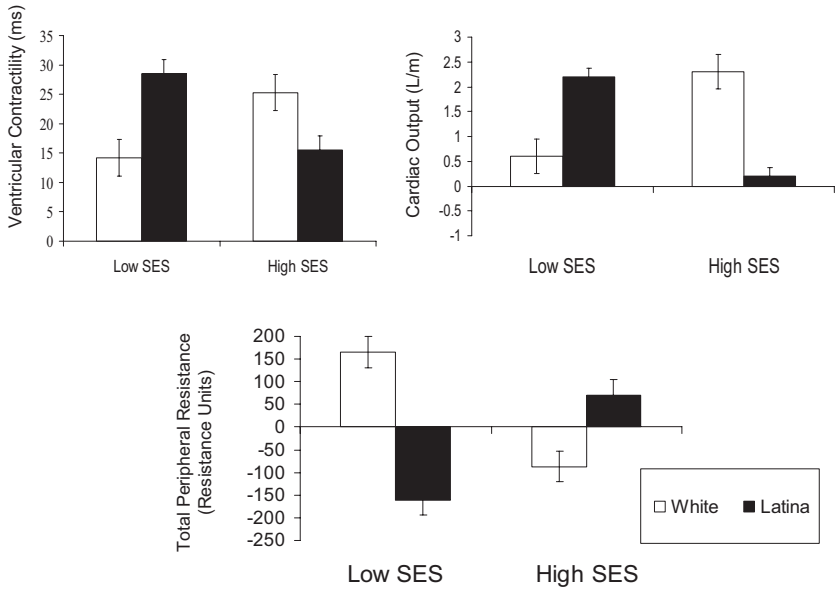

Figure 1. Means and standard errors of cardiovascular reactivity during the first minute of the word-finding task by ethnicity and socioeconomic status (SES) of the partner in Experiment 1. All values are expressed as change scores from the last minute of baseline.

Performance: Word-Finding Task

No significant differences were observed by partner's ethnicity, SES, or their interaction.

Self-Report Ratings

There were no differences in stress ratings following the speech. However, after the word-finding task and consistent with the physiological data, a significant interaction was observed, $F(1, 41) = 6.11$, $p < .02$. Participants paired with stereotypical partners rated the word-finding game as less stressful (Latina-low SES: $M = -0.8$; White-high SES: $M = -0.7$) than those paired with counterstereotypical partners (White-low SES: $M = 0.8$; Latina-high SES: $M = 1.2$).

Discussion

Results showed general support for the primary predictions. Participants paired with counterstereotypical partners exhibited lower CO and higher TPR compared with participants paired with stereotypical partners during the cooperative word-finding task. We interpret these data as indicative of threat responses resulting from interactions with counterstereotypical relative to stereotypical partners. Participants' stress ratings after the word-finding task were consistent with the CV data; those paired with stereotypical partners reported less stress after the word-finding task than participants paired with counterstereotypical partners.

Experiment 2: Men—High and Low SES and Ethnicity

In Experiment 2, we sought to replicate Experiment 1 with men and modify the procedures to create a face-to-face interaction, which we reasoned would create more intense affective responses (Aronson, Wilson, & Brewer, 1998; Stemmler, Heldmann, Pauls, & Scherer, 2001). The major change to the procedure in Experiment 2 was that during both the speech and word-finding task, we used video cameras and monitors to allow the dyad to see and hear each other during both tasks (as opposed to merely hearing each other during the word task). We disconnected the video and audio feed during baseline and recovery periods but allowed the dyad to see and hear each other during task instructions, preparations, and task periods.

Method

Participants

Healthy male undergraduates ($N = 63$), age 17–24 ($M = 18.9$), primarily White (73%) and Asian (10%), participated for course credit or $10. On average, participants rated their own socioeconomic background as significantly above middle class ($M = 1.0$), $t(60) = 5.10$, $p < .0001$.

Procedure

Experiment 2 was conducted exactly like Experiment 1 except for the changes described below. Two Latino and three White confederates were used in this study, and they all completed a training program together and dressed in a similar manner for the experiment (white t-shirt, jeans, black baseball cap). Confederates were named Tomás (Latino) or Thomas (White), and his background, which varied with respect to SES, was as described in Pilot Study 2. To gauge any overt reactions of the participants, confederates completed two identical rating forms (pre- and postinformation exchange) that consisted of three questions regarding the reactions of the participants to the confederate. The questions included the extent to which the participant made eye contact with him and how friendly and positive the participant was. Again, all responses ranged from −4 (*strongly disagree*) to + 4 (*strongly agree*).

After the information exchange, the two were taken to different participant rooms, and nonoperating sensors were applied to the confederate because he would be seen by the participant on the video monitor later in the experiment. After the baseline period, we connected the participant rooms with audiovisual equipment that allowed the dyad to see and hear each other on large 68.6-cm screen monitors, approximately 132 cm in front of the participant. We "randomly" assigned the participant to deliver a speech, as described in Experiment 1, and the "partner" was instructed to listen to the speech (we also shortened the speech to 2 min). After the speech, we disconnected the rooms, and the participant completed the postspeech questionnaire and sat for a 5-min rest/recovery period. We then connected the rooms again and explained the word-finding task. After instructions, the participant and confederate played the game for 4 min. Afterwards, we disconnected the audiovisual equipment, and the participant completed the final questionnaire. The participant completed ratings of the confederate that included questions regarding how unfriendly, likable, trustworthy, unhelpful, creative, independent, and unintelligent the participant perceived his partner to be and how well they performed "as a team." All responses ranged from −4 (*strongly disagree*) to + 4 (*strongly agree*). After the final questionnaire, participants were probed for suspicion, debriefed, and paid.

Results

Participant Attrition

No participants were lost to suspicion. Four participants' physiological data were lost because of initial problems with software interfacing. Other problems with physiological hardware and noisy signals resulted in a total of 57 participants with usable speech data and 54 with word-finding data, and 63 participants for analyses of self-reports and performance.

Manipulation Checks

Status and ethnicity were perceived as intended. Participants paired with high-SES partners rated his background as wealthier ($M = 2.6$) than those paired with low-SES partners ($M = -1.4$), $F(1, 61) = 144.64$, $p < .0001$. Ethnicity of the partner was never misidentified.

CV Measures

Baseline differences. A MANOVA examining baseline physiological responses by conditions revealed no main effects for partner's ethnicity or SES, $F(3, 52) = 1.98$, $p < .13$. An examination of the univariate tests revealed one significant main effect for cardiac output, $F(1, 56) = 4.09$, $p < .05$. Participants assigned

to partners who were either Latino-low SES or White-high SES had lower CO at baseline than participants assigned to Latino-high SES or White-low SES. Because we found some baseline evidence of differences between our manipulated conditions, we used baseline CV responses as covariates in all subsequent analyses.

Goal relevance. Changes in VC and HR by conditions were examined for both tasks. All analyses yielded significant changes from baseline to task for all conditions (speech: all ps $< .0001$; word finding: all ps $< .05$). We then tested our main predictions during the speech and word-finding task.

Challenge and threat: Speech delivery. We ran a 2 × 2 multivariate analysis of covariance (MANCOVA) using the independent variables of ethnicity (Latino or White) and SES (high or low) and the dependent variables of VC, CO, and TPR reactivity from the speech task. We included baseline CV responses as covariates to control for significant differences at baseline between groups, though the exclusion of the covariates yielded the same significant and nonsignificant findings reported with covariates. The multivariate interaction from the speech task was significant (Wilks's $\Lambda = .74$), $F(3, 48) = 5.71$, $p < .002$. All univariate analyses yielded significant differences in the predicted direction: VC, $F(1, 55) = 9.53$, $p < .003$; CO, $F(1, 55) = 14.92$, $p < .0003$; TPR, $F(1, 55) = 15.58$, $p < .0002$ (see Figure 2). As hypothesized, interactions with stereotypically consistent partners (White-high SES; Latino-low SES) resulted in significantly greater cardiac reactivity (VC and CO) and lower TPR than interactions with stereotypically inconsistent partners (White-low SES; Latino-high SES). Responses during the second minute of the speech were all significantly different as predicted and mirrored the patterns in the first minute, though responses were attenuated. Repeated measures ANOVA revealed a significant time effect; all cardiac responses during the second minute were lower than the first minute, and vascular differences were higher in the second minute than in the first minute (see Kelsey et al., 1999, for similar patterns and a discussion of the temporal effects of cardiac and vascular responding). More important, ethnicity and SES did not interact with time, so the same pattern remained throughout the speech.

Challenge and threat: Word-finding task. We then analyzed CV responses during the word-finding task. The multivariate interaction was significant, $F(3, 45) = 3.89$, $p < .02$, Wilks's $\Lambda = .78$. All univariate tests yielded significant interactions: VC, $F(1, 52) = 10.14$, $p < .003$; CO, $F(1, 52) = 7.88$, $p < .008$; TPR, $F(1, 52) = 8.57$, $p < .006$ (see Figure 3). Similar to the speech data, participants paired with counterstereotypical partners exhibited lower cardiac activity and higher vascular resistance than participants paired with stereotypical partners.

Performance: Word-Finding Task

We conducted an ANOVA on the number of words generated after correcting for repeated, misidentified, and nonsense words; a marginal interaction was obtained, $F(1, 57) = 3.37$, $p < .072$. Consistent with the CV data, participants paired with White-high SES and Latino-low SES partners found more words ($M = 15.3$ and $M = 15.2$) than did participants paired with Latino-high SES and White-low SES partners ($M = 13.2$ and $M = 12.6$). Given that stereotypical conditions did not differ from each other (nor did counterstereotypical conditions), we conducted contrasts comparing counterstereotypical with stereotypical conditions. This contrast yielded a near significant performance effect, $F(1, 57) = 3.82$, $p < .056$.

Self-Report Ratings

Participants' ratings. We then examined participants' ratings of their partner's traits, abilities, and the extent to which participants believed they made a "good team." Some significant interactions emerged, consistent with the predictions. Specifically, participants' ratings of how well they performed "as a team" yielded a significant Partner Ethnicity × SES interaction, $F(1, 57) =$

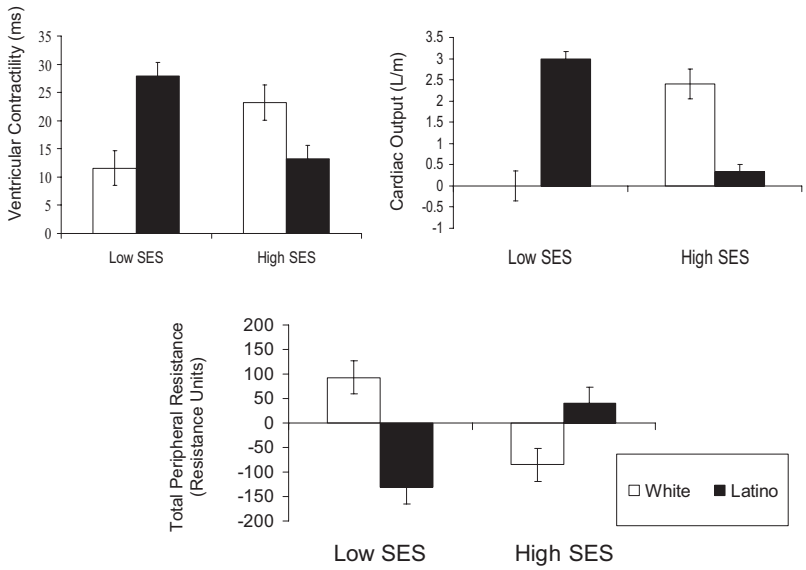

Figure 2. Means and standard errors of cardiovascular reactivity during the first minute of the speech task by ethnicity and socioeconomic status (SES) of the partner in Experiment 2.

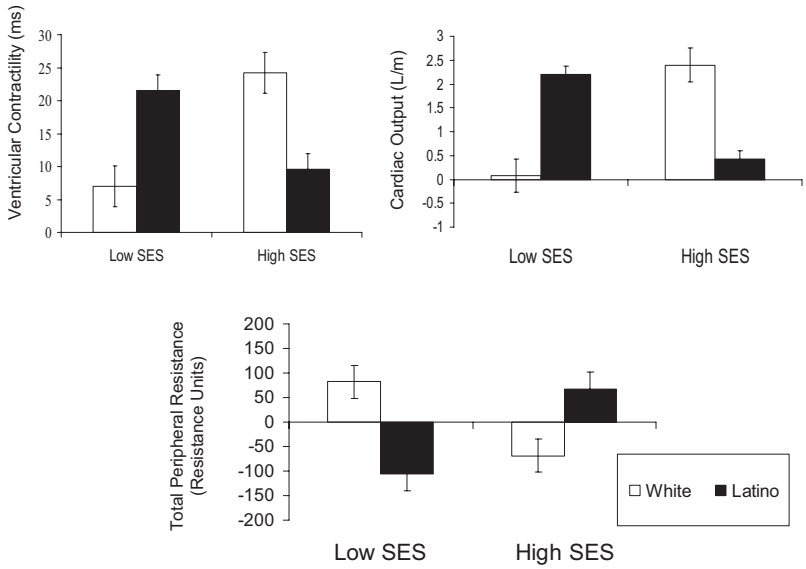

Figure 3. Means and standard errors of cardiovascular reactivity during the first minute of the word-finding task by ethnicity and socioeconomic status (SES) of the partner in Experiment 2.

10.81, $p < .002$, and no significant main effects. Participants rated team quality higher when paired with stereotypical partners (White-high SES $M = 2.3$; Latino-low SES $M = 1.9$) than participants paired with counterstereotypical partners (Latino-high SES $M = 0.8$; White-low SES $M = 0.3$). This same pattern was observed for participants' ratings of how likeable their partner was, $F(1, 57) = 3.85$, $p < .055$ (Latino-low SES $M = 2.5$; White-high SES $M = 2.4$ vs. Latino-high SES $M = 1.9$; White-low SES $M = 1.8$). A marginal interaction was observed for whether participants viewed their partner as unintelligent, $F(1, 57) = 3.07$, $p < .085$, and the pattern of means followed other trait ratings, such that participants paired with the stereotypical partners rated them as less unintelligent (White-high SES: $M = -3.3$; Latino-low SES: $M = -3.4$) than counterstereotypical partners (Latino-high SES: $M = -2.9$; White-low SES: $M = -2.7$). No other main effects or interactions were observed.

Confederates' ratings. Confederates' ratings of the participants were combined to form two indexes, one from preinformation exchange ($\alpha = .88$) and the second from postexchange ($\alpha = .95$). Results from the initial interaction (when the participant and confederate met each other in the hallway and no information was exchanged) yielded a large main effect for confederate ethnicity, $F(1, 53) = 24.58$, $p < .0001$. White confederates rated participants' behavior toward them more positively ($M = 2.7$) than Latino confederates ($M = 0.5$). Because the background information had yet to be exchanged (and the confederates did not know their SES assignment until after the initial interaction), it was not surprising that there was no effect of SES or interaction.

After the information exchange, we observed the same main effect for confederates' ethnicity and a significant interaction with SES background, $F(1, 53) = 4.38$, $p < .05$. Simple effects tests confirmed that White confederates who described their background as high SES rated the participants' behavior toward them significantly more positively than did White confederates assigned to the low-SES condition (Ms = 3.0 vs. 1.9), $F(1, 53) = 4.48$, $p < .04$. In contrast, Latino confederates who described their background as low SES rated the participants' behavior toward them more positively than Latino confederates who described their backgrounds as high SES (Ms = 1.4 vs. 0.9), though not significantly, $F(1, 53) = 1.13$, *ns*. Though obviously not blind to condition, the confederates were unaware of the study hypotheses and yet observed changes in participants' behavior on the basis of our predicted effects, significantly so for White confederates. When confederates' ethnicity and SES conditions matched cultural expectations, they judged the participants' behavior toward them to be more positive than when ethnicity and SES contradicted expectations.

Discussion

Experiment 2 yielded strong support for the main predictions regarding interacting with partners who were counterstereotypical versus stereotypical. Participants paired with counterstereotypical partners exhibited CV responses consistent with threat, rated their partners less positively, and performed worse than participants paired with stereotypical partners. In this study, the speech and word-finding tasks involved actual face-to-face interactions rather than imagined interaction (the speech in Experiment 1) or audio information alone (word-finding task in Experiment 1). We believe that these changes resulted in a more impactful interaction than in the previous experiment. Taken together, these experiments support the hypothesis that interactions with partners who violate expectations engender CV threat responses, poorer task performance, and are rated less positively than partners who are consistent with stereotypical expectations.

Though these findings are supportive of our predictions, because of the specific traits and characteristics that we chose (ethnicity and SES), they are also consistent with theories emphasizing the role of the status quo in social hierarchies. System justification theory, in particular, would predict the effects observed in these

experiments (see Jost et al., 2004, for a review). This theory challenges the notion that intergroup relations are necessarily defined by antipathy toward minority or subordinate groups (e.g., Kay & Jost, 2003) but rather that individuals are motivated to defend and justify the existing social order, whatever it is. In this respect, system justification theory would predict that people who conform to extant social ordering on the basis of ethnicity or status would satisfy the desire to believe in a just and predictable world. In contrast, when people defy the expected social ordering, they threaten the status quo; the resulting responses would include discomfort, anxiety, and possibly motivations to oppress the mavericks. It is in these cases of defying the social order that antipathy toward expectancy-violating out-group members (and nonconformist in-group members) may develop.

Another theory that is broadly consistent with these results is the social dominance perspective (Sidanius & Pratto, 1999; see Sidanius, Pratto, van Laar, & Levin, 2004, for a review). This perspective addresses the processes by which group-based hierarchies are established, maintained, and legitimized. The results here are consistent with a social dominance perspective in that members of subordinated groups who are counterstereotypical may threaten the hegemony of dominant group members.

Though our results are supportive of these theories, our primary goal of these experiments was to advance a more general argument regarding the effects of partners who violated expectations, not necessarily because they directly defied the social order but because they were simply unexpected in certain ways. Thus, Experiment 3 was designed to generalize beyond the specific ethnic and SES manipulations we used and be a more general test of our predictions. In designing Experiment 3, our goals were to develop a confederate profile that was (a) a combination of characteristics that would result in a truly novel interaction partner—one who would have been unlikely to be encountered previously—(b) less associated with negative stereotypes, and (c) not directly related to ideological or sociopolitical influences. Thus, our goal for Experiment 3 was to create a confederate profile that was atypical or surprising and not negative.

Experiment 3: Asians With Southern Accents

For Experiment 3, we selected two partner characteristics that, in conjunction, would create an unexpected or atypical profile. To accomplish this goal, we varied the interaction partner's ethnicity (White or Asian) and accent (U.S. southern accent or expected regional [Californian] accent—hereafter referred to as *local accent*).

Our main hypothesis was that participants who were paired with unexpected partners (Asians with southern accents) would exhibit threat relative to participants paired with expected partners (Whites with local accents). Two competing hypotheses relative to the remaining conditions (White southern accents and Asian local accents) were considered. The first was that threat would be observed only when characteristics (ethnicity and accent) came together in an unexpected way, leading to the prediction that all conditions would result in challenge except the Asian with a southern accent, which would elicit threat. The alternative prediction was that for each characteristic that differed from "typical," a monotonically additive effect would occur such that White partners with local accents would be the least threatening, followed by Asians with local accents, Whites with southern accents, and finally Asians with southern accents. This latter prediction is consistent with the interpersonal contact effects described in the introduction. Support for this prediction would result in two additive main effects, one for ethnicity and one for accent.

Method

Participants

Female participants ($N = 47$) received either course credit or $10 for their participation. The majority of the sample was White (75%), and the remaining participants were Latina (5), African American (4), and Asian/Filipino (3). Mean age was 19.6 years ($SD = 1.6$).

Confederates

We initially recruited four female Asian Americans (two Chinese Americans, one Korean American, and one Japanese American) and four female European American undergraduates as confederates. All were native Californians, with English as their first and primary language. Confederates were similar in age (19–21), average height and build, and wore the same type of clothes for the study (white t-shirt and jeans). For approximately 2 months prior to the beginning of the study, the confederates met weekly for speech training conducted by a native of Charleston, South Carolina. Confederates were trained together to optimize accent similarity. Two of the confederates—one Chinese American and one European-American—could not master the southern accent and were not used as confederates in the experiment. Once the southern accent was acceptable to the native speaker, we began the experiment.

The confederate's name was Jenny, and her background was almost identical across accent conditions. During the initial information exchange, she described herself as from Garden Grove, California (or Charleston, South Carolina), 20 years old, and a communications major. Her father worked for a grocery store chain, and her mother taught elementary school. Her older brother attended the University of Southern California (or University of South Carolina), and she had a younger sister in high school. Her hobbies included going to the movies and hiking, and her summer plans included getting a job as a peer counselor.

Measures

We added an item to the postword-finding questionnaire to assess perceptions of "typicality" of the partner, "My partner is a typical UCSB undergraduate," using the same response format as the other questions (-4 to $+4$). We omitted ethnicity and SES questions.

Behavioral Observation and Coding

Videotapes of the background information exchange were coded by six research assistants from a different university so that we could assess behaviors consistent with avoidance, freezing, inhibition, and affect. After training, 25% of the exchanges were scored by all of the coders to determine consistency. Once consistency was established, then each exchange was scored by two

coders. Interrater reliability was high ($\alpha > .90$ for all subcategories). We then averaged the two coders' responses for each participant. Coders rated the participant's body orientation, leaning, global movement, specific movement (hand, feet, and head), nodding, giggling, smiling, eye contact, questioning, and extent of verbal affirmation provided to the confederate (e.g., "That's great!"). These categories were coded separately for the time that the confederate was speaking and for the time that the participant was speaking. Coders also rated how much detailed information participants provided, how much it appeared the participant liked the confederate, and general affect that could be detected from the interaction using 15 items from the Positive and Negative Affect Schedule (PANAS; Watson, Clark, & Tellegen, 1988) on the basis of the most theoretically relevant emotions that were likely to appear during the interaction.[2]

Procedure

Other than the changes described above, Experiment 3 was conducted in the same manner as Experiment 2. When experimenters probed for suspicion at the end of the experiment, they also ascertained whether participants believed the confederate was an authentic participant and whether the accent was real. We use the following abbreviations to indicate conditions: Asians with southern accents (ASA), Asians with local accents (ALA), Whites with southern accents (WSA), and Whites with local accents (WLA).

Results

Participant Attrition

We excluded one exchange student participant who expressed in debriefing that she did not detect her partner's accent. Two participants were excluded because they expressed suspicion regarding whether their partners were confederates, and 1 participant had completed a similar study in our lab before (although during recruitment, she stated she had not). This attrition left 43 participants with video data that we coded for nonverbal behavior during the information exchange.

Additional participants were lost after the social interaction for various reasons. One was lost because the audiovisual equipment failed to connect prior to the speech, and 3 participants' physiological data were lost because of equipment problems. Hence, 39 participants remained with usable physiological data (42 with self-report and performance data).

CV Measures

Baseline differences. A MANOVA tested baseline physiological responses (VC, CO, and TPR) by condition to determine any initial effects of partner's ethnicity or accent. This analysis revealed no significant effects or interaction for partner's ethnicity or accent (all Fs < 1).

Goal relevance. Changes in VC and HR were analyzed by experimental conditions for all minutes of the speech and word-finding tasks. All changes in VC and HR were significantly different from zero (all ps $< .05$).

Challenge and threat: Speech delivery. Our primary hypothesis was that participants interacting with ASA partners would exhibit CV threat reactivity relative to participants interacting with WLA partners. Given that our primary hypothesis focused on the comparisons of these extreme conditions, we first conducted a priori contrast analyses comparing these two groups. These analyses revealed significant differences in the expected direction for both task minutes; Minute 1: Wilks's $\Lambda = .74$, $F(3, 33) = 3.80$, $p < .02$; Minute 2: Wilks's $\Lambda = .72$, $F(3, 28) = 3.62$, $p < .02$. Participants interacting with WLA partners relative to ASA partners exhibited larger increases in VC ($M = 24.6$ vs. $M = 9.8$), $F(1, 37) = 8.48$, $p < .01$, and CO ($M = 1.8$ vs. $M = 0.1$), $F(1, 37) = 10.49$, $p < .01$, and decreases in TPR ($M = -224.8$ vs. $M = 2.2$), $F(1, 37) = 6.67$, $p < .02$. These analyses confirm that participants interacting with ASA partners exhibited CV responses consistent with threat compared with participants interacting with WLA partners.

To test the secondary predictions, we examined the main effects and interaction of ethnicity and accent of the confederates. In general, we observed support for the additive prediction. All three univariate analyses revealed significant effects for ethnicity: VC, $F(1, 37) = 4.51$, $p < .05$; CO, $F(1, 37) = 5.65$, $p < .03$; and TPR, $F(1, 37) = 4.10$, $p < .05$, such that Asian (vs. White) confederates engendered less VC, lower CO, and higher TPR, consistent with the threat constellation. The effects for accent were weaker but in the direction of greater threat when confederates had a southern accent than when they did not: VC, $F(1, 37) = 3.48$, $p < .07$; CO, $F(1, 37) = 4.24$, $p < .05$; and TPR, $F(1, 37) = 2.28$, $p < .14$. The weaker effects for the accent condition are probably not surprising given that the confederates were merely passive observers during the speech and did not speak at all but merely watched the participant deliver her speech. The interactions between ethnicity and accent were not significant.

Challenge and threat: Word-finding task. We then examined CV reactivity during the word-finding task, focusing on planned contrasts. As predicted, and consistent with the data from the speech task, interactions with ASA confederates resulted in CV reactivity consistent with threat relative to the WLA condition, $F(3, 32) = 8.29$, $p < .0003$, Wilks's $\Lambda = .56$. Participants interacting with ASA relative to WLA confederates exhibited smaller increases in VC, $F(1, 37) = 9.83$, $p < .01$, and CO, $F(1, 37) = 19.94$, $p < .0001$, and greater increases in TPR, $F(1, 37) = 18.89$, $p < .0001$; that is, they showed a pattern consistent with threat (see Figure 4).

We then examined data from all conditions to test our secondary predictions. Consistent with the speech data, the effects appeared to be additive. Specifically, both multivariate main effects were significant; ethnicity, $F(3, 32) = 6.32$, $p < .002$, Wilks's $\Lambda = .63$; accent, $F(3, 32) = 3.23$, $p < .03$, Wilks's $\Lambda = .77$; and the multivariate interaction was not. Participants interacting with Asian (compared with White) partners and those interacting with partners with southern accents (compared with local accents) exhibited weaker VC, lower CO, and higher TPR. These physiological data provide evidence that participants who interacted with unexpected or atypical partners exhibited greater threat reactivity relative to individuals who interacted with expectancy-congruent partners. The notion that the partner characteristics have an addi-

[2] PANAS adjectives used were *distressed*, *excited*, *upset*, *scared*, *hostile*, *enthusiastic*, *interested*, *irritable*, *alert*, *ashamed*, *nervous*, *attentive*, *jittery*, *active*, and *afraid*.

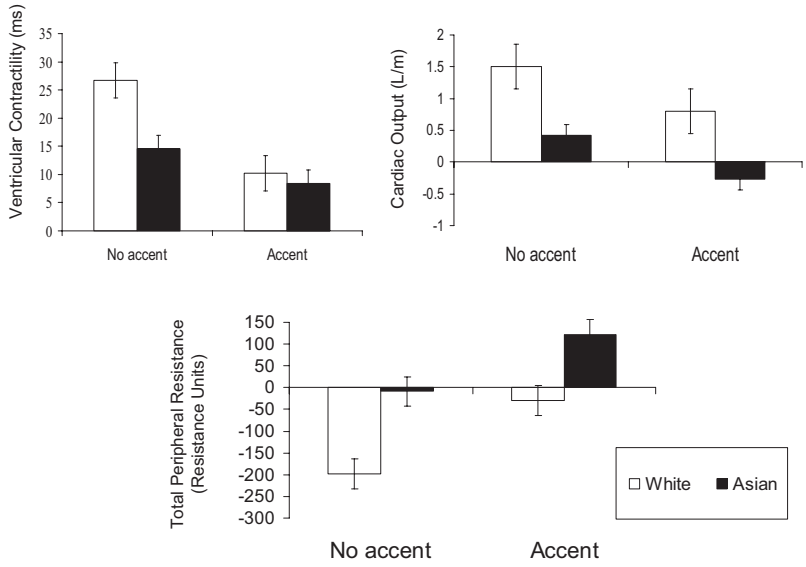

Figure 4. Means and standard errors of cardiovascular reactivity during the first minute of the word-finding task by ethnicity and accent of the partner in Experiment 3.

tive effect such that the presence of atypical characteristics leads to more threat, earned support.

Performance: Word-Finding Task

The number of words found during the word game was used as the primary dependent variable to examine the effects of partner's characteristics on performance. We first tested the prediction that participants paired with ASA partners would find fewer words than participants paired with WLA partners. The planned contrast yielded the significant and predicted effect, $F(1, 40) = 7.21$, $p < .01$. Participants paired with ASA partners found fewer words than participants paired with WLA partners ($M = 10.8$ vs. 16.3) (see Figure 5).

Regarding the other cells, the main effect for accent was significant, $F(1, 40) = 5.14$, $p < .03$, such that participants paired with southern accent partners found fewer words than those paired with local accent partners. Although participants paired with Asian partners found fewer words than those paired with White partners, the effect was not significant, $F(1, 40) = 2.88$, $p < .10$, nor was the interaction. In general, ASA and WLA conditions resulted in the largest difference between conditions, and the other conditions yielded values between these extremes (WSA: $M = 13.4$; ALA: $M = 14.2$).

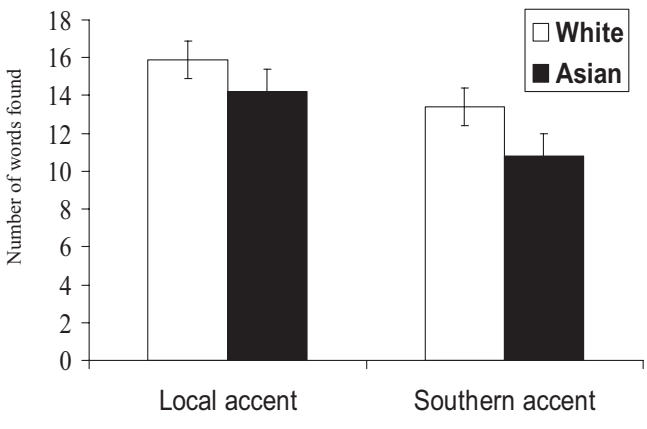

Figure 5. Means and standard errors of the performance measures from the word-finding task (number of correctly generated words) by ethnicity and accent of the partner in Experiment 3.

Behavioral Observation

We then turned to the behavioral data from the information exchange. First, to reduce the data, we conducted an exploratory factor analysis with varimax rotation, which yielded three clear factors, with eigenvalues exceeding 1.00. The first factor consisted of somatic movement items (e.g., global and specific body movement, nodding), the second factor represented observable positive behavior (e.g., giggling, smiling, verbal affirmations), and the last factor was body orientation (e.g., leaning, general body orientation relative to the confederate). Variables from the three factors were combined into three scales: Movement ($\alpha = .85$), Positive Behavior ($\alpha = .78$), and Body Orientation ($\alpha = .89$). We also created a measure of positive affect ($\alpha = .69$) and negative affect ($\alpha = .73$) from the PANAS items.

Our primary prediction was that interactions with ASA partners would result in behavior consistent with the threat constellation (e.g., freezing, absence of positive behavior, and avoidance) compared with WLA partners. The analysis of participants' body movements during the social interaction yielded a significant effect for the planned contrast, $F(1, 41) = 10.21$, $p < .003$. As expected, participants interacting with ASA partners had less somatic activity ($M = 2.4$) than did participants interacting with WLA partners ($M = 3.3$). We also observed a significant effect for positive behavior, $F(1, 41) = 6.14$, $p < .02$. Again, the ASA partner engendered less observable positive behavior ($M = .79$)

than did WLA partners ($M = 1.3$). Contrasts with respect to body orientation were not significant, though, admittedly, body orientation was greatly circumscribed during the information exchange. We placed the chairs at 90°angles to each other, the chairs were rigid—not allowing for easy manipulation of one's body beyond sitting in the chair facing forward—and difficult to move. Finally, negative affect was not significant, but video coders did rate participants' positive affect higher in interactions with WLA partners ($M = 2.8$) than interactions with ASA partners ($M = 2.4$), but this effect only approached significance ($p < .10$).

Next, we examined the secondary predictions regarding the effect of one discrete atypical category membership. Recall that we suggested two possible results: (a) a significant interaction between ethnicity and accent such that the ASA condition would differ from the other three categories or (b) two main effects in which confederates with accents and Asian confederates would engender relatively greater threat. In general, some support was garnered for the additive prediction, but the interactive prediction was also observed, albeit in a direction not hypothesized. We observed (one significant and one marginal) main effects for movement such that less movement was observed in the presence of Asian partners and partners with southern accents than White partners and partners with local accents: ethnicity, $F(1, 41) = 6.57$, $p < .02$, and accent, $F(1, 41) = 3.00$, $p < .10$. The interaction was not significant.

However, we also observed significant Ethnicity × Accent interactions for positive behavior and positive affect. WLA partners were associated with the greatest positive behavior, $F(1, 41) = 8.02$, $p < .01$, and positive affect, $F(1, 41) = 9.20$, $p < .01$, and the other three conditions (WSA, ALA, and ASA) did not differ from each other. There were no significant main or interaction effects for body orientation or negative affect.

Self-Report Ratings

We combined six questions that indicated how well the dyad performed (team quality) during the word-finding task ($\alpha = .76$). This index, with larger values indicating better perceived team performance, was then submitted to planned comparisons and ANOVAs. Participants paired with ASA confederates rated the team quality during the word-finding task lower ($M = 1.1$) than did participants in the WLA condition ($M = 2.3$), $F(1, 40) = 5.77$, $p < .03$ (see Table 1). The ANOVA revealed a significant interaction, $F(1, 40) = 4.58$, $p < .04$; participants with ASA confederates differed from the other three conditions.

We averaged participants' ratings of their partner's positive (e.g., trustworthy, likable, independent, $\alpha = .83$) and (reverse coded) negative traits (e.g., unattractive, unintelligent, unfriendly, $\alpha = .72$) and then averaged these indexes together to obtain a general measure of participants' positive ratings of their partner. Only a significant interaction emerged, $F(1, 40) = 4.72$, $p < .04$. ASA partners were rated the least positive and significantly differed from WLA and ALA partners. Participants' ratings of how cooperative the word-finding task was yielded a similar interaction, $F(1, 40) = 4.51$, $p < .04$. Again, ASAs were rated the least cooperative, but they only significantly differed from the ALA condition.

The final question asked participants to rate how typical her partner was, which yielded the expected significant difference between the ASA and WLA conditions, $F(1, 40) = 16.72$, $p < .0001$, with ASA partners rated as less typical ($M = -1.5$) than WLA partners ($M = 2.2$). This time, the overall pattern yielded two main effects rather than an interaction such that Asian partners and those with accents were rated as less typical than partners who were White and had no accent: ethnicity, $F(1, 40) = 5.97$, $p < .04$; accent, $F(1, 40) = 11.33$, $p < .002$.

Intercorrelations Among Dependent Variables

To better understand how the results varied on the basis of the specific measurement, we explored the relationships among our dependent variables. Table 2 presents the correlations among CV reactivity variables from the word task, behavioral observations, performance, and participants' self-reports. The same analyses were repeated using CV reactivity from the speech task, which yielded similar, albeit slightly lower magnitude, correlations. As can be seen, relationships among the CV data and the behavioral observation data are significant and moderately large. This is

Table 1
Participants' Mean and Standard Deviation Ratings After the Word-Finding Task by Confederates' Ethnicity and Accent in Experiment 3

Participants' ratings postword task	Confederate							
	White local accent		White southern accent		Asian local accent		Asian southern accent	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Team quality	$2.3_{a,1}$	1.0	2.2_{1}	0.5	2.4_{1}	1.0	$1.1_{b,2}$	1.3
Confederates' traits	$0.7_{1,2}$	1.8	1.5_{2}	2.1	1.6_{2}	2.1	-0.4_{1}	2.0
Cooperative	$1.2_{1,2}$	2.5	$2.0_{1,2}$	1.9	2.4_{1}	1.5	0.5_{2}	2.1
Competitive	−1.9	1.8	−1.2	2.4	−1.0	2.5	−0.5	1.9
"Partner is typical. . ."	$2.2_{a,1}$	1.6	$-0.1_{2,3}$	2.4	0.5_{2}	2.6	$-1.5_{b,3}$	1.5

Note. Planned comparisons between White-local accent and Asian-southern accent conditions were conducted first, followed by analysis of variance. Confederates' positive and (reverse-scored) negative traits are averaged so that higher scores indicate more positive trait ratings. Significant differences ($p < .05$) between White-local accent and Asian-southern accent confederates are indicated by different superscript letters across rows. Post hoc differences are indicated by different subscript numbers across rows. Letters and numbers indicating significant differences are provided only when differences existed with that variable. Scale ranges from −4 to +4.

Table 2
Correlations Among Study Dependent Variables in Experiment 3

	Observer rating			Cardiovascular reactivity				Self-report		
Dependent variable	1	2	3	4	5	6	Perform: 7	8	9	10
1. Move.	—									
2. Orient.	.06	—								
3. Positive behavior	.55***	.18†	—							
4. VC	.44**	.28*	.34	—						
5. CO	.47*	.19	.37*	.80***	—					
6. TPR	−.66*	−.12	−.45**	−.67***	−.71*	—				
7. Words	.06	.03	.21	.18	.26†	−.35*	—			
8. Quality	−.09	.17	.07	.16	.14	−.20	.29†	—		
9. Typical	.33†	−.03	.20	.46**	.50**	−.44**	.31†	−.13	—	
10. Traits	.02	.10	.18	.00	−.05	.00	.03	.58***	−.05	—

Note. Observer ratings are taken from the background information at the beginning of the study. All other data are taken from the word-finding task. Move. = movement; Orient. = body orientation; VC = ventricular contractility; CO = cardiac output; TPR = total peripheral resistance; Words = number of words found during the word-finding task; Quality = participants' ratings of team quality; Typical = participants' ratings of how "typical" their partner was; Traits = participants' ratings of positive and negative traits combined with negative traits reverse coded such that higher scores indicate more positive and less negative traits ascribed to their partners.
† $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

noteworthy not only because these measures are distinct in terms of the type of data, instrumentation, and quantification but also because of temporal factors (measures are separated in time by over 1 hr). The basic finding is that behavioral data from the information exchange were related to changes in CV responses during the word-finding task such that more freezing and less positive behavior during the face-to-face interactions were associated with CV threat responses during the cooperative task. There were other noteworthy relationships between partner's typicality ratings and CV responses and performance, which are explored using mediation models below.

Correlations among self-report responses (partner's traits and team quality) and the less consciously controlled measures (CV reactivity and behavioral observation data) were not significant. Within the intergroup domain, we and other researchers (Blascovich et al., 2002; Mendes et al., 2002; Vanman, Paul, Ito, & Miller, 1997) have noted these types of (non) relationships between less and more consciously controlled measures. Typically, it is suggested that physiological variables are related to more automatic or reflexive responses, and self-reports are related to more deliberate or consciously controlled responses.

Mediational Analysis

Given the bivariate correlations among CV responses, performance, and partner's typicality ratings, we explored whether the physiological responses mediated the relationship between partner's typicality ratings and subsequent performance. That is, could the physiological responses exhibited during the word-finding task explain the relationship between partners' typicality ratings and task performance?

To test this idea, we first created a single index of CV reactivity by standardizing (z score) TPR and CO reactivity then multiplying CO reactivity by −1 so that increases in this measure would indicate more threat. Finally, we added the two standardized scales together such that the higher the scores, the greater the threat response (increased TPR and decreased CO). We then conducted three regression analyses to test for mediation, as outlined by Baron and Kenny (1986) and more recently by MacKinnon and colleagues (MacKinnon, Lockwood, Hoffman, West, & Sheets, 2002). The first regression equation used participants' ratings of how typical their partners were to predict performance, which yielded a (near) significant effect, $t(34) = 1.94$, $p < .06$, suggesting that the less typical partners were perceived, the worse the participant performed (see Figure 6). The next regression equation estimated changes in CV threat using typical ratings and yielded a significant effect, $t(34) = -3.70$, $p < .001$—the less typical the partner was perceived, the greater the CV threat response. Finally, the last regression equation estimated performance using both typical ratings and CV threat responses. The link between CV threat and performance was just short of significance, $t(34) = -1.91$, $p < .065$, and the relationship between typical ratings and performance was reduced once CV threat was included in the model, $t(34) = 0.66$, *ns*, Sobel test = 1.75, $p < .08$.

These analyses yielded several paths short of significance, which is most likely because the path from the independent variable to the mediator (typical ratings to CV threat) is so large that there was little variance left over to predict an independent relationship from the mediator to the dependent variables (CV threat to

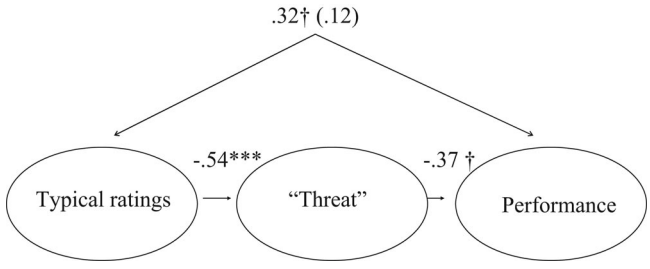

Figure 6. Test of mediation linking partner's typicality ratings, performance, and cardiovascular reactivity in Experiment 3 (Sobel test = 1.75, $p < .08$). † $p < .10$. *** $p < .001$.

performance). This type of multicollinearity is often seen when the independent variable and mediator are strongly related (D. P. MacKinnon, August 8, 2005, personal communication). Nevertheless, the basic relationship between these variables sheds some light on how performance can be impaired with out-group members, and this impairment may in part be attributed to physiological arousal experienced during the interaction.

Discussion

Experiment 3 yielded strong support for the primary hypothesis that partners who violated expectations because they were atypical (Asians with southern accents) would engender more threat, less somatic activity (i.e., freezing), less positive affect, and poorer performance compared with partners who were most typical (Whites with local accents). Mixed support was observed for the competing secondary hypotheses regarding whether partner's characteristics would manifest their effects additively or interactively. Physiological and performance data supported the additive model; that is, with each trait that differed from expectations, participants exhibited more threat responses. However, the self-report data yielded more support for the interactive model such that participants rated Asian southern accent partners least positively compared with all other conditions. In addition, the observational data suggested that White local accent partners were treated more positively than partners in the other conditions.

We also examined the correlations among our dependent variables and observed some noteworthy relationships. The most notable findings were the strong relationships between the less consciously controlled measures—behavioral observation data and physiological measures. The biased and defeat-related behaviors (freezing, avoidance posture, less positive affect, and less smiling) were related to lower cardiac responses and higher vascular responses (threat). Of the possible nine correlations across the CV measures and behavioral measures, seven were significant and in the predicted direction. As this finding is the first that we know of to demonstrate specific nonverbal behaviors relating to CV responses, indicating challenge and threat states, we believe that these results add to the constellation of CV responses that differentiate challenge and threat states and show important behavioral concomitants consistent with the phenomenology of challenge and threat.

The more consciously controlled measures (i.e., self-report measures) were only correlated with other self-report measures. For example, participants' ratings of the quality of the team performance were associated with positive attributes ascribed to their partner but no other primary dependent variables. These data demonstrate the importance of multiple measures that vary with respect to conscious control, specifically in domains in which self-report responses may be easily distorted or demand characteristics and social desirability concerns are high. The value of these less consciously controlled measures is best characterized by Brewer and Brown (1998), who point out that "in intergroup interactions, nonverbal behaviors are more likely to reflect negative affect than is the content of the exchange" (p. 575).

We observed some support for the mediating role of physiology linking person perception and task performance. Though several pathways linking these three variables were short of significance, there was some suggestive evidence that online physiological responses can partly explain the relationship between how participants perceived the unusualness of their partner and how they performed at a cognitive task. Future studies should continue to examine the links among person perception, physiology (or other online, less controlled measures), and performance as a way to understand the underlying processes that occur during social interactions.

It is important to note that although the study was designed to examine the effects of expectancy violations independent of extant negative stereotypes, the specific identity used, Southerners, may have inadvertently manipulated perceived lower SES or have evoked other negative stereotypes. Although we cannot completely rule out the possibility that the southern-accent condition evoked negative stereotypes, we do believe that from a comparative perspective, participants in our sample did not hold as many negative stereotypes for Southerners as they did for Latinos. In support of this, we surveyed undergraduates not involved in any of the above studies and had them generate three adjectives that came to mind when thinking of a 20-year-old female undergraduate who now lives in California but was either from the U.S. South or Mexico. From a comparative perspective, the data were clear. Participants generated more positive words when considering a 20-year-old southern female than a 20-year-old Latina (a score of 3 would indicate all positive words; U.S. Southerner: $M = 2.4$, $SD = .6$; Latina: $M = 1.4$, $SD = 1.0$), $t(32) = 4.60$, $p < .0001$. The most common positive traits were "pretty" and "polite" for Southerners; the most common positive trait for Latinas was "hardworking." We also asked in a paired choice task what group of people did participants believe had more negative stereotypes associated with them: U.S. Southerners or Latinos. Over 88% of the respondents chose Latinos as having more negative stereotypes associated with them. Although these data do not eliminate all ambiguity regarding how participants viewed the Southern confederate, these data strongly suggest that, compared with Latinos, participants perceived far fewer negative stereotypes associated with Southerners.

General Discussion

Three experiments supported the hypothesis that perceivers are threatened by partners who violate expectations. Experiments 2 and 3 also revealed that participants paired with expectancy-violating partners performed worse on a performance task and rated their partners more negatively than did participants paired with partners who did not violate expectations. Experiment 3 further demonstrated that participants paired with the former exhibited less somatic activity (i.e., motoric freezing) and displayed less positive behavior toward their partner than participants paired with the latter. These experiments demonstrate that atypical, and hence unfamiliar, interaction partners can engender more malignant physiological responses, decrements in cognitive performance, and defeat-related and negative behavior even during cooperative interactions.

We theorized that such unfamiliarity can increase uncertainty and required cognitive effort and ultimately result in greater demand evaluations and threat responses. In previous work, Bartholow et al. (2001) showed that expectancy-violating information required greater cognitive processing. To the extent that participants interacting with expectancy-violating partners also recruited more cognitive and attentional resources during the interaction

than those who were paired with the expectancy-consistent partners would we speculate that the sum total of their working memory and attentional resources was more taxed. The depleted resources are then likely to result in reduced resource evaluations leading to threat states and can partly explain the performance impairments observed. Also, in support of the contention that uncertainty leads to threat, we observed significant correlations between participants' ratings of how typical their partner was and CV responses. The less typical partners were perceived, the greater the threat responses exhibited during a cooperative social interaction.

Relationships to Previous Research

As described in the introduction, our previous research showed a strong effect of out-group bias on physiological responses, indicating threat. Specifically, African American confederates engendered CV responses consistent with threat (in White participants) compared with White confederates. However, a general out-group bias was not present with Latino confederates. Because in previous work we observed an effect of intergroup contact moderating CV responses with African Americans such that the more contact resulted in less threat, we reasoned that the large numbers of Latinos on campus and in the surrounding community would increase exposure and likelihood of intergroup contact with Latinos. Indeed, our pilot data confirmed that level of suspected contact. However, the contact levels reported in the pilot data were specifically with lower SES Latinos, leading us to the prediction that only when Latino partners were consistent with the stereotype or the well-known subtype, in this case low SES, did the interactions with them become routinized. However, when the Latino was unexpected or counterstereotypical, he or she would be unusual and hence threatening. From the survey data we presented, we argued that low-SES Latinos are likely not threatening because they were not particularly rare in this population. Indeed, Latinos in the city in which the experiments took place comprised over 34% of the population, and 21% of the student body at the time identified as Chicano or Latino (the highest minority group on campus).

What is the value of demonstrating that unfamiliar others engender psychological and physiological states of threat? We believe these results demonstrate the importance of familiarity, in general, and contact, in particular, when considering ways to decrease intergroup bias. The contact hypothesis articulates that positive interactions that are characterized by equal status and common goals are likely to reduce intergroup bias (see Dovidio, Gaertner, & Kawakami, 2003, for a review). Two related and important factors associated with reducing bias are familiarity and intergroup friendships. Familiarity can reduce uncertainty and provide a sense of shared reality and knowledge of a social script to follow during exchanges. Similarly, intergroup friendships are related to significantly lower levels of bias toward out-group members (Pettigrew & Tropp, 2006). The results reported here underscore the importance of expectancies in motivational and physiological states associated with interacting with strangers. To the extent that partners were counterstereotypical (or atypical), participants experienced more negative responses. Given the possibility of peripheral feedback of the autonomic signals, we might speculate that antipathy toward out-groups could possibly grow from these awkward or uncomfortable social interactions, particularly if individuals are more inclined to remember the "felt states" over any positive social exchanges.

Future Directions

Future research using the paradigm described here could test whether prior exposure to (or contact with) expectancy-violating partners could reduce threat as a way of testing the unfamiliarity–threat link. Another direction would be to examine the extent to which exceeding cognitive processing capacities (e.g., using cognitive busyness strategies) influences physiological responses with expectancy-violating and confirming partners as a way to test the contribution of increased cognitive effort. Initial evidence in our laboratory suggests that individuals delivering speeches while under cognitive load (e.g., rehearsing a nine-digit number) show the constellation of autonomic nervous system threat responses and increased cortisol responses in comparison with those giving speeches without cognitive load. These preliminary data, coupled with recent EEG data (Bartholow et al., 2001), begin to specify the cognitive components and neurological bases underlying interactions with unexpected or counterstereotypical partners.

These data may also be relevant to how discrimination affects physical and mental health of minorities and other stigmatized groups (Clark, Anderson, Clark, & Williams, 1999; Krieger, 2000; Link & Phelan, 2001). We have found majority group members exhibit threat responses, less positive behavioral responses, and performance impairments when interacting with counterstereotypical or atypical group members. To the extent that these responses then, in turn, affect perceptions, stress appraisals, and physiological reactivity of minority or stigmatized group members can researchers begin to demonstrate the cyclical effects of intergroup anxiety. By definition, minority group members interact more with majority group members than vice versa (see also Frable, Platt, & Hoey, 1998). If these intergroup interactions create more cumulative wear and tear on the body, then this could be a potential link to how discrimination gets under the skin to affect mental and physical health.

It is interesting to note that racial and ethnic differences in mental and physical health persist across the SES gradient (Williams, 1999). That is, even though significant variance in racial disparities in health can be attributed to differences in SES, racial/ethnic differences can still be observed at equivalent levels of SES. One possible mechanism suggested by this research may be the treatment and reception minority group members receive who do advance socioeconomically. In Experiments 1 and 2, when Latinos presented themselves as high SES, participants were threatened, performed worse, and rated them more negatively. Thus, the possibility of advancing in society in terms of the SES standards of income, education, and prestige did not translate into greater acceptance but rather into more negative reactions. It may be plausible to view the racial/ethnic differences even at higher levels of SES as partly a function of the negative treatment minorities may receive when they successfully climb the SES ladder.

Conclusions

We have demonstrated that participants interacting with partners who violate stereotypical expectations exhibit threat responses, as

assessed cardiovascularly, have poorer cognitive performance, evaluate their partners more negatively, and have perceptible defeat-related and negative behavior. Because the putative pathway through which these responses occur involves unfamiliarity with interaction partners, the research here provides groundwork for understanding the role of unfamiliarity—independent of negative affect and stereotypes—in intergroup conflict, antipathy toward out-groups, and preservation of the status quo. In addition, the findings suggest underlying cognitive and affective mechanisms as well as physiological consequences. To the extent that participants experience threat when interacting with expectancy-violating group members because of a lack of exposure, researchers can begin to understand the process by which unfamiliar or unusual others become the victims of discrimination.

References

Aronson, E., Wilson, T. D., & Brewer, M. B. (1998). Experimentation in social psychology. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology* (pp. 99–142). New York: McGraw-Hill.

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.

Bartholow, B. D., Fabiana, M., Gratton, G., & Bettencourt, B. A. (2001). A psychophysiological examination of cognitive processing of and affective responses to social expectancy violations. *Psychological Science, 12,* 197–204.

Bartholow, B. D., Pearson, M. A., Gratton, G., & Fabiani, M. (2003). Effects of alcohol on person perception: A social cognitive neuroscience approach. *Journal of Personality and Social Psychology, 85,* 627–638.

Bettencourt, B. A., Dill, K. E., Greathouse, S. A., Charlton, K., & Mulholland, A. (1997). Evaluations of ingroup and outgroup members: The role of category-based expectancy violations. *Journal of Experimental Social Psychology, 33,* 244–275.

Blanchard, R. J., Flannelly, K. J., & Blanchard, D. C. (1986). Defensive behavior of laboratory and wild Rattus norvegicus. *Journal of Comparative Psychology, 100,* 101–107.

Blascovich, J., & Mendes, W. B. (2000). Challenge and threat appraisals: The role of affective cues. In J. Forgas (Ed.), *Feeling and thinking: The role of affect in social cognition* (pp. 59–82). Paris: Cambridge University Press.

Blascovich, J., Mendes, W. B., Hunter, S. B., Lickel, B., & Kowai-Bell, N. (2001). Perceiver threat in social interactions with stigmatized others. *Journal of Personality and Social Psychology, 80,* 253–267.

Blascovich, J., Mendes, W. B., & Seery, M. D. (2002). Intergroup threat: A multi-method approach. In D. M. Mackie & E. R. Smith (Eds.), *From prejudice to intergroup emotions: Differentiated reactions to social groups* (pp. 89–109). New York: Psychology Press.

Blascovich, J., & Tomaka, J. (1996). The biopsychosocial model of arousal regulation. *Advances in Experimental Social Psychology, 28,* 1–51.

Brewer, M. B., & Brown, R. J. (1998). Intergroup relations. In D. T. Gilbert & S. T. Fiske (Eds.), *Handbook of social psychology* (4th ed., pp. 554–594). New York: McGraw-Hill.

Cannon, W. B. (1929). *Bodily changes in pain, hunger, fear, and rage*. Oxford, England: Appleton-Century-Crofts.

Clark, R., Anderson, N. B., Clark, V. R., & Williams, D. R. (1999). Racism as a stressor for African Americans: A biopsychosocial model. *American Psychologist, 54,* 805–816.

DePaulo, B. M., & Friedman, H. S. (1998). Nonverbal communication. In D. T. Gilbert & E. T. Fiske (Eds.), *Handbook of social psychology* (4th ed., pp. 3–40). New York: McGraw-Hill.

Devine, P. G., Monteith, M. J., & Zuwerink, J. R. (1991). Prejudice with and without compunction. *Journal of Personality and Social Psychology, 60,* 817–830.

Dienstbier, R. A. (1989). Arousal and physiological toughness: Implications for mental and physical health. *Psychological Review, 96,* 84–100.

Dovidio, J. F., Gaertner, S. L., & Kawakami, K. (2003). Intergroup contact: The past, present, and the future. *Group Processes & Intergroup Relations, 6,* 5–21.

Dunton, B. C., & Fazio, R. H. (1997). An individual difference measure of motivation to control prejudiced reactions. *Personality and Social Psychology Bulletin, 23,* 316–326.

Epel, E. S., McEwen, B. S., & Ickovics, J. R. (1998). Embodying psychological thriving: Physiological thriving in response to stress. *Journal of Social Issues, 54,* 301–322.

Folkman, S., & Lazarus, R. S. (1985). If it changes it must be a process: Study of emotion and coping during three stages of a college examination. *Journal of Personality and Social Psychology, 48,* 150–170.

Folkow, B. (1993). Physiological organization of neurohormonal responses to psychosocial stimuli: Implications for health and disease. *Annals of Behavioral Medicine, 15,* 236–244.

Frable, D. E. S., Platt, L., & Hoey, S. (1998). Concealable stigmas and positive self-perceptions: Feeling better around similar others. *Journal of Personality and Social Psychology, 74,* 909–922.

Gao, G., & Gundykunst, W. B. (1990). Uncertainty, anxiety, and adaptation. *International Journal of Intercultural Relations, 14,* 301–317.

Gilbert, D. T. (1998). Ordinary personology. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., pp. 89–150). New York: McGraw-Hill.

Goodwin, S. A., & Fiske, S. T. (1996). Judge not, unless . . .: Standards for social judgments and ethical decision-making. In D. M. Messick & A. Terbrunsel (Eds.), *Codes of conduct: Behavioral research and business ethics* (pp. 117–142). New York: Russell Sage Foundation.

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

Grieve, P. G., & Hogg, M. A. (1999). Subjective uncertainty and intergroup discrimination in the minimal group situation. *Personality and Social Psychology Bulletin, 25,* 926–940.

Gudykunst, W. B. (1984). *Communicating with strangers: An approach to intercultural communications*. Reading, MA: Addison-Wesley.

Hebl, M. R., Tickle, J., & Heatherton, T. F. (2000). Awkward moments in interactions between stigmatized and nonstigmatized individuals. In T. F. Heatherton & R. F. Kleck (Eds.), *Social psychology of stigma* (pp. 275–306). New York: Guilford Press.

Henry, J. P. (1986). Neuroendocrine patterns of emotional response. In R. Plutchik & H. Kellerman (Eds.), *Emotion: Theory, research, and experience: Vol 3. Biological foundations of emotion* (pp. 37–60). Orlando, FL: Academic Press.

Islam, M. R., & Hewstone, M. (1993). Dimensions of contact as predictors of intergroup anxiety, perceived out-group variability, and out-group attitude: An integrative model. *Personality and Social Psychology Bulletin, 19,* 700–710.

Jackson, L. A., Sullivan, L. A., & Hodge, C. N. (1993). Stereotype effects of attributions, predictions, and evaluations: No two social judgments are quite alike. *Journal of Personality and Social Psychology, 65,* 69–84.

Jones, E. E., Farina, A., Hastorf, A. H., Markus, H., Miller, D. T., & Scott, R. A. (1984). *Social stigma: The psychology of marked relationships*. New York: Freeman.

Jost, J. T., Banaji, M. R., & Nosek, B. A. (2004). A decade of system justification theory: Accumulated evidence of conscious and unconscious bolstering of the status quo. *Political Psychology, 25,* 881–920.

Jost, J. T., Pelham, B. W., & Carvallo, M. R. (2002). Non-conscious forms of system justification: Implicit and behavioral preferences for higher status groups. *Journal of Experimental Social Psychology, 38,* 586–602.

Jussim, L., Coleman, L. M., & Lerch, L. (1987). The nature of stereotypes: A comparison and integration of three theories. *Journal of Personality and Social Psychology, 53,* 536–546.

Kamarck, T. W., Jennings, J. R., Debski, T. T., Glickman-Weiss, E., Johnson, P. S., Eddy, M. J., & Manuck, S. B. (1992). Reliable measures of behaviorally-evoked cardiovascular reactivity from a PC-based test battery: Results from student and community samples. *Psychophysiology, 29,* 17–28.

Kay, A. C., & Jost, J. T. (2003). Complementary justice: Effects of "poor but happy" and "poor but honest" stereotype exemplars on system justification and implicit activation of the justice motive. *Journal of Personality and Social Psychology, 85,* 823–857.

Kelsey, R. M., Blascovich, J., Tomaka, J., Leitten, C. L., Schneider, T. R., & Wiens, S. (1999). Cardiovascular reactivity and adaptation to recurrent psychological stress: Effects of prior task exposure. *Psychophysiology, 36,* 818–831.

Kelsey, R. M., & Guethlein, W. (1990). An evaluation of the ensemble averaged impedance cardiogram. *Psychophysiology, 27,* 24–33.

Kernahan, C., Bartholow, B. D., & Bettencourt, B. A. (2000). Effects of category-based expectancy violation on affect-related evaluations: Toward a comprehensive model. *Basic and Applied Social Psychology, 22,* 85–100.

Krieger, N. (2000). Discrimination and health. In L. F. Berkman & I. Kawachi (Eds.), *Social epidemiology* (pp. 36–75). Oxford, England: Oxford University Press.

LeDoux, J. E. (2000). Emotion circuits in the brain. *Annual Review of Neuroscience, 23,* 155–184.

Link, B. G., & Phelan, J. C. (2001). Conceptualizing stigma. *Annual Review of Sociology, 27,* 363–385.

Lundberg, U., & Frankenhaeuser, M. (1980). Pituitary-adrenal and sympathetic adrenal correlates of distress and effort. *Journal of Psychosomatic Research, 24,* 125–130.

MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods, 7,* 83–104.

Marks, A. (1987). *Fears, phobias, and rituals: Panic, anxiety, and their disorders*. London: Oxford University Press.

McConahay, J. B. (1986). Modern racism, ambivalence, and The Modern Racism Scale. In J. F. Dovidio & S. L. Gaertner (Eds.), *Prejudice, discrimination, and racism* (pp. 91–125). San Diego, CA: Academic Press.

Mendes, W. B., Ayduk, O., Gyurak, A., & Akinola, M. (2006). *Neuroendocrine responses associated with appraisals of challenge and threat*. Manuscript in preparation.

Mendes, W. B., Blascovich, J., Lickel, B., & Hunter, S. (2002). Challenge and threat during social interactions with White and Black men. *Personality and Social Psychology Bulletin, 28,* 939–952.

Olson, J. M., Roese, N. J., & Zanna, M. P. (1996). Expectancies. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 211–238). New York: Guilford Press.

Pettigrew, T. F., & Tropp, L. R. (2006). A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology, 90,* 751–783.

Phelps, E. A., O'Connor, K. J., Cunningham, W. A., Funayama, E. S., Gatenby, J. C., Gore, J. C., et al. (2000). Performance on indirect measures of race evaluation predicts amygdala activation. *Journal of Cognitive Neuroscience, 12,* 729–738.

Richeson, J. A., & Shelton, J. N. (2003). When prejudice does not pay: Effects of interracial contact on executive function. *Psychological Science, 14,* 287–290.

Rotenberg, V. S., Sirota, P., & Elizur, A. (1996). Psychoneuroimmunology: Searching for the main deteriorating psychobehavioral factor. *Genetic, Social, and General Psychology Monographs, 122,* 329–346.

Scherer, K. R., Zentner, M. R., & Stern, D. (2004). Beyond surprise: The puzzle of infants' expressive reaction to expectancy violation. *Emotion, 4,* 389–402.

Sherwood, A., Allen, M. T., Fahrenberg, J., Kelsey, R. M., Lovallo, W. R., & van Dooren, L. J. P. (1990). Methodological guidelines for impedance cardiography. *Psychophysiology, 27,* 1–23.

Sidanius, J., & Pratto, F. (1999). *Social dominance: An intergroup theory of social hierarchy and oppression*. New York: Cambridge University Press.

Sidanius, J., Pratto, F., van Laar, C., & Levin, S. (2004). Social dominance theory: Its agenda and method. *Political Psychology, 2,* 845–880.

Srull, T. K., Lichtenstein, M., & Rothbart, M. (1985). Associative storage and retrieval processes in person memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11,* 316–345.

Srull, T. K., & Wyer, R. S., Jr. (1989). Person memory and judgment. *Psychological Review, 96,* 58–83.

Stangor, C., & Duan, C. (1991). Effects of multiple task demands upon memory for information about social groups. *Journal of Experimental Social Psychology, 27,* 357–378.

Stemmler, G., Heldmann, M., Pauls, C. A., & Scherer, T. (2001). Constraints for emotion specificity in fear and anger: The context counts. *Psychophysiology, 38,* 275–291.

Stephan, W. G., & Stephan, C. W. (2000). An integrated threat theory of prejudice. In S. Oskamp & W. Crano (Eds.), *Reducing prejudice and discrimination* (pp. 23–45). Mahwah, NJ: Erlbaum.

Stern, R. M., Ray, W. J., & Quigley, K. S. (2001). *Psychophysiological recording* (2nd ed.). Oxford, England: Oxford University Press.

Tomaka, J. W. (1994). Effects of cognitive appraisal on autonomic, affective, and behavioral reactions to stress (Doctoral dissertation, SUNY Buffalo, 1993). *Dissertation Abstracts International, 54,* 4968.

Tomaka, J., Blascovich, J., Kelsey, R. M., & Leitten, C. L. (1993). Subjective, physiological, and behavioral effects of threat and challenge appraisal. *Journal of Personality and Social Psychology, 18,* 616–624.

Tomaka, J., & Palacios-Esquivel, R. (1997). Motivational systems and stress-related cardiovascular reactivity. *Motivation and Emotion, 21,* 275–296.

van den Bos, K. (2001). Uncertainty management: The influence of uncertainty salience on reactions to perceived procedural fairness. *Journal of Personality and Social Psychology, 80,* 931–941.

Vanman, E. J., Paul, B. Y., Ito, T. A., & Miller, N. (1997). The modern face of prejudice and structural features that moderate the effect of cooperation on affect. *Journal of Personality and Social Psychology, 73,* 941–959.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54,* 1063–1070.

Williams, D. R. (1999). Race, socioeconomic status, and health: The added effects of racism and discrimination. In N. Alder, M. Marmot, B. McEwen, & J. Stewart (Eds.), *Socioeconomic status and health in industrial nations: Social, psychological, and biological pathways* (pp. 173–188). New York: NYAS.

Wright, C. I., Martis, B., Schwartz, C. E., Shin, L. M., Fischer, H. H., McMullin, K., et al. (2003). Novelty responses and differential effects of order in the amygdala, substantia innominata, and inferior temporal cortex. *NeuroImage, 18,* 660–669.

Wyer, N. A., Sherman, J. W., & Stroessner, S. J. (2000). The roles of motivation and ability in controlling the consequences of stereotype suppression. *Personality and Social Psychology Bulletin, 26,* 13–25.

Received November 13, 2005
Revision received September 2, 2006
Accepted September 3, 2006 ■