Contextual Moderation of Racial Bias: The Impact of Social Roles on Controlled and Automatically Activated Attitudes

Jamie Barden, William W. Maddux, Richard E. Petty, and Marilynn B. Brewer
Ohio State University

Three experiments tested the hypothesis that the social roles implied by specific contexts can attenuate or reverse the typical pattern of racial bias obtained on both controlled and automatic evaluation measures. Study 1 assessed evaluations of Black and Asian faces in contexts related to athlete or student roles. Study 2 compared evaluations of Black and White faces in 3 role-related contexts (prisoner, churchgoer, and factory worker). Study 3 manipulated role cues (lawyer or prisoner) within the same prison context. All 3 studies produced significant reversals of racial bias as a function of implied role on measures of both controlled and automatic evaluation. These results support the interpretation that differential evaluations based on Race × Role interactions provide one way that context can moderate both controlled and automatic racial bias.

Over the past decade and a half, researchers in social psychology have focused a great deal of effort on investigating automatic or implicit forms of racial bias. In general, differential racial attitudes (*bias*) are indicated when cues of one racial category elicit more positive and/or less negative automatically activated evaluations than cues from other racial categories. For White respondents, such measures generally reveal a pattern of in-group bias—more positive and/or less negative evaluations of White names or faces compared with Black names or faces. Because there is individual variance around this overall pattern of bias, results from these measures can also be used as measures of individual differences in automatically activated racial bias (Fazio, Jackson, Dunton, & Williams, 1995).

When researchers first reported methods to assess automatically activated attitudes, these procedures appeared to provide relatively pure measures of automatic racial bias (Dovidio, Kawakami, Johnson, Johnson, & Howard, 1997; Fazio et al., 1995; Greenwald, McGhee, & Schwartz, 1998). In particular, these measures have been viewed as an improvement over self-report measures, be-cause they are not subject to self-presentation concerns or inaccurate introspection (Nisbett & Wilson, 1977). Initial work indicated that racial bias in the form of stereotyping or prejudice tended to be prevalent at the automatic level, much more so than was indicated by responses on questionnaires and other explicit measures (e.g., Devine, 1989; Fazio et al., 1995; Greenwald et al., 1998; Wittenbrink, Judd, & Park, 1997). Early work suggested that such automatic biases were present in a large majority of people and that such biases, because of their automatic nature, were more or less immune from prevention or amelioration (see Bargh, 1999).

Recently, researchers have found that manipulation of the social context can moderate what was previously thought to be a ubiquitous automatic racial bias favoring the in-group (for a review, see Blair, 2002). In particular, findings have shown that White participants' typical pattern of bias in favor of Whites relative to Blacks can be either partially or completely attenuated under certain circumstances (Dasgupta & Greenwald, 2001; Lowery, Hardin, & Sinclair, 2001; Wittenbrink, Judd, & Park, 2001). Although there is convergent evidence supporting the contextual malleability of automatically activated racial biases, the particular contextual constructs that contribute to this malleability have not been clearly isolated (Devine, 2001). Indeed, different researchers have proposed very different contextual variables to explain similar operations and results. For example, Lowery et al. (2001) observed an attenuation of automatic racial bias when a Black as opposed to a White experimenter was used, and they explained this finding via *social tuning*, whereby social judgments become similar to those of individuals we expect to encounter. Dasgupta and Greenwald (2001) observed an attenuation of bias in a condition where participants were presented with positive Black exemplars (e.g., golfer Tiger Woods) and negative White exemplars (e.g., serial killer Ted Bundy). They argued that automatic evaluations are malleable to the exemplars that are available. Although these

accounts are plausible, the constructs highlighted in each approach could actually be used to predict the results observed by the other. That is, a Black experimenter certainly represents a positive exemplar of the Black category, suggesting that exemplars could have been the construct involved in Lowery et al.'s (2001) research. Similarly, the presentation of positive Black and negative White exemplars in Dasgupta and Greenwald's (2001) research could easily be interpreted as an expression of the biases of the experimenter, which could lead to social tuning effects.

In another demonstration of the malleability of automatic racial attitudes, Wittenbrink et al. (2001) modified an evaluative priming procedure by adding two visual contexts—a spray-painted wall (indicative of a ghetto) and a church—to the Black and White face primes. The ghetto context produced a pattern of in-group bias on the automatic measure, but a church context completely attenuated this bias. Wittenbrink et al. (2001) argued that contextual information combines with racial stereotypes in a gestaltlike fashion to produce patterns of evaluation that reflect shifts in the meaning of race. However, no specific contextual construct was isolated as the critical contributor to this gestalt. The attenuation in the church context could result from any number of specific contextual constructs (e.g., positive affect, the role of a churchgoer, general ideals of equality under God, etc.). In fact, recent evidence suggests that contextual affect can produce moderation of automatic measures of in-group bias. For example, DeSteno, Dasgupta, Bartlett, and Cajdric (in press) have shown that contexts eliciting certain negative emotions, such as anger, produced automatic in-group bias, whereas the absence of negative affect produced no bias. If the ghetto context elicited more negative emotion (e.g., anger or fear) than the church context, this could provide an alternative account for the observed contextual moderation. So again, although research has established that context can moderate patterns of automatic racial bias, it remains unclear what contextual variables are responsible for producing moderation (see Devine, 2001).

Understanding what constructs moderate measures of automatic racial attitudes is very important, because these measures correlate with behaviors toward members of racial out-groups, particularly those that fall outside of conscious control, such as nonverbal behaviors (Dovidio et al., 1997). Variations in these behaviors can be sensed by interaction partners from different racial out-groups (Fazio et al., 1995) and so represent a unique challenge in trying to improve interracial interactions. For these reasons, it is crucial to go beyond establishing that a variety of contexts can moderate automatic racial biases to isolate individual contextual variables that are capable of producing such moderation.

One variable that might logically lead to contextual moderation of automatic racial bias is social roles. Social roles provide an important source of information to observers attempting to judge individuals they encounter. Although observers use facial features to automatically categorize individuals according to important social categories such as race, gender, and age and evaluate them accordingly (Fazio et al., 1995), this information still leaves enormous room for ambiguity. Social role categories provide critical additional meaning, coloring and influencing the interpretation of these other important categorizations. Thus, the first goal of the current research was to isolate social roles as one important contextual construct capable of producing moderation of automatic racial bias.

SOCIAL ROLES AS A MODERATOR OF CATEGORY-BASED SOCIAL JUDGMENTS

The importance of social roles in forming social judgments has long been noted, especially in the area of gender research (e.g., Eagly & Steffen, 1984). One distinctive feature of much of the theorizing in the area of social roles is that roles may interact with other category information to determine the impression that is drawn of a specific individual. For instance, Brewer (1988) has suggested that

> *different* representations of the same social role are developed, depending on whether the role occupant is male or female. In other words, the concept of "woman doctor" is not simply the intersection of the categories "woman" and "doctor," but a specific type within the general category of women, and part of a configuration of traits and features distinct from those associated with "man doctor." (p. 12)

If that is the case, then social judgments result from the interaction of two category memberships, role category and gender category (for similar views, see Higgins & Rholes, 1976; Taylor, 1981). Thus, the interactionist viewpoint specifies that judgments of multiply categorized targets can result from an interaction between multiple categories.

Recently, Eagly and Karau (2002), in an attempt to explain prejudice against women in leadership roles, have provided evidence for one mechanism through which social roles and gender categories interact to produce bias. They argue that incongruities between social group stereotypes (i.e., women as communal and nonagentic) and certain social roles (i.e., leaders as individualistic and agentic) lead to prejudice against women relative to men within leadership roles (i.e., we do not like people who act in counterstereotypic ways). Although Eagly and Karau (2002) described their findings on the basis of a specific role congruity mechanism, their theory illustrates one way that social category and social role information can interact in producing controlled social judgments.

Although the interactionist viewpoint has been applied primarily to the interaction of social roles and gender, it could also be applied to other important social categories, including race. For example, research has indicated that Black subtypes essentially exist at the intersection of a racial category and a social role (i.e., Black businessman and Black athlete; Devine & Baker, 1991). In line with interactionist views, these subtypes are evaluated positively, whereas the global Black category is evaluated negatively. To further investigate the interactionist approach as it applies to race, the current research assessed whether the interaction of race and role information moderates controlled measures of racial bias.

Furthermore, if physical contexts imply social roles for individuals appearing within them, then this could provide one explanation for findings of contextual moderation of automatic racial bias (Wittenbrink et al., 2001). In particular, the second goal of the current investigation was to test our proposed differential role evaluation hypothesis, which holds that specific role information from the context can influence automatic racial bias within that context. This hypothesis has two components. First, it suggests that physical contexts (e.g., ghetto, classroom, prison, church, basketball court) activate social roles (e.g., gang member, student, prisoner, churchgoer, basketball player) on the basis of the type of individuals one would expect to encounter in that context. Con-

textually inferred roles are assigned to any individual in the context along with other categorizations, such as race, that are based on the physical cues and features of the specific individual. Second, the hypothesis holds that the pattern of automatic racial bias in a given context will reflect the interaction of race and role information when people hold different automatic evaluations of different races in the same role. Thus, according to this hypothesis, race and role cues are the key inputs, and the racial bias that results will reflect the interaction of these inputs.

In order to provide support for the interactionist hypothesis, the final goal for the current investigation was for each study to include one context where one pattern of racial bias is established and a second context where this pattern of racial bias is significantly reversed. Such a reversed pattern would establish that race information impacted evaluations in both contexts. In addition, a reversed pattern would establish that the impact of race on evaluations depended on context. This reversed pattern would be the first of its kind, because previous research on moderation of racial bias has been limited to attenuation of racial bias (Dasgupta & Greenwald, 2001; Lowery et al., 2001; Wittenbrink et al., 2001), and when reversals have occurred (Mitchell, Nosek, & Banaji, 2003), the targets were no longer being categorized on the basis of race, as explained more fully below (see also Bargh, 1999). Thus, such a reversed pattern of racial bias would provide the clearest support for the proposed differential role evaluation hypothesis, and it would establish the strength of context to moderate automatic racial biases.

OVERVIEW OF THE PRESENT RESEARCH

The differential role evaluation hypothesis was investigated in three studies. For each study, we first drew on the existing stereotyping and subtyping literature to select roles that were likely to be associated with differential implications of race. Controlled evaluative ratings of targets from each race and role combination were then obtained to establish that the pattern of racial bias in each role was as predicted. Automatic racial biases were then measured using an evaluative priming procedure that was modified to include contexts that implied the specific roles used in the controlled measure. The pattern of automatic racial bias was assessed using a single index, which summarized the magnitude and direction of the Race of Prime $\times$ Adjective Valence interaction on facilitation scores within each contextual role. If parallel patterns of moderation of racial bias were observed on controlled and automatic measures, this would support the notion that the interactionist framework underlies both types of judgment. In this way, the current approach parallels the extension of the role congruity mechanism from controlled judgments (Eagly & Karau, 2002) to automatic gender prejudice (Rudman & Kilianski, 2000).

Given two racial categories, there are three possible patterns of relative evaluation that can occur on a controlled or an automatic measure of racial bias: A bias can favor Group A relative to Group B, a bias can favor Group B relative to Group A, or an egalitarian pattern can occur where both groups are evaluated equally. For the current investigation, each study included one role where it was expected that the pattern of bias would favor Group A relative to Group B and a second where bias would favor Group B relative to Group A. Such a reversed pattern provides the clearest support for

the interactionist position, because it establishes that racial cues had an impact on evaluations in each role.

Notably, no prior research on the malleability of automatic racial bias has produced significant reversals in automatic bias of racial targets when the targets were still being clearly differentiated on the basis of the racial cues present. Previous reversals of automatic judgments have been obtained only when racial categories were either inhibited or replaced. For example, Macrae, Bodenhausen, and Milne (1995) demonstrated that when an individual with multiple potential categories acted in a way that was consistent with one category (e.g., Asian), this resulted in inhibition of the competing category stereotype (e.g., woman). Similarly, Mitchell et al. (2003) used two versions of the Implicit Association Test (IAT) to investigate relative evaluations for targets that differed both in terms of racial category and role category (i.e., White politicians, Black athletes). When the IAT required categorization based on race, the evaluation was the result of racial cues; thus, a bias favoring White politicians relative to Black athletes emerged. Conversely, when the IAT required categorization based on occupation, the evaluation was the result of occupation cues, leading to a bias favoring Black athletes relative to White politicians. In this case, there is no reason to believe that race cues had any impact on the pattern of automatic bias that was produced. In general, this finding is in line with the dominant view: that multiple categorizations often operate in a mutually inhibitory fashion for automatic social judgments.

In contrast with the view of category replacement, there is preliminary evidence that automatic social judgments can sometimes reflect the interaction of multiple categories much as conscious judgments can. For example, Rudman and Kilianski (2000) showed that when both genders were assigned to high-status roles (e.g., police officer, boss, doctor) on an evaluative priming measure, there was a bias favoring men relative to women; however, within low-status roles (e.g., cook, nurse, waiter or waitress), there was a nonsignificant trend in the opposite direction. If this type of interactive finding were shown to be a reliable and general phenomenon, then this would provide the first clear evidence that category replacement is not the only manner in which multiply categorized targets can produce automatic social judgments.

In the current investigation, parallel predictions were made for racial bias within a given role for both the controlled and automatic measures. Study 1 provided an initial test of the interactionist viewpoint as it applies to race and role categories by investigating an obvious circumstance for a reversal of bias. In this study, Black, White, and Asian targets were evaluated in either a basketball court or a classroom context. On the basis of stereotypes and subtypes held for each racial group, it was expected that racial bias should favor Blacks relative to Asians in the role of a basketball player and Asians relative to Blacks in the role of a student. Study 2 investigated whether other contextual roles could produce a reversal of the typical pattern of in-group bias that White participants exhibit with respect to Blacks. In this study, Black and White targets were evaluated in a prison, church, or factory context. On the basis of previous research, in-group bias was expected in the prisoner role, an egalitarian pattern was expected in the churchgoer role, and out-group bias was expected in the factory worker role. In Study 3, roles were manipulated directly through explicit instructions and the clothing in which Black and White targets appeared. In this study, in-group bias was expected in the prisoner

role, and out-group bias was expected in the lawyer role. Taken together, these studies addressed the primary goals of the current investigation: to establish that social roles represent one important contextual construct capable of producing moderation of automatic racial bias and, more specifically, to show that differential role evaluation can attenuate or reverse automatic racial biases, in line with the general interactionist viewpoint.

STUDY 1

Study 1 was designed to provide an initial demonstration that when a context implies specific roles for individuals present, the context can interact with target racial cues in producing moderation of racial bias. The opposing views that Whites hold of Blacks and Asians offer the possibility of obtaining a reversal of racial bias, which provides the clearest support for the interactionist viewpoint. Previous research has indicated that Blacks are viewed as athletic and not intellectual (Devine & Baker, 1991; Kao, 2000; Wittenbrink et al., 1997), whereas Asians are viewed as intellectual but unathletic (Kao, 2000). This suggests that in an athletic role, Blacks would be evaluated positively and Asians would be evaluated negatively, but in a student role, Asians would be evaluated positively and Blacks negatively. These predictions are in line with Eagly and Karau's (2002) role congruity mechanism, which holds that incongruities between certain social group stereotypes and social roles result in prejudice against those groups in those roles (i.e., Black students and Asian basketball players). These predictions are also in line with Devine and Baker's (1991) subtyping mechanism, because a positively evaluated Black athlete subtype has been established, and it seems likely that our White college student sample would hold a Black student subtype. Such a Black student subtype could certainly be evaluated negatively, given that stereotype expectancies for Blacks include that they are not intellectual (Wittenbrink et al., 1997).

A basketball court was selected as the context that would be expected to elicit the athlete role spontaneously (see Figure 1). This appeared to be a particularly good context for the current use, because it clearly implies the basketball player role and because this role epitomizes Black athleticism and the lack of Asian athleticism. A technical classroom was selected as the second context (see Figure 1), again because it clearly implies the student role and because this role epitomizes Asian academic success and Black academic underachievement. White targets were also included in the design in order to provide a comparison point based on the cultural default (Fiske, 1998). If anything, stereotypes held toward Whites suggest that they should fall between Asians and Blacks as both basketball players and students.

Method

Using the same race and context combinations, controlled and automatic evaluations were obtained from separate samples.

Controlled Evaluation

The controlled assessment for Study 1 was obtained from a sample of 199 undergraduate students at Ohio State University who volunteered for the study in partial fulfillment of course requirements.[1] Participants for all three studies were recruited via the psychology department's online research sign-up program. From this initial sample, 14 minority participants

were excluded from the analysis, leaving a final sample of 185 White participants for formal analysis.

Participants were randomly assigned to a 2 (role: student vs. basketball player) × 3 (race: Asian vs. Black vs. White) between-participants design. The questionnaire was entitled "Quick Responses," with instructions on the first page indicating that "we can often make very accurate responses to individuals we encounter if we follow our first gut instinct." The instructions also indicated that participants should pay attention to the situation in which the target was seen. The second page contained the image of a Black, White, or Asian face; role information; and evaluative items. Faces were superimposed onto either a basketball court or classroom (see Figure 1). The role information appeared below the image and indicated that the individual was either a basketball player or a student to correspond to the context in which the face appeared. Each target was evaluated on the basis of three 9-point semantic differential scale items (*strong–weak*, *wonderful–horrible*, *excellent–lousy*). These items corresponded directly to the valence adjectives used in the evaluative priming procedure for Study 1. The three-item scales were highly reliable for the student ($\alpha = .90$) and the basketball player ($\alpha = .96$) roles, so analyses were conducted on indexes made up of an average of each of the three items.

In addition to corresponding to the automatic measure, the assessment was designed to minimize the impact of self-presentational concerns, which often prevent participants from honestly reporting their race-based judgments. A between-participants approach prevented participants from making direct evaluative comparisons between races. Furthermore, the salience of racial category was minimized by manipulating race using the image of a face rather than explicitly mentioning any racial category. To further minimize awareness of the importance of race for the results, a White experimenter was used for both the controlled and automatic assessments in all three studies.

Automatic Evaluation

The evaluative priming procedure developed by Fazio et al. (1995) was used to measure automatically activated racial biases. Racial biases are measured by recording how quickly adjectives are categorized as positive or negative when they follow the race–face primes that serve as attitude objects. A pattern of in-group bias from White participants would be evidenced if positive adjectives were categorized more quickly following White primes and negative adjectives were categorized more quickly following Black primes. That is, racial biases are indicated by the presence of an interaction between the race of the prime and the valence of the adjective to be categorized. In the current studies, the standard procedure was modified to allow for the manipulation of physical context. Images of physical backgrounds were presented continuously on the computer screen during an entire block of trials. In this way, the face primes appeared as individuals normally do within settings.

Primes from three races (Asian, Black, and White) were presented in two separate blocks, one for each contextual background (a basketball court and a technical classroom; see Figure 1). According to the differential role evaluation hypothesis, the role implications of the context should be applied to the individuals appearing within that context. That is, individuals in the basketball court will be seen as basketball players, and individuals in the technical classroom will be seen as students. On the basis of the interactionist perspective, the role category information will combine with the race category information from the prime to produce moderation of racial biases as expected on the controlled assessment.

---

[1] Participant gender was not assessed for any of the studies. All samples for the current investigation were drawn from the Psychology 100 population at Ohio State University, which includes somewhat more women than men.

Figure 1. Images of physical contexts used as backgrounds for the automatic and controlled racial attitudes measures for Studies 1–3. The prison context image from Study 2 was identical to the prison context used in Study 3, except that the faces used for the priming measure of Study 2 were not shown with any clothing denoting role. For the automatic evaluation measures, these images were presented in vivid color, which enhanced both their realism and immediacy (e.g., navy pinstripe business suit vs. bright orange jumpsuit).

Participants

A total of 112 introductory psychology students at Ohio State University voluntarily participated in partial fulfillment of a course requirement. The data from 17 minority students were excluded from analysis, as were the data from 8 participants who had error rates in excess of 45%. This left the data from 87 White participants for formal analysis.

Independent Variables

Context. Two digital pictures (800 × 600 pixels) of a technical classroom and a basketball court served as background images (see Figure 1).

Race of prime. Race was manipulated using primes consisting of a total of nine photographs of college-aged Black, White, and Asian men. Photos were selected so that they exhibited neutral facial expressions and did not include any jewelry or facial hair.[2] Photo-editing software was used to remove the background and shirt from each photo, leaving the face and the neck. All race primes appeared in both contexts.

Adjective valence. The adjectives used for the critical trials were three positive evaluative adjectives (*strong*, *excellent*, *wonderful*) and three negative evaluative adjectives (*weak*, *horrible*, *lousy*) that could be applied to performance in both contexts.

Context order. Participants were randomly assigned to one of two block orders (basketball–classroom) or (classroom–basketball).

Dependent Variable: Response Time

The response time was the amount of time between the appearance of the valence adjective and the categorization of the adjective as recorded to the nearest millisecond by DirectRT software (Jarvis, 2000). These response times were used to compute facilitation scores relative to baseline response times to each valenced adjective alone. This computation is discussed in detail shortly. Facilitation scores were tabulated according to the three independent variables (e.g., context, target word, and race) collapsing within the three face primes of the same race and three target adjectives of the same valence. The results presented for all three studies in this investigation are based on raw (untransformed) response time scores.

Procedure

Participants were escorted into the laboratory by a White male experimenter and randomly assigned to one of several personal computers. The experimenter told participants that the experiment dealt with automatic tasks and that they should follow the instructions on the screen, and then left the room. The procedure followed the evaluative priming procedure (Fazio et al., 1995) designed to assess automatic activation of racial attitudes, with a few modifications in order to incorporate the manipulation of physical context, and three races of face primes. The experiment was presented using DirectRT software.

Phase 1 of the experiment was designed to obtain participants' baseline response rates to the six adjectives that would be used as targets in a subsequent priming procedure. The instructions indicated that an automatic word judgment task would come first. For each adjective that appeared, participants were instructed to press a key on the left labeled "bad" for negative adjectives and a key on the right labeled "good" for the positive adjectives. Participants were told to respond as quickly as possible while being as accurate as possible. Six practice trials consisting of filler adjectives not related to performance were presented in order to familiarize participants with the task. Each trial began with a 500-ms presentation of a row of asterisks in the center of the computer screen. The screen went blank for 100 ms and then the adjective was presented on the screen until the participant responded. An intertrial delay of 3 s was maintained for all phases of the procedure. After the practice trials, the six target adjectives that would later be used in the critical Phase 4 were randomly presented twice each in the same manner. Latencies were recorded for these trials to

the nearest millisecond in order to provide baseline response times for each participant.

Phase 2 of the experiment involved familiarizing participants with a face-recognition task, ostensibly the second automatic task of interest. Participants were instructed to pay attention to the faces that appeared, because they would later be tested for recall of these faces. Nine male faces of Whites, Blacks, and Asians were then presented one by one on the screen for 215 ms. All faces were different from those used during the critical Phase 4 of the procedure. The faces appeared in the bottom center of the screen, with the neck terminating on the bottom of the screen. Nine trials (one of each face) were presented in random order.

Phase 3 of the experiment consisted of familiarizing participants with the actual priming task. This involved the combination of the previous two tasks. As part of the cover story, the instructions indicated that the next portion would measure participants' ability to perform both the face recognition and word judgment tasks simultaneously. At the beginning of each trial, a row of asterisks appeared for 500 ms. The asterisks were followed by a blank screen for 100 ms, after which either a White or Black face was presented for 215 ms. After a 100-ms delay, a target adjective appeared on the screen. These presentation times resulted in a total stimulus onset asynchrony of 315 ms. Participants were instructed to remember each face for the future recall test while at the same time judging the meaning of each adjective. Six practice trials were then presented, consisting of filler faces and adjectives. Phase 4 was the most critical phase of the experiment. The instructions explained that in order to make the experiment more like real life, a picture of a context would be added on the screen. Participants were instructed to keep the context in mind as they performed the combined task and to imagine the sights, sounds, and smells that would be present. Next, the first context appeared on the screen as it would for the 54 trials making up that contextual block. During each trial, presentation times of the asterisks, primes, and target adjectives were identical to Part 3. Primes and adjectives appeared over each context, giving the impression that the faces were individuals within the context. The contexts remained on the screen more or less constantly, disappearing only for approximately 100 ms following participants' responses to the target adjectives (time required by the DirectRT software to load the next trial). The interruption was brief enough that it did not dispel the feeling that participants constantly remained in the context.

The contextual blocks were presented in one of two orders (basketball–classroom or classroom–basketball). Each context was presented as a single block of 54 trials for a total of 108 trials in Phase 4. These 54 critical trials per context resulted from every combination of six target adjectives with nine face primes. The total of nine face primes was judged to be sufficient to maintain the cover story of the facial recognition task, so no filler face primes were used. The order of the trials within each context was randomly determined. Participants were debriefed and thanked at the end of the evaluative priming procedure.

Results

Controlled Evaluation

Controlled judgments were computed separately for each race and role combination, and then these were compared within role in order to assess which of the three patterns of racial bias emerged

---

[2] Race primes for all three studies were judged by the authors to be prototypical of their race to ensure that they would provide sufficient race cue information. Furthermore, in selecting faces we attempted to control for obvious differences in information other than the race cue (e.g., attractiveness). Notably, any simple confound with race, such as attractiveness, cannot account for the expected interaction on attitudes of race and role.

in each context. Overall, the results supported the expected effects of the basketball player and the student roles paired with Asian, Black, and White targets, respectively. The two-way Role × Race interaction was significant, $F(2, 179) = 34.78$, $p < .001$, indicating that context did impact the pattern of racial bias. Further analysis showed that within the student role, Asians ($M = 6.68$, $SD = 0.93$) were evaluated more positively than Whites ($M = 5.87$, $SD = 0.96$), $F(1, 59) = 10.75$, $p = .002$, who were evaluated more positively than Blacks ($M = 5.41$, $SD = 0.59$), $F(1, 64) = 5.17$, $p = .026$. Within the basketball player role, this pattern was completely reversed so that Blacks ($M = 7.00$, $SD = 1.11$) were evaluated more positively than Whites ($M = 5.92$, $SD = 0.89$), $F(1, 56) = 16.45$, $p < .001$, who were evaluated more positively than Asians ($M = 4.95$, $SD = 1.53$), $F(1, 67) = 10.14$, $p = .002$. Thus, the controlled assessment indicated exactly the pattern of racial biases that was suggested by the stereotypes and subtypes that Whites have of Blacks and Asians (Devine & Baker, 1991; Kao, 2000; Wittenbrink et al., 1997).

Automatic Evaluation

Error Rates, Outliers, and Facilitation Scores

Across participants, relatively few errors were made in judging adjectives as positive or negative (4.2%), and these trials were excluded from further analyses. The data also exhibited few outliers (0.1%), defined as response times less than 300 ms or greater than 1,500 ms. Response times less than 300 ms indicate that participants responded more quickly than is considered physically possible (indicating a blind guess), so these trials were dropped from the analyses. Response times greater than 1,500 ms usually indicate temporary attention lapses. These response times were reset to 1,500 ms so that they would not overly influence the results.

Facilitation scores were computed for each participant by taking the average of the two baseline response times for each adjective and subtracting these averages from the corresponding response latencies of the critical fourth phase of the experiment. A total of 12 facilitation scores were calculated for each participant, one for

each of the 12 cells of the 2 × 3 × 2 (Context × Race × Adjective Valence) within-participants factors of the design.

In addition, a single index was calculated to summarize the Race of Prime × Adjective Valence interaction within each context. This type of index provides a summary of the direction and magnitude of the racial bias present (Dovidio et al., 1997; Fazio et al., 1995; Wittenbrink et al., 1997). The index was calculated for the critical Asian–Black comparison according to the following formula: Asian–Black bias index = (Asian prime, positive adjective − Asian prime, negative adjective) − (Black prime, positive adjective − Black prime, negative adjective). As a result, negative scores indicated bias favoring Asians relative to Blacks, zero indicated egalitarian responses, and positive scores indicated a bias favoring Blacks relative to Asians.

Race of Prime and Context Effects

The data initially were analyzed in a four-way, Context × Race of Prime × Adjective Valence × Context Order multivariate analysis of variance (MANOVA), which indicated that context order failed to produce a significant overall interaction, lower order interactions, or overall main effect (all $Fs < 1.8$). The only exception to this was a two-way, Context × Order interaction, $F(1, 85) = 20.06$, $p < .001$, $\eta^2 = .191$, indicating a training effect such that participants responded more quickly overall when a given context appeared later in the order. Because order did not interact with either race or word valence, further analyses were conducted collapsing across order. Thus the data were analyzed in a three-way, Context × Race of Prime × Adjective Valence, within-participants analysis of variance (ANOVA). This analysis revealed a marginally significant Context × Race interaction, $F(2, 172) = 2.90$, $p = .058$, $\eta^2 = .033$. This effect was qualified by a significant three-way Context × Race of Prime × Adjective Valence interaction, $F(2, 172) = 5.40$, $p = .005$, $\eta^2 = .058$, indicating that context significantly moderated automatic performance judgments that were based on race (see Figure 2).

A separate analysis of the two-way Race of Prime × Adjective Valence interaction in the two contexts revealed the nature of this
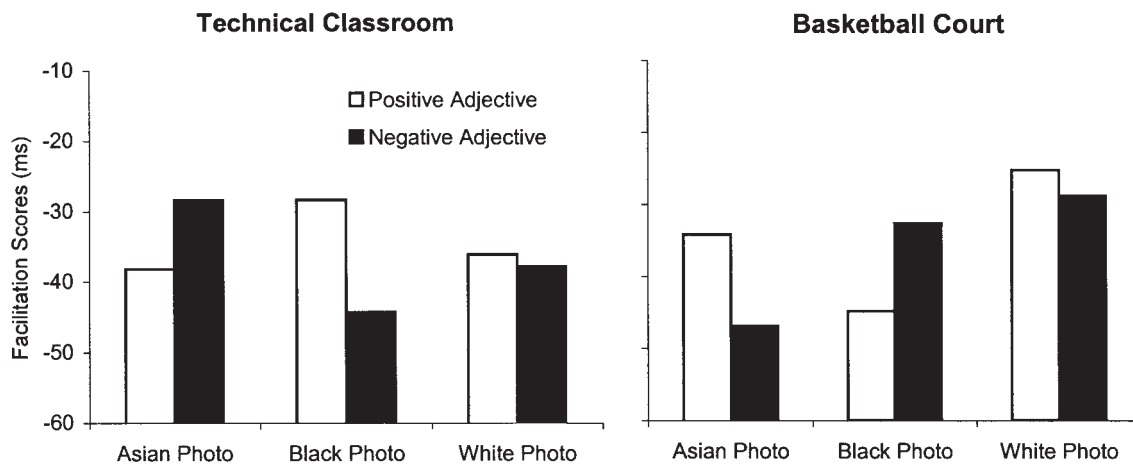

Figure 2. Contextual moderation of automatic racial biases across two contexts, technical classroom and basketball court, on the basis of facilitation scores.

interaction. A significant two-way Race of Prime × Adjective Valence interaction was obtained in both the basketball court context, $F(2, 172) = 3.14$, $p = .048$, $\eta^2 = .035$, and classroom context, $F(2, 172) = 3.44$, $p = .034$, $\eta^2 = .038$. To further investigate these effects, the two-way Race of Prime × Adjective Valence interaction was investigated for each pair of races in each context. These analyses indicated that as predicted, Blacks were evaluated more positively than Asians in the basketball court, $F(1, 86) = 6.59$, $p = .012$, $\eta^2 = .071$; however, Asians were evaluated more positively than Blacks in the technical classroom, $F(1, 86) = 6.68$, $p = .011$, $\eta^2 = .072$. In both contexts, Whites were not evaluated as significantly different from either Asians or Blacks. Finally, the apparent main effect indicating that White primes were responded to more slowly in the basketball court than in the technical classroom was not statistically reliable, $F(1, 86) = 2.26$, $p = .136$.

Additional analyses were conducted based on the single index in order to explore the critical pattern of moderation of the Asian–Black bias. Initial analyses, based on a one-way within-participants ANOVA, revealed that the Asian–Black index was significantly moderated by context, $F(1, 86) = 10.48$, $p = .002$, $\eta^2 = .109$. One-sample t tests were then conducted within each context to compare the index to the zero point, which corresponded to an egalitarian pattern of evaluation. This revealed that in the technical classroom, there was a bias favoring Asians relative to Blacks ($M = -25.77$, $SD = 93.01$), $t(86) = 2.584$, $p = .011$, Cohen's $d = .277$, whereas in the basketball court, there was a bias favoring Blacks relative to Asians ($M = 24.95$, $SD = 90.63$), $t(86) = 2.568$, $p = .012$, $d = .275$. Thus, the pattern of racial bias favoring Asians relative to Blacks in the classroom was completely reversed to a bias favoring Blacks relative to Asians in the basketball court. These analyses show that the key findings were identical using the Race of Prime × Adjective Valence interaction or the index of racial bias. In fact, the index is mathematically identical to the statistics generated for the interaction, producing the same p values (Dovidio et al., 1997). Because the index directly reflects the pattern of racial bias that reflects our hypotheses, it was also used in Studies 2 and 3.

Discussion

Study 1 provides initial evidence that role implications can produce contextual moderation of automatic racial attitudes by taking advantage of the opposing performance stereotypes held about Blacks and Asians. A controlled assessment verified an evaluative bias in favor of Asians relative to Blacks when both were assigned to the student role, and this pattern was completely reversed in the basketball player role, with the evaluative bias favoring Blacks relative to Asians. In both roles, Whites fell in between these two groups.

The same pattern of moderation was also reflected on the automatic measure. In the technical classroom (which implied the role of a student), the racial bias favored Asians relative to Blacks, with Whites falling in between. This pattern of bias was significantly reversed in the basketball court (which implied the role of a basketball player), where the racial bias favored Blacks relative to Asians, with Whites again falling in between. To our knowledge, this represents the first time that an automatic measure of racial bias has produced one pattern of bias in one context and a

significant reversal of that pattern of bias within a second context when the targets were still being clearly differentiated on the basis of racial cues. As such, these results provide strong evidence of the contextual malleability of automatic racial biases.

Previous research showing moderation of automatic bias toward racial targets has been open to the critique that the context somehow overwhelmed or distracted from the racial category cue information (for this point, see Bargh, 1999). Research that shows a racial bias in one context and an attenuation of that bias in another is certainly vulnerable to this alternative (Dasgupta & Greenwald, 2001; Lowery et al., 2001; Wittenbrink et al., 2001). Even results that exhibit a reversal of bias are not necessarily immune to this critique. For example, Mitchell et al. (2003) produced a reversal by using an occupation IAT and assigning Black targets to a positive occupation and White targets to a negative occupation (i.e., Black athletes and White politicians). As Mitchell et al. indicated, this reversal occurred because occupation category replaced racial category as the basis for the evaluation, which falls squarely within Bargh's (1999) critique. Because this explanation would lead to the same prediction if the athletes and politicians were of the same race, these results do not represent a reversal of an automatic race-relevant bias.

By contrast, Study 1 provides clear evidence of a reversal of automatic race-relevant bias, because it is not subject to the critique that context distracted or overwhelmed target racial cues. In this case, moderation occurred when both races were in the same context, so race was not confounded with any other variable (e.g., occupation). Put another way, the presence of an evaluative bias favoring Blacks relative to Asians on a basketball court and a bias favoring Asians relative to Blacks in a classroom shows that race cues had an impact on target evaluations. Furthermore, the results of Study 1 are consistent with the differential role evaluation hypothesis, which holds that physical contexts often imply roles that are applied to individuals within a context and that roles and race interact to produce different patterns of racial bias.

Taken together, these findings are consistent with the general interactionist perspective that targets can be categorized simultaneously on the basis of race and role, with target evaluations resulting from the conjoint categories. In addition, the exact pattern of racial bias within each context was consistent with Eagly and Karau's (2002) role congruity mechanism. That is, relative prejudice against one race was observed whenever the context implied a social role that was incongruent with the stereotype of that race (i.e., Black students and Asian basketball players). Similarly, these findings were consistent with Devine and Baker's (1991) subtyping mechanism, which predicted that bias would favor Blacks when the context activated the positive Black athlete subtype, and bias would go against Blacks when the context implied a Black student subtype. Thus, the current evidence is consistent with the notion that both the role congruity and the subtyping mechanisms could be underlying the interaction of race and role information in producing moderation of automatic racial bias.

From a more concrete standpoint, the results of Study 1 indicate that contextual moderation of automatic racial biases might be a very common occurrence. In the course of a 5-min walk across campus from class to a weekly game of pick-up basketball, White students may completely reverse their relative evaluation of a Black and an Asian target. Fluctuations in automatic racial biases

suggest that the Asian student would be picked before the Black student for a group project in class, but the same Black student would be picked before the same Asian student for a basketball team. Of course, such fluctuations in controlled evaluations are to be expected (Ajzen & Fishbein, 1977). What is more surprising is that it appears that role information can be extracted from physical contexts and integrated with race-category information to produce a complex pattern of biases at the automatic level as well.

STUDY 2

Study 1 provides support for the differential role evaluation hypothesis and the interactionist framework by taking advantage of an obvious instance where a reversal of racial bias could be predicted. These results provide a basis to approach other socially relevant prejudices. As a matter of social concern, researchers in the United States have focused much of their effort on the racial bias that Whites have, which favors Whites relative to Blacks. Over the past decade, this overall pattern of in-group racial bias has been replicated using different measures of automatic racial bias (Dovidio et al., 1997; Fazio et al., 1995; Greenwald et al., 1998). More recently, research on the malleability of automatic racial biases has shown that certain contexts can produce an attenuation (though not a reversal) of this automatic racial bias (Dasgupta & Greenwald, 2001; Lowery et al., 2001; Wittenbrink et al., 2001).

One critical question that is raised by the current investigation is whether social roles implied by physical contexts might explain previous evidence of moderation of automatic racial bias. In the most relevant research, Wittenbrink et al. (2001) showed that when racial primes were superimposed over a ghetto context, this produced in-group racial bias for White participants, whereas a church context produced a full elimination of that bias. Given that the ghetto context likely implied the role of a gang member, and the church context likely implied the role of a churchgoer, differential role evaluation provides one possible explanation for these findings. According to this interpretation, both Black and White targets were assigned to the gang member and churchgoer roles in each of the contexts, but in each case, the same role had different evaluative implications depending on the target's racial category. To test this possibility, Study 2 included a church context, which was intended to activate the churchgoer role and produce an egalitarian pattern of bias replicating previous findings. In addition, a prison context (see Figure 1) was used to activate the prisoner role expected to produce a prejudiced pattern of bias, much like the ghetto context did in previous research (Wittenbrink et al., 2001). Finally, Study 2 included a context that was intended to elicit a significant reversal of racial bias. Such a pattern of out-group bias would mean that White participants showed favoritism toward Blacks relative to Whites.

Study 1 indicates that a complete reversal of racial evaluations is possible when comparing two out-groups (i.e., Blacks and Asians). However, given that our participants were White, it is still unclear whether it is possible to reverse a pattern of automatic in-group bias. Previous research using contextual manipulations has only attenuated such racial bias, even under circumstances with the potential to produce reversals. For example, when the contextual manipulation consisted of activating both positive Black exemplars and negative White exemplars, this produced

only an attenuation of automatic racial bias (Dasgupta & Greenwald, 2001). In another study, the presence of a Black experimenter, which might lead Whites to bend over backwards, again elicited only an attenuation of automatic racial bias (Lowery et al., 2001). Thus it remains an open question whether it is possible to reverse a bias favoring a racial in-group on a measure of automatic evaluation.

From the current perspective, a significant reversal of automatic in-group bias should be possible if a role was found that had more positive or fewer negative implications for Black targets than for White targets. Furthermore, these implications must be strong enough to both overwhelm any default prejudice present and create a significant out-group bias. Previous research in controlled judgments has shown that when both races are portrayed within a role that is hardworking, Whites are more willing to help out Blacks than Whites (Sniderman & Piazza, 1993). Accordingly, a role implying gainful employment was sought. In order to increase the differential impact of the same role on the two races, the goal was to find a context implying a blue-collar job. Such a role might be viewed in a positive light for Black targets (i.e., hardworking blue-collar Black subtype) while at the same time implying a relatively unattractive job for White targets, at least in the eyes of our college student sample. On the basis of this reasoning, the image that was selected for the third context was of the parking lot outside of a dilapidated but functioning factory (see Figure 1). Given that it is daytime and there is smoke coming from the factory, this context should clearly imply the role of a worker at the factory. Thus, Study 2 used three contexts (prison, church, factory), which imply corresponding roles (prisoner, churchgoer, factory worker) that previous research has suggested are likely to produce three corresponding patterns of racial bias (in-group, egalitarian, out-group).

Method

Controlled Evaluation

As with Study 1, a controlled evaluation was conducted in order to verify that these initially selected contexts would produce the expected pattern of biases on the controlled measure before proceeding to the evaluative priming procedure.

Participants

A sample ($N = 333$) of introductory psychology students from Ohio State University participated in partial fulfillment of a course requirement. Twenty-nine minority participants were excluded, leaving a total of 304 White participants for further analysis.

Procedure

Participants were randomly assigned to a 4 (role: factory worker vs. churchgoer vs. prisoner vs. lawyer) × 2 (race: Black vs. White) between-participants design. The factory, church, and prisoner roles were to be used for Study 2, and the prisoner and lawyer roles were to be used for Study 3. The assessments were collected simultaneously for the purpose of efficiency. Participants were given explicit role information (factory worker, churchgoer, prisoner, or lawyer) below the image portraying the target in the context. For the factory worker and churchgoer images, Black and White target faces were superimposed on the factory and church contexts with no further embellishment. In the case of the prisoner and lawyer roles,

targets were portrayed wearing clothing corresponding to their role (an orange jumpsuit or a formal suit), and both appeared in the prison context (see Figure 1). Each participant again responded to only one target, which minimized the salience of race and prevented any direct comparisons between responses to multiple targets.

Unlike Study 1, where the evaluative adjectives were related to performance beliefs, Study 2 used only general evaluative adjectives for both controlled and automatic measures. For the controlled assessment, targets were evaluated using three 9-point semantic differential scales (*like–dislike*, *favorable–unfavorable*, *positive–negative*). These three scale items were highly reliable for the factory worker ($\alpha = .88$), churchgoer ($\alpha = .90$), prisoner ($\alpha = .81$), and lawyer ($\alpha = .86$). Thus, further analyses were conducted on indexes made up of the average of the three items.

Automatic Evaluation

Participants

Fifty-eight introductory psychology students at Ohio State University participated in partial fulfillment of a course requirement. Because this study was concerned with automatic racial attitudes held by Whites, the data from 2 minority students were excluded, as were data from 8 participants who had previously participated in similar reaction-time experiments earlier in the quarter. Furthermore, 1 participant was excluded for providing no correct responses during the baseline phase of the task, which precluded the calculation of accurate baseline response times. This left 47 participants for formal analysis. Each participant was entered into a 3 (context: factory vs. church vs. prison) $\times$ 2 (race of prime: Black vs. White) $\times$ 2 (adjective valence: positive vs. negative) $\times$ 2 (context order: factory–church–prison, church–factory–prison) mixed-design ANOVA. The first three factors were within participants, and the last was between participants.

Independent Variables

Race of prime. Race was manipulated using primes consisting of six color photographs of college-aged White and Black men with neutral expressions. Photos were selected so that they did not include any jewelry or facial hair.

Context. The three digital images (800×600 pixels) used as contexts included the image of a church sanctuary, a prison cell, and a factory parking lot (see Figure 1).

Adjective valence. Six adjectives were selected from Anderson's (1968) list of trait adjectives to serve as targets for the priming task. Three positive adjectives (*pleasant*, *likeable*, *wonderful*) and three negative adjectives (*horrible*, *repulsive*, *annoying*) were used for the critical trials. In order to obtain an accurate measure of general evaluation, target adjectives were selected to have no association with either Black or White stereotypes.

Dependent Variable: Response Time

Response time to categorize the valence adjective again served as a dependent variable. As in Study 1, this was measured to the nearest millisecond by DirectRT software.

Procedure

The sequential priming procedure was very similar to Study 1, with a few adaptations to accommodate the use of three contexts (factory, church, prison), the use of two race prime categories (Black and White), and the use of general evaluative targets rather than performance-related target adjectives. During Phase 1, participants again practiced the word judgment task with filler adjectives as targets. The 12 baseline trials that followed presented each of the six evaluative adjectives twice to provide response

times to calculate facilitation scores from the critical Phase 4. During Phase 2, filler face primes were again flashed on the screen so participants could practice the face-recognition task. Phase 2 was modified to include 8 face-memory practice trials instead of 9, so that half the trials could be Black faces and half White faces. During Phase 3, participants again practiced both word judgment and facial recognition simultaneously. Again, filler Black and White faces were used with filler adjectives.

During the critical Phase 4, each of the three contexts was presented as a single block of 42 trials for a total of 124 trials. Each block included the 36 critical trials produced by every combination of six target adjectives with six face primes. Each contextual block also included six additional filler trials that were intended to maintain interest in the task, because Study 1 included only six face primes for critical trials rather than the nine used in Study 2. Filler trials used four faces (two White and two Black) and four target adjectives (two positive and two negative). As in Study 1, the order of the trials within each context block was randomly determined, and a break was provided midway through each contextual block.

The contextual blocks were presented in one of two orders (church–factory–prison or factory–church–prison) with a distractor task always occurring before the prison context. The orders of the contexts were not fully counterbalanced, because previous results showed that contexts that typically elicit a bias favoring Whites relative to Blacks when they preceded the prison context showed a significant decrease in that bias when they followed the prison context (Barden, Maddux, Petty, & Brewer, 2000). This was of particular concern in the current experiment, because a decrease in bias in the factory context would be in the direction of our predicted reversal in that context. There was also the possibility that the negativity of the factory context could impact responses in the prison context. Therefore, a distractor task was introduced prior to the presentation of the prison context in order to clear participants' minds. As a cover story for the distractor task, participants were told that the main priming task was finished and that they were now to complete a second, unrelated experiment on verbal ability. This distractor task consisted of seven scrambled sentences containing five words that participants were instructed to unscramble. Sentences were created to be neutral and innocuous in content (e.g., "Mark walked to work"). Following the completion of this task, participants were instructed that for pretesting purposes, they were to do one final contextual block. Participants were instructed that the computer was going to choose an additional context at random, and they were to perform the priming task as they had done earlier in the experiment. The trials for the prison context were then presented.

Following the priming task, participants were informed that it would not be necessary for them to perform the face-recall test that they had been led to expect, and that this modification to the usual evaluative priming procedure was made because of time constraints. None of the participants inquired about this omission. Participants were then debriefed and thanked for their time.

Results

Controlled Evaluation

A 3 (role: factory worker vs. churchgoer vs. prisoner) $\times$ 2 (race: Black vs. White) ANOVA was conducted initially. This analysis indicated that role did impact racial bias overall, $F(2, 226) = 4.64$, $p = .011$. Judgments were again compared between race and within role in order to provide an assessment of which of the three patterns of racial bias were obtained within each context. Within the prisoner role, White targets ($M = 4.00$, $SD = 1.03$) were evaluated more positively than Black targets ($M = 3.51$, $SD = 1.00$), $F(1, 70) = 4.07$, $p = .047$. Within the churchgoer role, Black targets ($M = 6.04$, $SD = 1.24$) were evaluated equal to Whites ($M = 6.05$, $SD = 1.14$), $F(1, 78) = 0.004$, $p = .951$. The

factory worker role elicited a reversal of the traditional pattern of racial bias, so that Blacks ($M = 6.01$, $SD = 0.81$) were favored relative to Whites ($M = 5.46$, $SD = 0.99$), $F(1, 78) = 7.38$, $p = .008$. Thus, the results from the explicit evaluation measure indicated that all three patterns of racial bias (in-group bias, no bias, and out-group bias) could be shown in the context of different social role information.

Automatic Evaluation

Error Rates, Outliers, and Facilitation Scores

Error rates, outliers, and facilitation scores were handled in the same manner as Study 1 with slight modifications to incorporate two prime races and three contexts. Across participants, relatively few errors were made in judging adjectives as positive or negative (3.7%), and these trials were excluded from further analyses. The data also exhibited few outliers (0.04%), defined as response times less than 300 ms or greater than 1,500 ms. A total of 12 facilitation scores were calculated for each participant, one for each of the 12 cells of the $3 \times 2 \times 2$ (Context $\times$ Race of Prime $\times$ Adjective Valence) within-participants factors of the design. Then, as in Study 1, a single index was calculated on the basis of the four facilitation scores from each context: White–Black bias index = (White prime, positive adjective – White prime, negative adjective) – (Black prime, positive adjective – Black prime, negative adjective). Thus, negative scores indicated in-group bias, the neutral point indicated egalitarian responding, and positive scores indicated out-group bias.

Race of Prime and Context Effects

The index of White–Black racial bias was initially analyzed based on a two-way, Context $\times$ Order MANOVA, which indicated that context order failed to produce a significant interaction or

main effect. The analyses were then collapsed across order condition. A one-way, within-participants ANOVA revealed a significant main effect of context, $F(2, 92) = 6.01$, $p = .004$, $\eta^2 = .116$, indicating that context significantly moderated automatic racial bias.

One-way analyses based on the context condition were then conducted in order to compare each pair of contexts. These analyses revealed that the effects in the prison and factory contexts were significantly different, $F(1, 46) = 11.06$, $p = .002$, $\eta^2 = .194$, as were the effects in the church and factory contexts, $F(1, 46) = 5.26$, $p = .026$, $\eta^2 = .103$. The effects in the prison and church contexts were not significantly different, $F(1, 46) = 1.49$, $p = .229$, $\eta^2 = .031$.

Separate analyses were then conducted using a single-sample t test comparing the index in each context to the zero point, which corresponds to no racial bias. These analyses revealed a significant pattern of in-group bias in the prison context ($M = -28.20$, $SD = 80.68$), $t(46) = -2.396$, $p = .021$, Cohen's $d = .350$. An egalitarian pattern of preference was observed in the church context ($M = -7.78$, $SD = 82.55$), $t(46) = -.646$, $p = .521$, $d = .094$. Finally, a significant pattern of out-group bias was observed in the factory context ($M = 28.29$, $SD = 89.01$), $t(46) = 2.179$, $p = .035$, $d = .318$, so that bias favored Blacks relative to Whites. Thus, participants showed a significant in-group bias in the prison context, no automatic evaluative biases in the church context, and a significant out-group bias in the factory context (see Figure 3).

Discussion

The results of Study 2 demonstrate that it is possible for all three potential patterns of group bias (in-group bias, no bias, and out-group bias) to occur even when one of the two groups is a racial in-group. As in Study 1, the controlled assessment verified that each specified role produced the expected pattern of racial bias.
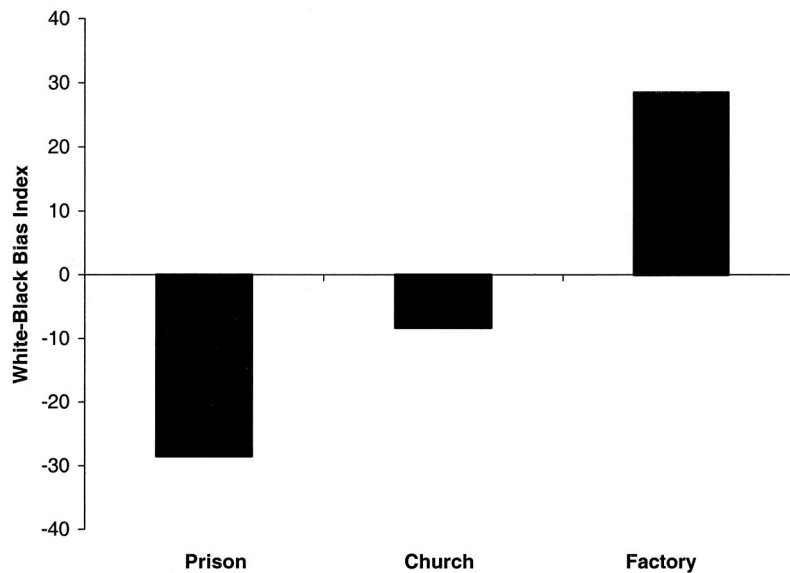

Figure 3. Contextual moderation of the index of White–Black automatic racial group bias across three contexts: prison, church, and factory. Values below the origin indicate in-group bias from White participants, values above indicate out-group bias, and values near the origin indicate the absence of group bias.

That is, the prisoner role produced in-group bias, the churchgoer role produced equal evaluation, and the factory worker role produced out-group bias.

The automatic results clearly reproduced the patterns of bias obtained on the controlled assessment. The prison context elicited in-group bias, so there was a bias favoring Whites relative to Blacks. In the church context, Whites and Blacks were evaluated equally, showing a complete attenuation of bias, replicating previous research (Wittenbrink et al., 2001). Finally, the factory context elicited a significant pattern of out-group bias, so that there was a bias favoring Blacks relative to Whites. These results indicate the strength of contextually implied roles to influence automatic racial biases, because this out-group bias occurred only moments before the same participants produced a pattern of in-group bias to the identical racial faces in the prison context. Furthermore, the magnitude of the out-group bias in the factory context (Cohen's $d = .318$) was comparable to the magnitude of the in-group bias in the prison context ($d = .350$).

The observed out-group bias in the factory context represents the first instance we have encountered in which White participants evaluated images of prototypic Black targets (Livingston & Brewer, 2002) more positively than White targets on a measure of automatic racial bias when race was not confounded with other variables (e.g., Mitchell et al., 2003). This is also the first time that one pattern of automatic racial-based bias has been significantly reversed when the groups included an in-group. Most importantly, the current results are consistent with the proposed differential role evaluation hypothesis. That is, all three contexts (prison, church, and factory) were carefully selected to imply clear roles (prisoner, churchgoer, and factory worker) for individuals appearing within them. Furthermore, the pattern of racial bias on the automatic measure, where role should have been assigned on the basis of context, was the same as the pattern of bias on the controlled measure where role was explicitly assigned.

Study 2 also had implications for the mechanisms underlying differential role evaluation. According to the role congruity mechanism (Eagly & Steffen, 1984), prejudice against a group in a given role is expected when the role is inconsistent with the stereotypically proscribed roles for that group (e.g., female in a leadership role). Given that the prisoner role is inconsistent with White stereotypic roles but consistent with Black stereotypic roles, this mechanism would predict a prejudice against Whites relative to Blacks. Thus, the finding of a bias favoring Blacks relative to Whites in the prison was inconsistent with this mechanism. On the other hand, the subtyping mechanism (Devine & Baker, 1991) was consistent with the out-group bias in the factory context, because a blue-collar Black subtype would likely be evaluated positively. In addition, the observed in-group bias in the prison could derive from a Black criminal subtype, which would likely be evaluated negatively. Thus the results from Study 2 suggest that role congruity does not provide a general mechanism to explain all of the current findings, whereas the current findings are still consistent with the subtyping mechanism.

Although the results of Study 2 provide additional support for the importance of roles in producing moderation of automatic racial biases, the evidence provided in Studies 1 and 2 was somewhat indirect. An inference was required to conclude that roles were responsible for the observed pattern of moderation on the automatic measure. Accordingly, Study 3 was designed to provide direct evidence that roles cause moderation of automatic racial biases.

STUDY 3

In Studies 1 and 2, physical contexts were selected based on the roles they implied for individuals appearing within them. The results support differential evaluation on the basis of race within the same role as the most plausible explanation of the findings, given the parallel results on the controlled measure, where roles were explicitly assigned, and on the automatic measure. Nevertheless, there is no direct evidence that roles are responsible for producing the observed relationship between context and moderation of automatic racial biases. If contexts produce moderation of automatic racial biases by activating social roles, then direct manipulation of roles even within the same surrounding context should also produce moderation.

Thus, the primary goal of Study 3 was to provide direct evidence that roles can in fact cause moderation of automatic racial biases by manipulating roles against the background of the prison context used in Study 2. Roles were manipulated by digitally placing clothing on each prime (either a formal suit or an orange jumpsuit) that corresponded to explicit instructions indicating that lawyers would appear in formal suits and prisoners would appear in orange jumpsuits (see Figure 1). The lawyer role was selected on the basis of a number of criteria, starting with the fact that a lawyer is a reasonable individual to observe visiting clients in prison. Secondly, being a lawyer implies employment, which may lead Whites to evaluate Blacks more positively (Devine & Baker, 1991; Sniderman & Piazza, 1993). Finally, as in the case of the factory worker role, there is reason to believe that the typical White lawyer may not be generally liked. For example, polls indicate that only 15% of respondents rated lawyers high or very high in honesty, just above car salesmen (Brehm, Kassin, & Fein, 2002). Thus, the lawyer role seems likely to have more positive implications for Black targets than White targets.

In addition to providing a direct test of whether roles can cause moderation of automatic biases, Study 3 provided a critical test of two alternative hypotheses for the way that contexts related to roles in producing the results of Studies 1 and 2. Our differential *role evaluation* hypothesis proposes that contexts lead both Blacks and Whites to be assigned the same role; however, the evaluative implications of the same role can differ on the basis of target racial cues. According to this account, the prison context elicited in-group bias in Study 2, because there was a bias favoring Whites as prisoners relative to Blacks as prisoners. This leads to the prediction that directly assigning both Whites and Blacks to the prisoner role would produce the same in-group bias. Similarly, when Black and White targets were both presented as lawyers, we expected a bias favoring Black targets relative to White targets (i.e., out-group bias).

Study 3 was also intended to provide a direct test of the differential *role assignment* hypothesis, which provides an alternative explanation for the automatic evaluation results of the two previous studies. This account suggests that roles moderate automatic racial biases when a contextual role is more frequently applied to one race than the other, which is analogous to the social role theory of gender stereotypes (Eagly & Steffen, 1984). For example, if the prisoner role were seen as more typical of Blacks than Whites, then

it would be applied to Black targets more readily than White targets when both were portrayed within a prison context. Thus, Blacks would be evaluated as Black prisoners and would be evaluated more negatively than Whites who were only categorized as Whites. Given that being a prisoner is a negative role, this would produce the pattern of in-group bias observed on the automatic measure of racial biases in Study 2. By the same token, Blacks would be favored relative to Whites in the factory context if Whites were categorized as low-status White factory workers whereas Blacks were only categorized as Blacks.

Although differential role assignment is unable to account for results obtained in Study 2 on the controlled ratings (where roles were explicitly indicated), it could have produced the results on the measure of automatic racial bias. To provide a critical test of the two different role hypotheses, Study 3 directly manipulated roles through both verbal instructions and a clothing manipulation, making it unavoidable for both races. When the same role is clearly applied to both races, the role assignment hypothesis predicts that no racial biases will be observed in either role, producing an egalitarian response pattern in both role conditions. By contrast, our proposed role evaluation hypothesis predicts that either in-group or out-group bias can occur depending on the role. By directly manipulating social roles in both the controlled and automatic measurements, Study 3 was designed to provide (a) direct support for the notion that social roles are a contextual variable capable of causing moderation and (b) a critical test pitting the differential role evaluation hypothesis against the differential role assignment hypothesis.

Method

Controlled Evaluation

As in the previous studies, a controlled assessment was conducted to verify the expected patterns of bias in the Role × Race design. The method for this assessment is described under Study 2 (above), because assessments for both studies were collected simultaneously. Study 3 made use of the lawyer and prisoner roles, which were both portrayed in the prison context wearing clothing corresponding to each role (as presented in Figure 1). Thus, the image presented for the controlled assessment was identical to images presented in the automatic measure.

Automatic Evaluation

Participants

Eighty-one introductory psychology students at Ohio State University voluntarily participated in partial fulfillment of a course requirement and were recruited in the same manner as in the previous studies. The data from 13 minority students were excluded from analysis, as were the data from 8 participants who had error rates in excess of 45%. This left the data from 60 White participants for formal analysis. Each participant was entered into a 2 (role: prisoner vs. lawyer) × 2 (race of prime: Black vs. White) × 2 (adjective valence: positive vs. negative) × 2 (role order: prisoner first vs. lawyer first) × 2 (prime counterbalancing: Primes 1–4 prisoner and 5–8 lawyer vs. Primes 1–4 lawyer and 5–8 prisoner) mixed design, with the first three factors being within participants and the last two factors being between participants.

Independent Variables

Role. Our manipulation of social role was accomplished through verbal instructions and by digitally manipulating the clothing the primes were

wearing. Before one block of critical trials, participants were informed that the subsequent block would include prisoners appearing in orange jumpsuits, and before the other block they were told that the block would include lawyers appearing in formal suits. Accordingly, digital imaging software was used to fit an orange jumpsuit and a formal suit onto 8 Black and 8 White face primes, making a total of 32 prime images (see Figure 1). To preserve the sense that prisoners and lawyers are separate groups of people, participants saw four of the Black primes as lawyers and the other four Black primes as prisoners. The clothing was obtained from actual digital images of attorneys and prisoners found on the Internet. The prison context from Study 2 appeared in the background during both the prisoner and the lawyer blocks.

Race of prime. Race was manipulated using primes consisting of 16 color photographs of college-aged White and Black men. As in previous studies, photos were selected so that they exhibited neutral facial expression and did not include any jewelry or facial hair.

Adjective valence. The adjectives used for the critical trials were identical to Study 2, with three positive adjectives (*pleasant*, *likeable*, *wonderful*) and three negative adjectives (*horrible*, *repulsive*, *annoying*).

Role order. Participants were randomly assigned to one of two role orders (lawyer–prisoner, prisoner–lawyer).

Prime counterbalancing. Half of the participants were randomly assigned to see Primes 1–4 for each race as prisoners and Primes 5–8 for each race as lawyers. The other condition saw Primes 5–8 for each race as prisoners and Primes 1–4 for each race as lawyers. This design meant that memory detection for each role could be assessed separately, because a given participant saw a given face in only one role.

Dependent Variables

Response time. As in Studies 1–2, response time to categorize the valence adjectives was measured to the nearest millisecond by DirectRT software.

Face detection. Participants were presented with a total of 32 Black and White faces in random order, including the 16 that had been previously presented and 16 fillers that had never been presented. Participants responded on the basis of whether they had or had not seen the face earlier in the experiment, and responses were recorded and compared with the correct answers. For this task, all faces were shown to the neck without any clothing visible.

Procedure

Up until the critical Phase 4 of the priming measure, Study 3 was identical to Study 2. Several changes were made to Phase 4 of the evaluative priming procedure in order to implement the manipulation of social role.

The manipulation of social role included two components: an explicit verbal instruction, indicating which type of prime (prisoner or lawyer) would appear in the subsequent block, and a manipulation of the clothing the primes were wearing. At the outset of Phase 4, participants were told that two components would be added to the procedure to make it more true to real life. The instructions read, "In general, when we encounter people they appear wearing certain clothing, and they appear within certain physical surroundings." Participants were then given instructions corresponding to the block that followed:

> In the next set of trials, you will be shown lawyers [prisoners] who will appear in formal suits [orange jumpsuits]. In addition to trying to remember their faces, you will need to remember which faces appeared wearing formal suits [orange jumpsuits], because you will be tested on this later.

This memory instruction was intended to ensure that participants paid attention to the clothing the prime was wearing, although no clothing-memory task was used.

The two blocks of critical trials were presented in one of two orders (lawyer–prisoner or prisoner–lawyer) with no distractor task in between. There were a total of 96 critical trials, with each block comprising 48 trials, which came from fully crossing the six target adjectives and the eight race primes assigned to that role. Because these additional face primes increased the number of face primes seen in this study, no filler faces were used in Study 3. The sequence of each trial was identical to that used in Study 2, with the prison context image remaining in the background throughout both blocks of trials.

Following the priming procedure, participants performed a face-detection task. Participants indicated whether they had or had not seen each face appearing on the screen during the previous trials. Then, 16 Black and White faces that had appeared as primes and 16 faces that had not appeared previously were presented. All faces were shown without the role clothing. After that, participants were debriefed and thanked for their time.

Results

Controlled Evaluation

In order to test whether role had any impact on racial biases overall, an initial 2 (role: prisoner vs. lawyer) $\times$ 2 (race: Black vs. White) ANOVA was conducted, which indicated that role did impact racial bias overall, $F(1, 140) = 7.82$, $p = .006$. As reported for Study 2, White targets in the prisoner role were evaluated more positively ($M = 4.00$, $SD = 1.03$) than Black targets ($M = 3.51$, $SD = 1.00$), $F(1, 70) = 4.07$, $p = .047$. Furthermore, this was significantly reversed within the lawyer role, where Black targets ($M = 6.13$, $SD = 1.20$) were favored relative to Whites ($M = 5.47$, $SD = 1.55$), $F(1, 70) = 3.96$, $p = .050$, indicating that direct manipulation of clothing, and thus an even stronger role implication, produced the predicted pattern of results on a controlled measure.

Automatic Evaluation

Error Rates, Outliers and Facilitation Scores

Error rates, outliers, and facilitation scores were handled in the same manner as in Studies 1 and 2, with slight modifications to incorporate two races and two roles. Across participants, relatively few errors were made in judging adjectives as positive or negative (2.9%), and these trials were excluded from further analyses. The data also exhibited few outliers (0.2%), defined as response times less than 300 ms or greater than 1,500 ms. Eight facilitation scores were calculated to fill out the $2 \times 2 \times 2$ (Role $\times$ Race of Prime $\times$ Adjective Valence) cells of the within-participants portion of the design. Then, the same White–Black bias index was calculated as it was for Study 2. Negative scores on this index indicated in-group bias, the neutral point indicated egalitarian responding, and positive scores indicated out-group bias.

Race of Prime and Role Effects

The index of White–Black racial bias was initially analyzed on the basis of a two-way, Role $\times$ Order MANOVA, which indicated that context order failed to interact with racial preferences. The analyses were then collapsed across order condition. A one-way, within-participants ANOVA revealed a significant main effect of role, $F(1, 59) = 8.96$, $p = .004$, $\eta^2 = .132$, indicating that role significantly moderated automatic racial bias (see Figure 4).

Separate analyses were then conducted using a single-sample t test comparing the index in each role to the zero point, which corresponds to no bias. A significant pattern of in-group bias was observed in the prisoner role ($M = -17.94$, $SD = 63.88$), $t(59) = -2.176$, $p = .034$, Cohen's $d = .281$, indicating a bias favoring White prisoners relative to Black prisoners, as was observed in Study 2. However, a significant pattern of out-group bias was observed in the lawyer role ($M = 18.38$, $SD = 67.99$), $t(59) = 2.094$, $p = .041$, $d = .270$, indicating a bias favoring Black lawyers relative to White lawyers.[3]

Facial Recognition

The facial detection results indicated that significant attention was paid to the individuating features of targets in each race and face combination and that the level of individuation was uniform across all four race and role combinations. Face-recognition task results were analyzed by subtracting the proportion of false alarms on filler trials (filler not previously presented that the participant incorrectly identified as having been presented) from the proportion of correctly identified hits (previously presented faces correctly identified as such; see Fazio et al., 1995). A zero on this detection index would indicate performance at chance levels. Overall detection was significantly better than chance ($M = .59$, $SD = .20$), $t(59) = 21.83$, $p < .001$. When the detection index was calculated separately for each Race $\times$ Role condition, facial detection was significantly better than chance in all four conditions (all Ms $> .52$), all ts$(59) > 15$, all ps $< .001$, indicating that attention was paid to the individuating features of all four types of facial prime. Furthermore, an overall 2 (race: Black vs. White) $\times$ 2 (role: lawyer vs. prisoner) within-participants ANOVA showed

---

[3] The data from the automatic measures used in all three of the current studies were reanalyzed using an inverse transformation. The results of these analyses show that the critical patterns of results remain the same under these conditions. In all three studies, the instance of contextual moderation remained significant when response times for both the baseline and critical trials were inverse transformed. That is, when one context produced significant preference and a second produced a significant reversal of that preference, the transformed data reflected significant contextual moderation in the same direction. In addition, analyses were conducted to show that both the pattern of automatic bias and the pattern of reversal of bias were reliable across the three studies. Participants from all three studies ($N = 194$) were entered into one simultaneous analysis. This was done so that contexts that elicited typical patterns of bias were collapsed across (including Asians and Blacks in the classroom context and Whites and Blacks in the prison context and prisoner role). Similarly, contexts that elicited reversals were collapsed across (including Asians and Blacks in the basketball court and Whites and Blacks in the factory context and lawyer role). Thus, Asians in Study 1 were viewed as functionally equivalent to Whites in Studies 2 and 3. Results indicate that there was a significant overall Context $\times$ Race of Prime $\times$ Adjective Valence interaction, $F(1, 193) = 17.90$, $p < .001$, $\eta^2 = .085$. Furthermore, the typical pattern of bias was reliable when collapsing across the classroom, prison, and prisoner contexts, as indicated by a significant Race of Prime $\times$ Adjective Valence interaction, $F(1, 193) = 11.11$, $p = .001$, $\eta^2 = .054$. This pattern of bias was significantly reversed when collapsing across the basketball court, factory, and lawyer contexts, $F(1, 193) = 5.53$, $p = .020$, $\eta^2 = .028$. It is important to note that a variable coded for study failed to qualify any of these three interactions. Thus, the critical findings remained when the response time data were inverse transformed.

Figure 4. Contextual moderation of the index of White–Black automatic racial group bias across two roles: lawyer and prisoner. Values below the origin indicate in-group bias from White participants, and values above indicate out-group bias.

no significant main effects for either role, $F(1, 59) = 2.40$, $p = .126$, or race, $F(1, 59) = 1.65$, $p = .203$. Furthermore, the Race $\times$ Role interaction was not significant, $F(1, 59) = 1.21$, $p = .275$, indicating that differential attention to individuating features of the primes did not parallel the moderation of automatic racial attitudes by roles.

Discussion

The results of Study 3 provide direct evidence that social roles can moderate automatic racial biases. When face primes were explicitly portrayed as prisoners, via both instructions and relevant clothing, an evaluative bias favoring White targets relative to Black targets emerged. This was consistent with the in-group bias shown in the prison context in Study 2. However, when face primes were portrayed as lawyers via instructions and clothing, an evaluative bias favoring Black targets relative to White targets emerged. This finding was consistent with the out-group bias obtained in the factory context in Study 2. Thus, in Study 3, the direct manipulation of social role produced a third significant reversal of the typical pattern of automatic racial bias. It is important to note that the lawyer role influenced evaluations enough to overwhelm the pattern of in-group racial bias typically obtained in the prison context and conceptually similar contexts such as a ghetto (Wittenbrink et al., 2001).

The second goal of Study 3 was to provide a critical test of two alternative explanations for how contextual roles influence automatic racial biases. The evidence clearly supported the differential role evaluation hypothesis, because the reversal of racial bias occurred when both Black and White targets were assigned to the same social role. This reversed pattern could only have occurred if racial bias resulted from the interaction of race and role. Through the direct manipulation of roles, Study 3 provides the first direct evidence that the same role can have differential implications for automatic evaluations based on race. That is, the evaluative implications of being a lawyer depended on whether the target was

Black or White. Thus, Study 3 provides the most definitive support for the differential role evaluation hypothesis to date.

GENERAL DISCUSSION

A review of the literature in the area of moderation of automatic racial biases revealed that existing explanations are somewhat interchangeable, and thus in each case it is unclear what contextual variable is truly responsible for producing the observed moderation. Accordingly, the current investigation provides three advances: (a) It introduces the differential role evaluation hypothesis as one way to account for how different contexts (e.g., prison, factory) can moderate automatic racial biases (i.e., different contexts imply different roles that are evaluated differently for different races), (b) it demonstrated a reversal pattern in racial biases that was not present in previous research and is predicted by the proposed role evaluation hypothesis, and (c) it demonstrated that manipulating social roles produces reversals on both controlled and automatic evaluations.

The primary goal of this research was to isolate roles as one important contextual variable that can produce moderation and even reversals of automatic racial biases. Contexts that clearly implied roles produced moderation in Studies 1 and 2, and direct manipulation of roles based on instructions and clothing produced moderation in Study 3. In each case, the automatic measure of racial bias paralleled the patterns of racial bias that were produced on the controlled assessments, where social roles were explicitly specified. These results occurred across a variety of contexts invoking different roles (basketball player, student, factory worker, churchgoer, prisoner, and lawyer) and included biases between two out-groups (Study 1) and between an in-group and an out-group (Studies 2 and 3).

A related goal was to test the two components of the differential role evaluation hypothesis. The first component holds that roles are automatically applied to individuals appearing within a context when they are not already assigned to a role. This was supported

by three pieces of evidence: (a) When targets were not already assigned a role, being perceived within the prison context elicited automatic in-group bias (Study 2), the same bias that occurred when the prisoner role was directly manipulated through clothing and instructions (Study 3); (b) when the target was already assigned to the lawyer role on the basis of clothing and instructions, the assigned role rather than the context determined the pattern of racial bias (Study 3); and (c) in Studies 1 and 2, contexts that were selected to clearly imply roles produced moderation of automatic racial biases in the same pattern as controlled assessments where the same roles were explicitly provided. Therefore, evidence from both the controlled and automatic racial biases across three studies supports the notion that roles can be automatically applied to individuals appearing within a context when they are not explicitly assigned to a role by instruction.

The second component of the proposed hypothesis, which is based on the interactionist viewpoint (Brewer, 1988; Higgins & Rholes, 1976; Taylor, 1981), states that being assigned to the same role can have different evaluative implications for targets from different races. This was supported by the following evidence: (a) The controlled assessments in all three studies showed that one role elicited one pattern of racial bias and a second role reversed that bias, extending the interactionist hypothesis from gender to race; (b) in Studies 1 and 2, one context established a pattern of automatic bias and a second significantly reversed that bias, which could only occur if race interacted with a contextual variable to produce evaluations; and (c) in Study 3, both races were explicitly assigned to the same roles, and role produced a significant reversal. Study 3 was also inconsistent with the alternative that moderation occurred because the same context led different roles to be assigned based on race. Thus, overall, the evidence supports the differential role evaluation hypothesis.

Underlying Mechanisms

According to the differential role evaluation hypothesis, contextual moderation of automatic racial bias can occur on the basis of two pieces of information: social roles from the context and racial cues from the target. Furthermore, the hypothesis holds that when the same role is assigned to two different races, this results in a pattern of racial bias that reflects an interaction between the race and role information. It is important to note that this general hypothesis does not specify any specific mechanism to explain what occurs between the extraction of race and role information and the interactive evaluative outcome. Because this remains a critical issue, the relative merits of three candidate mechanisms are discussed here in light of the current results: role congruity, shifting standards, and subtyping.

If contextual roles cause prejudice when they are incongruent with roles that are stereotypical of a given race, this could provide one process mediating the current findings. Eagly and Karau's (2002) role congruity theory holds that prejudice against female leaders occurs because gender role stereotypes are viewed as incongruent with leadership roles. Just as this general mechanism has been extended to automatic gender prejudice (Rudman & Kilianski, 2000), it provided useful predictions when extended to automatic racial bias in the current Study 1. That is, racial prejudice was observed whenever racial targets assumed stereotype-incongruent roles (e.g., Asian basketball players, Black students).

However, predictions were less accurate for Studies 2 and 3. Given that the lawyer role is stereotype incongruent for Blacks, this should have elicited prejudice against Blacks relative to Whites. Similarly, given that the prisoner role is stereotype incongruent for Whites, this should have elicited prejudice against Whites relative to Blacks. Taken as a whole, the current investigation suggests that role congruity may be one mechanism that underlies the interactive effects observed under the differential role evaluation hypothesis, at least under some circumstances.

The shifting-standards approach to group-based judgments (Biernat & Manis, 1994; Biernat, Manis, & Nelson, 1991) provides a second process that might be considered as underlying our findings. According to this model, there are two kinds of judgments, objective (e.g., dollars and liters) and subjective (e.g., traits and evaluations; see also Upshaw, Ostrom, & Ward, 1970). For the subjective judgments used in the current investigation, this model predicts that target evaluations will be made relative to the standard set by the target's group. The model predicts that compared with objective measures, subjective measures can show either attenuated or reversed stereotype judgments. Although this is not a very specific prediction, it could be used to explain the reversals in the factory worker and lawyer roles, because both are good roles for a Black person relative to other Blacks but not as good roles for a White person relative to other Whites. However, this approach fails to explain the in-group bias observed in the prisoner role, because a prisoner should be a worse role for Whites relative to other Whites than it is for Blacks relative to other Blacks. It also fails to explain either the bias favoring Asians relative to Blacks as students or the bias favoring Blacks relative to Asians as basketball players.

Finally, if contextual roles cause differential activation of racial subtypes, this would provide a process capable of mediating the current findings (Brewer & Lui, 1984; Devine & Baker, 1991). The single investigation into racial subtypes isolated two Black subtypes that were both evaluated positively: the Black businessman and the Black athlete (Devine & Baker, 1991). These subtypes provide a good account for the reversals of racial bias observed in the basketball player and lawyer roles, particularly because subtypes are viewed as isolated from global racial categories. The addition of three Black subtypes (blue-collar Black, Black student, and Black prisoner) would provide a more complete account for the current results, assuming that they could be established empirically and would receive the expected evaluations. This is not necessarily a given; for example, a Black student subtype should be evaluated negatively, according to the stereotype of Blacks as not intellectual (Wittenbrink et al., 1997), but positively according to the shifting-standards approach (Biernat & Manis, 1994). Nevertheless, because none of the current findings run counter to established racial subtypes, subtyping provides a plausible general mechanism underlying all of our findings.

One final point is that a subtyping account does not reduce the importance of roles for the moderation of automatic racial bias. It could be that racial categorization occurs first, in which case roles would determine which subtypes were activated within the race. Conversely, it could be that role categorization occurs first and that racial category determines which racial subtypes are activated within the role. In either case, roles would critically impact the activation of subtypes.

In summary, the shifting-standards approach fails to predict the results in any one of the current studies, so it does not provide a likely mechanism underlying our findings. On the other hand, role congruity appears to be one mechanism capable of producing interactive effects under the differential role evaluation hypothesis, at least in some cases. Finally, racial subtypes provide the most promising general mechanism, because they predict two of the current reversals and are not inconsistent with any of the current findings. Future research should investigate the existence of other racial subtypes in order to establish whether this approach can provide a complete account for findings in this area.

Putting Attitude Measurement in Context

The current findings raise a number of critical questions regarding the methods typically used to measure automatic racial biases. One might suppose that the blank screen typically used in evaluative priming procedures (Fazio et al., 1995), lexical decision tasks (Dovidio et al., 1997), and the IAT (Greenwald et al., 1998) would produce a decontextualized measure of racial bias. The current results suggest that when faced with an impoverished blank screen, participants may respond on the basis of temporary contextual influences rather than more stable chronic associations. A student who comes to the experiment after a game of basketball may provide very different responses from a student who comes straight from class. Because the current investigation did not incorporate a blank context condition, it is unclear what pattern of bias this would produce in our samples. Nevertheless, the large contextual fluctuations observed in the current investigation suggest that the low test–retest reliability of evaluative priming measures (Fazio & Olson, 2003) may reflect meaningful variance in temporarily activated roles. Of course, a substantial portion of this low reliability undoubtedly derives from measurement noise (Cunningham, Preacher, & Banaji, 2001). On the other hand, the IAT may have better test–retest reliability (Cunningham et al., 2001) because the measure itself requires participants to categorize stimuli solely on the basis of one categorization (Fazio & Olson, 2003). This allows for moderation of prejudiced responses only under more limited circumstances, such as when the representation of the global category is changed (Dasgupta & Greenwald, 2001) or one single categorization is replaced by another (Mitchell et al., 2003).

The current results suggest that multiple simultaneous categorization is a reality that should be incorporated into both controlled and automatic measures of racial bias. The highly contextual nature of racial biases may also contribute to the difficulty in finding relationships between automatic and controlled measures when care is not taken to ensure that the attitude object, the appropriate contextual factors, and the measurement approach correspond (Ajzen & Fishbein, 1977). Overall, the current results indicate that abstracting race stimuli from any physical context exaggerates the importance of global attitudes and obscures the importance of substantial contextual variation.

Previous research on the impact of category assignment on priming measures has focused on demonstrating that categories at the same level of specificity can exhibit inverse relationships (Kunda & Sinclair, 1999; Macrae et al., 1995; Mitchell et al., 2003; Rudman & Borgida, 1995). The current findings emphasize the point that automatic social judgments may result from not one but multiple categorizations of the same individual (e.g., role and race). The implication is that there may be no upper limit to the number of categories assigned on the basis of the target's facial appearance (e.g., race, gender, facial expression, age, etc.) and the broader context (e.g., clothing, role, status, nationality, etc.). The current evidence clearly shows that the evaluations that result from multiple simultaneous categorizations reflect emergent properties of these combined categories that violate adding or averaging. Appearing in front of a dingy factory actually has positive consequences for Blacks appearing in the context relative to the impact it has on Whites. The current investigation speaks to how adept people are at extracting information from multiple sources and combining it together in complex ways to produce both controlled and automatic patterns of bias.

References

Ajzen, I., & Fishbein, M. (1977). Attitude–behavior relations: A theoretical analysis and review of empirical research. *Psychological Bulletin, 84,* 888–918.

Anderson, N. H. (1968). Likableness ratings of 555 personality–trait words. *Journal of Personality and Social Psychology, 9,* 272–279.

Barden, J., Maddux, W. W., Petty, R. E., & Brewer, M. B. (2000). [Response latencies from evaluative priming procedure]. Unpublished raw data.

Bargh, J. A. (1999). The cognitive monster: The case against the controllability of automatic stereotype effects. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 361–382). New York: Guilford Press.

Biernat, M., & Manis, M. (1994). Shifting standards and stereotype-based judgments. *Journal of Personality and Social Psychology, 66,* 5–20.

Biernat, M., Manis, M., & Nelson, T. E. (1991). Stereotypes and standards of judgment. *Journal of Personality and Social Psychology, 60,* 485–499.

Blair, I. V. (2002). The malleability of automatic stereotypes and prejudice. *Personality and Social Psychology Review, 6,* 242–261.

Brehm, S. S., Kassin, S. M., & Fein, S. (2002). *Social psychology*. Boston: Houghton Mifflin.

Brewer, M. B. (1988). A dual-process model of impression formation. In T. K. Srull & R. S. J. Wyer (Eds.), *Advances in social cognition* (Vol. 12, pp. 1–36). San Diego, CA: Academic Press.

Brewer, M. B., & Lui, L. (1984). Categorization of the elderly by the elderly: Effects of perceiver's category membership. *Personality and Social Psychology Bulletin, 10,* 585–595.

Cunningham, W. A., Preacher, K. J., & Banaji, M. R. (2001). Implicit attitude measures: Consistency, stability, and convergent validity. *Psychological Science, 121,* 163–170.

Dasgupta, N., & Greenwald, A. G. (2001). On the malleability of automatic attitudes: Combating automatic prejudice with images of admired and disliked individuals. *Journal of Personality and Social Psychology, 81,* 800–814.

DeSteno, D., Dasgupta, N., Bartlett, M. Y., & Cajdric, A. (in press). Prejudice from thin air: The effect of emotion on automatic intergroup attitudes. *Psychological Science.*

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology, 56,* 5–18.

Devine, P. G. (2001). Implicit prejudice and stereotyping: How automatic are they? Introduction to the special section. *Journal of Personality and Social Psychology, 81,* 757–759.

Devine, P. G., & Baker, S. M. (1991). Measurement of racial stereotype subtyping. *Personality and Social Psychology Bulletin, 17,* 44–50.

Dovidio, J. F., Kawakami, K., Johnson, C., Johnson, B., & Howard, A. (1997). On the nature of prejudice: Automatic and controlled processes. *Journal of Experimental Social Psychology, 33,* 510–540.

Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review, 109,* 573–598.

Eagly, A. H., & Steffen, V. J. (1984). Gender stereotypes stem from the distribution of women and men into social roles. *Journal of Personality and Social Psychology, 46,* 735–754.

Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fide pipeline? *Journal of Personality and Social Psychology, 69,* 1013–1027.

Fazio, R. H., & Olson, M. A. (2003). Implicit measures in social cognition research: Their meaning and uses. *Annual Review of Psychology, 54,* 297–327.

Fiske, S. T. (1998). Stereotyping, prejudice, and discrimination. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vol. 2, pp. 357–411). New York: McGraw-Hill.

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

Higgins, E. T., & Rholes, W. S. (1976). Impression formation and role fulfillment: A "holistic reference" approach. *Journal of Experimental Social Psychology, 12,* 422–435.

Jarvis, B. (2000). DirectRT (Version 2000.4.9) [Computer software]. New York: Empirisoft.

Kao, G. (2000). Group images and possible selves among adolescents: Linking stereotypes to expectations by race and ethnicity. *Sociological Forum, 15,* 407–430.

Kunda, Z., & Sinclair, L. (1999). Motivated reasoning with stereotypes: Activation, application, and inhibition. *Psychological Inquiry, 10,* 12–22.

Livingston, R. W., & Brewer, M. B. (2002). What are we really priming? Cue-based versus category-based processing of facial stimuli. *Journal of Personality and Social Psychology, 82,* 5–18.

Lowery, B. S., Hardin, C. D., & Sinclair, S. (2001). Social influence effects on automatic racial prejudice. *Journal of Personality and Social Psychology, 81,* 842–855.

Macrae, C. N., Bodenhausen, G. V., & Milne, A. B. (1995). The dissection of selection in person perception: Inhibitory processes in social stereotyping. *Journal of Personality and Social Psychology, 69,* 397–407.

Mitchell, J. P., Nosek, B. A., & Banaji, M. R. (2003). Contextual variations in implicit evaluation. *Journal of Experimental Psychology: General, 132,* 455–469.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84,* 231–259.

Rudman, L. A., & Borgida, E. (1995). The afterglow of construct accessibility: The behavioral consequences of priming men to view women as sexual objects. *Journal of Experimental Social Psychology, 31,* 493–517.

Rudman, L. A., & Kilianski, S. E. (2000). Implicit and explicit attitudes toward female authority. *Personality and Social Psychology Bulletin, 26,* 1315–1328.

Sniderman, P. M., & Piazza, T. (1993). *The scar of race.* Cambridge, MA: Belknap Press.

Taylor, S. E. (1981). A categorization approach to stereotyping. In D. L. Hamilton (Ed.), *Cognitive processes in stereotyping and intergroup behavior* (pp. 83–113). Hillsdale, NJ: Erlbaum.

Upshaw, H. S., Ostrom, T. M., & Ward, C. D. (1970). Content versus self-rating in attitude research. *Journal of Experimental Social Psychology, 6,* 272–279.

Wittenbrink, B., Judd, C. M., & Park, B. (1997). Evidence for racial prejudice at the implicit level and its relationship with questionnaire measures. *Journal of Personality and Social Psychology, 72,* 262–274.

Wittenbrink, B., Judd, C. M., & Park, B. (2001). Spontaneous prejudice in context: Variability in automatically activated attitudes. *Journal of Personality and Social Psychology, 81,* 815–827.