

Research Report

Color Blindness and Interracial Interaction

Playing the Political Correctness Game

Michael I. Norton,¹ Samuel R. Sommers,² Evan P. Apfelbaum,² Natassia Pura,² and Dan Ariely³

¹Harvard Business School, ²Tufts University, and ³Massachusetts Institute of Technology

ABSTRACT—*Two experiments explored the ramifications of endorsing color blindness as a strategy for appearing unprejudiced. In Study 1, Whites proved adept at categorizing faces on the basis of race, but understated their ability to do so. In Study 2, Whites playing the Political Correctness Game—a matching task that requires describing other individuals—were less likely to use race as a descriptor when paired with a Black partner than when paired with a White partner, a strategy that impaired communication and performance. In addition, avoidance of race was associated with Whites making less eye contact with and appearing less friendly toward Black partners.*

In many conversations, people are called upon to describe others solely on the basis of their appearance: Who was it that showed up late for the meeting? Which character in the movie said that? The new guy—what does he look like again? In such cases, the best information to use is that which is most diagnostic in distinguishing the target from other individuals—for example, hair color, height, gender, and race. We propose (and anecdotal experience suggests) that people vary in their willingness to use certain descriptors: Although having red hair and being Black might be equally diagnostic in a particular setting, we suggest that many individuals are more reluctant to use race than hair color in their descriptions. The present investigation explored how White individuals behave in situations in which the desire to appear unprejudiced leads to efforts to appear color-blind.

Why would simply mentioning someone's race serve as evidence of bias? There is, after all, nothing inherently racist about

noticing race. But in a culture where motivations to avoid appearing prejudiced are increasingly pervasive (Gaertner & Dovidio, 1986; Plant & Devine, 1998) and few labels are as aversive as that of “racist” (Crandall, Eshleman, & O'Brien, 2002; Sommers & Norton, 2006), color blindness can serve as a useful stratagem: *If I do not notice race, then I cannot be a racist.* Certainly, noticing race is a necessary precursor to racism, but we propose that noticing race can be perceived as a sufficient indication of racism: *People who do not notice race are not racist, whereas those who do notice race probably are.* Indeed, noticing race does lead to the activation of stereotypic associations (Devine, 1989; Greenwald & Banaji, 1995), which can in turn lead to prejudicial behavior (Correll, Park, Judd, & Wittenbrink, 2002; Dovidio, Kawakami, & Gaertner, 2002). Thus, one mechanism for decreasing bias could be simply not to notice someone's race, thereby forestalling the associations that result in prejudicial behavior. In most cases, however, perceivers do encode the race of other people relatively effortlessly and rapidly (Ito & Urland, 2003; Montepare & Opeyo, 2002; see Cosmides, Tooby, & Kurzban, 2003), and such information does affect judgments, though decision makers are reluctant to acknowledge this influence (Norton, Vandello, & Darley, 2004; Sommers & Norton, in press). We propose that the incongruity between trying to appear color-blind while automatically noticing color complicates strategic efforts to appear unbiased, creating an inevitable tension between efforts to achieve color blindness and actual success at doing so.

The present research constitutes an empirical investigation of Whites' efforts to fulfill their desire to appear unprejudiced by attempting to be color-blind. In Study 1, we examined Whites' reluctance to admit to their facility at categorizing other people on the basis of race. In Study 2, we examined the consequences of this reluctance by introducing the Political Correctness Game, a two-player photograph-identification task that contrasts efficiency of communication with the desire to appear color-blind.

Address correspondence to Michael I. Norton, Harvard Business School, Soldiers Field Rd., Boston, MA 02163, e-mail: mnorton@hbs.edu, or to Samuel Sommers, Department of Psychology, Tufts University, 490 Boston Ave., Medford, MA 02155, e-mail: sam.sommers@tufts.edu.

STUDY 1: CLAIMING COLOR BLINDNESS

Though research suggests that individuals exhibit an own-race bias in memory for faces (e.g., Malpass & Kravitz, 1969), merely categorizing faces by race is a simpler task (see Stangor, Lynch, Duan, & Glass, 1992). In this study, we sought to demonstrate that Whites' reports of difficulty in classifying other people on the basis of race are not borne out by their actual ability to do so.

Method

Fifty-seven White students (36 females) participated in exchange for \$10. Photographs of volunteers were taken at a mall. We created a set of 24 photographs for each of seven categorical dimensions: race (Black/White), gender (male/female), age (over 30/under 25), color of the background in the photo (blue/red), hair color (light/dark), facial expression (smiling/not smiling), and facial hair (present/absent).

In the *sorting task*, participants completed seven blocks of a computer task; in each block, they categorized 24 photos on one of the seven dimensions. When a photo appeared, participants used the mouse to click on the correct classification (e.g., "White" or "Black" when the photos were to be categorized by race).

In the *hypothetical task*, participants read a questionnaire that included sample photos. The questionnaire asked them to imagine themselves performing the computer-based sorting task. Participants ranked the seven dimensions in terms of how quickly they believed they would be able to categorize the photos on the basis of that dimension.

Results and Discussion

Participants in the sorting task were at least 95% accurate in their categorizations across all seven dimensions, and there was no difference in accuracy across dimensions, $F(6, 114) = 1.65$, n.s.; accuracy for the race categorization was 99.1%. In order to compare results across the two tasks, we converted response times in the sorting task to rank scores from 1 through 7, similar to the self-reported ranks from the hypothetical task.

Participants in the sorting task were quickest to categorize faces by background color; categorizing by gender came next, followed by categorizing by race ($M = 3.60$). In the hypothetical task, in contrast, participants estimated that race would be their second slowest dimension ($M = 4.43$), ranking only age lower. Whites thus underestimated the speed with which they would be able to categorize by race, $t(47.3) = 2.07$, $p_{\text{rep}} = .93$, $d = 0.55$, via Welch's separate-variance t test. The only other dimension to vary by task was gender: Participants overestimated their speed in categorizing by gender ($M = 2.19$ in the hypothetical task) compared with their actual performance ($M = 3.05$), $t(31.4) = 2.09$, $p_{\text{rep}} = .92$, $d = 0.60$. In essence, participants substituted a less controversial dimension—gender—for a more controversial one, race (Rodin, Price, Bryson, & Sanchez, 1990).

Of course, it is possible that individuals are simply generally unaware of their facility at racial categorization. We addressed this alternative explanation by conducting a follow-up study with Black individuals ($N = 22$). Although Black participants' ability to categorize by race ($M = 3.67$) was comparable to that of Whites in Study 1, $t < 1$, their estimates of this ability ($M = 2.78$) were significantly higher than Whites', $t(36.7) = 3.68$, $p_{\text{rep}} = .99$, $d = 1.04$. Rather than supporting a general tendency for individuals to lack awareness of their facility with racial categorization, these data support the conjecture that Whites' underestimation of this ability results from a specific desire—one not shared by our Black participants—to appear unprejudiced. Finally, the fact that no underestimation emerged in Whites' reports of ability to categorize by background color (red/blue) suggests that it is not color itself but rather the meaning of color when associated with ethnicity that led to the observed discrepancies among Whites.

STUDY 2: THE POLITICAL CORRECTNESS GAME

Having demonstrated Whites' reluctance to admit the extent to which they use race to differentiate other people, we examined some of the possible consequences of this reluctance. In Study 2, White participants engaged in a two-player game (see Weber & Camerer, 2003) in which they asked a partner questions in order to identify a target face in a set of photos that varied on the same dimensions used in Study 1. We manipulated whether participants worked with a White or Black confederate, expecting participants to be particularly likely to underutilize race as a descriptor when interacting with a Black partner because of increased pressure to appear unprejudiced. We further expected this avoidance of race to impede communication, resulting in less efficient performance on the task. We videotaped the sessions and coded participants' behavior to capture the interpersonal costs of efforts to appear color-blind.

Method

Thirty White students (20 females) participated in exchange for \$10. Participants were randomly assigned to work with either a White or a Black confederate.

Individuals arrived at the lab expecting to participate in a study with a partner, who was in fact a confederate.¹ The confederate arrived 2 min after the participant, at which point a White male experimenter explained the two roles for the study: questioner and answerer. An ostensibly random drawing was rigged so that the participant became the questioner and the confederate the answerer. The questioner received an array of 32 photographs from Study 1 that varied orthogonally on three

¹Confederates were female undergraduates instructed to answer only "yes" or "no." In a pretest ($N = 16$), participants' ratings of their partner's performance were similar for Black and White confederates ($M_s = 8.25$ and 8.63), $t = 1.21$, n.s.

TABLE 1
Task Performance by Confederate's Race in Study 2

Performance measure	Black confederate (<i>n</i> = 15)		White confederate (<i>n</i> = 15)	
	Mean	<i>SD</i>	Mean	<i>SD</i>
Percentage of trials in which racial descriptors were used	64%	48%	93%	26%
Percentage of trials in which "Black" or "African American" was used	21%	37%	57%	37%
Number of questions asked	6.69	0.55	6.28	0.42

Note. For each dependent measure, the difference between conditions was statistically significant at $p < .05$ in a *t* test.

dimensions: gender, background color (red or blue), and race (White or Black). The answerer received a six-page photo album with one photo per page (three White and three Black faces) and was told to open it to the first page.

Questioners were told that on each trial their objective was to identify the photo the answerer was looking at by asking as few yes/no questions as possible. Once a photograph was identified, the answerer flipped to the next page, and a new trial started (for a total of six trials).

Results and Discussion

Performance

As expected, participants were more likely to mention race when interacting with a White confederate (doing so, on average, in 93% of trials) than when interacting with a Black confederate (64%), $t(28) = 2.05$, $p_{\text{rep}} = .92$, $d = 0.75$ (see Table 1 for a summary of the results). The confederate's race also affected how participants talked about race. Participants could choose to use either "Black" and "African American" or "White" and "Caucasian" when describing a face. Participants working with White confederates referenced "Black" or "African American" on 57% of trials; participants working with Black confederates, in contrast, used these terms on just 21% of trials, $t(28) = 2.65$, $p_{\text{rep}} = .97$, $d = 0.97$.²

Was this tendency to avoid using race associated with performance on the task? Asking about the target's race was clearly an efficient strategy, halving the number of possible targets remaining in the array. Therefore, we expected the tendency to avoid asking about race to translate into less efficient performance. As expected, participants paired with a Black partner were less efficient than participants paired with a White partner, asking more questions per trial ($M = 6.69$ vs. $M = 6.28$), $t(28) = 2.27$, $p_{\text{rep}} = .94$, $d = 0.83$.

²Focusing solely on questions about race, we also found a significant difference between participants with White partners and participants with Black partners. In interactions with White partners, 62% of such questions included "Black" or "African American," but in interactions with Black confederates, just 33% of such questions included either term, $\chi^2(1, N = 30) = 11.1$, $p_{\text{rep}} = .99$.

Nonverbal Behavior

Two independent coders watched the first 3 min of each videotaped session without audio and with the confederate cropped out. We created a composite measure of friendliness by asking coders to rate participants on the dimensions used by Dovidio et al. (2002): pleasant, cruel, unfriendly, unlikable, and cold (intraclass reliability = .73). When participants were paired with a Black confederate, the tendency to avoid asking whether the target was Black or African American was negatively correlated with perceived friendliness, $r = -.50$, $p_{\text{rep}} = .92$. Coders also rated the extent to which participants attempted to make eye contact with the confederate by looking across the table (intraclass reliability = .69). When the confederate was Black, the tendency to avoid asking whether the target was Black or African American was negatively correlated with attempted eye contact, $r = -.69$, $p_{\text{rep}} = .97$. No similar correlations were observed for participants with White confederates ($r_s = .02$ and $-.03$). These results suggest that when Whites interact with a Black partner, their attempts to appear color-blind—by avoiding race—are accompanied by additional costs: less friendly nonverbal behaviors.

Follow-Up Studies

To identify individual difference predictors and situational moderators of these effects, we conducted two additional paper-and-pencil studies. First, we presented 30 White participants with the photo array from Study 2 and asked them to write down the first five questions they would ask were they to play the game. We created a composite measure ($r = .69$) of two items that assessed endorsement of a color-blind ideology: "When I interact with other people, I try not to even notice the color of their skin" and "If everyone paid less attention to race and skin color, we would all get along much better." As expected, the more participants endorsed color blindness, the greater was their reluctance to use race in the hypothetical game, $r = .51$, $p_{\text{rep}} = .93$.

We then explored whether increasing the diagnosticity of race would cause Whites to forgo color blindness. We showed 130 White participants eight photos from Study 1, described the Political Correctness Game to them, and asked them which face

they would choose to describe first were they to play the game. Participants saw one of two sets of photos: Each included six White male faces, but one included two Black males and the other included two White females. Despite the fact that race and gender were equally—and highly—diagnostic characteristics, 83% of participants chose to describe one of the females when gender was diagnostic but just 61% chose to describe one of the Black faces when race was diagnostic, $\chi^2(1, N = 75) = 3.81$, $p_{\text{rep}} = .92$. Finally, we showed some participants the set of photos with six White and two Black males, but identified the individuals in the photos as criminals and labeled the task “FBI’s Ten Most Wanted.” The specter of racial profiling in this condition made participants even less likely to use race as a cue; just 36% chose to describe one of the Black faces, $\chi^2(1, N = 106) = 6.32$, $p_{\text{rep}} = .95$.

Taken together, these results are consistent with our prediction that concerns about appearing prejudiced during interracial interactions predict efforts to appear color-blind even when race is diagnostic, resulting in inefficient dyadic performance. In addition, avoidance of race has apparent social costs: Reluctance to ask Black confederates about race was associated with decreased eye contact and decreased perceived friendliness. Ironically, those Whites who tried hardest to appear color-blind by avoiding the use of race were the individuals who appeared least friendly when interacting with Black partners.

GENERAL DISCUSSION

Though Whites are adept at identifying other people on the basis of race, their ability to do so outstrips their reports of that ability (Study 1). Such avoidance of race when interacting with a Black partner led Whites to perform poorly on a dyadic task; assessment of nonverbal behavior suggested that the more reluctant Whites were to use race in the presence of Black confederates, the more unfriendly they appeared (Study 2). Thus, strategic efforts to appear color-blind have, at best, mixed results (see Plaut, 2002; Pollock, 2005; Richeson & Nussbaum, 2004; Wolsko, Park, Judd, & Wittenbrink, 2000). The present studies capture the complicated impact of the desire to appear unprejudiced on intergroup relations. Although we have focused on the negative impact for Whites—on communication, performance, and self-presentation—these costs may ultimately extend to both Whites and Blacks, in the form of strained interracial interactions.

Is, then, the motivation to appear unprejudiced maladaptive? Not necessarily. Like most social norms that guide interactions, the desire to avoid prejudice produces mixed outcomes; the wisdom of generosity toward another person, for example, depends on whether the recipient of one’s generosity honors reciprocity norms (Fehr & Fischbacher, 2004). Given that countless studies continue to reveal bias directed against members of stigmatized groups in many real-world settings (e.g., Bertrand & Mullainathan, 2004; Turner et al., 2002), the discomfort arising

from seeking to appear unprejudiced, although aversive, may be precisely the motivator needed to decrease bias. Whites’ efforts to make interracial interactions go smoothly, for example, can lead Black partners to like them more (Shelton, Richeson, Salvatore, & Trawalter, 2005), and the tension of anticipating interracial interaction has been found to improve information processing and group decision making (Sommers, 2006). To the extent that concerns about appearing biased reflect a desire among Whites to be more egalitarian, they represent a step in the right direction toward the amelioration of racial bias. We suggest, however, that the end results of pursuing this goal through mere color blindness may be more complicated than previously thought.

Acknowledgments—The first two authors contributed equally to this work. We thank Benoît Monin for his suggestion to criminalize the game, Colin Camerer and Roberto Weber for their advice on game design, and Robbi Behr, Amanda Bell, Sudanè del Valle, Laura Fruitman, Austin Neudecker, Ben Skolnik, Carrie Sun, Yvonne Wakeford, and Anjuli Willmer for their assistance.

REFERENCES

- Bertrand, M., & Mullainathan, S. (2004). Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *American Economic Review*, *94*, 991–1013.
- Correll, J., Park, B., Judd, C.M., & Wittenbrink, B. (2002). The police officer’s dilemma: Using ethnicity to disambiguate potentially threatening individuals. *Journal of Personality and Social Psychology*, *83*, 1314–1329.
- Cosmides, L., Tooby, J., & Kurzban, R. (2003). Perceptions of race. *Trends in Cognitive Sciences*, *7*, 173–179.
- Crandall, C.S., Eshleman, A., & O’Brien, L. (2002). Social norms and the expression and suppression of prejudice: The struggle for internalization. *Journal of Personality and Social Psychology*, *82*, 359–378.
- Devine, P.G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, *56*, 5–18.
- Dovidio, J.F., Kawakami, K., & Gaertner, S.L. (2002). Implicit and explicit prejudice and interracial interaction. *Journal of Personality and Social Psychology*, *82*, 62–68.
- Fehr, E., & Fischbacher, U. (2004). Social norms and human cooperation. *Trends in Cognitive Sciences*, *8*, 185–190.
- Gaertner, S.L., & Dovidio, J.F. (1986). The aversive form of racism. In J.F. Dovidio & S.L. Gaertner (Eds.), *Prejudice, discrimination, and racism* (pp. 61–89). Orlando, FL: Academic Press.
- Greenwald, A.G., & Banaji, M.R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, *102*, 4–27.
- Ito, T.A., & Urland, G.R. (2003). Race and gender on the brain: Electrocortical measures of attention to the race and gender of multiply categorizable individuals. *Journal of Personality and Social Psychology*, *85*, 616–626.

- Malpass, R.S., & Kravitz, J. (1969). Recognition for faces of own and other race. *Journal of Personality and Social Psychology*, *13*, 330–334.
- Montepare, J.M., & Opeyo, A. (2002). The relative salience of physiognomic cues in differentiating faces: A methodological tool. *Journal of Nonverbal Behavior*, *26*, 43–59.
- Norton, M.I., Vandello, J.A., & Darley, J.M. (2004). Casuistry and social category bias. *Journal of Personality and Social Psychology*, *87*, 817–831.
- Plant, E.A., & Devine, P.G. (1998). Internal and external motivation to respond without prejudice. *Journal of Personality and Social Psychology*, *75*, 811–832.
- Plaut, V.C. (2002). Cultural models of diversity in America: The psychology of difference and inclusion. In R.A. Shweder, M. Minnow, & H.R. Markus (Eds.), *Engaging cultural differences: The multicultural challenge in liberal democracies* (pp. 365–395). New York: Russell Sage Foundation.
- Pollock, M. (2005). *Colormute: Race talk dilemmas in an American school*. Princeton, NJ: Princeton University Press.
- Richeson, J.A., & Nussbaum, R.J. (2004). The impact of multiculturalism versus color-blindness on racial bias. *Journal of Experimental Social Psychology*, *40*, 417–423.
- Rodin, M.J., Price, J.M., Bryson, J.B., & Sanchez, F.J. (1990). Asymmetry in prejudice attribution. *Journal of Experimental Social Psychology*, *26*, 481–504.
- Shelton, J.N., Richeson, J.A., Salvatore, J., & Trawalter, S. (2005). Ironic effects of racial bias during interracial interactions. *Psychological Science*, *16*, 397–402.
- Sommers, S.R. (2006). On racial diversity and group decision-making: Identifying multiple effects of racial composition on jury deliberations. *Journal of Personality and Social Psychology*, *90*, 597–612.
- Sommers, S.R., & Norton, M.I. (2006). Lay theories about White racists: What constitutes racism (and what doesn't). *Group Processes and Intergroup Relations*, *9*, 117–138.
- Sommers, S.R., & Norton, M.I. (in press). Race-based judgments, race-neutral justifications: Experimental examination of peremptory use and the *Batson* challenge procedure. *Law and Human Behavior*.
- Stangor, C., Lynch, L., Duan, C., & Glas, B. (1992). Categorization of individuals on the basis of multiple social features. *Journal of Personality and Social Psychology*, *62*, 207–218.
- Turner, M.A., Freiberg, F., Godfrey, E., Herbig, C., Levy, D.K., & Smith, R.R. (2002). *All other things being equal: A paired testing study of mortgage lending institutions*. Washington, DC: The Urban Institute.
- Weber, R.A., & Camerer, C. (2003). Cultural conflict and merger failure: An experimental approach. *Management Science*, *49*, 400–415.
- Wolsko, C., Park, B., Judd, C.M., & Wittenbrink, B. (2000). Framing interethnic ideology: Effects of multicultural and colorblind perspectives on judgments of groups and individuals. *Journal of Personality and Social Psychology*, *78*, 635–654.

(RECEIVED 12/23/05; REVISION ACCEPTED 2/22/06;
FINAL MATERIALS RECEIVED 2/24/06)