Journal of Experimental Psychology: General
2006, Vol. 135, No. 4, 501–512

0096-3445/06/$12.00 DOI: 10.1037/0096-3445.135.4.501

Distortions in the Perceived Lightness of Faces: The Role of Race Categories

Daniel T. Levin
Vanderbilt University

Mahzarin R. Banaji
Harvard University

Although lightness perception is clearly influenced by contextual factors, it is not known whether knowledge about the reflectance of specific objects also affects their lightness. Recent research by O. H. MacLin and R. Malpass (2003) suggests that subjects label Black faces as darker than White faces, so in the current experiments, an adjustment methodology was used to test the degree to which expectations about the relative skin tone associated with faces of varying races affect the perceived lightness of those faces. White faces were consistently judged to be relatively lighter than Black faces, even for racially ambiguous faces that were disambiguated by labels. Accordingly, relatively abstract expectations about the relative reflectance of objects can affect their perceived lightness.

Keywords: lightness perception, social categories, face perception, adjustment methodology, attitudes

A wide range of research traditions have emphasized the effects of cognition and knowledge on even the most basic of perceptual processes. For example, the familiarity of a contour can determine whether it is assigned to a figure or to the ground (Peterson, 1994), knowledge about the constraints of the human body affects how people perceive the motion of limbs (Shiffrar & Freyd, 1993), and the identification of ambiguous images is strongly affected by expectation (see, e.g., Steinfield, 1967). Other research demonstrates that placing two stimuli in the same category versus different categories can affect similarity ratings (Tajfel & Wilkes, 1963) and accuracy of perceptual discrimination between the stimuli (for a review, see Harnad, 1987). In this article, we extend this tradition by presenting experiments showing that social categories such as race can affect the perception of the lightness of faces. In particular, we show that the relative associations between lightness and White faces and between darkness and Black faces seem to make White and Black faces appear lighter and darker, respectively, than they actually are. This finding demonstrates that perception of a fundamental property such as lightness is affected not only by the immediate perceptual context provided by surface or form as has been shown but also by a top-down influence previously unstudied in the context of high-level vision. Further, because relative light–dark judgments of skin color are shown to be associated with beliefs about deeper psychological qualities (e.g., aggression; Dasgupta, Banaji, & Abelson, 1999), these findings may be viewed as having relevance for a broad range of questions concerning social judgment.

Lightness Perception

Research on lightness perception has a long history, starting with a tradition in early psychophysics that emphasized the degree to which physical parameters of surface luminance were not perceived objectively. For example, the well-known distortion illustrated by Mach bands demonstrates that the perceived difference in reflectance (e.g., lightness) on the left and right sides of each band (see Figure 1) departs from the physically uniform reflectance within each band.

One explanation of such distortions attributes them to very early visual processes. The most widely familiar of these rests on properties of the receptive fields of retinal ganglion cells. Some of these cells are excited by light in the center of their receptive field but are inhibited by light in the region surrounding the center receptive field (and others show the reverse of this pattern). These cells could produce the illusion described above if one assumes that a cell would respond most strongly when its excitatory center is stimulated while its inhibitory center escapes stimulation. This is precisely what would happen at the edge of a Mach band if the excitatory center is stimulated by a relatively light region while the inhibitory surround partially escapes stimulation because it is positioned in a part over a darker region (Cornsweet, 1970). Similar positioning of center-surround receptive fields can be used to understand the Hermann grid illusion.

The key to these explanations is that lightness distortions are fundamental to the earliest stages of visual processing. Thus, they reflect context effects that are localized across single early-visual receptive fields and that, presumably, are immune to top-down effects. However, another research tradition emphasizes a more inferential approach to lightness that includes a wide range of contextual factors, including shadows and transmittal cues (for a review, see Purves, Williams, Nundy, & Lotto, 2004). The key is

This research and writing were supported by grants from the National Institute of Mental Health, the Third Millennium Foundation, and a Rockefeller Foundation fellowship at the Bellagio Study and Conference Center to Mahzarin R. Banaji. We are grateful to Andrew Baron, Randolph Blake, Jeff Ebert, Michael Kubovy, Jamaal McDell, Ken Nakayama, Kristina Olson, Andrew Rossi, Kristin Lane, Roy Ruhling, and Michael Tarr for helpful comments on previous versions of this article.

Correspondence concerning this article should be addressed to Daniel T. Levin, Department of Psychology and Human Development, Vanderbilt University, Peabody College #512, 230 Appleton Place, Nashville, TN 37203-5701, or to Mahzarin R. Banaji, Department of Psychology, William James Hall, Harvard University, 33 Kirkland Street, Cambridge, MA 02138. E-mail: daniel.t.levin@vanderbilt.edu or mahzarin_banaji@harvard.edu

Figure 1. Mach bands in which a series of bands appear to be lighter on the right and darker on the left because of contrast with the darker bands on the right and lighter bands on the left.

that these cues operate over a much broader range than individual early vision receptive fields (and even early vision accounts presume input from outside of a given cell's receptive field; Rossi & Paradiso, 2003). For example, Wolfe (1984) observed that the Hermann grid illusion is stronger when the number of grid elements is increased. Because early receptive fields are much smaller than the entire grid, this finding suggests that the illusion is influenced by a broader context that is generated later in the chain of visual processing. In a similar vein, research in lightness perception has emphasized the impact of shadows, image articulation, and 3-D form and grouping cues on lightness perception (for a review, see Adelson, 1993, 2000; Gilchrist & Annan, 2002). These inputs to lightness perception probably reflect sophisticated visual processing that is well beyond the capability of the retina and the earliest cortical visual maps.

All such research assumes that higher level processing influences lightness perception, but such influences are typically assumed to reflect global processes for encoding form and shadow. Therefore, the assumption remains that lightness perception is influenced by an immediate visual context that may be quite broad and sophisticated but that is nonetheless nonconceptual—it emanates from the perceptual features of the current percept itself (see, e.g., Williams, McCoy, & Purves, 1998). The idea that lightness perception can be influenced by factors outside the realm of ongoing visual processing has, as far as we know, been untested. In other words, it is not known whether and in what ways knowledge or assumptions about the category to which the perceived object belongs influences perception of its lightness.

However, one report hints at the possibility that the category a stimulus belongs to can influence brightness perception. Goldstone (1994) trained subjects to categorize continua of stimuli that varied in luminance and size. He then tested their ability to discriminate between stimulus pairs along the continua that either fell within one of the categories or straddled the category boundaries such that each member of the pair had been associated with different categories during the training. He reported that subjects who were trained to subdivide the grayscale continua were more accurate at discriminating the grayscale-continuum category-straddling pairs than were subjects who had not received this training. This result implies that the training influenced subjects' perception of the brightness of the stimuli. For example, it is possible that subjects perceived the stimuli just inside the dark end of the continuum as being darker than they were and those at the light end of the continuum as being lighter then they were, a distortion that would be similar to that implicated in Mach bands. However, the finding is ambiguous because the improved discrimination might have been the result of the subjects' ability to represent category-straddling values more precisely as opposed to shifting the representations in a specific direction. Therefore, it remains an open question whether category membership itself can influence brightness.

The Impact of Social Categories on Face Perception

In previous studies, researchers have examined the degree to which social categorization influences face perception. For example, Levin (1996) showed that White subjects searched for a target with a Black face more quickly than they searched for a White target. On the basis of research exploring the impact of feature-present–feature-absent relationships on visual search (e.g., Treisman & Gormican, 1988), this result suggests that Black faces are detected on the basis of the presence of a race-specifying feature, whereas White faces are detected on the basis of the absence of that feature. Levin (1996, 2000) has argued that the observed feature definition is caused by a social–cognitive context in which out-group individuals are processed at a group level at the expense of individuating among them. Thus the tendency to make social categorical inductions becomes embodied in a visual feature marking that group.

The present collaboration begins with a shape and texture-map continuum between a prototypical White face and a prototypical Black face (see Figure 2) used by Levin (2000). In viewing these faces, Mahzarin R. Banaji noticed an interesting illusion. First, even though Levin had stated that White and Black faces were controlled for reflectance, the White face appeared to Banaji to be lighter than the Black face. Moreover, attending to the set more deeply in an attempt to remove the illusion only seemed to increase it. This suggested that the more social category (race) was attended to, the harder it was to view the reflectance of the faces objec-

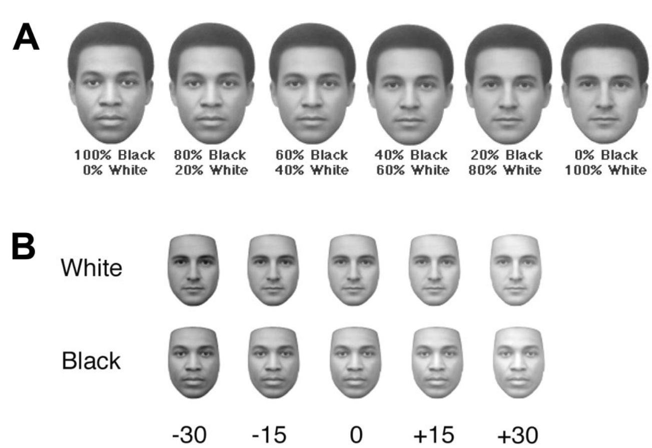

Figure 2. A: An equal-reflectance continuum between a Black average face and a White average face. B: Samples of Black and White average faces at different lightness levels.

tively. This observed distortion, it turned out, was not unique: On seeing these faces, many observers have asked why the reflectance of the faces was not controlled, only to be told that all the faces were, indeed, equally bright. These observations led to the present collaboration on a series of studies designed to systematically measure a possible bias in lightness perception and thereby to understand aspects of both color perception and social cognition.

Only one set of studies to date has explored the degree to which these social classifications can affect perceived lightness. MacLin and Malpass (2001, 2003) showed subjects a series of faces that had identical racially ambiguous features except for their hair, which was diagnostic of either Black or Hispanic faces. Subjects each saw a series of these faces, classified them by race, and rated the faces on a number of dimensions, including the lightness of each face's complexion. The race-diagnostic hair appeared to be effective in leading subjects to categorize the faces by race and, moreover, it caused subjects to rate the Black-hair faces as darker than the Hispanic-hair faces.

This finding provides the interesting possibility of categorization-induced distortion in lightness, although it does not do so directly. The lightness judgments were ratings on a response scale with text anchor points (*light* and *dark*). Therefore, subject responses were more akin to estimates of how they would match an actual lightness sample to the faces rather than the direct matches typical of the lightness perception literature. We addressed this concern in these experiments by eliminating the rating scale and using a dial-adjustment method similar to that used in more traditional lightness perception research.

In the present experiments, we asked subjects to adjust one stimulus (either a face or a patch of gray) to match a face they classified racially as Black or White. In Experiment 1, this entailed matching one face with another. Specifically, subjects adjusted either a Black or a White face to match either a Black or a White face. In Experiment 2, subjects adjusted the luminance of a gray patch to match a face and, in addition, matched unambiguous and ambiguous faces to control for potential stimulus effects. In Experiment 3, line-drawing faces filled with a single gray level were used in a further attempt to control for stimulus effects.

In all three experiments, a measure of explicit attitude toward social groups was included to test whether there was any association between favorability toward the groups and basic color perception. One possibility is that those who are more negatively predisposed toward African Americans and willing to express it explicitly will show a stronger dark bias than will those who are not. Alternatively, if a bias in face color perception is caused by knowledge of race differences in color perception and equally present in all individuals, no such effect should be observed. Because this was not a primary concern and to avoid underpowered analyses, we present all of these distortion–attitude correlations collapsed across all three experiments in the *Results* section of Experiment 3.

Experiment 1: Testing the Lightness Distortion Effect

In Experiment 1, we chose to use the method of adjustment to obtain brightness judgments for the faces. This method is common in the brightness literature, and one of its strengths is that it requires far fewer trials than do matching procedures. We felt this was important here, both for the sake of efficiency and to avoid the possibility that the effect would be eliminated over a large number of trials as subjects learned unusual attentional strategies. Subjects viewed two faces on each trial. One was defined as a reference face and the other was defined as an adjustable face. Subjects were told that their task was to adjust the adjustable face until it matched the lightness of the reference face. On half of the trials, the faces were of the same race (both were Black or both were White), and, on the remaining half, they were of different races. On the basis of our previous informal observations, we tested whether subjects would choose relatively dark versions of the adjustable face for the Black reference face and relatively light versions for the White reference face.

Method

Subjects. A total of 67 undergraduates from Kent State University completed Experiment 1 in exchange for credit in their General Psychology course. Of these, 48 indicated that they were women, 58 that they were White, 7 that they were African American, and 2 that they were Asian. Subjects' mean reported age was 19.6 years (range 18–34 years).

Apparatus. Stimuli were presented using MacOS computers attached to 15-in. monitors (Sony fx-100 using Mac standard gamma) set at a resolution of 640 × 480 pixels. Subjects responded using the computer's keyboard, and presentation was controlled by a program written in True BASIC.

Stimuli. Subjects rated variants of two prototype faces, one White and one Black. We chose to use these computer-generated faces because they are nondistinctive, have no marks or other image defects, and can be easily adjusted for luminance while avoiding clipping (i.e., the tendency for adjustments in brightness to cause the highest or lowest brightness values to go out of range, thereby compressing the range of high values to a single value). Each race prototype face was created by blending a set of 16 faces from the race the prototype was to represent. All images were based on full-frontal photographs of male faces with neutral expressions that were digitized into 256-level grayscale images (see Levin, 1996, for more details). The hair was then removed from the prototypes, and these base images were then matched for both mean luminance and contrast, as measured by the mean and standard deviation, respectively, of the gray-level histogram of the faces as wholes. As part of the contrast-matching process, the eye whites of the Black prototype were reduced slightly in luminance. It is important to note that matching for overall reflectance required relatively little manipulation because the original set of 16 faces that constituted each prototype were themselves matched for mean luminance between the races.

Once the initial luminance-matched faces were created, 12 new variants of each prototype were created, 6 of which were progressively lighter than the original and 6 of which were progressively darker. Each step of variation in lightness corresponded to 5 out of 256 possible levels of grayscale. All lightness levels in these experiments are reported using these 8-bit gray-level units. The base prototypes (which are referred to here as being at a *level of 0 lightness*) had a mean gray level of 141 out of 256. This was equivalent to 47.5 cd/m^2, as measured by a Minolta LS-110 luminance meter on one of the four monitors used to present the stimuli. The darkest faces (anchored at –30 lightness) had a mean gray level of 112 (29.4 cd/m^2), and the brightest faces (anchored at +30 lightness) had a mean gray level of 170 (68.3 cd/m^2). The contrast range of the original faces was kept within a relatively narrow range to avoid clipping in the brightest and darkest faces, and so contrast was almost identical throughout the range of mean luminances (the standard deviations of the gray levels ranged from 48.71 to 49.25). All faces were presented at a size of 58 (horizontal) × 73 (vertical) pixels at 72 dots per inch.

Procedure. Subjects completed the experiment in small groups ranging in size from 1 to 5 on separate computers in the same room. Before completing the task, subjects entered their age, sex, and race into their computer. They then read task instructions along with the experimenter.

There was no deception in these studies, with subjects being told that the experiment was about “how people perceive the shading of faces of different races.” They were made aware that they would experience a series of trials in which they would see a reference face next to an adjustable face and that their task was to manipulate the adjustable face so that its shading matched that of the reference face as closely as possible. For each trial, subjects increased the luminance of the adjustable face using the up-arrow key and decreased it using the down-arrow key. Subjects were free to adjust the luminance of the face up and down as much as they liked, and they pressed the space bar to indicate when they had made their final judgment.

Each subject experienced trials with each possible combination of Black and White adjustable and reference faces. So subjects adjusted Black faces to match White faces and the reverse; these are referred to as *mixed-race* trials. They also adjusted one copy of the Black face to match the other and did the same for the White face; these are referred to as *same-race* trials. Therefore, each subject completed trials in four conditions: (a) Black face reference, White face adjustable; (b) White face reference, Black face adjustable; (c) Black face reference and adjustable; and (d) White face reference and adjustable.

The lightness of the reference face and initial lightness of the adjustable face was systematically rotated over trials. The reference face was set at one of five different lightness levels (−10, −5, 0, +5, +10). The initial lightness of the adjustable face was offset from the reference face by 2 or 4 levels above or below the lightness of the reference face. For example, if the reference face was set at –5 lightness, the initial lightness of the adjustable face was set at a lightness of –25, −15, +5, or +15. Therefore, within each of the four conditions, there were 20 trials (for a total of 80 trials): 5 trials at each level of reference image lightness and, within these, 4 offsets for the adjustable image starting point.

Finally, the relative positions of the adjustable and reference images were counterbalanced between subjects. Half of the subjects saw the adjustable image to the right of the reference image, and half saw it to the left of the reference face.

After completing the lightness judgment trials, subjects completed a brief measure of attitudes toward each race using a variant of the “feeling thermometer.” For this measure, subjects rate their attitudes toward 12 different groups that might be represented on a college campus. Among these are Black or African American students and White students. Consistent with previous versions of the task, subjects rated their attitudes by using a thermometer, entering a number between 0 and 100, with 0 corresponding to the anchor point *very coolly* and 100 corresponding to *very warmly*. The measure was presented on the computer, and subjects typed in their ratings. The 12 social groups were presented in a different random order for each subject.

Results and Discussion

Subjects consistently chose less luminant samples for the Black reference faces than for the White faces. In the most basic analysis, we tested whether Black reference faces were more likely to be matched with a darker face than were White reference faces. Therefore, for each subject, the mean luminance of the face chosen for Black reference faces was subtracted from the mean luminance of the faces chosen to match White reference faces. Note that this entailed averaging over trials for which the reference and adjustable faces were of the same and of different races. On average, subjects chose a Black reference that was 2.9 gray levels darker than the corresponding White face, $t(66) = 6.19$, $SE = 0.095$, $p < .001$, $d = 0.75$, which corresponds to 1.94 cd/m^2 (see Figure 3). It is important to note that although this effect is small in brightness units, it was consistent across subjects: 52 out of 67 subjects chose darker Black faces, $\chi^2(1, N = 67) = 20.43$, $p < .001$.

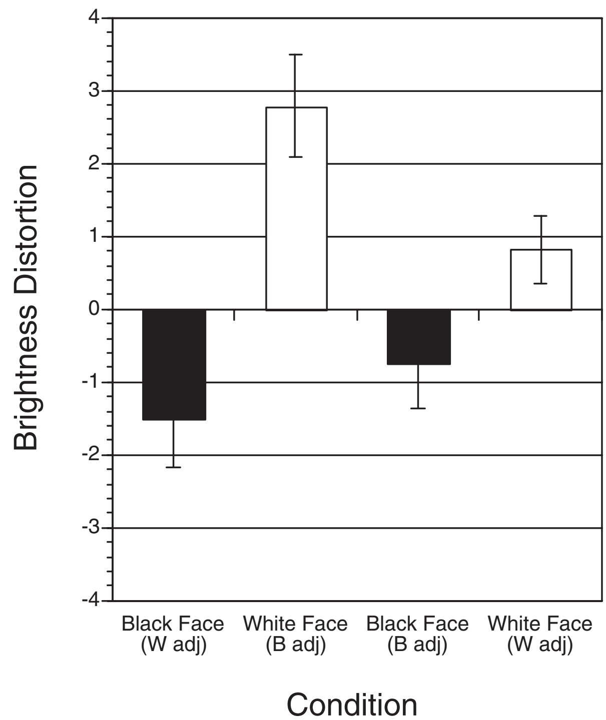

Figure 3. Results of Experiment 1. The two bars on the left represent trials where subjects adjusted a face of one race to match a face of the other race. The two bars on the right represent trials in which the adjusted face and the target face were the same. Positive numbers represent overlightening of the face. Adj = adjustable.

Data from the 7 Black subjects in this experiment were also analyzed separately. Of these, 6 chose relatively darker samples for Black faces, and, on average, these 7 subjects chose Black samples that were 3.45 gray levels darker than the White samples, $t(6) = 2.63$, $SE = 0.263$, $p = .039$, $d = 0.99$.

The relative darkness of Black reference faces was greater for the mixed-race trials than for the same-race trials, but the distortion was significant for both. For mixed-race trials, the Black faces were judged to be 4.35 levels darker, $t(66) = 5.02$, $SE = 0.172$, $p < .001$, $d = 0.61$, and for same-race trials they were judged to be 1.55 levels darker, $t(66) = 4.82$, $SE = 0.065$, $p < .001$, $d = 0.75$. A paired t test comparing mixed and same-race trials was significant, $t(66) = 3.12$, $SE = 0.178$, $p = .003$, $d = 0.58$. In addition, there was no hint of a correlation between attitudes toward the races and the degree of lightness distortion (an analysis combining data across studies is presented later).

In Experiment 1, subjects showed a consistent tendency to choose a relatively darker adjustable face when matching a Black reference face than when matching a White reference face. This effect was significant for both same-race and mixed-race trials, although the effect was significantly larger when the two faces were of different races.

Although these results reveal a distortion that is consistent across conditions, this consistency is, in a sense, surprising. Why should subjects show a distortion effect in the same-race conditions, when the adjustable face and the reference face are identical? The most likely explanation is that the lightness of the adjustable face is perceived more objectively because it is not static, and subjects see it change as they adjust it. Therefore, the distortion is relatively stronger in the reference face because its shading is perceived as being more integral to the identity of that object. For example, when subjects adjust the Black face to match a reference copy of the same face, they might perceive the reference face as 5 levels darker than it is, whereas they perceive the adjustable face more objectively, perhaps seeing it as only 2.5 levels darker than it actually is. Thus, to match the two, they will need to darken the more objectively perceived adjustable face to match the darker appearing reference face. This is particularly interesting because it suggests that subjects perceive properties that they can easily manipulate as less inherently bound to the stimulus they are associated with. However, we reserve further exploration of this question for future studies.

One benefit of the presence of the same-race distortion effect is that it helps reduce the plausibility of a particularly troublesome alternative hypothesis about the cause of the effect we have observed here. Perhaps subjects adjust Black faces to be relatively darker not because the faces are globally perceived as darker but rather because they tend to focus their attention on relatively light parts of Black faces, whereas they focus their attention on relatively dark parts of White faces. For example, assuming that noses tend to be light (because they protrude from the face) and eye sockets tend to be dark (because they are concave), it would be problematic if subjects tend to focus on the relatively light noses of Black faces whereas they focus on the relatively dark eyes of White faces. Accordingly, subjects might adjust a Black face to be too dark because they are trying to darken a nose that is inherently lighter than the eyes they are looking at in the White face. However, this is not a problem with a pair of identical faces if it is assumed that subjects focus on the same parts of the faces when comparing two faces from the same group, and hence we do not have to contend with this possibility.

Although a selective attention explanation for the darkness effect seems unlikely, it is still possible that some sort of stimulus artifact causes the effect. Such a hypothesis turns on the possibility that some incidental property of the specific faces we used might make them seem relatively dark or light. For example, if the shape of the eyes in the White face happens to make them appear dark, then subjects will lighten the face artificially. We therefore designed Experiment 2 to avoid this problem by creating a face that was racially ambiguous and having subjects judge it in the context of other unambiguous faces. Accordingly, for half of the subjects, the ambiguous face was labeled *Black* and presented in the context of unambiguous White faces; for the other half, the same face was labeled *White* and presented in the context of unambiguous Black faces. If subjects match the ambiguous face with a dark standard when they believe it is Black and match it with a lighter standard when they believe it is White, then the effect is probably not bound to the stimulus.

In addition to using ambiguous faces in Experiment 2, we made several other changes to the experiment. Most important, instead of having subjects adjust one face to match another, we asked subjects to adjust the lightness of a square gray region to match a face. This was necessary because in the two-face procedure, it is not completely clear which of the two faces (reference or adjustable) is being distorted. Although we argued above that the reference face is more likely to be distorted, we cannot with confidence rule out the possibility there is some distortion in the adjustable face. This is problematic if we intend to isolate the distortion effect to an ambiguous reference face.

Experiment 2: Removing Stimulus Artifacts

Method

Subjects. A total of 27 Kent State University undergraduates completed Experiment 2 in exchange for credit in their General Psychology course. Of these, 24 indicated that they were women, 24 indicated that they were White or Caucasian, and 3 indicated that they were Black or African American. Subjects' mean age was 18.8 years (range 18–30 years). Thirteen subjects completed the *B/BW* condition (in which an ambiguous face was paired with an unambiguous Black face), and 14 completed the *BW/W* condition (in which an ambiguous face was paired with an unambiguous White face).

Apparatus. The apparatus was identical to that used in Experiment 1.

Stimuli. In addition to the Black and White average faces used in Experiment 1, an ambiguous face was created (Figure 4). This was done by first creating a continuum of 21 faces between the Black average and the White average in 5% increments. Then, a group of 15 pilot subjects classified each of the faces by race. In this classification experiment, subjects viewed each of the 21 faces a total of 4 times in random order and were instructed to hit one key if they thought the face was White and another if they thought it was Black. On the basis of these classifications, the most ambiguous face was the intermediate face that represented a 50-50 blend of Black and White. These distortions represented very close matches with the respective Black and White unambiguous face distortions.

Procedure. Procedures were similar to those used in Experiment 1, with several exceptions. First, for half of the subjects, the ambiguous face replaced the Black face, and, for the other half, the same face replaced the White face. Before beginning the task, all subjects saw an instruction screen that included the ambiguous face next to one of the unambiguous faces underneath the labels *Black* and *White.* Subjects in the BW/W condition saw the ambiguous (BW) face paired with the White face, and therefore the BW face was labeled *Black.* In contrast, subjects in the B/BW condition saw the same ambiguous face labeled *White* and paired with the unambiguous Black face. Otherwise the instruction screens were the same as in Experiment 1.

On each trial, instead of presenting two faces, we presented only one face adjacent to an adjustable gray region. The region was rectangular and filled with a uniform gray shade. It measured 80 (horizontal) × 100 (vertical) pixels. The starting gray levels of this region were matched to the mean starting gray levels of the original adjustable faces from Experiment 1, and they were adjusted during each trial using the same procedure as in Experiment 1.

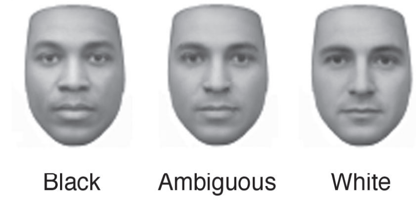

Figure 4. The ambiguous face.

Results and Discussion

As in Experiment 1, subjects chose darker samples for Black faces than for White faces and, in this case, chose a darker standard for the ambiguous face when it was labeled *Black*. The mean lightness chosen for the ambiguous face was 0.465 levels darker than the sample when it was labeled *Black* and 15.95 levels lighter when it was labeled *White*, $t(25) = 4.14$, $SE = 0.794$, $p < .001$, $d = 1.65$ (Figure 5), a difference that corresponds to approximately 11.0 cd/m^2. In addition, the within-subjects lightness distortion effects were significant both overall (Black relative lightness = −0.195, White relative lightness = 16.95), $t(26) = 6.52$, $SE = 0.526$, $p < .001$, $d = 1.26$, and for each of the two conditions. In the B/BW condition, the chosen-sample lightness was 0.10 levels lighter for the Black face, compared with 15.95 levels lighter for the BW face, $t(12) = 5.16$, $SE = 0.615$, $p < .001$, $d = 1.47$. In the BW/W condition, the chosen-sample lightness was 0.465 levels darker for the BW face, compared with 17.85 levels lighter for the White face, $t(13) = 4.28$, $SE = 0.856$, $p = .001$, $d = 1.15$.

Again, the effect was consistent across subjects, with 27 out of 27 subjects choosing a darker sample for the Black face.

Because only 3 Black subjects participated in the present experiment, an analysis that combines the results from Black subjects in Experiments 2 and 3 is presented in the *Results* section of Experiment 3.

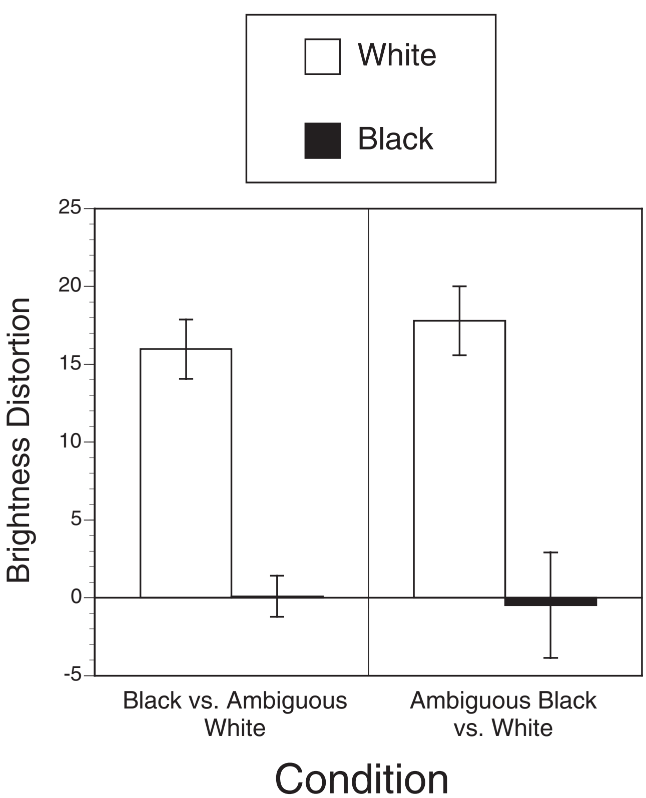

Figure 5. Results of Experiment 2. The left section represents a group of subjects who judged the lightness of an unambiguous Black face on some trials and an ambiguous face labeled *White* on other trials. The right section represents a different group of subjects who judged the same ambiguous face labeled *Black* on some trials and the unambiguous White face on other trials. Positive numbers represent overlightening of the face. Error bars reflect standard errors.

Once again, a darker sample was chosen for the Black face than for the White face, even when the two were the same and only labeled differently. This reduces the probability that stimulus artifacts such as differing between-region contrasts cause the distortion effect. In addition, it is interesting to note that the absolute magnitude of the effect was larger here than it was in Experiment 1. These between-experiment differences in magnitude will be covered in the General Discussion section so that Experiment 3 can be included for comparison.

Experiment 3: Eliminating the Attentional Focus Alternative Hypothesis

Although we attempted to control for the impact of spurious shape and texture features on lightness judgments in Experiment 2, the control is not quite complete. Notwithstanding the demonstration that the classification of an ambiguous face affected which final shading sample subjects selected for each face, it still may be the case that subjects do not misperceive the lightness of the faces. Instead, it is possible that subjects adjusted the faces differently because they focus their attention on different parts of a face when they believe it represents one race instead of another. For example, if subjects believe a face to be Black, they may focus on the eyes (a darker region), and, when they believe it is White, they may focus on the nose (a lighter region). If so, they might adjust the sample patch to be relatively dark for the Black face because they are matching it with the face's eyes. Conversely, White faces would be overbrightened because subjects are focused on the relatively light nose. Thus, attentional focus could cause our effect instead of lightness perception.

Experiment 3 was designed to eliminate this alternative by using faces with consistent luminance throughout. Therefore, we used line-drawing stimuli that were filled with a given level of gray so, no matter where subjects focus their attention, a consistent fill tone is present. However, even this is not sufficient, because areas with a relatively large number of lines would, assuming the lines are darker than the fill, be darker overall, assuming some spatial integration of lightness (or brighter overall, assuming lightness contrasts with the dark lines). For example, if we argue that the eyes are darker, the line-drawing stimulus still has darker eyes because of the relative density of dark lines that represent the eye's detail lines. To eliminate this concern, we used two kinds of drawings: one with lines brighter than the fill and one with lines darker than the fill. Across stimuli, then, attentional focus on a given region should lead to canceling effects of local and integrated lightness (and lightness contrast).

Method

Subjects. A total of 45 Vanderbilt University undergraduates completed Experiment 2 in exchange for credit in their General Psychology course or for a $5.00 payment. Of these, 18 were women. Thirty-five subjects indicated that they were White or Caucasian, 4 indicated that they were Black or African American, 5 indicated that they were Asian, and 1 indicated that he or she was Indian. Subjects' mean age was 20.8 years (range 18–39 years). Twenty-three subjects completed the B/BW condition, and 22 completed the BW/W condition.

Stimuli. A new set of line-drawing faces was created (Figure 6). The drawings were derived from the averaged keypoint maps of the same 16 faces of each race that went into creating the average faces from Experiments 1 and 2. The keypoint maps were based on a set of 232 individual data points connected into 39 lines to represent the major features of faces (see Levin, 1996; Rhodes, Brennan, & Carey, 1987). The line drawings were rendered at a resolution of 151(horizontal) × 202 (vertical) pixels, then one was created with relatively dark lines and another with relatively light lines. Both were then filled with the same sets of shades such that the difference in the percentage of gray value between the fill (the base level of lightness was 136 on a 0–255 dark–light scale) and the lines was the same for the light lines (having a base level of lightness of 187 out of 255) and dark lines (having a base level of lightness of 68 out of 255). The onscreen luminance values of the light lines, dark lines, and 0-level fill were 91.9 cd/m^2, 23.9 cd/m^2, and 59.2 cm/m^2, respectively. The +25-level light faces had a fill luminance of 72.2 cd/m^2, and the –25-level dark faces had a fill luminance of 43.2 cd/m^2.

Apparatus. Faces were presented on eMac computers set at a resolution of 1024 × 768 pixels (89-Hz refresh) and 256-level grayscale mode using eMac gamma.

Procedure. Procedures were similar to those used in Experiment 2. All subjects judged faces with light and dark lines presented in a single block of trials. For each of the four faces (Black with dark lines, White with dark lines, Black with light lines, White with light lines), subjects completed ratings with five different reference face start points. Within each of these five reference face start points, there were two trials of each of four sample start points representing +15-, +5-, −5-, and –15-level reference-sample lightness differences. Therefore, subjects completed a total of 160 trials (4 faces × 5 reference face start points × 4 sample start points × 2 trials).

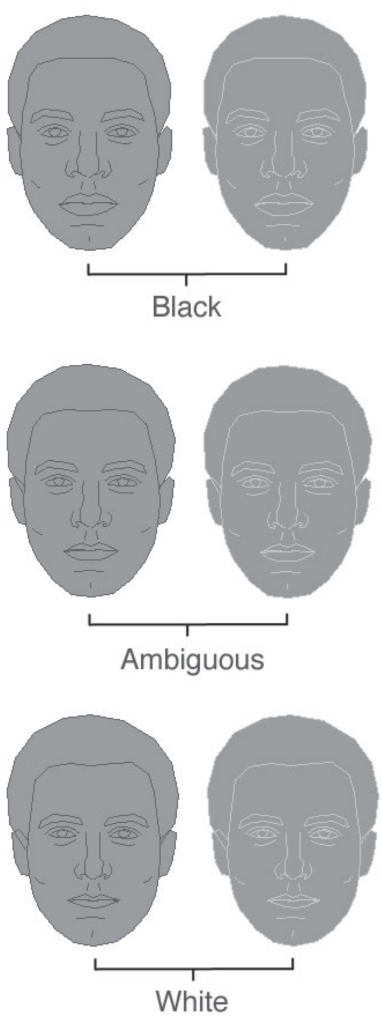

Figure 6. Line drawing faces.

Results and Discussion

Once again, subjects chose darker samples for the Black faces overall. Data were entered into a three-factor Condition (B/BW vs. BW/W) × Line Level (light lines vs. dark lines) × Race (Black vs. White) mixed-factors analysis of variance (ANOVA) with condition as the between-subjects factor. The race main effect was significant: The mean lightness chosen for Black faces was 4.45 levels lighter than the reference face, compared with 8.00 levels lighter for the White faces, $F(1, 43) = 4.67$, $MSE = 120.87$, $p = .036$, $d = 0.33$ (which corresponds to a difference of 2.06 cd/m^2). The line level main effect was nonsignificant, $F < 1$, but there was a significant interaction between line level and race, $F(1, 43) = 6.93$, $MSE = 1.46$, $p = .012$. The race effect was stronger in the dark-line faces (Black distortion was 3.90, White distortion was 7.95), $t(44) = 2.46$, $p = .018$, $d = 0.37$, than in the light-line faces (Black distortion was 5.05, White distortion was 8.10), $t(44) = 1.89$, $p = .066$, $d = 0.28$. The condition main effect and interactions were all nonsignificant. The distortion effect was similar in magnitude for the between-subjects comparison of ambiguous faces (Black face distortion was 4.70, White distortion was 8.00) but was nonsignificant, $t(43) = 1.24$, $p = .221$ (see Figure 7).

Twenty-seven of the 45 subjects chose darker samples overall for the Black faces than for the White faces, $\chi^2(1, N = 45) = 1.80$, *ns*.

Results from the 7 Black subjects in Experiments 2 and 3 were combined and tested for a distortion. Of these 7 subjects, 6 chose relatively darker samples for the Black face, and their mean distortion level was 19.80, which was significantly different from 0, $t(6) = 2.86$, $SE = 1.388$, $p = .029$, $d = 1.08$.

In this experiment, subjects chose relatively darker samples for the Black face and relatively lighter samples for the White face overall, and the effects were similar in magnitude even when the faces were ambiguous drawings and the normal dark lines with light fill relationship was reversed. This suggests that the race-based lightness distortions are not caused by attentional focus. That is, even if subjects tend to focus attention on different parts of Black versus White faces, local differences in brightness and contrast cancel out across the different versions of these stimuli. Therefore, the remaining distortion effect is more likely due to the perceived race of the faces and not due to the tendency to look at relatively light parts of Black faces and relatively dark parts of White faces. In addition, we again observed no correlation between attitudes toward the races and lightness distortions. However, several aspects of the present experiment warrant discussion.

Most important, the effects were smaller ($d = 0.33$) than those observed in Experiments 1 and 2 (*d*s ranging from 0.75 to 1.26 for the primary Black–White comparisons). A look at the stimuli suggests that the race-specifying information they contained was subtle and, for the ambiguous faces, almost nonexistent. Indeed, a number of the subjects noted that at least one of the faces they saw “didn’t really look Black or White.” Accordingly, we may have

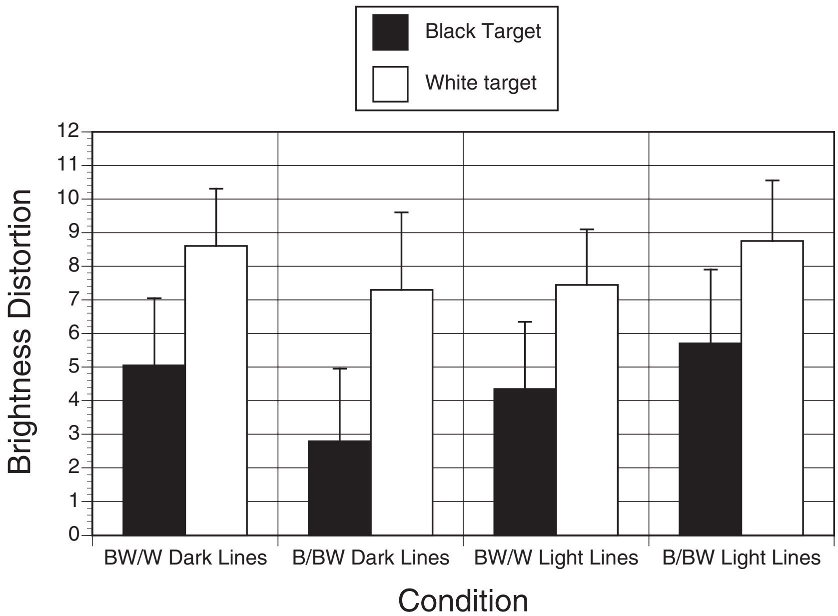

Figure 7. Results from Experiment 3. The left two panels represent results from the dark-line faces, and the right two from the light-line faces. Error bars reflect standard errors. BW/W = ambiguous face paired with an unambiguous White face; B/BW = ambiguous face paired with an unambiguous Black face.

traded a fair amount of validity for control. We would, therefore, not recommend using these faces in future demonstrations and note that they have served their purpose of demonstrating a lightness distortion in stimuli with consistent skin tone and inverted line–fill lightness relationships. In retrospect, we might have avoided the ambiguous faces because the contrast inversion and one-level fill inherent to the unambiguous faces eliminated just about all of the possible stimulus confounds we have discussed.

There was no correlation between attitudes, as measured by the feeling thermometer, and the degree of lightness distortion. Attitudes were operationalized as the rating for Black students subtracted from the rating for White students. Therefore, positive numbers indicate relative favoritism toward White students. To allow a maximally powerful test of the relationship, we normalized and combined data from all three experiments. Across a total of 139 subjects, the correlation between the brightness distortion and attitudes was .073 ($p = .38$). In addition, none of the individual experiments produced a significant correlation (for Experiment 1, $r = .054$; for Experiment 2, $r = .093$; for Experiment 3, $r = .093$).

Experiment 4

In Experiments 1–3, we tested distortions in brightness judgments using an adjustment method in which subjects manipulated a face or color patch sample to match a reference face. Although we consistently observed that subjects selected relatively darker samples to match Black faces, all of these experiments depend on a conscious, deliberate report that could be contaminated by demand characteristics or by related postperceptual decision processes. Finally, it is important to note that in these experiments, assumptions are made about the subjects' phenomenology. At some level, we assume that the results of our adjustment methodology reflect the perceptual experience of subjects that one face looks darker than it actually is. However, this is not necessarily the case: Subjects may experience the same level of gray on both faces but adjust the sample to be darker on the basis of a conceptual reinterpretation of perceptual experience that may or may not be conscious. So, in the most deliberate case, subjects may see the faces as being the same but decide that they had better lighten the White face because they "know" it should be brighter or because they are conforming to their interpretation of the experimenter's wishes. The same postperceptual editing may even take place outside of awareness if subjects do not fully realize that they are adjusting the samples to be consistent with their understanding of faces rather than what they currently see.

Although it may be difficult to resolve this issue fully (Dennett, 1991), we point out that a similar objection can be made about much of the lightness literature. Indeed, recent research has suggested that brightness perception includes a postsensory processing based on a very general anchoring effect (Logvinenko, 2002). However, to reinforce these findings, we develop evidence against the hypothesis that our adjustment method caused a deliberative modification of subjects' final choice. For example, it is possible that subjects initially adjusted the face to match an objective lightness but, before quitting, made a final adjustment to fit the demands of the experiment. Another alternative is simply that the time necessary to make multiple adjustments caused deliberative postperceptual processes to affect subjects' decisions. The current data can eliminate both of these alternatives because we recorded the number of adjustments subjects made before settling on a final decision. Most important, for a small proportion of trials, subjects made no adjustments, judging that the starting face exactly matched the target. Therefore, if the distortion effect is present in these trials, then it is reasonable to assume that the effect does not reflect the kind of long-term deliberations inherent to making the adjustments. To test this, we reanalyzed only the zero-adjustment trials from Experiment 2 (this experiment was chosen because it showed the strongest effect). Subjects had an average of 7.1 (out of 80) zero-adjustment trials, and, of the 27 subjects in this experiment, 22 had at least 1 zero-adjustment trial for each race. Across these subjects, lighter samples were chosen for the White faces (12.70 levels lighter than the target face) than for the Black faces (0.395 units lighter), $t(21) = 3.845$, $p = .001$, $d = 0.82$, on the zero-adjustment trials.

Although this post hoc analysis helps, it cannot completely rule out the possibility that when subjects focus attention on making a judgment, they recruit a range of processes that do not generally affect perception. Therefore, in Experiment 4, we tested whether the brightness distortion would affect performance in a same–different task that does not require an explicit judgment. In this experiment, subjects viewed pairs of faces that were either of the same race (in which case they were a pair of identical White or Black prototypes from Experiment 1) or were different races (the White prototype and the Black prototype) and were asked to respond as quickly as possible about the match in race while also being explicitly told that any mismatches in brightness were not to be used as the basis for a *different* judgment. Across trials, we varied the relative luminances of the faces and predicted that when the luminances are effectively different, subjects would perform the task more quickly, and when the luminances are effectively the same, their performance would be slowed. In particular, subjects should be slower to classify face pairs as different when the Black face is actually more luminant than the White face (leading to a match in perceived brightness), as compared with pairs for which the two are the same luminance (and therefore are different in brightness). Thus, the effect would run counter to the physical similarity of the form and luminance of the stimuli.

Method

Subjects. A total of 15 Vanderbilt University students (all women) completed the experiment in exchange for course credit in their General Psychology class or for a $5 payment. Of these, 13 were White and 2 Black, and their mean age was 19.1 years (range: 18–21 years).

Apparatus. This was the same as that used in Experiment 3.

Stimuli and procedure. The unambiguous Black and White faces used in Experiment 1 were used here. The faces were placed onscreen in pairs: Same-race pairs consisted of two Black faces or two White faces, and different-race pairs consisted of one face of each race. Among the pairings, the faces varied relative to each other in luminance such that they were the same or were different by 5, 10, or 15 gray levels in either direction. For convenience, we will retain the reference–adjustable face nomenclature. On each trial, the reference face varied among five levels (-10, -5, 0, $+5$, or $+10$ gray levels relative to the base level referred to in Experiment 1), and the adjustable face was rotated among seven luminances relative to the reference face (-15, -10, -5, 0, $+5$, $+10$, or $+$ 15 gray levels). This corresponds to approximately $\pm$ 7.79 cd/m^2. Eight of these sets of stimuli were created: 4 conditions (Black face reference, White face adjustable; White face reference, Black face adjustable; Black face reference and adjustable; and White face reference and adjustable) $\times$ 2 sets (reversing screen sides). This set of 280 stimuli was repeated twice for a total of 560 trials.

Subjects were told they would see pairs of faces that were either the same race or different races and that they were to respond *same* by hitting the *1* key and *different* by hitting the *2* key on the computer keyboard using two fingers from their dominant hand. They were instructed to respond as quickly as possible without sacrificing accuracy. Subjects were explicitly told that the faces would also vary in lightness and that they were to ignore this variation in responding. They were also told that they would be responding to only one Black face and one White face.

On each trial, subjects saw a pair of faces separated by 10.1 cm from a viewing distance of approximately 50 cm (10.1° visual angle). This display was response terminated and followed immediately by feedback in the form of a "+" for correct responses, and a "−" for incorrect responses. The feedback was onscreen for 400 ms, and this was followed by a 300-ms blank-screen interstimulus interval. Subjects completed the entire experiment in one sitting of approximately 15–20 min with a break after the first 280 trials.

Results

On average, subjects had an error rate of 6.81%, and these trials along with trials with reaction times (RTs) greater than 3 standard deviations above each subject's grand mean were eliminated from the final analysis. The RT cutoff resulted in removing 1.71% of trials.

Different-race trials. RT data for different-race trials were entered into a 7-level one-factor repeated measures ANOVA (with luminance difference between the Black and White faces as the factor) to test the initial hypothesis that luminance differences affected RTs. This ANOVA revealed a significant effect of luminance difference on RTs, $F(6, 84) = 3.73$, $MSE = 3{,}094$, $p = .002$ (see Figure 8), and pairwise comparisons demonstrate that subjects classified pairs for which the Black face was 10 levels lighter significantly more slowly (by 42 ms) than they classified pairs for which the Black and White faces were the same luminance, $t(14) = 3.07$, $p = .008$. The pairs for which the Black face was 5 or 15 levels lighter were not significantly more slowly classified than the equiluminant pair, $t(14) < 1$ and $t(14) = 1.07$, $p = .303$, respectively. Comparisons between the equiluminant pair and the pairs for which the Black face was darker were nonsignificant for the 5- and 10-level darker Black pairs, $t < 1$ and $t(14) = 1.35$, $p = .199$, respectively, and significant for the 33-ms difference between the 15-level darker Black pair and the equiluminant pair, $t(14) = 2.37$, $p = .033$.

Twelve of the 15 subjects classified pairs in which the Black face was 10 levels lighter more slowly than they did matching pairs.

Same-race trials. An analysis of the same-race trials revealed a nonsignificant effect suggesting that subjects classified pairs with luminance differences as *same* more slowly as the difference increased. A four-level one-factor repeated-measures ANOVA with luminance difference as the factor (0-, 5-, 10-, and 15-level differences) produced a marginally nonsignificant luminance difference effect, $F(3, 42) = 2.24$, $MSE = 1{,}400$, $p = .098$. Mean classification times for the 0-level, 5-level, 10-level, and 15-level differences were 989 ms, 995 ms, 1,002 ms, and 1,022 ms, respectively. The difference between the equiluminant pair and the 15-level difference pair approached significance, $t(14) = 1.88$, $p = .081$, whereas the other two comparisons with the equiluminant pair did not, $t < 1$.

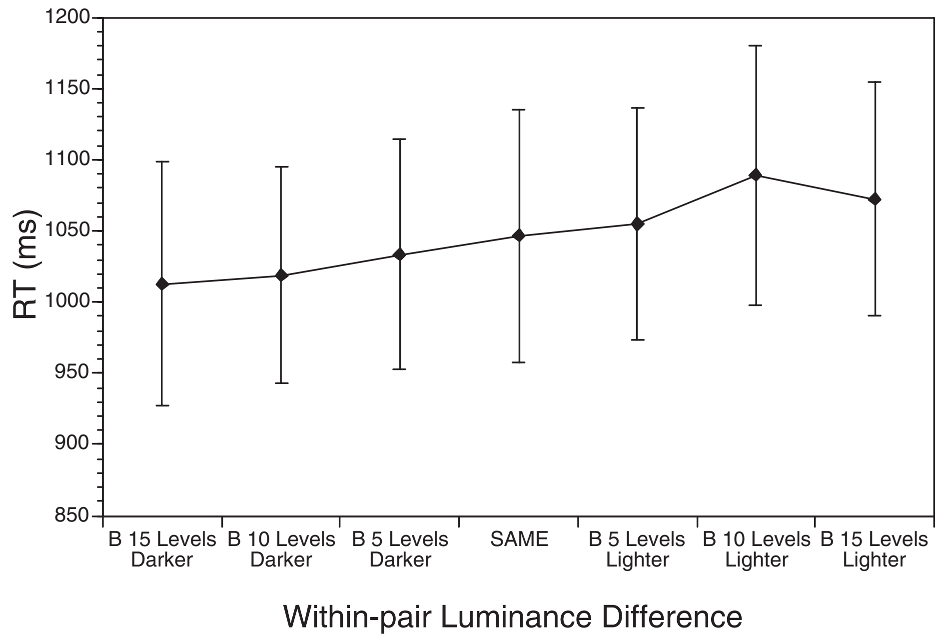

Figure 8. Mean reaction times (RTs) for different-race trials. Pairs for which the Black (B) face was darker than the White face are on the left, ranging from 15 levels darker to 5 levels darker. Pairs for which the Black face was lighter than the White face are on the right, and pairs for which there was no difference in luminance are in the middle, labeled *SAME*. Error bars reflect standard errors.

Discussion

Experiment 4 demonstrates that subjects are slower to classify face pairs as *different* when they match in apparent brightness versus when they match in actual luminance. The direction of the effect is again consistent with the hypothesis that subjects perceive Black faces to be darker than they are and/or White faces to be lighter than they are: Pairs for which the Black faces were more luminant than the White faces were matched for perceived brightness, causing them to be less effectively different and therefore more slowly classified as different. Thus, for a speeded classification task in which the average reaction time is about 1 s, we again see evidence of a brightness distortion. As such, the distortion effect appears to satisfy at least some of the criteria for a perceptual process: It occurs quickly and without extensive deliberation, and information is used in a relatively automatic way even in a context where the task explicitly excludes it. This latter point is particularly important because subjects were told that the mismatches in lightness were not part of the task—successful performance of the task demanded that subjects ignore brightness. The fact that they did this successfully means that subjects' explicit, deliberate responses were matched to the form categories represented by the different race faces, even as the RTs revealed a lightness distortion. This finding, therefore, provides additional evidence against the demand characteristic explanation of the results.

General Discussion

In four experiments, we have consistently demonstrated a distortion in lightness perception when the reference was a Black versus White face. Subjects choose a relatively lighter sample for perceptually unambiguous White faces, for ambiguous faces that were only differentiated by race on the basis of their context and/or a label, and for line-drawing faces with consistent fill. These findings therefore extend MacLin and Malpass's (2001, 2003) observations that people rate Black faces to be darker than White faces. In those initial studies, subjects simply rated the faces to be darker using a labeled light–dark response scale, whereas in this case, subjects actually compared the faces with visual samples that varied in luminance. These data also suggest that the relationship between lightness and faces of a given race may be independent of attitudes. Clearly, however, this conclusion must be considered tentative because it rests on negative findings in three relatively small populations and only an explicit measure of attitude. Studies by Hugenberg and Bodenhausen (2003) demonstrate the association of implicit attitude (Black–White association with good–bad) and adjustments of emotion on the face, with those who show greater implicit negative associations to Black also perceiving threat in a Black face more rapidly. Future studies will look at this issue directly.

One issue that deserves some exploration is the between-experiment differences in the perceptual magnitude of the lightness distortion. In Experiment 2, the distortion effect was the largest: 15.49 levels (d = 1.65; a reflectance difference of approximately 11.0 cd/m^2) for the between-subjects comparison between the ambiguous faces and 17.15 levels (d = 1.26; a reflectance difference of approximately 11.5 cd/m^2) for the within-subjects comparison between the ambiguous face and the unambiguous face. This compares with a 2.94-level difference (d = 0.75; a reflectance difference of approximately 1.9 cd/m^2) between the unambiguous faces in Experiment 1 and a 3.55-level overall difference (d = 0.33; reflectance difference of approximately 2.38 cd/m^2) in Experiment 3. At this point, a definitive explanation for these differences is not possible, but we suspect that the match-to-gray method used in Experiment 2 was in large part responsible. When matching between face stimuli, a number of factors may tend to reduce the size of the effect. First, this technique may encourage a focus on more than just matching absolute levels of gray. For example, subjects may compare the degree to which corresponding regions contrast with nearby regions. To the extent they do this, the effect might be eliminated, which might account not only for the smaller effect size but also for the fact that it was less consistent across subjects for Experiment 2. Of course, Experiment 3 also involved adjusting a uniform sample and it produced a similarly small and less than universal effect. However, it seems likely that the stimuli themselves caused this. As we mentioned above, these highly controlled stimuli do not strongly invoke the social categories they correspond to and might therefore have diluted the effect.

Summary and Conclusion

These findings add to not only the literature documenting the impact of social categorizations on the perception of basic features but also the lightness perception literature. As reviewed above, contributors to this literature have focused on early and midlevel visual inputs to lightness perception and have yet to explore the degree to which most abstract expectations might influence lightness. However, current theory in lightness and brightness perception clearly suggests this kind of influence. For example, recent research has emphasized the idea that lightness perception is based on probabilistic associations between different contexts and lightness percepts (see, e.g., Purves, Williams, Nundy, & Lotto, 2004). This idea is learning based and therefore provides precedence for our demonstration that knowledge about the typical luminance associated with a given category can impact lightness perception. This distortion probably reveals an underlying process that initially appears to be quite distinct from those causing other lightness distortions and contrast effects. In the case of the contrast effects demonstrated by Mach bands and the Hermann grid illusion, the distortions reveal a process that has a basic perceptual purpose: increasing the salience of edges. Similarly, the perceptual context effects that cause lightness distortions probably stem from a visual heuristic that helps the perceptual system sort out the relative luminances of difference surfaces to facilitate perceptual processes that might depend on this information (e.g., assignment of surfaces to the same or different objects).

In contrast, it would be difficult to argue that distortions in the lightness perception of faces is the result of a more basic process that is necessary to disambiguate downstream perceptual hypotheses. However, if one only slightly expands the nominal utility of lightness as a perceptual feature, then these effects represent an extension of existing traditions rather than a different process altogether. This distortion might be considered a case where a set of correlated features mutually facilitate each other such that the presence of most members of the set causes activation of representations of the missing members. So, the correlation between form and shading causes shading features to be activated in the presence of form features. Such a process would be similar to a wide range of hypothesized processes aimed at decisively settling the competition between two perceptual alternatives in a winner-take-all fashion. Thus, Black faces might appear relatively dark not because people are able to see them better that way but as the result of feature activations resulting from a perceptual classification.

To conclude, we argue that these data represent an extension of the lightness perception literature to include broad contextual and knowledge-based influences that go beyond the current emphasis on surface properties and other basic aspects of form and lighting. Conversely, they demonstrate the impact of categorical social knowledge on perception, and they offer an opportunity to explore the degree to which subjects have assimilated between-feature correlations in stimuli that underlie perceptual social categories. It is interesting to consider that the link between form and lightness represented by different races of faces may be relatively rare if one assumes that few other form-based categories are strongly associated with characteristic reflectances. However, we suspect that it will not be long before someone comes up with additional examples, and we look forward to more research exploring how this most basic of percepts is modulated by a broad range of knowledge.

References

Adelson, E. H. (1993, December 24). Perceptual organization and the judgment of brightness. *Science, 262,* 2042–2044.

Adelson, E. H. (2000). Lightness perception and lightness illusions. In M. Gazzaniga (Ed.), *The new cognitive neurosciences* (2nd ed., pp. 339–351). Cambridge, MA: MIT Press.

Cornsweet, T. N. (1970). *Visual perception.* New York: Academic Press.

Dasgupta, N., Banaji, M. R., & Abelson, R. P. (1999). Group entitativity and group perception: Associations between physical features and psychological judgment. *Journal of Personality and Social Psychology, 77,* 991–1003.

Dennett, D. C. (1991). *Consciousness explained.* Boston: Little, Brown.

Gilchrist, A. L., & Annan, V. (2002). Articulation effects in lightness: Historical background and theoretical implications. *Perception, 31,* 141–150.

Goldstone, R. L. (1994). Influences of categorization on perceptual discrimination. *Journal of Experimental Psychology: General, 123,* 178–200.

Harnad, S. (1987). Psychophysical and cognitive aspects of categorical perception: A critical overview. In S. Harnad (Ed), *Categorical perception: The groundwork of cognition* (pp. 1–28). New York: Cambridge University Press.

Hugenberg, K., & Bodenhausen, G. V. (2003). Facing prejudice: Implicit prejudice and the perception of facial threat. *Psychological Science, 14,* 640–643.

Levin, D. T. (1996). Classifying faces by race: The structure of face categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 1364–1382.

Levin, D. T. (2000). Race as a visual feature: Using visual search and perceptual discrimination tasks to understand face categories and the cross-race recognition deficit. *Journal of Experimental Psychology: General, 129,* 559–574.

Logvinenko, A. D. (2002). The anchoring effect in lightness perception in humans. *Neuroscience Letters, 334,* 5–8.

MacLin, O. H., & Malpass, R. S. (2001). Racial categorization of faces. *Psychology, Public Policy, and Law, 7,* 98–118.

MacLin, O. H., & Malpass, R. (2003). The ambiguous-race face illusion. *Perception, 32,* 249–252.

Peterson, M. A. (1994). Object recognition processes can and do operate before figure–ground organization. *Current Directions in Psychological Science, 3,* 105–111.

Purves, D., Williams, S. M., Nundy, S., & Lotto, R. B. (2004). Perceiving the intensity of light. *Psychological Review, 111,* 142–158.

Rhodes, G., Brennan, S., & Carey, S. (1987). Identification and ratings of caricatures: Implications for mental representations of faces. *Cognitive Psychology, 19,* 473–497.

Rossi, A. F., & Paradiso, M. A. (2003). Surface completion: Psychophysical and neurophysiological studies of brightness. In L. Pessoa & P. De Weerd (Eds.), *Filling-in: From perceptual completion to cortical reorganization* (pp. 59–80). New York: Oxford University Press.

Shiffrar, M., & Freyd, J. J. (1993). Timing and apparent motion path choice with human body parts. *Psychological Science, 4,* 379–384.

Steinfield, G. J. (1967). Concepts of set and availability and their relation to the reorganization of ambiguous pictorial stimuli. *Psychological Review, 74,* 505–525.

Tajfel, H., & Wilkes, A. L. (1963). Classification and quantitative judgment. *British Journal of Psychology, 54,* 101–113.

Treisman, A., & Gormican, S. (1988). Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review, 95,* 15–48.

Williams, S. M., McCoy, A. N., & Purves, D. (1998). An empirical explanation of brightness. *Proceedings of the National Academy of Sciences USA, 95,* 13301–13306.

Wolfe, J. M. (1984). Global factors in the Hermann grid illusion. *Perception, 13,* 33–40.

Received September 9, 2004
Revision received April 28, 2006
Accepted April 28, 2006

■