January 2010

Implicit And Explicit Racial Attitudes: Moderation Of Racial Typicality Evaluations

Elena Stepanova
Washington University in St. Louis

WASHINGTON UNIVERSITY IN ST. LOUIS

Department of Psychology

Dissertation Examination Committee:
Michael J Strube, Chair
Randy J. Larsen
William Lowry
Michael Mueller
Simine Vazire
Desireé White

IMPLICIT AND EXPLICIT RACIAL ATTITUDES: MODERATION OF RACIAL TYPICALITY EVALUATIONS

by

Elena V. Stepanova

A dissertation presented to the
Graduate School of Arts and Sciences
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

August 2010

Saint Louis, Missouri

ABSTRACT OF THE DISSERTATION

Implicit and Explicit Racial Attitudes: Moderation of Racial Typicality Evaluations

by

Elena V. Stepanova

Doctor of Philosophy in Psychology

Washington University in Saint Louis, 2010

Professor Michael J Strube, Chairperson

Previous research has shown that racial images representing more typical Afrocentric phenotypic characteristics result in more negative evaluations, whether assessed by explicit or implicit attitudes measures. However, the factors that define and moderate the perception of racial typicality have not been sufficiently explored. The current research investigated additive and interactive influences of skin tone and facial physiognomy on racial typicality evaluations, as well as the degree to which those effects were moderated by explicit and implicit racial attitudes, ethnicity of participants, and availability of cognitive resources. Using a 6-point scale ranging from *very African American* to *very Caucasian*, participants (N = 250) judged faces varying on 10 levels of facial physiognomy (from very Afrocentric to very Eurocentric) and 10 levels of skin color (from very dark to very light). Additionally, time constraints were manipulated by having participants complete the racial typicality judgments three times--without a response deadline, with a deadline equal to their median response during the no-deadline condition, and with a deadline equal to their 25^{th} percentile response during the no-deadline condition. Skin color and facial physiognomy interacted to influence racial typicality ratings, and this interaction was

further qualified by the time constraint manipulation. Under time constraints, participants primarily relied on skin color when rating faces of extreme levels of facial physiognomy, whereas they relied on both skin color and facial physiognomy when rating faces of intermediate levels of facial physiognomy. Other results indicated that the relationship between skin color and participants' ratings of racial typicality was stronger for those with higher implicit racial attitudes. European American and Asian American participants relied upon skin color more than African American participants, and African American participants relied upon facial physiognomy more than European American and Asian American participants. Conceptual, methodological and practical implications for race-relevant decisions are discussed.

Acknowledgments

This dissertation research was partially funded by the 2008-2009 Psi Chi Graduate Research Grant. First and foremost, I am very grateful for all the guidance and support that my adviser, Dr. Michael Strube, provided throughout the course of my graduate training and this dissertation project. I extend my sincere appreciation to all the members of my dissertation committee for their valuable time and advice. I would like to express my gratitude to all staff members and volunteers from *HealthStreet*, the Center for Community-Based Research, at the Institute of Clinical and Translational Sciences, Washington University School of Medicine, for their assistance in recruiting participants. I offer a special thanks to Randy Larsen, who secured my funding for the fall semester of 2009. I would also like to thank all my friends and colleagues at the Department of Psychology, and elsewhere in the world. My sincere appreciation goes to Tatiana Pollo, who opened her house for me during my frequent trips to Saint Louis over the past two years. This dissertation is dedicated to my family: my parents Valeriy and Lidia, my husband David, and my son Alexander. They made many sacrifices for me to undertake doctoral training, and the completion of this dissertation would have been impossible without their constant love and support. Dear Alexander, thank you for surviving graduate school with me! You are my true inspiration.

Table of Contents

Abstract ii

Acknowledgments iv

Table of Contents v

List of Tables vii

List of Figures viii

Introduction 1

Use of Facial Stimuli in Research on Prejudice and Stereotyping 2

Within-group Variability, Structure of Concepts, and Maddox Model of Racial Phenotypicality Bias 3

Skin Color and Facial Physiognomy 8

Categorical Route and Potential Moderators 11

Moderators of Racial Typicality Ratings 13

The Current Research 20

Methods 24

- Participants 24
- Facial Stimuli 27
- Explicit Measures 30
- Implicit Measure 32
- Outcome Measure 33
- Procedure 36

Results 38

Discussion 70

References.....84

Appendices.....96

Footnotes.....102

List of Tables

Table 1. Gender, Political Outlook, and Political Affiliation by Participants' Ethnicity (in Percentages). 26

Table 2. Interrcorrelations, Means, Standard Deviations and Scale Reliabilities of Individual Differences Variables for the Entire Sample. 41

Table 3. Interrcorrelations, Means, and Standard Deviations of Individual Differences Variables for European American, African American, Asian American and Other Participants.....43

Table 4. Means and Standard Errors for Racial Typicality Ratings as a Function of Skin Color and Facial Physiognomy.....46

Table 5. Mean Differences Between Adjacent Levels of Facial Physiognomy Within Each Level of Skin Color. 49

List of Figures

Figure 1. A Model of racial phenotypicality bias. Adapted from "Perspectives on Racial Typicality Bias," by K. B. Maddox, 2004, *Personality and Social Psychology Review, 8,* p. 395. Copyright 2004 by Sage Publications. Reprinted by permission. 7

Figure 2. Category-based racial typicality judgments (Expansion of Maddox, 2004 Model).....12

Figure 3. The effect of skin color and facial physiognomy on ethnic typicality judgments (from Stepanova et al., 2008).....16

Figure 4. The effect of skin color and facial physiognomy on ethnic typicality judgments among participants with high implicit prejudice (from Stepanova et al., 2008).....17

Figure 5. The effect of skin color and facial physiognomy on ethnic typicality judgments among participants with low implicit prejudice (from Stepanova et al., 2008).....18

Figure 6. Facial stimuli used in Stepanova et al. (2008) and current research (a sample set).....29

Figure 7. The effect of skin color and facial physiognomy on racial typicality judgments (entire sample) 48

Figure 8. The effects of skin color and implicit racial attitudes on racial typicality judgments (entire sample). 51

Figure 9. The effect of skin color and facial physiognomy on racial typicality judgments in the No Response Deadline condition.....53

Figure 10. The effect of skin color and facial physiognomy on racial typicality judgments in the Median Response Deadline condition. 54

Figure 11. The effect of skin color and facial physiognomy on racial typicality judgments in the 25th Percentile Response Deadline condition. 55

Figure 12. The effect of skin color and participants' ethnicity on racial typicality judgments, when data from Hispanic Americans are excluded. 59

Figure 13. The effect of facial physiognomy and participants' ethnicity on racial typicality judgments, when data from Hispanic Americans are excluded, in no response deadline condition. 61

Figure 14. The effect of facial physiognomy and participants' ethnicity on racial typicality judgments, when data from Hispanic Americans are excluded, in Median Response Deadline condition. 62

Figure 15. The effect of facial physiognomy and participants' ethnicity on racial typicality judgments, when data from Hispanic Americans are excluded, in 25th Percentile Response Deadline condition 63

Figure 16. The effects of implicit racial attitudes, explicit racial attitudes and facial physiognomy on racial typicality judgments. 66

Figure 17. The effects of implicit racial attitudes, explicit racial attitudes and skin color on racial typicality judgments 68

Implicit and explicit racial attitudes: Moderation of racial typicality evaluations

Introduction

In studies of prejudice and stereotyping, participants are frequently presented with African American and European American faces as stimuli intended to activate a racial concept. Research has found that the more that racial images represent typical[1] Afrocentric phenotypic characteristics, the more negative are evaluations whether assessed by traditional explicit measures or by more automatic implicit measures (for a review, see Maddox, 2004). Despite the general consistency of these racial effects, however, the factors that *define* and *moderate* the perception of racial typicality have largely been ignored in categorization and impression formation research.

Accordingly, the current research had three major goals. First, two factors—skin color and facial physiognomy—have dominated attempts to manipulate faces used as racial stimuli, but little is known about the relative weighting of these factors in perceptions of racial typicality. The current study manipulated skin color and facial physiognomy independently to determine their separate and combined effects. Second, perceptions of racial typicality—and especially the role played by skin color and facial physiognomy—are likely moderated by implicit and explicit racial attitudes. These individual differences were examined in this study. Finally, the influence of skin color and facial physiognomy in racial categorization may depend on available cognitive resources or time demands. When a "racial decision" must be made quickly, for example, skin color might be expected to dominate the decision process because it is salient and easily viewed at a distance. The current research examined how the relative weighting of skin color and facial physiognomy changed as cognitive resources changed.

In the discussion that follows, I will first describe how facial stimuli have played a central role in research on racial prejudice and stereotyping. Next, I will describe how different models of categorization address within-group variability and the implications these models have for social categorization. Then, I will describe the role that skin color and facial physiognomy of facial stimuli play in racial typicality evaluations, including recent research on moderators of racial typicality. I will conclude with a description of the major hypotheses tested in the current study.

Use of Facial Stimuli in Research on Prejudice and Stereotyping

Generally, the activation of "race" in studies of implicit and explicit racial prejudice and stereotyping is achieved by employing two types of stimuli: lexical (e.g., by using ethnic labels such as "White," "Black," or by using ethnic names such as "Malic" or "Chip") or pictorial (e.g., by presenting African American or European American faces). An especially prominent example of work using facial stimuli is recent research employing automatic or implicit measures of racial prejudice and stereotyping. This research generally finds substantial evidence of automatic negative evaluations and stereotypic associations of African Americans by White participants (e.g., Dovidio, Kawakami, Johnson, Johnson, & Howard, 1997; Fazio, Jackson, Dunton, & Williams, 1995; Payne, 2001; Wittenbrink, Judd, & Park, 2001).

Pictorial race activation in particular has relied on diverse methods to create the stimuli. Sometimes the images are composites (e.g., Payne, 2001), sometimes they are pictures of actual people (Fazio et al., 1995) and sometimes they are generated schematic images (e.g., Dovidio et al., 1997) or animations (e.g., Hugenberg & Bodenhausen, 2004). Furthermore, the images sometimes are presented in color (e.g., Plant, Peruche &

Butz, 2005) but mostly in black and white or gray-scale (e.g., Bargh, Chen & Burrows, 1996). The stimuli vary on "relevant" dimensions of skin tone and facial physiognomy, but also vary along a number of other potentially important dimensions (age, attractiveness, emotional display, etc.). This variability raises two important questions. First, which specific dimensions or features are central to the activation of racial categories when stimuli varying on multiple dimensions are used? Second, to what extent is variability within a particular dimension important? Some guidance is provided by views of conceptual structure in general categorization research.

Within-group Variability, the Structure of Concepts, and the Maddox Model of Racial Phenotypicality Bias

Researchers generally try to construct or select facial stimuli so that the faces are clear representations of one particular racial category. This approach relies on the assumption that we simplify our complex social world through categorization (cf. Allport, 1954) such that presentation of a target stimulus triggers categorization along salient dimensions or features such as age, sex and race. The use of prototypical stimuli, however, belies the fact that, outside the laboratory, faces vary along many dimensions in often subtle ways. For example, human skin color and facial physiognomy vary along a wide continuum, even within any one racial or ethnic group (Farkas et al., 2005; Jablonski, 2004; Parra, 2007). This within-group variability may be especially important in social categorization because the categorization of less typical members of a racial group may be especially sensitive to situational factors (e.g., cognitive resources) and individual differences (e.g., implicit racial prejudices). In the discussion that follows, I will review how general categorization research addresses issues of stimulus variability

and its implications for social categorization, especially racial categorization. Three views will be described, each addressing how different features of stimuli and combinations of those features might influence the categorization process.

According to a somewhat older view of conceptual structure, the *classical view*, described by Smith and Medin (1981) and traced back to Aristotle, mental representations of categories have a set of necessary and sufficient features that determine category membership; as a result, all category members are good examples of that category if they possess those features (for review, see Kunda, 1999; Medin, 1989; Medin & Smith, 1984). In that case, a face can be categorized as either "African American" or "European American," as long as it possesses all necessary and sufficient features to be placed into one category or the other. A set of necessary and sufficient features for the "African American" category might be limited to a specific skin tone or a combination of skin tone and one or more facial features. This approach, however, does not account well for the categorization of highly variable stimuli (e.g., a face with light skin tone but very Afrocentric facial features).

The newer *probabilistic view*, foundations of which can be traced to Wittgenstein (1953), argues that category membership is probabilistic; that is, it is a matter of a degree. Members sharing more attributes or properties of a category are more typical than those sharing fewer attributes (for review, see Kunda, 1999; Medin, 1989; Medin & Smith, 1984; Smith & Medin, 1981), and a clear demarcation of category boundaries is not possible. To determine whether a target stimulus belongs to a category, one compares it (a) to a prototype or ideal summary representation of a category that possesses all characteristic features (prototype view), or (b) to a series of exemplars (exemplar view)

of specific category members (for review, see Kunda, 1999; Medin, 1989). Applied to racial judgments, this view argues that a perceiver compares a target face to a category prototype face (e.g., African American) or to a series of exemplar faces and makes a decision about the target face's membership in the category based on similarity. This approach assumes that stimuli might have different features, but that does not preclude these stimuli from being placed in the same category; instead, some stimuli will simply be considered more "typical" members of the category than others.

A third view, the *theory-based view of concepts*, argues that concepts also include causal knowledge. Categorization is not based on simple matching of example attributes and concept attributes, but on the correct "explanatory relationship" between an example and "the theory organizing the concept" (Medin, 1989, p. 1474). Medin gives as an example inferences that might be made about a person who has dived into a swimming pool totally clothed: "One might believe that having too much to drink impairs judgment and that going into the pool shows poor judgment" (p. 1474). However, the presence of other cues (e.g., knowledge that this person is particularly awkward or cannot swim) will alter those inferences and the categorization judgment itself (i.e., this person is a victim of her clumsiness rather than a drunkard). In other words, causal knowledge is used to resolve conflict among traits and categorization. Rather than relying on simple summation of attributes, individuals rely upon the "explanatory principle" common to category members and relations between attributes and concepts (Kunda, 1999; Medin, 1989). To return to the categorization problem central to the current research, if a person is presented with an ambiguous face (e.g., European facial features and dark skin color), it is possible to invoke causal reasoning (e.g., it is a European American person with a

tan) rather than simply comparing a set of facial attributes to a prototype or series of exemplars. This view addresses well how within-group variability can affect the categorization process and in particular implies that the same stimulus might be categorized quite differently under different situations or by different perceivers who make different causal assumptions (perhaps driven by racial attitudes).

To summarize, in the general categorization literature, two recent views (probabilistic and theory-based) suggest that within-group variability affects the general categorization process. However, these views are generic and do not address the complexities of within-group facial variability involved in facial processing—a matter of person perception. Some guidance is offered, however, by a recent model of racial phenotypicality bias developed by Maddox (2004). Maddox (2004) proposes two routes by which phenotypic features can affect racial evaluations (see Figure 1). First, facial features can lead to racial categorization directly through a *category-based route*. These judgments likely rely on a single strong cue such as skin tone. If additional phenotypic features (e.g., variations in facial features) influence judgments, they likely lead to subcategorization (e.g., Maddox & Gray, 2002) if this route is taken. Second, target attributes (e.g., a facial feature such as size of the nose) might directly affect racial evaluations (e.g., along a positive-negative dimension) even if no categorization occurs—a *feature-based route*. Maddox (2004) argues that these routes are largely independent and operate simultaneously, helping account for how within-group facial variability affects racial prejudice and stereotyping. In the *category-based route*, facial features activate a particular category (e.g., Black) and that category, in turn, leads to inferences

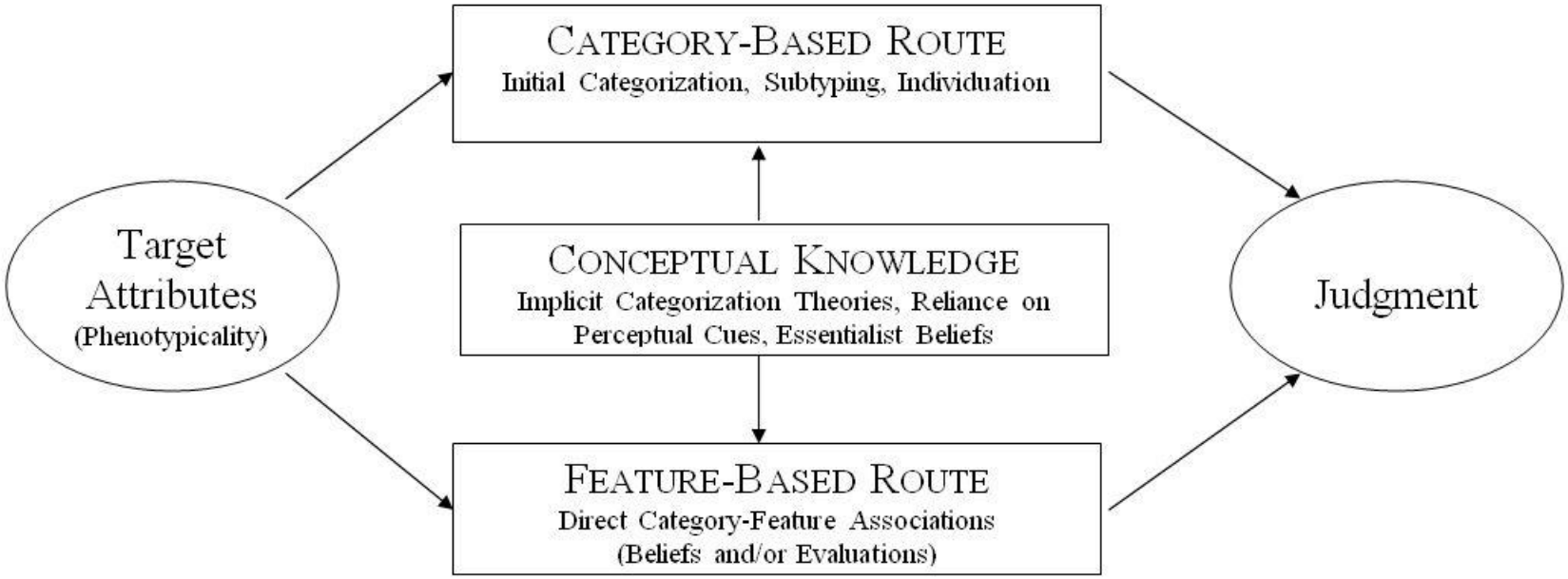

Figure 1. A Model of racial phenotypicality bias. Adapted from "Perspectives on Racial Typicality Bias," by K. B. Maddox, 2004, *Personality and Social Psychology Review, 8,* p. 395. Copyright 2004 by Sage Publications. Reprinted by permission.

of stereotypic traits or evaluative responses. In the *feature-based route*, certain facial features (e.g., width of the nose, size of the lips) lead directly to stereotypic or evaluative interferences—the feature itself, and not the category to which the face might belong, is the source of the inference. Additionally, conceptual knowledge might moderate processing for both of these routes. Maddox suggests that factors such as implicit categorization theories, reliance on perceptual cues and essentialist beliefs (e.g., naïve theories of racial category membership postulating that skin color reflects a person's genetic and/or cultural heritage) can influence judgments made through either route, an assumption shared with the theory-based view of concepts described previously. Of particular importance is the emphasis that the Maddox model places on differences within the same class of stimuli and the role that other factors—external to the stimulus attributes—may play in moderating the impact of stimulus attributes on racial categorization.

Skin Color and Facial Physiognomy

Empirical research has recently begun to address within-group variability and its implications for perceptions of racial categorization and typicality, but the efforts have not been especially impressive. There is little consistency in the facial features studied, with different researchers examining different numbers and combinations of features (e.g., size and fullness of nose, lips, specific hair structure, darkness of skin color, etc.; Livingston & Brewer, 2002; skin color only; Dasgupta, Banaji, & Abelson, 1999). The impression that one gets from this literature is that the features used in stimuli are interchangeable markers of race or that the differences among them are simply trivial.

Of the stimulus features used, skin color is usually assumed to be of primary salience in defining racial typicality judgments about non-White groups (for review, see Maddox, 2004). It is certainly the most visually salient cue for identifying members of many racial groups. When White participants are asked to give verbal descriptions of a Black face, they tend to mention darkness of the face, kinkiness of the hair, and brown eyes more than other features (Deregowski, Ellis, & Shepherd, 1975). When White participants are explicitly asked what features are important in racial categorization, they rate skin color as the most important criterion in the categorization of Black targets (Brown, Dane, & Durham, 1998). Interestingly, pre-categorized African American faces are perceived to be darker-skinned than European American faces even when their skin color is identical (see Levin & Banaji, 2006; MacLin & Malpass, 2001; MacLin & Malpass, 2003). Importantly, whether the perception of skin tone is accurate or not, darker-skinned African Americans are evaluated more negatively, judged more often as possessing stereotypic traits, and discriminated against more than lighter-skinned African Americans (for review, see Blair, Judd, Sadler, & Jenkins, 2002; Maddox, 2004; Maddox & Dukes, 2008).

In contrast, the role of facial physiognomy has rarely been addressed as an independent factor contributing to racial typicality judgments (but see Gitter & Satow, 1969). Previous research has shown that face pigmentation and shape contribute to face recognition independently (Russell & Sinha, 2007; Russell, Sinha, Biederman, & Nederhouser, 2006), suggesting that facial features can contribute independently to categorization and encoding. Facial physiognomy may be crucial because judgments of race cannot always depend reliably on skin color (it varies *within* groups as well as

between groups), and race-relevant judgments (e.g., eyewitness identification) may depend on finer distinctions than are provided by skin color alone. In one early study, Gitter and Satow (1969) manipulated physiognomy and skin color separately, albeit using dolls presented as stimuli in a study of racial misidentification in children. They found that physiognomy and skin color of stimuli were independent factors in children's judgments of their own racial identification. These results suggest that skin color and facial physiognomy might contribute independently to judgments of racial typicality for others as well, but recent research has only begun to investigate this possibility.

Stepanova and Strube (2009) followed up on Gitter and Satow's research in an effort to untangle the effects of skin tone and facial physiognomy on perception of racial typicality and racial categorization. They independently manipulated facial physiognomy,[2] skin tone and color presentation mode (grayscale vs. color) of computer generated faces. To manipulate physiognomy, Stepanova and Strube created three faces: a high physiognomy Afrocentric face, a low physiognomy Afrocentric face, and a Eurocentric face. Each of those faces was presented in two different skin tones: light and dark. These factors were manipulated orthogonally, resulting in six different faces. Results showed that both skin color and facial physiognomy contributed *independently* to racial typicality evaluations and were moderated by the mode of presentation (i.e., grayscale or color). Eurocentric faces were perceived as more European American in the grayscale presentation mode than in the color mode. Independent of facial physiognomy, dark skin tone faces were perceived as more African American than light skin tone faces, and this was especially true when faces were presented in color rather than in grayscale. Note, however, that this research had a limited number of levels of both skin color and

facial physiognomy. It may be that with finer manipulations of skin tone and facial physiognomy, these two factors might interact to determine racial typicality ratings, suggesting that people rely upon a *combination* of visual cues in racial typicality judgments. The current study addresses this limitation.

The Categorical Route and Potential Moderators

The current research focuses on the category-based route described by Maddox (2004) and concentrates on how exposure to facial stimuli affects racial typicality ratings. Of particular interest here is the relative weighting given to skin tone and facial features, and the role that pre-existing implicit and explicit racial attitudes play as moderators of the weight given to these cues. Furthermore, the weighting of facial features and skin tone might depend on the resources available to make decisions about race (e.g., time pressure). These moderating effects are depicted in Figure 2, an expansion of the categorization route from Maddox's (2004) model. The discussion that follows describes them in more detail.

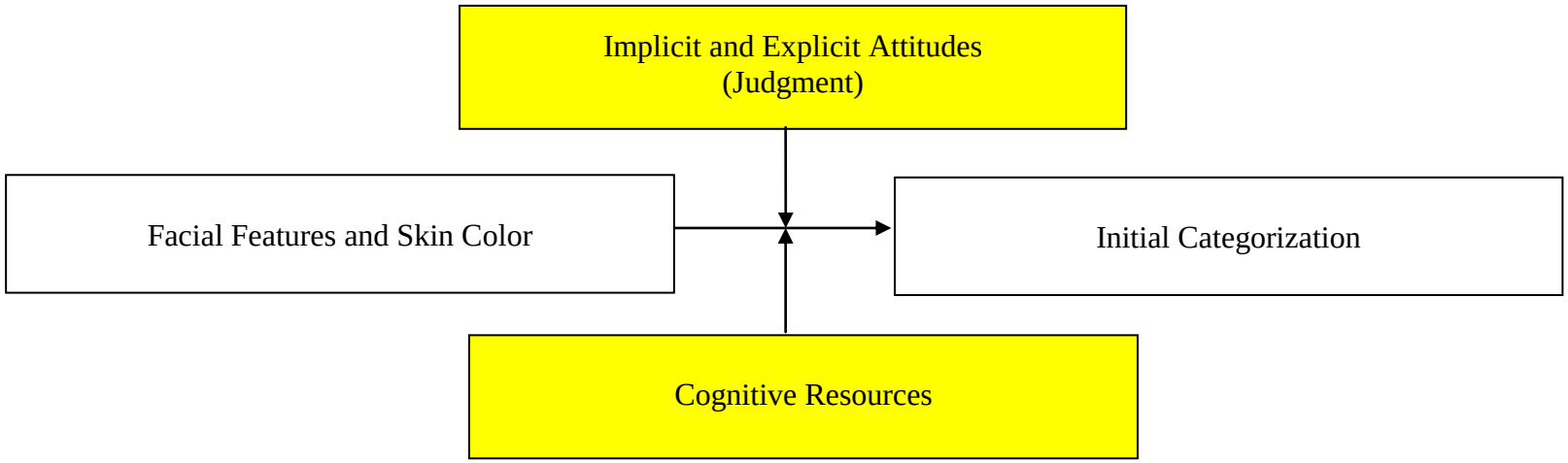

Figure 2. Category-based racial typicality judgments (Expansion of Maddox, 2004 Model).

Moderators of Racial Typicality Ratings

Explicit measures of prejudice. Racial and ethnic categorization research has explored some important potential moderators using explicit measures of prejudice, although usually conceptualizing it as in-group versus out-group identification. For example, Castano, Yzerbyt, Bourguignon, and Seron (2002) asked participants to categorize morphed faces as northern Italian or southern Italian when the target was composed of a varying percentage of a northern African and a northern European face. Consistent with the *in-group overexclusion hypothesis* (see Leyens & Yzerbyt, 1992), those participants who highly identified themselves with northern Italians (in-group) classified more target faces as out-group members (southern Italians) in comparison to those who did not highly identify with an in-group. Categorization latencies indicated that high in-group identifiers took longer to categorize those target faces that were more likely to be in-group members (higher percentage of northern European features in a morph) than out-group faces; low in-group identifiers took longer to categorize any ambiguous faces, regardless of their ethnic make-up.

Other studies have investigated moderators of racial categorization of ambiguous faces among South Africans and White Americans (Blascovich, Wyer, Swart, & Kibler, 1997; Pettigrew, Allport & Barnett, 1958). Although Pettigrew, Allport and Barnett did not explicitly measure racial attitudes of White Afrikaners, they assumed this ethnic group to be prejudiced against other groups ("Colored," Indians and Africans). Their findings indicated that White Afrikaners, when presented with racially mixed photographs, tended to include ambiguous faces in the extreme "African" group rather than in intermediate "Indian" or "Colored" groups. Blascovich, Wyer, Swart, and Kibler

(1997) found that highly prejudiced individuals took longer to categorize racially ambiguous faces. These results suggest that to protect their identity, perceivers are very careful about whom to include in their in-group. Relevant to the present research is the conclusion that the same faces can be judged quite differently depending on the racial attitudes of the perceiver, and that this variation is likely largest when facial cues are mixed or ambiguous, allowing wide latitude in how the features are weighted and combined.

Implicit measures of prejudice. Recent research has also begun to explore implicit moderators of racial categorization and racial typicality judgments based on facial cues. For example, Hutchings and Haddock (2008) and Hugenberg and Bodenhausen (2003) found that participants high in implicit racial prejudice were more likely to categorize angry (but not happy or neutral) ambiguous-race faces as Black. Stepanova, Strube, Yablonsky, Pehrson and Shuman (2008) also examined the role of implicit racial attitudes in racial typicality judgments using an expanded set of faces varying in skin color and facial physiognomy. Instead of only two levels of skin color and three levels of facial physiognomy, skin color and facial physiognomy were represented by ten levels each, with skin color varying from very dark to very light in gradual increments and facial physiognomy varying from very Afrocentric to very Eurocentric in gradual increments. Crossing these dimensions produced a stimulus set of 100 faces.

The research was conducted in a culture with a relatively low exposure to Afrocentric facial physiognomy (the Russian Federation). Participants were asked to rate faces on a 7-point scale: 1 (*Very non-Russian*), 2 (*Moderately non-Russian*), 3

(*Somewhat non-Russian*), 4 (*Not Clearly non-Russian or Russian*), 5 (*Somewhat Russian*), 6 (*Moderately Russian*) and 7 (*Very Russian*). Stepanova et al. assessed implicit ethnic attitudes, measured by the IAT (Greenwald et al, 1998). Results showed that skin color and facial physiognomy each affected racial typicality judgments, but unlike the earlier research, these two factors also interacted. Dark faces were rated consistently as non-Russian (the out-group for these participants), regardless of facial physiognomy. However, light colored faces received more varied racial typicality judgments and depended on facial physiognomy as well (see Figure 3). Most importantly, it was found that implicit attitudes also affected ethnic typicality ratings. Participants with high implicit ethnic prejudice tended to make finer distinctions when judging ethnic typicality in comparison to individuals with low implicit ethnic prejudice (see Figures 4 and 5).

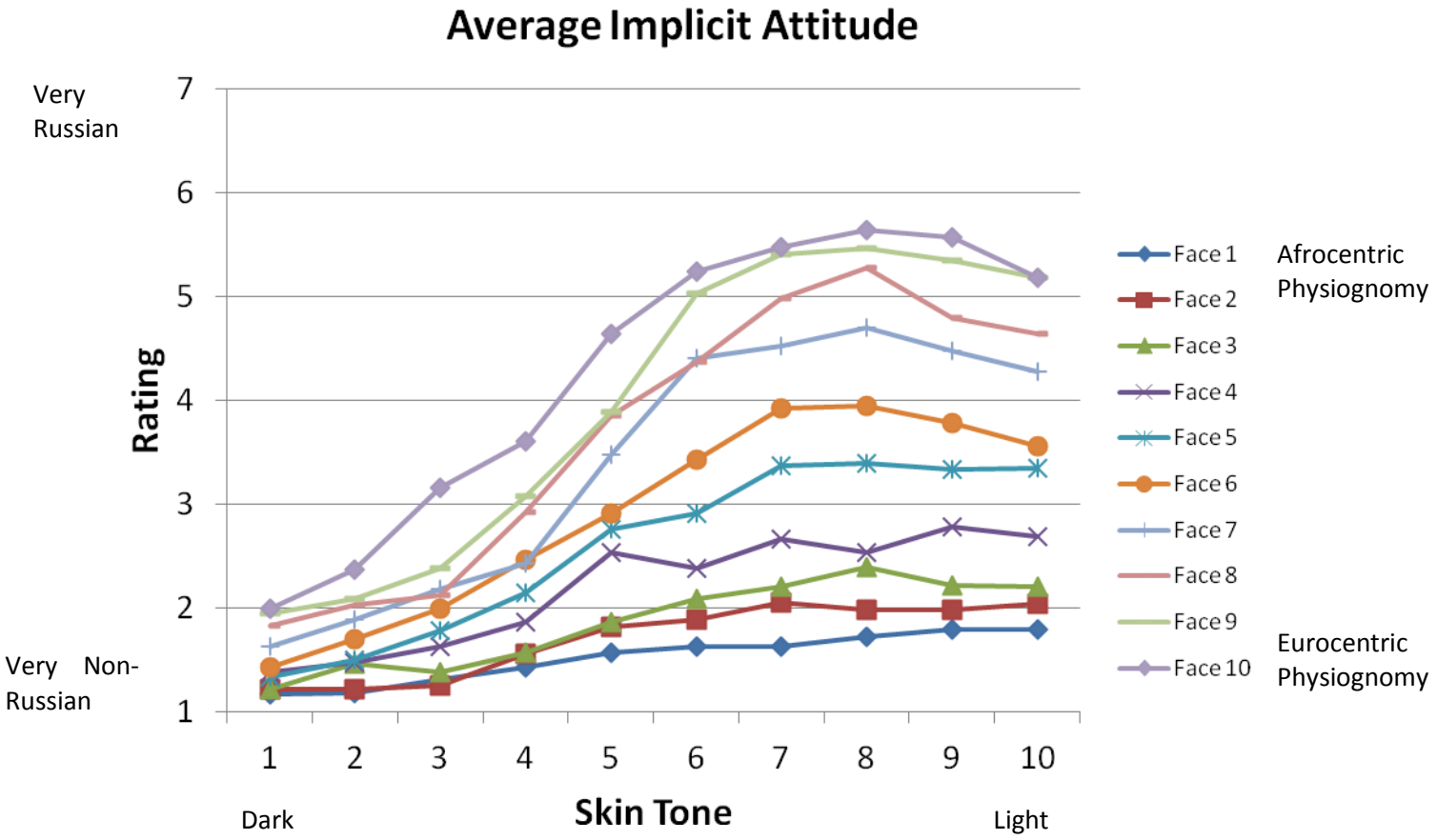

Figure 3. The effect of skin color and facial physiognomy on ethnic typicality judgments (from Stepanova et al., 2008). Note: Face 1= highest Afrocentric physiognomy and Face 10 = highest Eurocentric physiognomy.

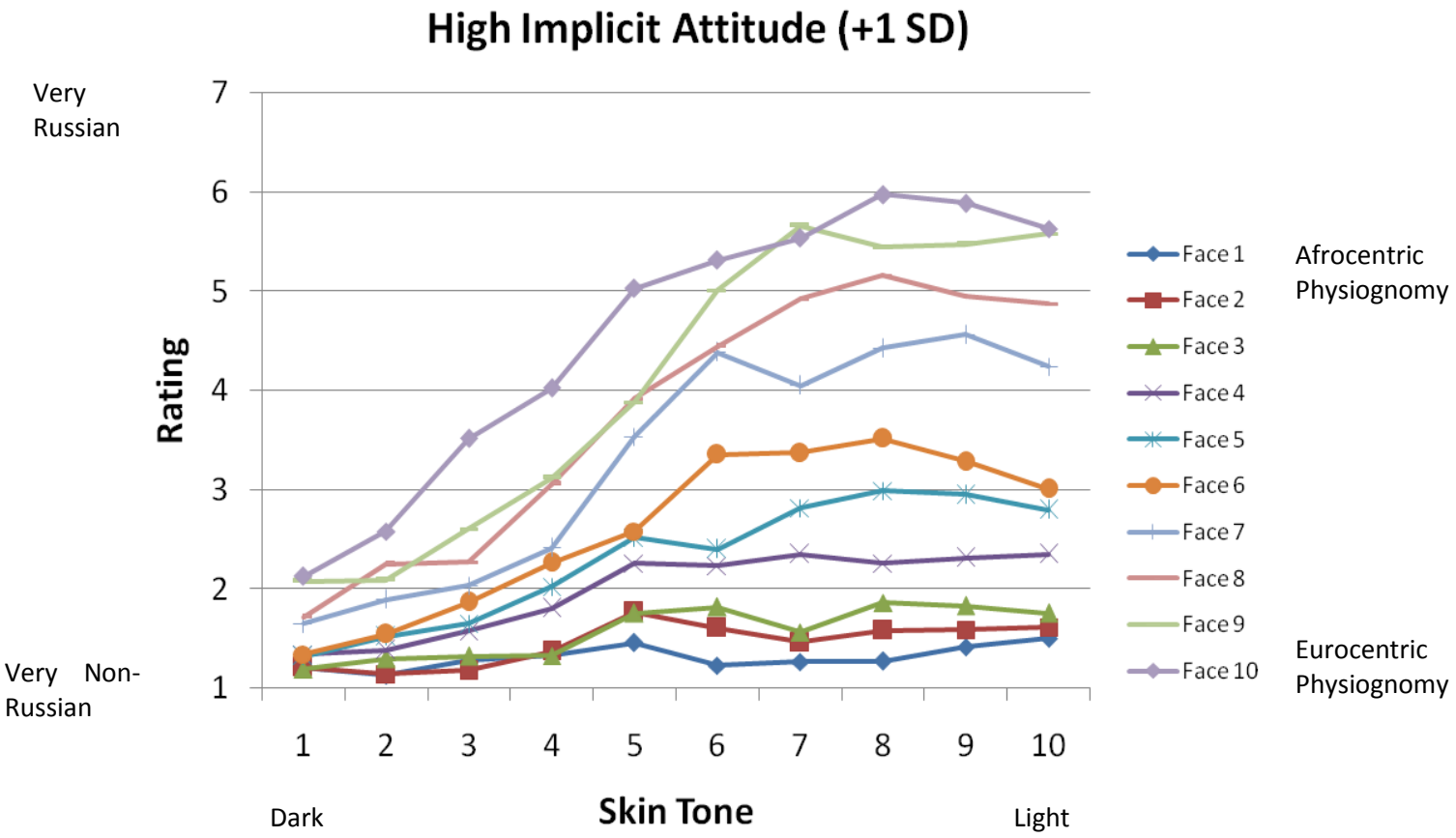

Figure 4. The effect of skin color and facial physiognomy on ethnic typicality judgments among participants with high implicit prejudice (from Stepanova et al., 2008).Note: Face 1 = highest Afrocentric physiognomy and Face 10 = highest Eurocentric physiognomy.

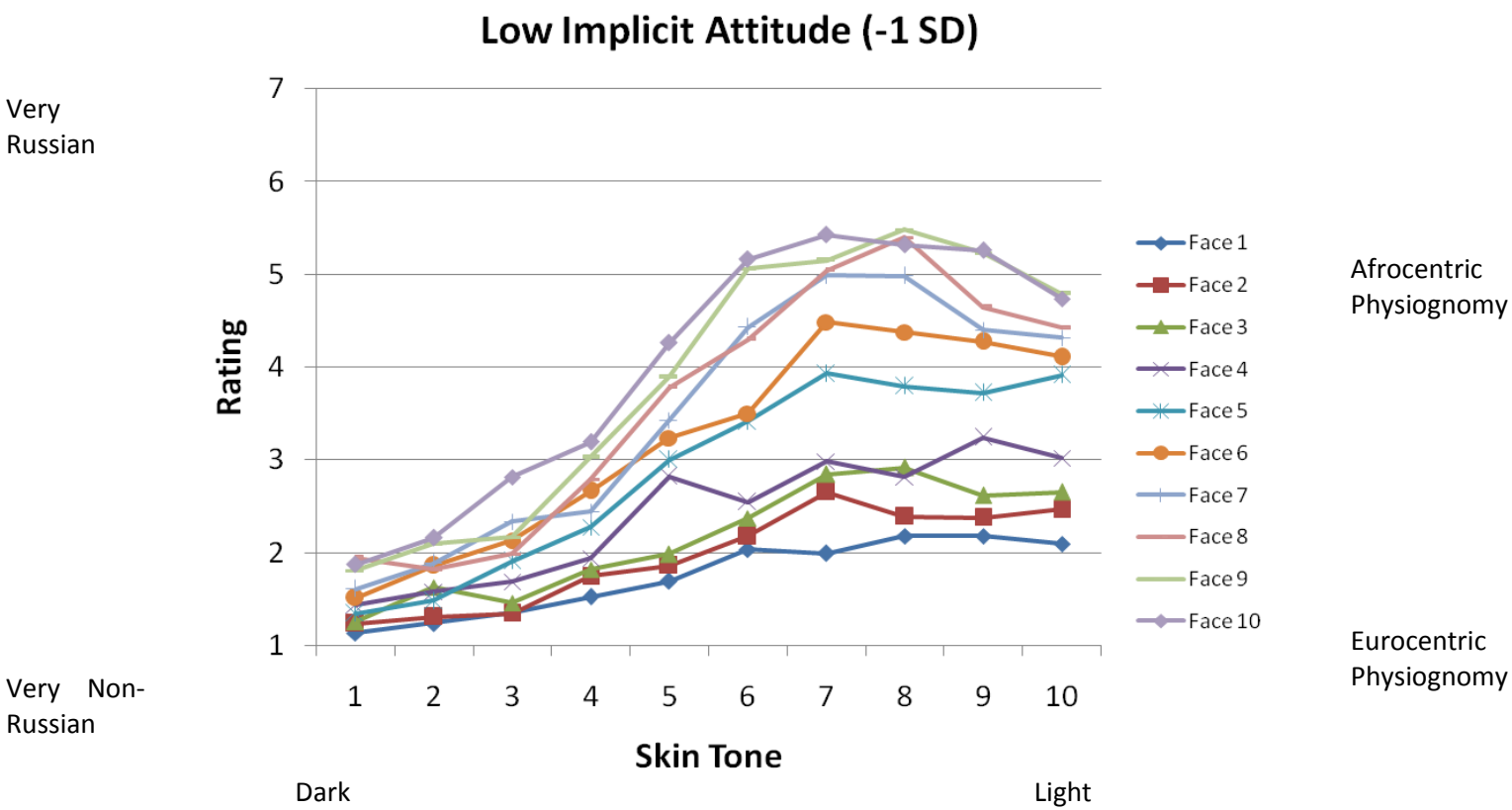

Figure 5. The effect of skin color and facial physiognomy on ethnic typicality judgments among participants with low implicit prejudice (from Stepanova et al., 2008).Note: Face 1 = highest Afrocentric physiognomy and Face 10 = highest Eurocentric physiognomy.

Both Stepanova and Strube (2009) and Stepanova et al. (2008) show that racial judgments are far more complex than previously thought and depend on much more than skin color alone. However, an important limitation of Stepanova et al. (2008) is that it was not conducted in the United States. The Russian Federation sample is very different from an American sample on a variety of dimensions, most notably exposure to African Americans. One key purpose of the current research was to establish these results in an American sample that resembles those on which most stereotyping research has been conducted. An additional purpose of this research was to determine if limited cognitive resources are an important moderator of the impact of facial cues on racial typicality judgments, as suggested by the extension of categorical route from Maddox's (2004) model.

Cognitive control and availability of cognitive resources. When cognitive resources are limited (e.g., under time constrains or cognitive load) people are unable to exert as much self-control over their judgments (Bodenhausen, 1990; Govorun & Payne, 2006; Richeson et al., 2003) and responses then are more likely to be based on automatic associations. When participants are categorizing faces, lack of cognitive control might fundamentally change how skin color and facial physiognomy influence racial categorization. With fewer cognitive resources available, more weight is likely to be given to salient features such as skin color, which require less processing, than to subtler facial details that require more processing. Moreover, when cognitive resources are limited, individual differences in levels of implicit and explicit prejudice might also produce quite different outcomes for racial typicality evaluations, with implicit attitudes perhaps playing a more prominent role. Therefore, another important purpose of this

research was to investigate how limited self-control through limited availability of cognitive resources affects racial typicality evaluations.

The Current Research

As Maddox and Dukes (2008) argue in their review, identifying the process by which facial features affect racial judgments and the particular features that affect social categorization are crucial to limiting racial biases. Specifically, if certain features (e.g., dark skin color) primarily drive social categorization effects, additional attentional resources and increased perceptual familiarity with those features might reduce racial biases. For example, one recent study by Lebrecht, Pierce, Tarr and Tanaka (2009) suggests that when Caucasian participants are trained to perceptually discriminate among various African American faces, their implicit racial biases decline. Consistent with the Maddox and Dukes' appeal, the current research examined skin color and facial physiognomy in greater detail than has been accomplished in previous research and examined individual differences and situational moderators that may alter the weighting of these facial features in determining judgments of racial typicality. The role of individual differences is suggested by research reviewed earlier—implicit and explicit racial attitudes can be expected to play important roles when facial features are used to categorize targets into racial groups. The availability of cognitive resources, however, can also be expected to play a part and is likewise implicated by the models described previously.

I examined the following questions in this research: (a) What are the additive and interactive (i.e., weighting) influences of skin tone and facial physiognomy on judgments of racial typicality when both skin color and facial physiognomy are varied independently

in gradual increments?; (b) Are those effects moderated by implicit and explicit racial attitudes?; and (c) Do time constraints alter the attention paid to (or weighting of) skin color and facial physiognomy?

Experiment Overview

This study was modeled after that conducted by Stepanova et al. (2008). Participants were presented with 100 computer-generated faces varying in skin tone and facial physiognomy. Each face was rated on a 6-point scale ranging from *very African American* to *very Caucasian*. The faces were rated three times. In the first block, participants performed the racial typicality task without any time constraints. In the second block, they performed the task under a modest time constraint. In the third block they performed the task under a stringent time constraint. Participants' implicit racial attitudes and explicit racial attitudes were also assessed.

Predictions

Hypothesis 1. I expected to replicate some of the findings that were obtained by Stepanova and Strube (2009). I expected that both skin color and facial physiognomy would independently influence racial typicality evaluations: Darker faces were expected to be judged as more African American than lighter faces, and faces with Afrocentric facial physiognomy were expected to be judged as more African American than faces with Eurocentric facial physiognomy.

Hypothesis 2. Given the greater sensitivity of the racial typicality task used in this study, I expected an interaction between facial physiognomy and skin color: Darker faces were expected to be rated consistently as African American with little influence from facial physiognomy. Lighter colored faces, however, were expected to receive

more varied racial typicality judgments and depend to a greater extent on facial physiognomy (see Figure 3).

Hypothesis 3. I expected to replicate the Skin Color x Facial Physiognomy x Implicit Racial Attitudes interaction obtained by Stepanova et al. (2008). Participants with higher implicit racial prejudice were expected to make finer distinctions when judging racial typicality in comparison to individuals with lower implicit ethnic prejudice. In other words, the pattern described for Hypothesis 2 was expected to be more pronounced for participants with high implicit racial prejudice than for participants with lower implicit racial prejudice (see Figures 4 and 5).

Hypothesis 4. I expected that, compared to no-time-constraint trials, time constraints (speeded categorization) would produce categorization decisions that would be more affected by skin tone than by facial physiognomy (e.g., Response Deadline x Skin Color x Facial Physiognomy interaction).

Hypothesis 5. I also expected to find a Response Deadline x Skin Color x Facial Physiognomy x Implicit Racial Attitudes interaction. Under time constraints, the effect described in Hypothesis 4 would be even more prominent in participants with higher implicit racial prejudice than in participants with lower implicit racial prejudice.

Even though past research has indicated moderation of racial categorization by explicit racial attitudes, explicit racial attitudes did not strongly moderate the impact of skin tone and facial features on racial typicality judgments in Stepanova et al. (2008). Accordingly, I offer no specific predictions for this study. Likewise, past research has not extensively examined racial categorization among non-White participants, so I offer no predictions for this potential moderator as well. Both explicit racial attitudes and

participant ethnicity were, however, examined in this study to explore their potential role in racial typicality judgments.

Methods

Participants

A sample of undergraduates (N = 207) from Washington University in Saint Louis was recruited through standard subject pool procedures. Data from all participants under 18 (N = 4) were excluded from the analysis in accordance with university and IRB regulations. Additionally, participants (N = 49) from the general population were recruited through the Volunteers for Health program at the Washington University School of Medicine, a program that recruits healthy volunteers from the community to participate in research conducted at Washington University. Some community participants were directly recruited by *HealthStreet*, the Center for Community-Based Research at the Washington University School of Medicine. Participants recruited from the general population received $10 for their participation. Data from two participants (one from the general population and one from the Washington University student population) were excluded from the analysis because they did not complete the experiment, resulting in 250 participants (48 from the general population and 202 from the student population). In the final sample used in this study, ethnicity of the participants was the following: 137 (54.8%) European American, 39 (15.6%) African American, 4 (1.6%) Hispanic American, 47 (18.8%) Asian American, and 23 (9.2%) "Other". Participants' mean age was 21.92 years, SD = 7.28 with an age range of 18-62 years. Participants' age varied across five ethnic groups: the oldest group was African Americans (M = 31.54, SD = 12.98), followed by Hispanic Americans (M = 21.75, SD = 6.18), Other (M = 20.30, SD = 3.62), European Americans (M = 20.29, SD = 3.82), and Asian Americans (M = 19.53, SD = 1.59). Forty percent of the total sample were men.

Along with demographic information, participants reported their "political outlook" and political affiliation. The distribution of participants for political outlook suggested a fairly liberal sample: "very liberal" (10.8%), "moderately liberal" (40.8%), "slightly liberal (14.0%), "neither liberal nor conservative" (14.4%), "slightly conservative" (11.2%), "moderately conservative" (5.6%), and "very conservative" (3.2%). Participants' political affiliation was primarily democratic: "strongly democratic" (17.2%), "moderately democratic" (31.2%), "slightly democratic" (14.4%), "neither democratic nor republican" (22.8%), "slightly republican" (7.6%), "moderately republican" (5.2%), and "strongly republican" (1.6%). Table 1 shows participants' gender, "political outlook" and political affiliation distributions for each ethnic group.

Table 1

Gender, Political Outlook, and Political Affiliation by Participants' Ethnicity (in Percentages)

Variable	European American (*N* = 137)	African American (*N* = 39)	Hispanic American (*N* = 4)	Asian American (*N* = 47)	Other (*N* = 23)
Gender					
Male	38.7	33.3	25.0	51.1	34.8
Female	61.3	66.7	75.0	48.9	65.2
Political Outlook					
Very Liberal	11.7	7.7	25.0	4.3	21.7
Moderately Liberal	43.1	28.2	50.0	44.7	39.1
Slightly Liberal	12.4	23.1		14.9	8.7
Neither Liberal Nor Conservative	12.4	20.5		14.9	17.4
Slightly Conservative	10.2	10.3		17.0	8.7
Moderately Conservative	5.8	7.7	25.0	2.1	4.3
Very Conservative	4.4	2.6		2.1	
Political Affiliation					
Strongly Democratic	16.1	33.3	25.0	4.3	21.7
Moderately Democratic	35.0	33.3	25.0	21.3	26.1
Slightly Democratic	13.9	15.4	25.0	14.9	13.0
Neither Democratic Nor Republican	16.8	15.4		44.7	30.4
Slightly Republican	8.8		25.0	10.6	4.3
Moderately Republican	6.6	2.6		4.3	4.3
Strongly Republican	2.9				

Facial Stimuli

Facial stimuli presented were the same as used in Stepanova et al. (2008) and featured faces created with Poser 6™ software. The faces were designed to be equivalent for affective expressions but to vary systematically in skin color and facial physiognomy (see Figure 6 for sample stimuli). Skin color varied from very light to very dark (10 levels) and facial physiognomy varied from very Afrocentric to very Eurocentric (10 levels). Note that the facial physiognomy manipulation encompassed several phenotypic characteristics (e.g., width of the nose, fullness of the lips, bone structure, etc.). A set of those characteristics was manipulated simultaneously in the Poser 6™ software using a control that globally modified the faces to make them "less/more African" (for European faces) or "less/more European" (for African faces). Two sets of stimuli were used to insure generalizability, with each set consisting of 100 faces. Both sets were pre-tested and matched on several characteristics by 321 Washington University students recruited from the Psychology Department Human Subject Pool. These participants rated the facial stimuli on the following 9-point scales: attractiveness (from *1=not at all attractive* to *9=very attractive*), racial typicality (from *1=very African American* to *9=very European American*), happiness (from *1=not at all happy* to *9=very happy*), anger (from *1=not at all angry* to *9=very angry*), and sadness (from *1=not at all sad* to *9=very sad*). Both sets (1 and 2) received very similar ratings on these characteristics: attractiveness (M_1=4.83, M_2=4.71), racial typicality (M_1=4.80, M_2=4.75), happiness (M_1=4.65, M_2=4.30), anger (M_1=4.28, M_2=4.64) and sadness (M_1=4.70, M_2=4.83). Each of the sets was constructed by beginning with one original face, with a neutral affective expression and unique facial features that then were manipulated by the software to produce the

other faces in the set. For one of the sets, I took an Afrocentric face as the starting point and for the other set I took a Eurocentric face as the starting point. A subset of these same stimuli representing light-skinned individuals with Eurocentric features and dark-skinned individuals with Afrocentric features was also used in the IAT (see Figure 6).

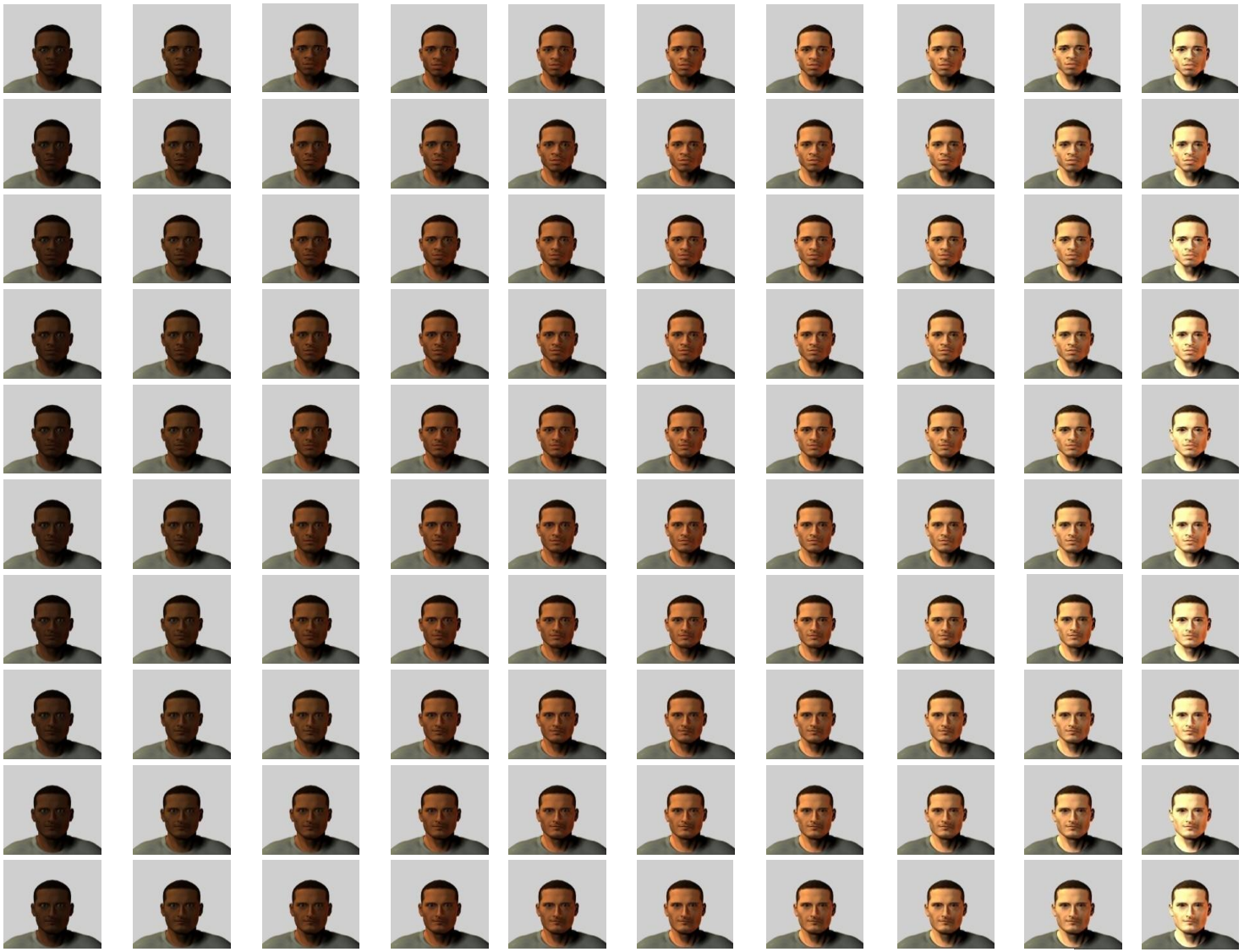

Figure 6. Facial stimuli used in Stepanova et al. (2008) and current research (a sample set).

Explicit Measures

Symbolic Racism. The Symbolic Racism Scale (Henry & Sears, 2002) contains 8 items that measure explicit racial attitudes (i.e., anti-Black racism; see Appendix A). A sample item is: *It's really a matter of some people not trying hard enough; if blacks would only try harder, they could do just as well as whites.* This scale does not employ a consistent response scale across items. The majority of items are rated on a 4-point scale but one item uses a 3-point scale. The scale is internally consistent, with Cronbach's αs ranging from .59 to .86. It assesses two highly correlated factors (i.e., individual versus structural attributions for Blacks' disadvantage; Henry & Sears, 2002; Tarman & Sears, 2005) but is generally used as a single-dimension construct. The scale possesses good construct validity, predictive validity and discriminant validity, indicating that symbolic racism is a "blend" of conservative values and racial antipathy, rather than just a combination of these two additive parts (see Henry & Sears, 2002).

Social Dominance. The Social Dominance Orientation Scale (Pratto, Sidanius, Stallworth, & Malle, 1994) contains 14 items measuring endorsement of societal hierarchy (i.e., social dominance orientation; see Appendix B). A sample item is: *Some people are just more worthy than others.* Each item is rated on a scale ranging from 1 (*Very Negative*) to 7 (*Very Positive*). The scale has high internal consistency (coefficient $\alpha = .90$), high test-retest reliability (when participants are tested in a 3-month period, *rs* ranging from .81 to .84), and high construct and discriminant validity (Pratto et al., 1994). The scale has been found to be significantly related to negative racial attitudes (e.g., Lowery, Hardin & Sinclair, 2001; Pratto et al., 1994), including, but not limited to, biological racism, symbolic racism, ethnocentrism, and aversive racism (e.g., Van Hiel &

Mervielde, 2005). It predicts well a host of preferences for hierarchical roles, and a variety of socio-political ideologies promoting group inequalities and support for policies promoting social inequality (Pratto et al., 1994).

Modern Racism. The Modern Racism Scale (McConahay, Hardee, & Batts, 1981) contains 7 items that measure explicit racial prejudice toward Blacks (see Appendix C). A sample item is: *Over the past few years, the government and news media have shown more respect to blacks than they deserve.* Each item is rated on a scale ranging from 1 (*Strongly Disagree*) to 7 (*Strongly Agree*). The scale has acceptable internal consistency, with Cronbach's αs varying from. 81 to .86 (Biernat & Crandall, 1999), and high tests-retest reliability (when participants are tested over a six week period, *r*s ranging from .87 to .93). Although this scale is widely used, critics have identified important measurement problems (see Henry & Sears, 2002, for review; see also Migetz, 2004). For example, it was constructed over 25 years ago, and some of the items might not have high relevance now. I included this scale because it has been used often in past research, but I included the Symbolic Racism and Social Dominance Orientation scales as well to insure adequate measurement of explicit racial attitudes.

Feeling Thermometers. Participants were asked to indicate how favorably they viewed different social and ethnic groups (see Appendix D). I included these measures in part to replicate the methodology used by Stepanova et al. (2008) and because these are the most explicit measures of racial affect--participants are asked directly how warm or cold they feel towards a variety of groups. These measures are known for exacerbating participants' tendency to express their attitudes in consistently negative or positive ways (i.e., give extreme ratings) (e.g., Wilcox, Seligman & Cook, 1989).

Implicit Measure

Implicit Association Test (IAT). The IAT task consists of seven blocks of trials (Greenwald et al., 2003). In the first block, participants are presented with targets (faces) and asked to categorize them as "European American" or "African American". Participants do so by pressing corresponding keys on the keyboard, with one category assigned to a response by one hand and the other assigned to a response by the other hand. In the second block, participants are asked to categorize words (e.g., Joy, Wonderful, Pleasure, Happy, Love, Terrible, Horrible, Evil, Awful, Agony) as being either "good" or "bad." In the subsequent third and fourth blocks, participants are presented with both of the categorization tasks simultaneously: They are presented with words and faces alternating on different trials, and are asked to press one key on the keyboard when the target is a pleasant word or an African American face and to press another key when the target is an unpleasant word or a European American face. The fifth block is analogous to the first block, but switches the side corresponding to a particular racial category. In the sixth and seventh blocks participants are presented with both words and faces again, but the pairing of the stimuli is the opposite of that used on blocks three and four. That is, participants are asked to press one key on the keyboard when the target is a pleasant word or a European American face and to press another key when a target is an unpleasant word or an African American face. For half of the participants (determined randomly), the positions of blocks 1, 3, and 4 were switched with blocks 5, 6, and 7 correspondingly. The side on which the key presses were required for "good" versus "bad" words and "African American" versus "European American" faces was likewise determined randomly for each participant. Facial stimuli

employed in this task were a subset of faces used during the racial typicality ratings task and included 5 faces with high Afrocentric physiognomy and dark skin color and 5 faces with high Eurocentric physiognomy and light skin color. The IAT score was derived according to the procedures described by Greenwald et al. (2003) and represents a standardized response time difference. Higher scores indicate more favorable implicit attitudes toward Whites compared to Blacks.

Note that the IAT measures differences between two target concepts (e.g., African American and European American) rather than differences between exemplars' of two target concepts (De Houwer, 2001) and procedurally asks participants to categorize those two concepts. Thus, the concepts assessed with the IAT are explicitly available to participants, but the implicit prejudice measured with the IAT is not sensitive to exemplar typicality because it measures associations to category labels (see Olson and Fazio, 2003).

Outcome Measure

Racial Typicality Ratings Task. Participants were asked to rate 100 facial stimuli on the dimension of racial typicality (a continuum of Afrocentric-Eurocentric typicality) using a 6-point scale: 1 (*Very African American*), 2 (*Moderately African American*), 3 (*Somewhat African American*), 4 (*Somewhat Caucasian*), 5 (*Moderately Caucasian*) and 6 (*Very Caucasian*). The choice of the one-dimensional typicality scale is substantiated by previous research.[3] Participants rated faces presented in a random order. The particular set of faces (from 2 sets) presented to each participant in the racial typicality task was randomly determined.

The racial typicality ratings task consisted of three separate blocks. In the first

block, participants performed the task without any response deadline and were given the following instructions:

> For this task, you will rate faces according to how representative they are of two racial groups (i.e., Caucasian and African American). Some of the faces will not fit neatly into one racial group or the other. Examine each face carefully and then decide how African American or Caucasian the face looks. Use any standards you like when making this judgment. Then choose a number between 1 (Very African American) and 6 (Very Caucasian) to indicate your judgment and press the appropriate key on the keyboard. The next face will appear automatically. Take as much time as you need to make your judgments. Do not rush through so you can get done quickly! You will find the task easier to perform if you position the middle three fingers of each hand over the numbers 1 through 6 on the top row of the keyboard. When you are ready to begin, press "Continue".

Most cognitive categorization research has employed only dichotomous judgments and short response deadlines when studying categorization under time pressure (600-1600 ms, Lamberts, 1995; Lamberts, 2000). Because I used a multiple-category response scale, it was not initially clear what response deadline should be employed to limit cognitive resources. To determine response deadlines, I examined how fast participants performed the racial typicality rating task without time pressure in the first block. In the second block I took each participant's median response time from the first block and made it the deadline for the second block. In the third block, I took the 25th percentile reaction time from the first block as an even shorter deadline. Therefore, in blocks two and three, I required participants to respond more quickly than they

normally would on 50% of the trials and 75% of the trials in block 1 (as suggested by K. Lamberts, personal communication, February 25, 2009).

In the second and third blocks of the racial typicality task, participants were given the following instructions:

> Again, for this task you will rate the faces according to how representative they are of the two racial groups (i.e., Caucasian and African American). Some of the faces will not fit neatly into one racial group or the other. This time, however, you will have to make a decision about each face as fast as possible because you will be given a limited time to rate each face. If you take longer than allowed, you will receive a warning ("Too Slow! Go Faster!"). Examine each face carefully and then quickly decide how much African American or Caucasian the face looks. Use any standard you like when making this decision. Choose a number between 1 (*Very African American*) and 6 (*Very Caucasian*) and press the appropriate key on the keyboard. The next face will appear automatically. You will find the task easier to perform if you position the middle three fingers of each hand over the numbers 1 through 6 on the top row of the keyboard. When you are ready to begin, press 'Continue'.

In all three blocks, after a participant made a decision on each trail, there was a blank screen and brief interval of 1.5 s. between displays of faces, so that when two adjacent faces were similar a participant would know that a new face was displayed. Also, in all three blocks, if participants responded in less than 250 ms, a warning was issued: "You are responding too quickly to be giving any though to your judgments. Please take enough time to provide a careful judgment. Press OK to continue." After the

subject pressed the "Ok" button to clear the warning message, a blank screen appeared for 1.5 seconds and then the same face was shown again.

Procedure

The study was conducted on computers and took approximately one hour. At the beginning of the experimental session, participants received the following instructions:

> During this session you will be asked to rate the ethnic typicality of various faces, perform a word-face classification task, and fill out several questionnaires assessing attitudes towards several groups and demographic information. Each task will be preceded by a set of instructions. Please read the instructions carefully and then complete each task as honestly as possible.

Participants were randomly assigned to one of two task order conditions. Half of the participants performed the racial typicality ratings of the faces first, then the IAT (Greenwald, McGhee & Schwartz, 1998; Greenwald, Nosek & Banaji, 2003), and then completed the explicit individual difference measures. The other half of the participants performed the IAT first, then the racial typicality ratings, and then the explicit measures. The explicit measures were collected last so that they would not sensitize participants or produce inadvertent priming effects.

After completing the IAT and racial typicality ratings, participants were asked to complete the explicit racial attitude measures: Symbolic Racism Scale (Henry & Sears, 2002), Social Dominance Scale (Pratto et al., 1994), Modern Racism Scale (McConahay, 1986), and Feeling Thermometers. These questionnaires were randomly ordered for each participant. When completing explicit measures, if participants responded too quickly (< 500 ms) they were warned: "Please take your time and make the judgments carefully.

Press 'OK' to continue with the next statement." If participants took more than 10 s to respond, they were likewise warned: "There is no need to take so long to make each judgment. Your first impression is sufficient. Press 'OK' to continue with the next statement." At the end of the study, participants were asked to provide demographic information (see Appendix E).

Results

Overview of Analyses

First, I tested all stated hypotheses. I examined the racial typicality ratings in a 10 (Skin Color) x 10 (Facial Physiognomy) x 3 (Response Deadline) x (Implicit Racial Attitudes) repeated measures multiple regression. Skin color, facial physiognomy and response deadline were treated as repeated measures.

Second, I explored the additional moderating influence of participant ethnicity by including ethnicity as a between-subjects predictor in a 10 (Skin tone) x 10 (Facial Physiognomy) x 3 (Response Deadline) x (Implicit Racial Attitudes) x (Participants' Ethnicity) repeated measures multiple regression. The interaction of participants' ethnicity and implicit racial attitudes was examined by including their product in the model.[4]

Third, I explored the moderating effects of explicit racial attitudes by including, in separate analyses, one of the racial attitude questionnaire composites as a between-subjects predictor in a 10 (Skin tone) x 10 (Facial Physiognomy) x 3 (Response Deadline) x (Implicit Ethnic Attitudes) x (Explicit Attitude Measure) repeated measures multiple regression. These analyses included only one questionnaire measure (e.g., Social Dominance Orientation [SDO], Modern Racism Scale [MRS], Symbolic Racism Scale [SRS], and Feeling Thermometers for Blacks [FT-B] or Whites [FT-W]) at a time. In models with two between-subjects predictors (e.g., implicit and explicit racial attitudes), the interaction of the two predictors was tested by entering their product. These analyses revealed many duplicate effects (e.g., Physiognomy x SDO and Physiognomy x MRS, Physiognomy x D x MRS and Physiognomy x D x SRS) involving three of the following

measures: SDO, MRS, and SRS scales. In fact, there were no unique effects that were not lower order effects qualified by duplicate interactions. In light of these findings, I formed a single composite variable based on a principal components analysis of the SDO, MRS, and SRS. Specifically, one principle component was extracted; all three scales loaded highly on it (MRS = .88; SDO =.81, and SRS = .88). This composite (the principal component score) is referred to as the Explicit Racial Attitudes index. For the sake of brevity, I only present analyses using this index. Thus I conducted a 10 (Skin Color) x 10 (Facial Physiognomy) x 3 (Response Deadline Condition) x (Implicit Ethnic Attitudes) x (Explicit Racial Attitudes Index) analysis. The product of the IAT score and explicit attitudes index was entered to examine the interaction of these two between-subjects predictors.

Fourth, I explored the joint effect of explicit racial attitudes and participants' ethnicity on racial typicality evaluations. In these analyses, I examined racial typicality ratings in a 10 (Skin Color) x 10 (Facial Physiognomy) x 3 (Response Deadline Condition) x (Participants' Ethnicity) x (Explicit Racial Attitudes) repeated measures multiple regression. The product of participant ethnicity and explicit racial attitudes was included to test their interaction.

More complex statistical designs were conducted (e.g., including implicit attitudes, an explicit attitude measure, and ethnicity), but they produced few effects of substantive interest or effects of such complexity (e.g., five-way interactions) that they defied interpretation. For the sake of brevity, those analyses are not reported here.

In all analyses, when the assumption of sphericity was violated (as indicated by Mauchly's test of sphericity), I used the Greenhouse-Geisser correction for *F* values and

associated degrees of freedom. All follow-up comparisons were conducted with Bonferroni corrections to control inflation of the Type I error rate.

Preliminary Analyses

Table 2 provides means, standard deviations, reliabilities, and correlations for the full sample (N = 250) for each of the measures (SDO, MRS, SRS, FT-B, FT-W, and *D*, the IAT score) used in the analyses to follow. The results show the expected high intercorrelations between SDO, MRS, and SRS. The Feeling Thermometer for Blacks was negatively correlated with the Modern Racism Scale, the Social Dominance Orientation Scale, and the Symbolic Racism Scale; that is, the higher were participants' ratings on the modern racism, the symbolic racism, and the social dominance orientation scales, the lower were their scores on the Feeling Thermometer for Blacks, indicating less positive attitudes towards Blacks.

Additionally, participants' scores on the Feeling Thermometers for Blacks and Whites were also positively correlated; the more positive attitudes participants expressed towards Whites, the more positive attitudes they expressed towards Blacks. There were two significant correlations between explicit and implicit measures: a positive correlation between the Symbolic Racism Scale and *D* (the IAT score) and a positive correlation between the Feeling Thermometer for Whites and *D*. Additionally, there was a significant negative correlation between Feeling Thermometers for Blacks and *D*.

Table 2

Interrcorrelations, Means, Standard Deviations and Scale Reliabilities of Individual Differences Variables for the Entire Sample.

Measure	1	2	3	4	5	6	Mean	SD	α
1. Modern Racism Scale	1.00						16.26	6.20	.80
2. Social Dominance Orientation Scale	.55**	1.00					32.69	12.98	.89
3. Symbolic Racism Scale	.68**	.56**	1.00				2.68	1.22	.73
4. Feeling Thermometer for Blacks	-.23**	-.23**	-.25**	1.00			5.14	1.24	
5. Feeling Thermometer for Whites	.03	.11	.12	.19**	1.00		5.37	1.25	
6. *D*, the IAT score	.09	.11	.13*	-.23**	.22**	1.00	.62	.60	

*$p < .01$. **$p < .05$.

These results are consistent with past research that has shown that there is a complex relationship between implicit and explicit measures of racial prejudice (Dovidio, Kawakami, Smoak, & Gaertner, 2009). The internal consistency reliabilities for the three explicit scales are acceptable and similar to those that have been previously reported.

Descriptive statistics calculated separately for different ethnicity groups are presented in Table 3. The pattern of intercorrelations between all three explicit scales was similar among participants of all ethnic groups (e.g., positive correlations among all three scales). In general, the pattern of intercorrelations for other scales was similar in all groups, with differences in statistical significance largely reflecting the substantial differences in sample size. A few isolated correlations differed between groups. As would be expected, African American participants exhibited the lowest average *D* score (*M* = .07, *SD* = .59) and European American participants exhibited the highest average *D* score (*M* = .81, *SD* = .53).

Table 3

Interrcorrelations, Means, and Standard Deviations of Individual Differences Variables for European American African American, Asian American and Other Participants

Participants' Ethnicity	European American (N = 137)								African American (N = 39)							
Measure	1	2	3	4	5	6	Mean	SD	1	2	3	4	5	6	Mean	SD
1. Modern Racism Scale	1.00						16. 02	6.13	1.00						13.80	5.39
2. Social Dominance Orientation Scale	.57**	1.00					32.53	12.57	.64**	1.00					29.92	13.10
3. Symbolic Racism Scale	.73**	.64**	1.00				2.61	1.23	.47**	.46**	1.00				2.28	1.17
4. Feeling Thermometer for Blacks	-.20*	-.25**	-.22**	1.00			4.95	1.16	-.30	-.25	-.22	1.00			6.08	1.22
5. Feeling Thermometer for Whites	.21*	.32**	.28**	.33**	1.00		5.60	1.15	-.19	-.08	-.04	.06	1.00		4.79	1.61
6. *D*, the IAT score	.15	.09	.12	-.06	.05	1.00	.81	.53	.002	.21	.18	-.14	.47**	1.00	.07	.59

**$p < .01$. *$p < .05$.

Participants' Ethnicity		Asian American (N = 47)								Other (N = 23)							
Measure		1	2	3	4	5	6	Mean	SD	1	2	3	4	5	6	Mean	SD
1.	Modern Racism Scale	1.00						17. 98	5.91	1.00						18.96	7.00
2.	Social Dominance Orientation Scale	.50**	1.00					36.77	13.84	.42**	1.00					31.43	12.99
3.	Symbolic Racism Scale	.65**	.47**	1.00				2.93	1.06	.77**	.52**	1.00				3.12	1.36
4.	Feeling Thermometer for Blacks	-.12	-.12	-.38**	1.00			4.91	1.27	-.11	-.04	-.04	1.00			5.13	1.10
5.	Feeling Thermometer for Whites	-.11	-.10	-.08	.31**	1.00		5.26	1.15	-.30	-.23	-.14	.38	1.00		5.13	1.04
6.	*D*, the IAT score	.10	-.04	.17	-.18	.05	1.00	.60	.51	-.31	.22	-.09	-.23	.-01	1.00	.45	.55

**$p < .01$. *$p < .05$.

Note. Data for Hispanic American participants are not reported in this table due to a very low sample size (n = 4).

Tests of Stated Hypotheses

Racial typicality ratings were analyzed in a 10 (Skin Color) x 10 (Facial Physiognomy) x 3 (Response Deadline Condition) x (Implicit Racial Attitudes) repeated measures multiple regression, with the last factor a between-subjects continuous variable. The analysis was collapsed across ethnicity of participants (separate analyses with ethnicity of participants as a between-subjects factor will be presented later). The analysis revealed a significant main effect for skin color, Greenhouse-Geisser $F(1.46, 361.56) = 230.46$, $p < .001$, $\eta_p^2 = .48$ and a significant main effect for facial physiognomy, Greenhouse-Geisser $F(1.34, 331.31) = 248.26$, $p < .001$, $\eta_p^2 = .50$. As predicted by Hypothesis 1, darker faces were rated as more African American than lighter faces, and faces with Afrocentric physiognomy were rated as more African American than faces with Eurocentric facial physiognomy (see Table 4).

Table 4

Means and Standard Errors for Racial Typicality Ratings as a Function of Skin Color and Facial Physiognomy

Skin Color	Mean	Std. Error	Facial Physiognomy	Mean	Std. Error
1 (dark)	1.98	.05	1 (Afrocentric)	2.20	.04
2	2.23	.05	2	2.35	.04
3	2.52	.04	3	2.52	.04
4	2.86	.04	4	2.78	.03
5	3.18	.03	5	3.14	.03
6	3.45	.03	6	3.48	.04
7	3.64	.03	7	3.74	.04
8	3.83	.03	8	3.88	.04
9	4.04	.04	9	3.97	.05
10 (light)	4.42	.05	10 (Eurocentric)	4.09	.05

Note. Skin color levels varied from dark (1) to light (10) and facial physiognomy levels varied from Afrocentric (1) to Eurocentric (10). Racial typicality ratings could range from 1 (*Very African American*) to 6 (*Very Caucasian*). Within columns, every mean is significantly different from every other mean ($p < .05$, Bonferroni adjustment).

As predicted by Hypothesis 2, a Skin Color x Facial Physiognomy interaction emerged as well, Greenhouse-Geisser $F(48.61, 12053.93) = 7.73$, $MSE = 7.09$, $p < .001$, $\eta_p^2 = .03$. Figure 7 illustrates this interaction and Table 5 provides mean differences between adjacent facial physiognomy levels (e.g., Facial Physiognomy 1 and Facial Physiognomy 2, Facial Physiognomy 2 and Facial Physiognomy 3) within each level of skin color. As predicted, when skin color was darker, participants' ratings were less variable, but as skin tone became lighter, racial typicality ratings depended more strongly on facial physiognomy.

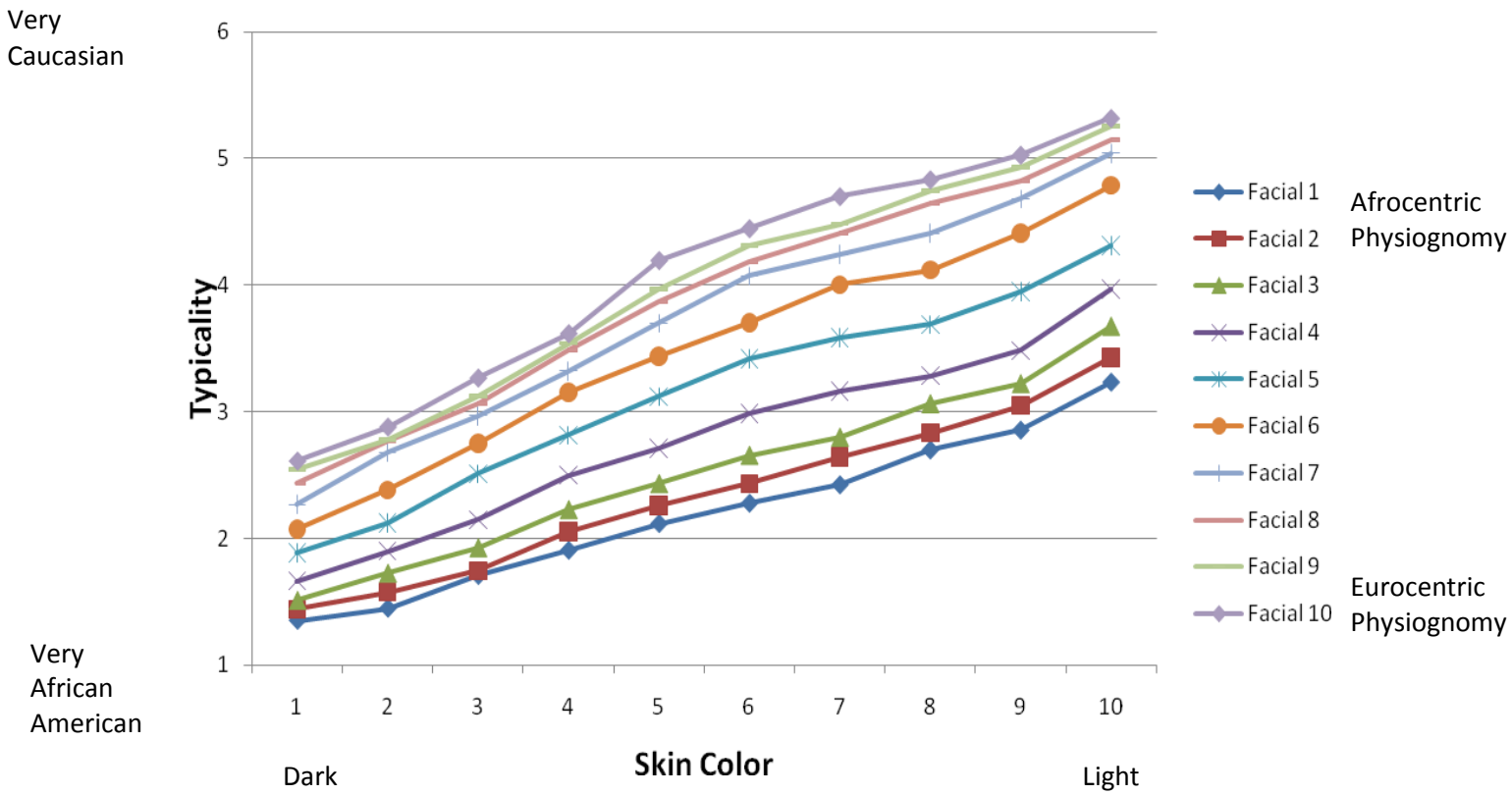

Figure 7. The effects of skin color and facial physiognomy on racial typicality judgments (entire sample). Note: Facial 1 = highest Afrocentric physiognomy and Facial 10 = highest Eurocentric physiognomy.

Table 5

Mean Differences Between Adjacent Levels of Facial Physiognomy Within Each Level of Skin Color

	Adjacent Facial Physiognomy Levels								
Skin Color	1-2	2-3	3-4	4-5	5-6	6-7	7-8	8-9	9-10
1 (Darkest)	-.09	-.07	-.15*	-.22*	-.19*	-.19*	-.17*	-.11	-.07
2	-.13*	-.15*	-.18*	-.22*	-.27*	-.30*	-.09	-.02	-.10
3	-.04	-.18*	-.22*	-.36*	-.24*	-.22*	-.09*	-.06	-.14
4	-.15	-.17*	-.27*	-.32*	-.34*	-.17*	-.17*	-.05	-.08
5	-.14*	-.17*	-.28*	-.41*	-.32*	-.26*	-.17*	-.10	-.23*
6	-.16*	-.23*	-.33*	-.44*	-.28*	-.37*	-.11	-.13	-.13
7	-.22*	-.16*	-.36*	-.42*	-.42*	-.24*	-.16	-.08	-.22*
8	-.14*	-.23*	-.22*	-.40*	-.43*	-.29*	-.24*	-.10	-.09
9	-.19*	-.17*	-.26*	-.47*	-.47*	-.27*	-.14*	-.11	-.10
10 (Lightest)	-.19*	-.25*	-.29*	-.34*	-.48*	-.25*	-.11	-.10	-.07

Note. *$p < .05$. Bonferroni adjustment for multiple comparisons is applied within each level of skin color.

Hypothesis 3 predicted that participants with higher implicit ethnic prejudice would make finer distinctions when judging racial typicality in comparison to individuals with lower implicit ethnic prejudice. Contrary to this prediction, the Skin Color x Facial Physiognomy x Implicit Racial Attitudes interaction did not reach significance, Greenhouse-Geisser $F(48.61, 12053.93) = 1.10$, $p = .302$. However, there was a significant Skin Color x Implicit Racial Attitudes interaction, Greenhouse-Geisser $F(1.46, 361.56) = 3.98$, $MSE = 210.85$, $p = .03$, $\eta_p^2 = .02$. Specifically, participants with higher implicit racial prejudice rated faces as less Caucasian (and more African American) than participants with lower implicit racial attitudes for darker levels of skin color; that pattern reversed for lighter levels of skin color (see Figure 8). Stated differently, skin color was more related to typicality ratings for participants with higher implicit racial prejudice than for participants with lower implicit prejudice. These results are consistent with Hypothesis 3 in showing that participants with higher implicit racial prejudice rely more on skin color in racial typicality judgments than participants with lower implicit racial prejudice.

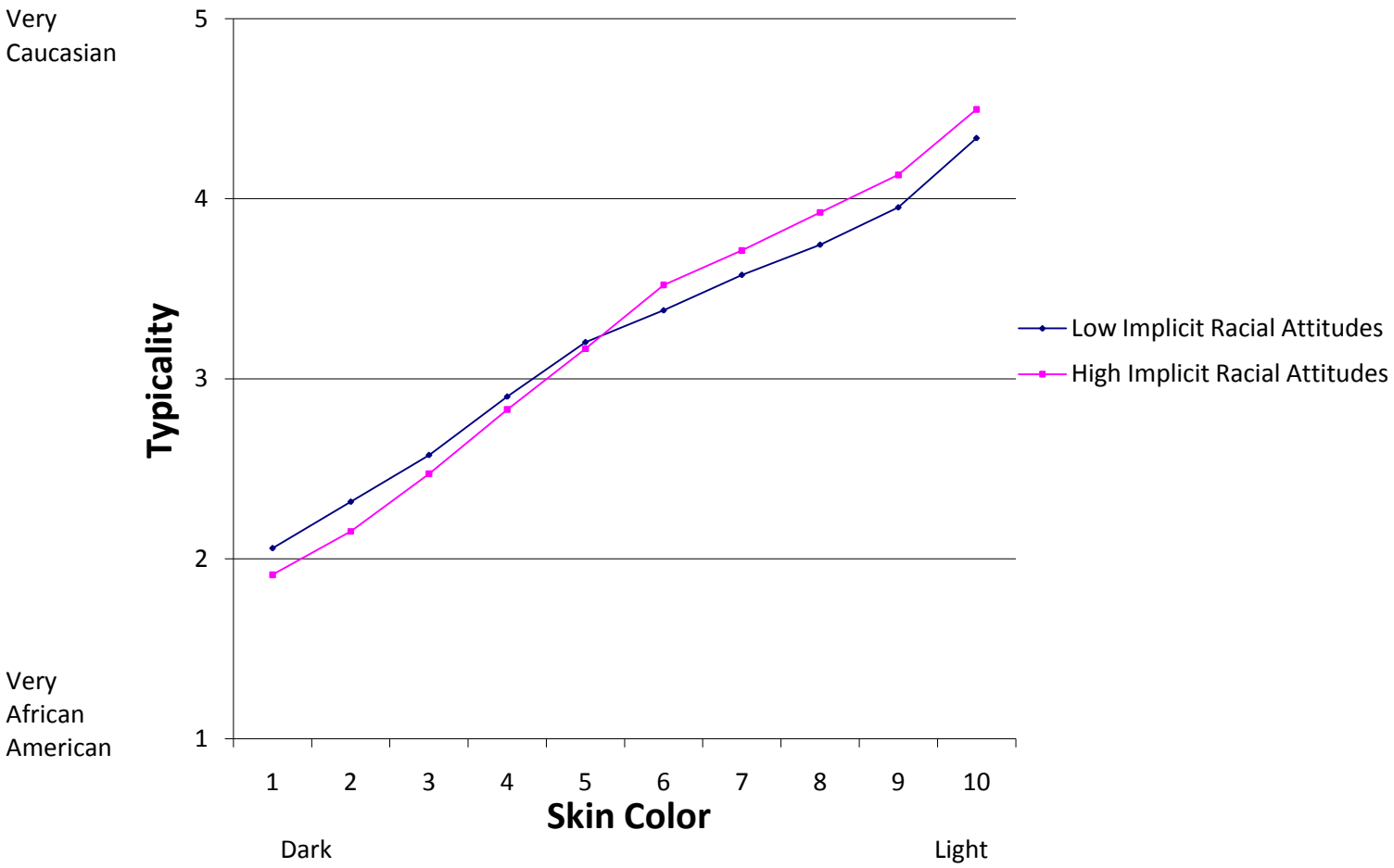

Figure 8. The effects of skin color and implicit racial attitudes on racial typicality judgments (entire sample).

Hypothesis 4 predicted that, compared to no-time-constraint trials, time constraints (speeded categorization) would produce categorization decisions that are more affected by skin tone than by facial physiognomy. This would be revealed in interactions involving the response deadline variable and skin tone. Analyses revealed a Response Deadline x Skin Color interaction, (Greenhouse-Geisser F[7.70, 1910.71] = 8.10, p <. 001, η_p^2 = .032) that was qualified by a Response Deadline x Skin Color x Facial Physiognomy interaction, Greenhouse-Geisser F(82.73, 20516.64) = 1.29, p = .041, η_p^2 =.005. To determine the nature of this interaction, the Skin Color x Facial Physiognomy interaction was examined separately for each of the Response Deadline conditions (i.e., No Response Deadline condition, Median Response Deadline condition, and 25th Percentile Response Deadline condition). In the No Response Deadline condition, there was a significant Skin Color x Facial Physiognomy interaction, Greenhouse-Geisser F(50.42, 12503.56) = 4.82, p < 0.001, η_p^2 =.019. In the Median Response Deadline condition, there was also a significant Skin Color x Facial Physiognomy interaction, Greenhouse-Geisser F(52.02,12899.95) = 3. 37, p < 0.001, η_p^2 =.013, although it was less significant and accounted for less variance than in the No Response Deadline condition. In 25th Percentile Response Deadline condition, there was also a significant Skin Color x Facial Physiognomy interaction, F(52.63, 13050.95) = 2.94, p < 0.001, η_p^2 = .012 of a magnitude similar to the Median Response Deadline condition. The separate Skin Color x Facial Physiognomy interactions for each Response Deadline condition are illustrated in Figures 9, 10, and 11.

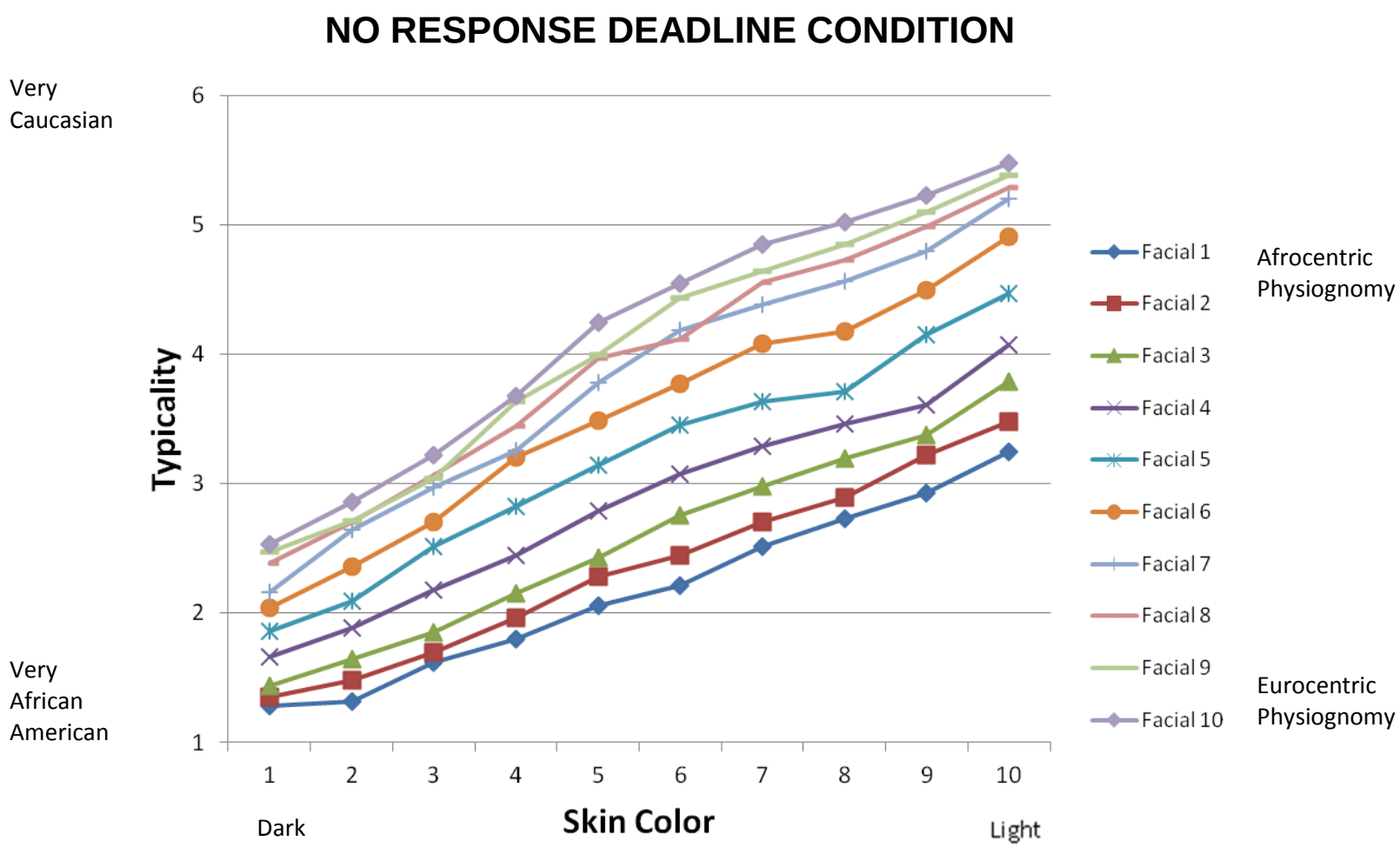

Figure 9. The effect of skin color and facial physiognomy on racial typicality judgments in the No Response Deadline condition.

Note: Facial 1 = highest Afrocentric physiognomy and Facial 10 = highest Eurocentric physiognomy.

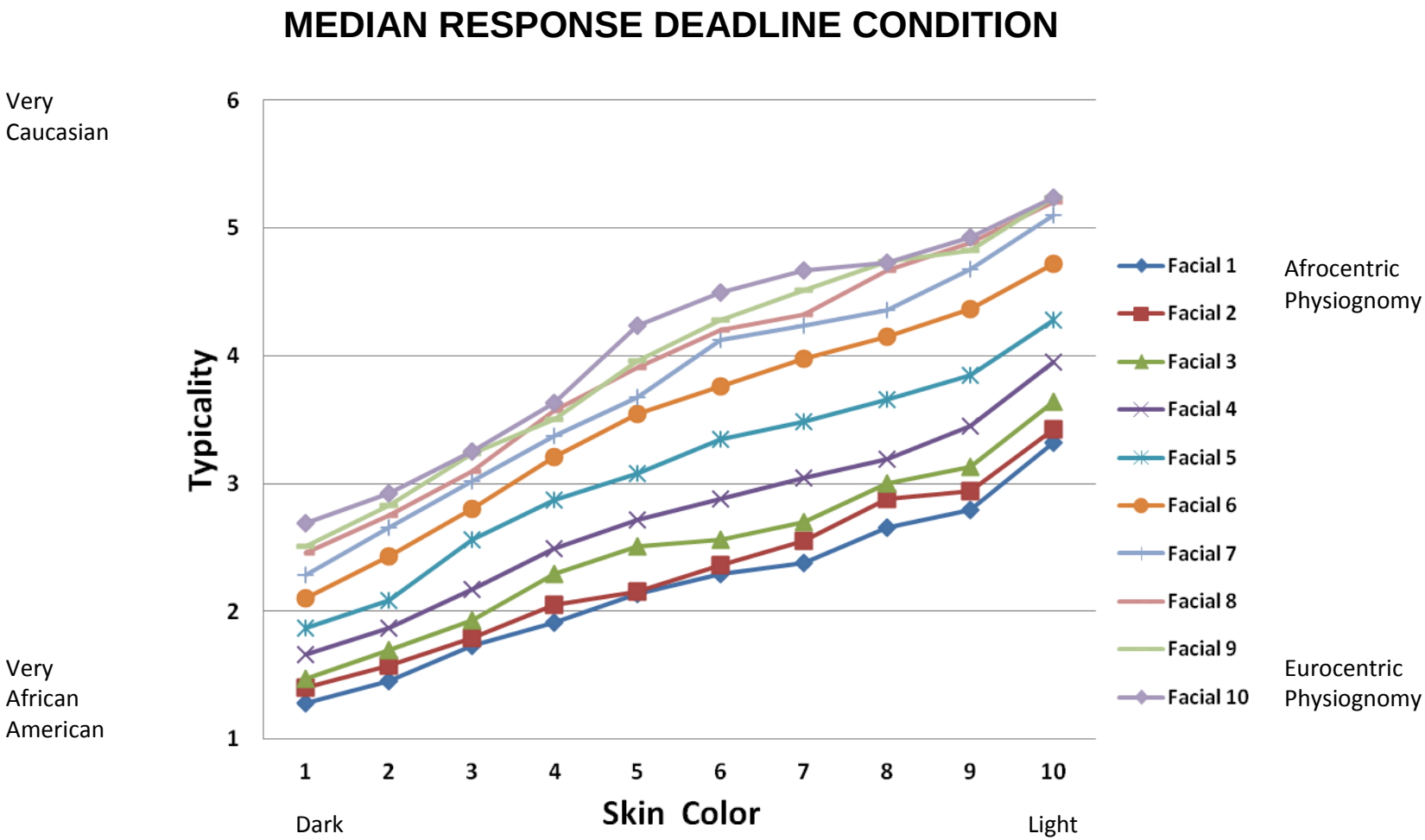

Figure 10. The effect of skin color and facial physiognomy on racial typicality judgments in the Median Response Deadline condition.

Note: Facial 1 = highest Afrocentric physiognomy and Facial 10 = highest Eurocentric physiognomy.

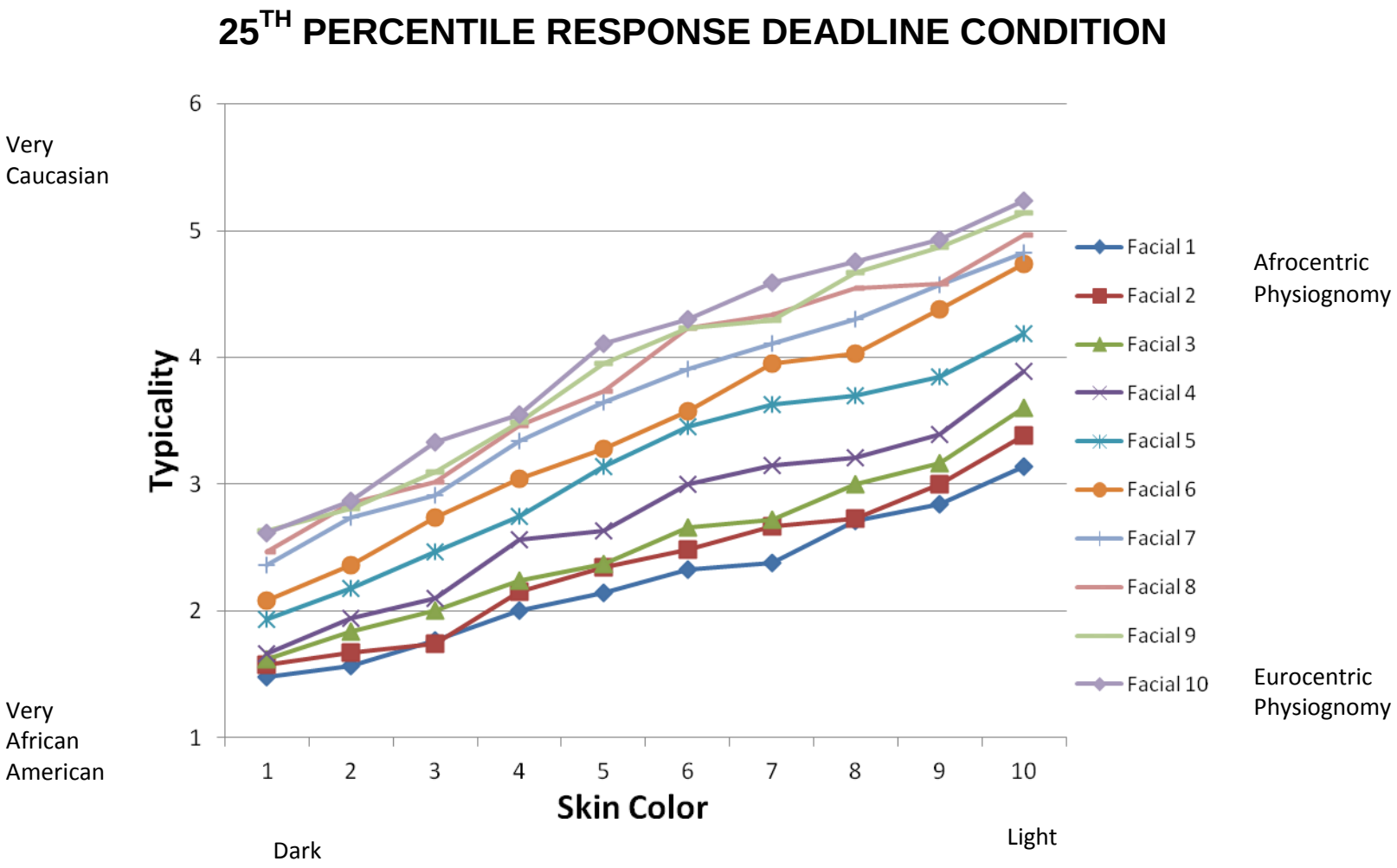

Figure 11. The effect of skin color and facial physiognomy on racial typicality judgments in the 25th Percentile Response Deadline condition. Note: Facial 1 = highest Afrocentric physiognomy and Facial 10 = highest Eurocentric physiognomy.

In general, Figures 9, 10, and 11 provide support for Hypothesis 4 in that the effect of facial physiognomy diminishes under time constraint. To further explore this interaction, I examined the pair-wise comparisons of each physiognomy level within each level of skin color in each response deadline condition. The majority of these pair-wise comparisons was significant and so highlighting the comparisons that were *not significant* is more useful in that these indicate where typicality ratings were particularly low in variability. Of note are the following patterns: (a) overall, in the No Response Deadline condition, there were fewer non-significant pair-wise comparisons (60 out of 450) than in the Median Response Deadline (85 out of 450) or 25th Percentile Response Deadline conditions (95 out of 450), (b) in the Median Response Deadline and 25th Percentile Response Deadline conditions, there were more non-significant pair-wise comparisons in nonadjacent physiognomy levels (e.g., 7 and 9) (22 and 26 correspondingly) than in the No Response Deadline condition (8), and that pattern was mostly present at extreme levels of facial physiognomy (e.g., low and high), and (c) in the Median Response Deadline and 25th Percentile Response Deadline conditions, there were more non-significant pair-wise comparisons than in the No Response deadline condition for the two most extreme dark and two most extreme light skin color levels (38 and 49 versus 28).

This interaction suggests that participants' ability to attend to both skin color and facial physiognomy when rating faces of extreme levels of physiognomy declines under time constraints. Participants in the No Response Deadline condition responded to all levels of facial physiognomy in a more orderly and pronounced fashion at all skin color levels than in the Median Response Deadline and 25th Percentile Response Deadline

conditions. When cognitive resources were limited, participants primarily relied upon skin color when rating faces of extreme Eurocentric and Afrocentric physiognomy. For faces with middle levels of physiognomy, racial categorization was driven by both skin color and facial physiognomy, although the influence of skin color was less orderly in the presence of the most severe time constraints.

Hypothesis 5 predicted that, under time constraints, the effect described in Hypothesis 4 would be even more prominent in participants with higher implicit racial prejudice than in participants with lower implicit racial prejudice. However, the Response Deadline x Skin Color x Facial Physiognomy x Implicit Racial Attitudes interaction did not achieve significance, Greenhouse-Geisser $F(82.73, 20516.64) = 1.17$, $p = .14$.

Exploratory Analyses: Ethnicity as Moderator

Because I was able to recruit a number of non-White participants, I examined participants' ethnicity as a between-participants factor in an exploratory analysis. There was a limited number of Hispanic participants ($N = 4$) so I excluded that group from the analyses. Data from the following ethnic groups were included in the analysis: European Americans ($N = 127$), African Americans ($N = 39$), Asian Americans ($N = 47$), and Other ($N = 23$).

Racial typicality ratings were analyzed in a 10 (Skin Color) x 10 (Facial Physiognomy) x 3 (Response Deadline Condition) x (Implicit Racial Attitudes) x 4 (Participants' Ethnicity) repeated measures multiple regression that mirrored the previously reported tests for the stated hypotheses. Skin color and facial physiognomy were treated as repeated measures, implicit racial attitudes were treated as a continuous between-participants predictor, and Participants' Ethnicity was treated as a categorical between-

participants factor. To examine the interaction of Implicit Racial Attitudes and Participants' Ethnicity, their product was entered in the model. Due to the exploratory nature of this analysis, I report only effects significant at $p < .01$ and interactions only up to third-order inclusively. Because of the duplication of effects from previously described analyses, only new effects involving Participant Ethnicity are described.

A significant Skin Color x Participants' Ethnicity interaction emerged, Greenhouse-Geisser $F(4.48, 360.37) = 4.87$, $p < .001$, $\eta_p^2 = .06$. Follow-up tests examined ethnicity group differences within each skin color level. Significant differences were found at skin color levels 1, 2, 3, 4, 5, and 9 (all Fs > 3.41, ps < .018). Pairwise comparisons showed that in the darker skin color levels (1, 2, 3, 4, and 5), African American participants rated faces as more Caucasian (less African American) than did European American and Asian American participants. However, that pattern reversed for lighter skin colors (9 and 10, although pairwise comparisons did not reach significance for Skin Color level 10) (see Figure 12). Stated differently, the relationship between skin color and racial typicality ratings was stronger for European American and Asian American participants than for African American participants.

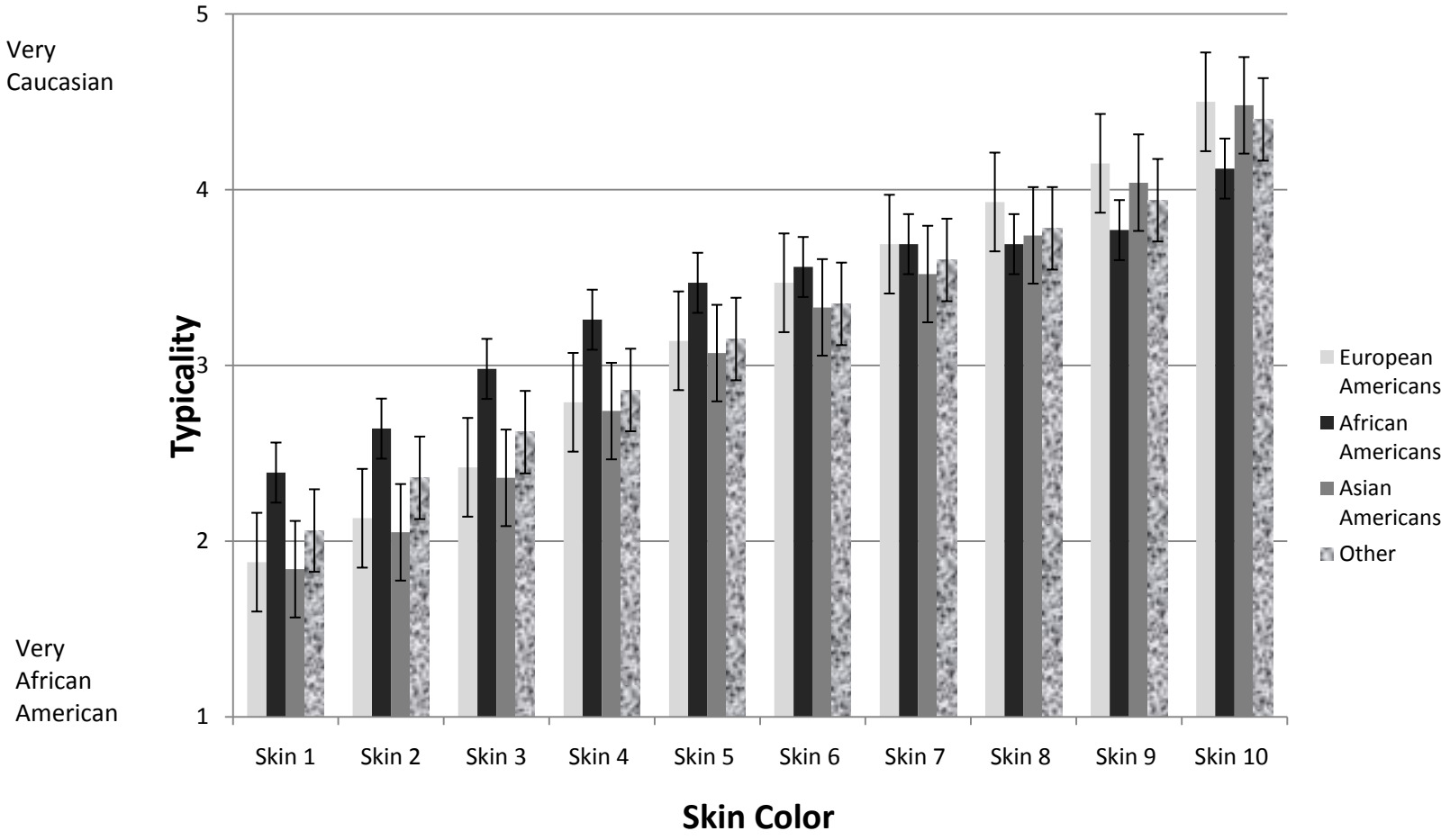

Figure 12. The effect of skin color and participants' ethnicity on racial typicality judgments, when data from Hispanic Americans are excluded. Standard errors are represented in the figure by the error bars.

The analysis also revealed a significant Response Deadline x Participants' Ethnicity interaction, Greenhouse-Geisser $F(5.43, 436.53) = 3.74$, $p = .002$, $\eta_p^2 = .04$; and a significant Physiognomy x Participants' Ethnicity interaction, Greenhouse-Geisser $F(4.1, 329.68) = 7.39$, $p < .001$, $\eta_p^2 = .08$. These two interactions were qualified by a significant Response Deadline x Physiognomy x Participants' Ethnicity interaction, Greenhouse-Geisser $F(35.46, 2848.5) = 1.69$, $p = .006$, $\eta_p^2 = .02$. To examine this interaction further, I examined the Physiognomy x Participants' Ethnicity interaction within each of the response deadline conditions.

As Figures 13, 14, and 15 illustrate, the relationship of facial physiognomy to racial typicality ratings was more pronounced for African American participants than for European American and Asian American participants, and this effect became even stronger with more stringent response deadlines. Follow-up tests within response deadline conditions revealed significant Physiognomy x Participants' Ethnicity interactions in each response deadline condition, but they were stronger in the two response deadline conditions than in the no response deadline condition: Greenhouse-Geisser $F(5.54, 444.79) = 4.60$, $p < .001$, $\eta_p^2 = .05$ for the No Response Deadline condition; Greenhouse-Geisser $F(5.26, 422.90) = 6.87$, $p < .001$, $\eta_p^2 = .08$ for the Median Response Deadline condition; Greenhouse-Geisser $F(6.21, 498.95) = 6.45$, $p < .001$, $\eta_p^2 = .07$ for the 25th Percentile Response Deadline condition.

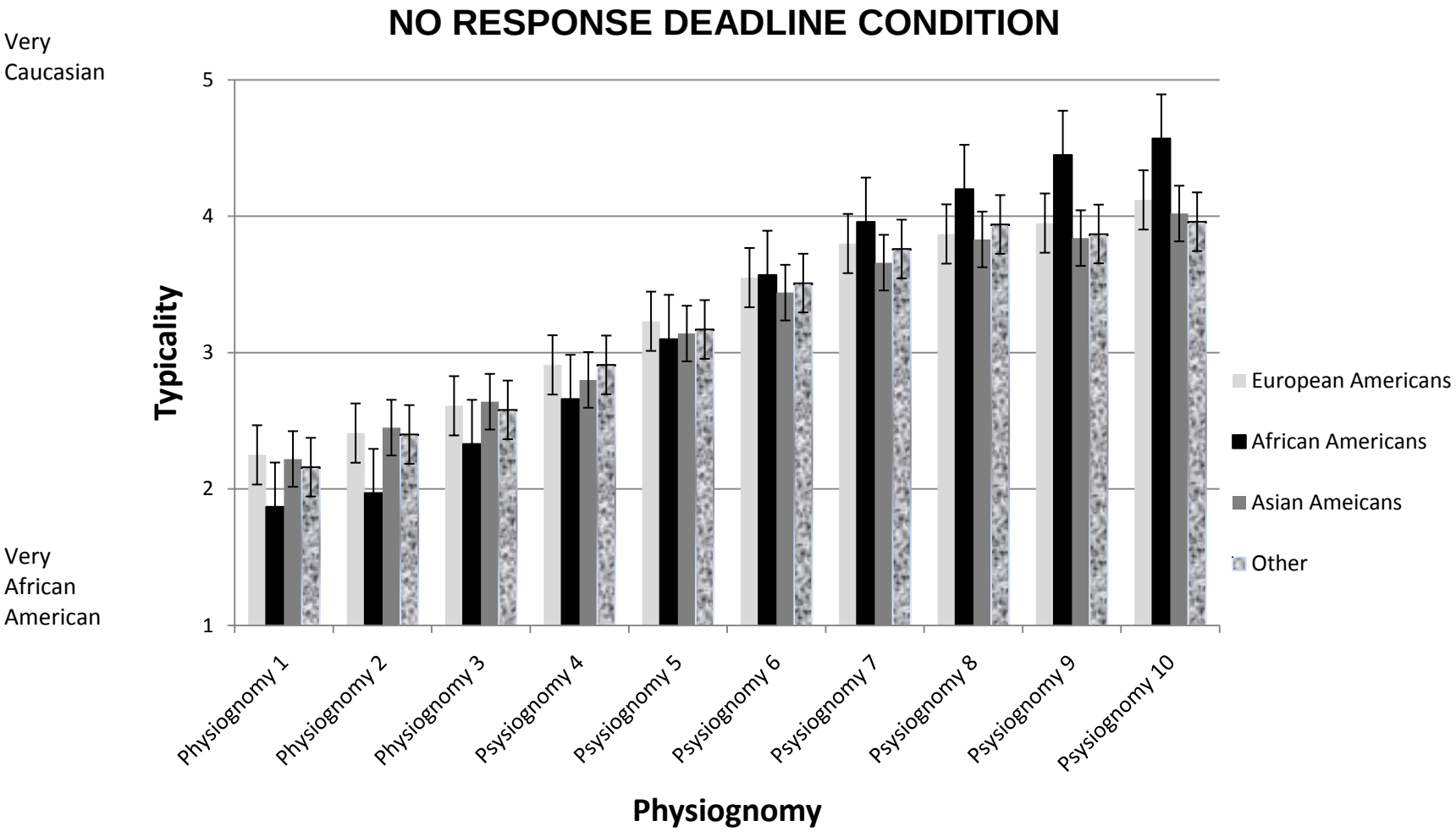

Figure 13. The effect of facial physiognomy and participants' ethnicity on racial typicality judgments, when data from Hispanic Americans are excluded, in no response deadline condition. Standard errors are represented in the figure by the error bars.

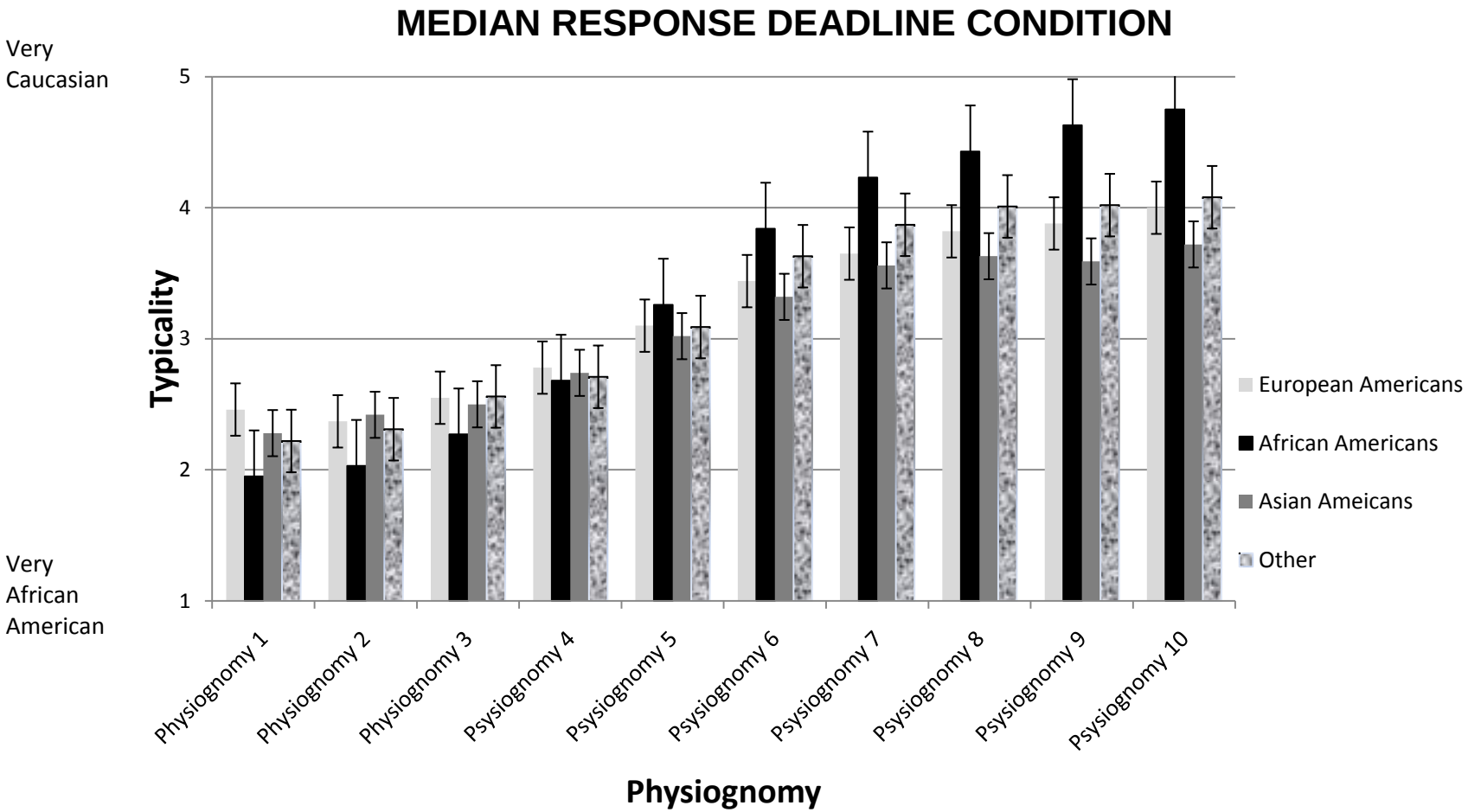

Figure 14. The effect of facial physiognomy and participants' ethnicity on racial typicality judgments, when data from Hispanic Americans are excluded, in Median Response Deadline condition. Standard errors are represented in the figure by the error bars.

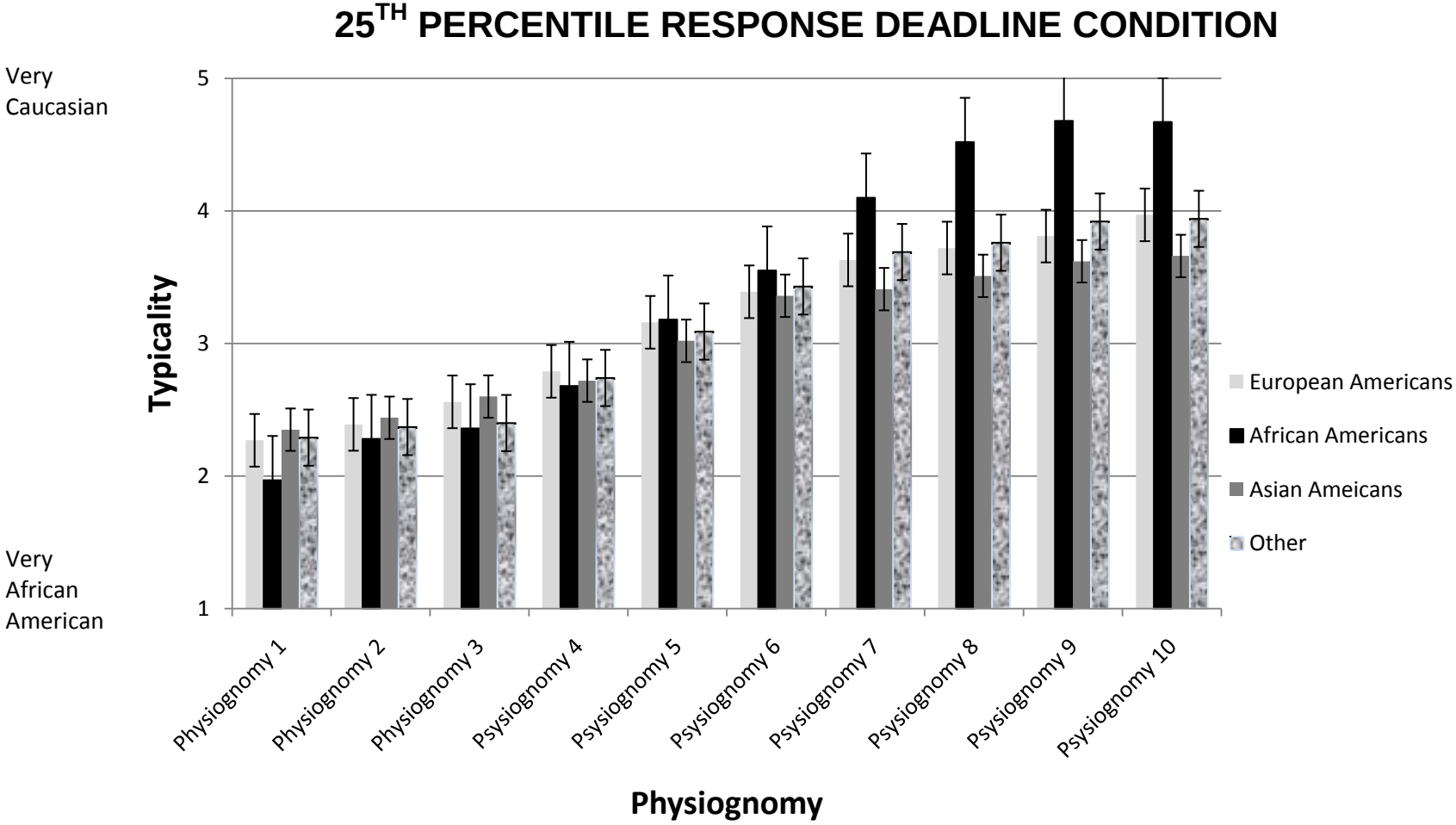

Figure 15. The effect of facial physiognomy and participants' ethnicity on racial typicality judgments, when data from Hispanic Americans are excluded, in 25th Percentile Response Deadline condition. Standard errors are represented in the figure by the error bars.

Note that ethnic groups differed in their median response times, $F(3, 242) = 5.13$, $p = .002$, $\eta_p^2 = .06$. Follow-up comparisons revealed that African Americans took longer to respond ($M = 2789.56$ ms, $SD = 1604.17$) than Asian Americans ($M = 1886.93$ ms, $SD = 610.99$), $p = .001$; and European Americans ($M = 2151.39$ ms, $SD = 1039.39$), $p = .01$. However, when response times were entered as a covariate in a 10 (Skin Color) x 10 (Facial Physiognomy) x 3 (Response Deadline) x (Implicit Racial Attitudes) x (Participants' Ethnicity) repeated measures multiple regression, the following significant interactions still emerged: Skin Color x Ethnicity ($p = .004$), Response Deadline x Participants' Ethnicity ($p = .001$), Physiognomy x Participants' Ethnicity ($p = .001$), and Response Deadline x Physiognomy x Participants' Ethnicity ($p = .015$). This signified that these interactions were not simply due to a variation in response times across ethnic groups.

Exploratory Analyses: Explicit Racial Attitudes as a Moderator

To explore the role of explicit racial attitudes, I conducted a 10 (Skin Color) x 10 (Facial Physiognomy) x 3 (Response Deadline Condition) x (Implicit Racial Attitudes) x (Explicit Racial Attitudes Index) repeated measures multiple regression. The product of the IAT score and explicit attitudes index was entered to examine the interaction of these two between-subjects predictors. As in the previously described exploratory analysis, I report effects only if they were significant at $p < .01$ and only report up to the third-order (three-way) interactions. Using these criteria, a Physiognomy x Explicit Racial Attitudes Index interaction emerged, Greenhouse-Geisser $F(1.35, 332.54) = 7.31$, $p = .003$, $\eta_p^2 = .03$. It was qualified by a Physiognomy x Explicit Racial

Attitudes x Implicit Racial Attitudes interaction, Greenhouse-Geisser $F(1.35, 332.60) = 6.05$, $MSE = 298.44$, $p = .008$, $\eta_p^2 = .02$.

As Figure 16 indicates, participants lower in both implicit and explicit racial attitudes showed the most pronounced responsiveness to facial physiognomy and participants lower in implicit racial attitudes but higher in explicit racial attitudes showed the least responsiveness to variation in facial features. The other participants showed sensitivity to facial physiognomy that fell between these two groups. Possible interpretations for this interaction will be deferred to the discussion section.

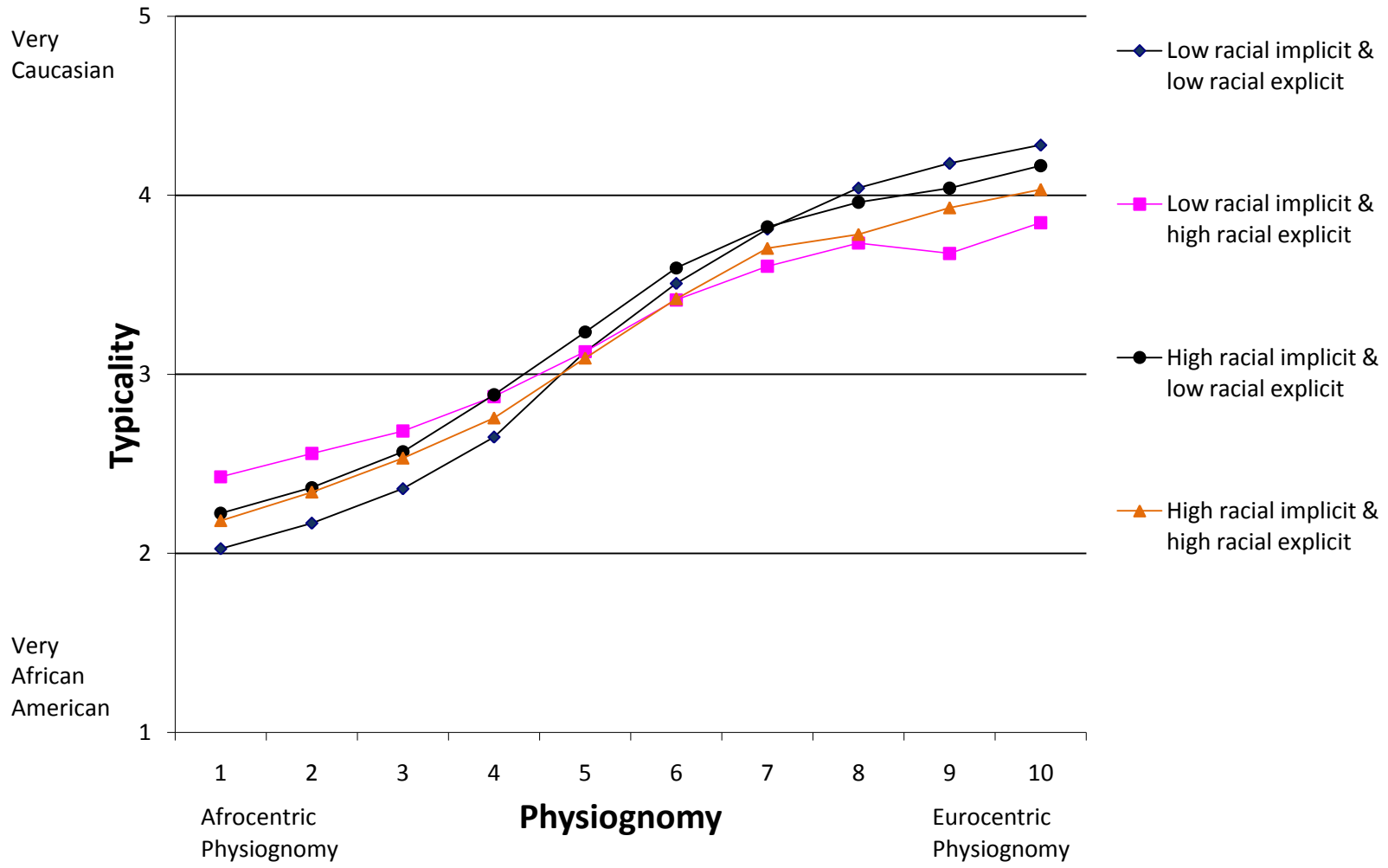

Figure 16. The effects of implicit racial attitudes, explicit racial attitudes and facial physiognomy on racial typicality judgments.

I used the same approach to explore effects involving the Feeling Thermometers for Blacks and for Whites. These analyses revealed a significant Skin Color x Feeling Thermometer for Whites interaction, Greenhouse-Geisser $F(1.48, 365.53) = 10.35$, $p < .001$, $\eta_p^2 = .04$, and a significant Skin Color x Implicit Racial Attitudes x Feeling Thermometer for Whites interaction, Greenhouse-Geisser $F(1.49, 367.32) = 5.88$, $MSE = 287.94$, $p = .007$, $\eta_p^2 = .02$. The Skin Color x Implicit Racial Attitudes x Feeling Thermometer for Whites interaction is illustrated on Figure 17 and shows that participants with higher implicit racial attitudes and more favorable feelings towards Whites exhibited the most pronounced sensitivity to variation in skin color. Potential explanations will be addressed in the discussion section.

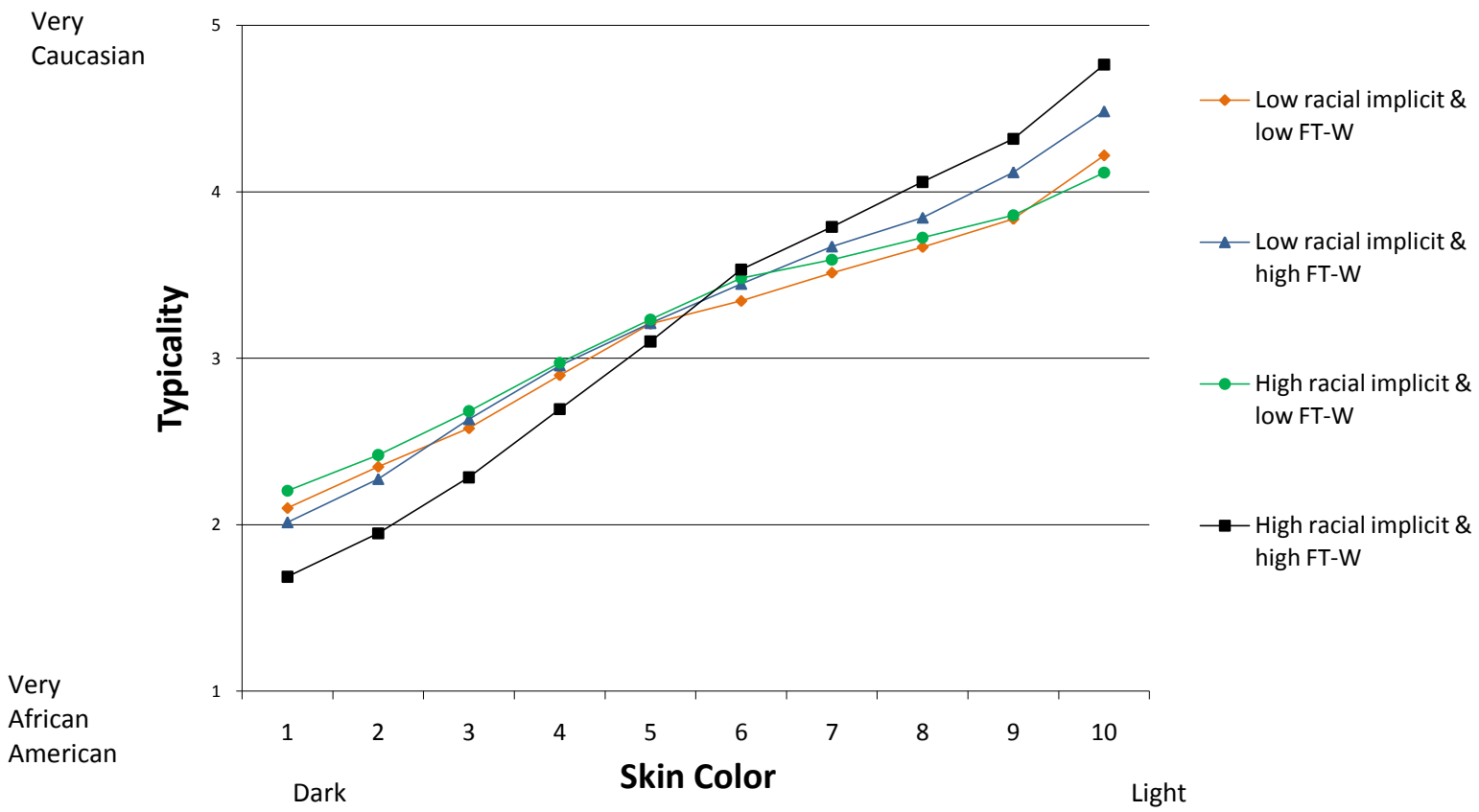

Figure 17. The effects of implicit racial attitudes, explicit racial attitudes and skin color on racial typicality judgments. Low FT-W indicates low ratings of Feeling Thermometer for Whites, indicating less positivity/more negativity; high FT-W indicates high ratings of Feeling Thermometer for whites, indicating more positivity/less negativity.

Exploratory Analyses: Ethnicity and Explicit Racial Attitudes

I also examined the joint effects of explicit racial attitudes and participants' ethnicity in 10 (Skin Color) x 10 (Facial Physiognomy) x 3 (Response Deadline Condition) x (Participants' Ethnicity) x (Explicit Racial Attitudes) repeated measures multiple regressions. The Explicit Racial Attitudes Index, Feeling Thermometer for Whites, and Feeling Thermometer for Blacks were examined in separate analyses. Using the exploratory criteria described earlier, no significant effects were found.

Discussion

This research investigated the following questions: (a) What are the additive and interactive effects of skin color and facial physiognomy on racial categorization when those variables are manipulated independently and vary in gradual increments?; (b) Are those effects moderated by implicit racial attitudes?; and (c) Do time constraints alter the weighting of skin color and facial physiognomy in racial typicality judgments? Exploratory analyses further examined the moderating role of participant ethnicity and explicit racial attitudes. In the sections that follow I will describe the support for the hypotheses, offer possible explanations for unanticipated findings, describe the limitations of this study, and suggest the conceptual implications and directions for future research.

Skin Color and Facial Physiognomy in Racial Categorization

Consistent with Hypothesis 1, I found that darker faces were rated as more African American than lighter faces, and faces with Afrocentric physiognomy were rated as more African American than faces with Eurocentric physiognomy. This was not a surprising finding in light of previous experiments (e.g., Gitter & Satow, 1969; Stepanova & Strube, 2009; Stepanova et al., 2008). Interestingly, however, facial physiognomy and skin color played equally important roles in racial categorization, as evidenced by their effect sizes.

Consistent with Hypothesis 2, skin color and facial features also interacted to influence racial typicality ratings (Figure 7). This interaction indicates that facial features are a more important judgment cue when skin color is intermediate and lighter than when it is very dark. This interaction suggests that classification of faces into the African

American category can be done largely on skin tone alone and that facial features do not offer especially diagnostic information. As faces become lighter, however, a simple skin tone heuristic is no longer as useful and the additional information supplied by facial features is relied upon more heavily. The interaction suggests that classification of faces into racial groups may follow a two-step process, with an initial evaluation of skin color that terminates in a classification into the African American category if the face is dark but that is followed by an evaluation of facial features if skin tone is lighter. An important implication is that equally atypical faces are not treated the same way. That is, a very dark Eurocentric face is, objectively, as unusual as a very light Afrocentric face—at least in the mismatch of skin tone and physiognomy. Yet, the former faces are not distinguished as much from very dark Afrocentric faces—all are viewed as African American faces—whereas light Afrocentric faces are more clearly distinguished from light Eurocentric faces.

Although the aforementioned interaction suggests that skin tone may play a primary role in racial classification, other evidence from this study suggests that reliance on skin tone and facial physiognomy varies by ethnicity. For example, White participants and Asian American participants relied more on skin tone than did African American participants (Figure 12). On the other hand, the use of facial physiognomy was more pronounced for African American participants than for European American and Asian American participants, especially when decisions had to be made quickly (Figures 13-15). These findings suggest that skin color and facial physiognomy play somewhat different roles or are weighted differently in racial categorization for different ethnic groups. Skin color is a more salient out-group marker for European Americans and Asian

Americans than for African Americans, whereas facial physiognomy is a more meaningful in-group marker for African Americans. There are several ethnic groups that include individuals with dark skin tone such as Hispanics, Asian Indians or American Indians. Therefore, skin color is not a very meaningful in-group marker for African Americans and they rely upon other cues such as facial features when making important in-group classification decisions.

The results of this experiment are also important to compare to those reported by Stepanova et al. (2008) based on a sample from the Russian Federation. Stepanova et al. found a much larger role of skin tone than was found in the current study—dark faces were almost exclusively rated as non-Russian and categorization of light faces depended on facial physiognomy (see Figure 3). One clear explanation is the different socio-cultural context and daily exposure to racial cues in the United States and Russian Federation. In the United States, most of my participants are exposed to both ethnic markers—skin color and facial physiognomy, and equally often. In the Russian Federation, especially in the Yarolslavl region, people of African ancestry are not common, but other ethnic groups that have dark skin tone are encountered frequently. Thus, dark skin tone allows perceivers to distinguish between in-groups and out-groups quite easily. Taken together, the Stepanova et al. results and the current findings underscore the role of socio-cultural context in determining the weighting of various visual markers in ethnic and racial categorization.

The interactive influence of skin color and facial features on racial categorization judgments also depended on availability of cognitive resources (Figures 9, 10, and 11). Somewhat consistent with Hypothesis 4, this interaction indicated that, when participants

made judgments under time constraints, they relied on facial physiognomy and skin color when rating faces of middle levels of physiognomy but relied on skin color primarily when rating faces of extreme levels of physiognomy. Another way to describe this interaction is that, as time constraints became more severe, fewer distinctions were made along the facial physiognomy dimension for extreme physiognomy faces, with faces being more likely to be clustered into larger perceptual groups. An important implication of this interaction is that the relative weighting or importance of skin tone and facial physiognomy is not fixed but varies with the processing demands imposed by the situation.

Collectively, these results are important in relation to previous research indicating that facial cues associated with African/Black features are more salient than European/White features in social categorization (Smith & Zarate, 1992). This research suggests that the process of categorization depends upon the degree of Black/African features in a face (Freeman, Pauker, Apfelbaum, & Ambady, 2010). Importantly, however, skin color and facial physiognomy are not independently manipulated in this previous work. The results of the current research thus offer an important qualification. Facial cues certainly are important in racial judgments, but their importance depends on the nature of other cues (skin tone) and on the degree to which the situation provides the opportunity to process those cues completely. Therefore, it appears that not all features are always salient, and some disproportionally affect racial categorization. These findings underscore the role of within-stimuli variability in racial categorizations.

Implicit Racial Attitudes as Moderators of Racial Categorization

Hypothesis 3 predicted that implicit racial attitudes would further moderate the impact of skin tone and facial physiognomy on racial typicality judgments. Partial support was found for this prediction, in the form of a Skin Tone x Implicit Racial Attitudes interaction (Figure 8). Skin color was more related to racial categorization for individuals with higher implicit racial prejudice than for those with lower implicit racial prejudice. Clearly, skin color plays an important role in racial identification for those who have negative implicit associations with African Americans and positive implicit associations with European Americans. Note that European American had the highest levels of implicit racial prejudice (Table 3) and that it was European Americans whose racial categorization ratings were more related to skin color (in comparison to African Americans, Figure 12). Taken together, these results suggest that individuals with high implicit racial prejudice, many of whom are European Americans, rely upon the most salient out-group marker in racial categorizations. Previous research (for review, see Dukes and Maddox, 2008) has established that certain facial cues such as skin color lead to negative attitudes; this study adds to that work by showing that implicit attitudes might influence weighting of certain visual cues in racial categorization.[6] On the other hand, contrary to Hypothesis 5, the influence of implicit attitudes on racial typicality judgments was not further qualified by the time constraint manipulation, suggesting that implicit attitudes influence such judgment regardless of limitations on cognitive resources.

Implicit and Explicit Racial Attitudes as Moderators of Racial Categorization

I also explored moderation of racial categorization by both explicit and implicit racial attitudes, and found an unusual pattern of results. Participants lower in both implicit and explicit racial attitudes showed the most pronounced responsiveness to facial

physiognomy, whereas participants lower in implicit racial attitudes but higher in explicit racial attitudes showed the least responsiveness to variation in facial features (Figure 16). This interaction was not anticipated and does not lend itself to an obvious interpretation. One feature of it, however, is intriguing and points to questions that future research will need to resolve. Specifically, these findings might indicate that some individuals are more aware of the racial implications of using particular cues in judging others and so may actively avoid those cues in favor of less salient markers. Individuals with low explicit and implicit racial attitudes, for example, may consciously ignore skin color, as they are aware that skin color is a more obvious marker for those who have negative attitudes towards African Americans. Instead, they may rely more heavily on facial physiognomy. Other features of this interaction, however, are not so easily explained. For example, it is not clear why individuals who are lower in implicit racial attitudes and higher in explicit racial attitudes are least responsive to variations in facial features. It is perhaps wise to withhold judgment on this interaction until future research can establish its reliability.

A second interaction between implicit and explicit attitudes is a bit more intuitive. Participants with higher implicit racial attitudes and more favorable feelings towards Whites exhibited the most pronounced sensitivity to variations in skin color (Figure 17). These participants rely upon the most salient marker, skin color, to make categorization judgments. This group of people is the most "White"-oriented; they have the most pro-White/anti-Black implicit racial attitudes, as well as the most positive explicit attitudes towards Whites. As I discussed previously, skin color is an especially important marker in racial categorization for European American participants, as well for those who have

high implicit racial attitudes. The same appears to be true for "White"-oriented participants, regardless of their ethnicity.

Overall, these findings are consistent with previous research indicating that the same faces can be judged quite differently depending upon racial attitudes of a perceiver (Blascovich et al., 1997; Castano et al., 2002; Pettigrew et al., 1958), especially when there is within-group variability. These findings add an important caveat: different facial markers carry different weight for individuals with various levels of implicit and explicit racial prejudice.

Additional Findings

A few additional findings are worth mentioning because they help shed some light on other results I have described. For example, the mean IAT score was unusually high in this sample (M = .62, SD = .6). This is higher than the maximum reported by Lane et al. (2007) for a variety of IAT tests. Furthermore, European American participants were characterized by very high levels of implicit prejudice (D = .81, SD =.53); whereas African American participants were characterized by an absence of implicit racial bias (D = .07, SD = .59). The latter finding is consistent with previous research that has indicated that African-Americans (and other lower-status groups) do not typically show in-group bias on the IAT (e.g., Jost, Pelham, & Carvallo, 2002; Nosek, Banaji, & Greenwald, 2002; Olson, Crawford, & Devlin, 2009; Richeson, Trawalter, & Shelton, 2005, but see Livingston, 2002). The higher-than-typical bias shown by White participants may be due to one unique characteristics of this study--the use of computer generated images as IAT stimuli. The stimuli used in this study may have allowed participants to identify race better because the faces were extreme on each end of the racial spectrum (e.g., images of

individuals with the lightest skin tone and most Eurocentric features were chosen to represent White faces; images with darkest skin tone and most Afrocentric physiognomy were chosen to represent Black faces). By comparison, most racial IAT studies employ a set of stimuli developed for the web-based IAT (Nosek, Banaji, & Greenwald, 2002; available at http://projectimplicit.net/nosek/stimuli/ under *Race Faces*). The stimuli consist of morphed young faces cropped at the forehead and chin. Each face has a neutral expression and peripheral features (e.g., hair, clothes) are not visible. These stimuli are in grayscale and cropped in a way that the mouth is generally not visible (e.g., only a small part of the upper lip is present). Although race differences are generally clear in each set, they are perhaps more obvious in those used in the present study.

The differences between the two types of stimuli are important because the IAT measures differences between two target concepts rather than differences between exemplars' of two target concepts (De Houwer, 2001, see also De Houwer, 2003). The task procedurally asks participants to categorize those two concepts, thus making them explicitly available. Accordingly, IAT effects depend on the valence of the categories (e.g., Black or White) rather than the properties of the exemplars (e.g., faces of Black and White individuals) (see De Houwer, 2009). Evidence that suggests some influence of exemplar properties on the IAT effects is limited to studies that employ *lexical* and not *pictorial* stimuli (for review, see De Houwer, 2009). However, there is some emerging evidence that pictorial characteristics (e.g., grayscale cropped vs. non-cropped colored primes) can influence the IAT effects (see Smith-McLallen, Johnson, Dovidio, & Pearson, 2006). Previously I noted that presenting facial stimuli in color increases perceptions of racial typicality (Stepanova & Strube, 2009). Perhaps the stimuli used in

the current study, in addition to activating the racial categories "Black" and "White," also made certain racial cues more salient than is true for the Nosek et al. stimuli. This may in turn magnify the sensitivity of the IAT and potentially explain such unusually high IAT scores. This suggests that the stimuli used in the IAT may play an important role in the nature of the score distributions that are obtained. This is certainly worth future investigation because the IAT is the most commonly used implicit prejudice measure.

Theoretical, Methodological and Practical Implications

This research has conceptual implications for views of categorization, impression formation and the Maddox (2004) model of racial typicality bias. Specifically, variability in judgments as a function of skin color and facial physiognomy provide support for the probabilistic view of racial categories. Faces with more Afrocentric features and darker skin color are considered to be more African American than lighter faces and faces with more Eurocentric physiognomy, indicating that some faces are more "typical" members of the category than the others. More importantly, however, this research showed that how individuals process the concept of race varies with characteristics of the situation (time demand) and with characteristics of the person (racial attitudes and ethnicity). Moderation of racial typicality ratings by implicit and explicit racial attitudes, ethnicity, and availability of cognitive resources supports theory-based accounts of racial categorization.

Perhaps the most important implication of this research is that it expands the Maddox model by identifying factors that moderate the categorization route. Specifically, this research found evidence that implicit racial attitudes moderate racial categorization by influencing one specific visual cue, skin color only, whereas limiting

cognitive resources moderates racial categorization by affecting the use of both skin color and facial physiognomy. Furthermore, participants' ethnicity moderated racial categorization as well.[7] Future research might explore other factors that moderate racial categorization (e.g., other individual differences such as motivation to control prejudice or intolerance for ambiguity) and investigate the role of perceivers' ethnicity in greater detail.

This study also has important methodological implications because research on racial prejudice and stereotyping depends critically on adequate choice of the facial stimuli used. As demonstrated by the variability in typicality judgments driven by skin color and facial physiognomy, the characteristics of facial stimuli do matter, and effect sizes in studies employing these stimuli might be influenced by the visual properties of facial racial stimuli. Specifically, a choice of extremely dark faces to represent African American primes would potentially produce stronger racial effects, especially so in European Americans and those whose implicit racial attitudes are high, because they rely more than other groups on skin color in racial categorization. This study further suggests that some of the stimuli in the current body of racial prejudice and stereotyping research exaggerate the typicality of European American and African American faces, which, in turn, influences the size of the effects obtained.

The practical implications are even more important. Race-relevant decisions occur in many important contexts with considerable potential for bias. Determining the factors that drive those decisions has substantial applied importance, and the design of interventions will depend critically on an understanding of the underlying mechanisms. For example, most of the research dealing with cross-racial eyewitness identification

addresses issues of recognition rather than initial racial categorization and encoding (e.g., Ayuk, 1990; Ellis & Deregowski, 1981; MacLin, MacLin, & Malpass, 2001; Meissner & Brigham, 2001, but see MacLin & Malpass, 2001) and is almost always limited to European American participants categorizing other races' target faces. Furthermore, under some circumstances, individuals might be exposed to faces for a very brief period of time, which might affect the process of categorization as well. Payne (2001) has discussed the case of Amadou Diallo, an African immigrant who was shot by New York City police officers who had mistaken his wallet for a gun; that research investigated stereotypic associations people have with African Americans and weapons and showed how easily misclassifications (e.g., of the wallet as a weapon) can occur and be moderated by ethnicity of targets. However, similar mistakes might be made for categorization of people into racial categories when a fast decision is required. I suggest that, under such conditions, the categorization of a target face might be altered, and factors that drive this categorization are weighed differentially. For example, under time constraints, faces with very Eurocentric or Afrocentric facial features will be more likely to be categorized on the basis of skin color only, regardless of facial physiognomy. At the same time, a person briefly exposed to a dark-skinned individual will weigh skin tone more than facial features and report seeing "African American"-- this might be even more prominent for individuals with high implicit racial prejudice and prompt especially high behavioral discrimination.

Limitations

This study explored racial categorization through explicit categorization judgments but does not allow inferences about the process of social categorization. There

is emerging research attempting to address the process of categorization using new computer-mouse tracking methods, in which participants' hand movements en route to dichotomous racial category alternatives are recorded by tracking the spatial coordinates of the computer mouse (see Freeman, Ambady, Rule, & Jonson, 2008; Freeman et al., 2010). Social neuroscience research also shows promise by employing ERPs (event-related brain potentials) to investigate temporal effects of racial categorization (Ito, Thompson, & Cacioppo, 2004; Ito & Urland, 2005; Ito, Urland, Willadsen-Jensen, & Corell, 2006; Kubota & Ito, 2007), including studies with racially ambiguous faces (Willadsen-Jensen & Ito, 2006). Application of such methodologies will allow investigating the temporal sequence of categorization and the role played by moderators at different stages.

Although the temporal sequence of categorization assumes a somewhat step-wise categorization process, facial perception and recognition research has stressed that faces are processed holistically in a Gestalt-like manner (e.g., Farah, Wilson, Drain, & Tanaka, 1998; Goffaux & Rossion, 2006; Hole, 1994; Homa, Haver, & Schwartz, 1976; Sergent, 1984; Tanaka & Farah, 1993; Young, Hellawell, & Hay, 1987), and my emphasis on separation of skin color and facial features in racial categorization might seem somewhat artificial. However, experimental evidence suggests that other-race faces are perceived less holistically than own-race faces (e.g., Michel, Caldara, & Rossion, 2006; Michel, Rossion, Han, Chung, & Caldara, 2006). Likewise, new research investigating separate effects of skin color and facial features on modulation of neural responses (Balas & Nelson, 2010), face recognition (Bar-Haim, Saidel, & Yovel, 2009) and racial

categorization (Stepanova & Strube, 2009) shows that separation of these two factors is justified.

The facial physiognomy manipulation in the current research did not center on any specific facial feature, rather it included a collection of facial features manipulated as a combination. This facilitated the demonstration that facial physiognomy has an important impact on categorization, but it will be important for future research to determine the particular facial features that are especially important. That task was beyond the scope of the present work. There are other procedural limitations as well. The stimuli that I employed are artificially constructed computer generated images. In that regard, they undoubtedly deviate from real faces in some respects. However, initial questions about how the nature of facial stimuli affects racial judgments are best answered when the features of those stimuli can be carefully controlled and manipulated—the distinct advantage of the approach taken in this research. Nonetheless, it will be important for future research to verify the key findings from the proposed research using real faces.

Analogously, this research only examined racial typicality judgments along an African American-European American continuum. It is possible that racial typicality judgments for other groups (e.g., Asian-European continuum) do not follow the same pattern involving these two types of cues—facial physiognomy and skin color. This area is a potential new direction for future research that can examine racial typicality judgments employing other racial groups and judgment tasks.

Other methodological aspects of the current study are worth exploring in future research. I employed a 6-point racial typicality scale, but some racial categorization

studies use a dichotomous judgment (e.g., “Black/ not Black” or “White/not White; Peery & Bodenhausen, 2008). Other studies use more extensive categorical scales (“Colored,” Indians and Africans; Pettigrew, Allport, & Barnett, 1958), and some use both categorical scales as well as continuous scales to investigate racial judgments (MacLin & Malpass, 2001). When Stepanova et al. (2008) employed a less variable racial categorization scale (e.g., Russian, non-Russian, not clear), results were similar to those obtained with a 7-point scale, although the effects were less pronounced. Because the results were clearer with the 7-point scale, I inferred that the racial categorization judgment itself might be more dimensional rather than binary. Nonetheless, future studies might investigate effects of within-group variability with dichotomous categorization decisions.

Concluding Remarks

Despite these limitations, the questions that I attempted to answer here regarding the factors that are important for racial categorization are a key step to a greater understanding of how variability within a group can affect race perception. Given the substantial variability in the facial features that exists within racial groups outside the laboratory, this work is also an important step toward linking laboratory work to the settings in which research and theory should apply.

References

Allport, G. W. (1954). *The nature of prejudice*: Addison-Wesley Cambridge, Mass.

Ayuk, R. E. (1990). Cross-racial identification of transformed, untransformed, and mixed-race faces. *International Journal of Psychology, 25*, 509-527. doi:10.1080/00207599008247880

Balas, B., & Nelson, C. (2010). The role of face shape and pigmentation in other-race face perception: An electrophysiological study. *Neuropsychologia*. *48*, 498–506, doi:10.1016/j.neuropsychologia.2009.10.007

Bar-Haim, Y., Seidel, T., & Yovel, G. (2009). The role of skin colour in face recognition. *Perception, 38*, 145–148. doi:10.1068/p6307

Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology, 71*(2), 230-244. doi:10.1037/0022-3514.71.2.230

Biernat, M. & Crandall, C.S. (1999). Racial attitudes. In J.P. Robinson & P.R. Shaver (Eds.), *Measures of political attitudes* (pp. 297-411). San Diego, CA: USnc.

Blair, I. V., Judd, C. M., Sadler, M. S., & Jenkins, C. (2002). The role of Afrocentric features in person perception: Judging by features and categories. *Journal of Personality and Social Psychology, 83*, 5-25. doi:10.1037/0022-3514.83.1.5

Blascovich, J., Wyer, N., Swart, L., A., & Kibler, J. L. (1997). Racism and racial categorization. *Journal of Personality and Social Psychology, 72*, 1364-1372. doi:10.1037/0022-3514.72.6.1364

Bodenhausen, G. V. (1990). Stereotypes as judgmental heuristics: Evidence of circadian variations in discrimination. *Psychological Science, 1,* 319-322. doi:10.1111/j.1467-9280.1990.tb00226.x

Brown, T. D., Jr, Dane, F. C., Durham, M. D. (1998). Perception of race and ethnicity. *Journal of Social Behavior and Personality, 13 (2),* 295-306.

Castano, E., Yzerbyt, V., Bourguignon, D., Seron, E. (2002). Who may enter? The impact of in-group identification on in-group/out-group categorization. *Journal of Experimental Social Psychology, 38,* 315-322. doi:10.1006/jesp.2001.1512

Dasgupta, N., Banaji, M. R., & Abelson, R. P. (1999). Group entitativity and group perception: Associations between physical features and psychological judgment. *Journal of Personality & Social Psychology, 77*(5), 991-1003. doi:10.1037/0022-3514.77.5.991

De Houwer, J. (2001). A structural and process analysis of the implicit associations test. *Journal of Experimental Social Psychology 37***,** 443–451. doi:10.1006/jesp.2000.1464

De Houwer, J. (2003). A structural analysis of indirect measures of attitudes. In J. Musch & K.C. Klauer (Eds.), *The Psychology of Evaluation: Affective Processes in Cognition and Emotion* (pp. 219-244). Mahwah, NJ: Lawrence Erlbaum.

De Houwer, J. (2009). Comparing measures of attitudes at the functional and procedural level: Analysis and implications. In R. E. Petty, R. H. Fazio, & P. Brinol (Eds.), *Attitudes: Insights from the new implicit measures* (pp. 361-390). Psychology Press: NY.

Deregowski, J., Ellis, H., & Shepherd, J. (1975). Descriptions of White and Black faces by White and Black subjects. *International Journal of Psychology*, *10*(2), 119-123. doi:10.1080/00207597508247317

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. doi:10.1006/jesp.1997.1331

Dovidio, J. F., Kawakami, K., Smoak, N., & Gaertner, S. L. (2009). The roles of implicit and explicit processes in contemporary prejudice. In R. E. Petty, R. H. Fazio, & P. Brinol (Eds.), Attitudes: Insights from the new implicit measures (pp. 165-192). New York: Psychology Press.

Ellis, H. D., & Deregowski, J. B. (1981). Within-race and between-race recognition of transformed and untransformed faces. *The American Journal of Psychology, 94* (1), 27-35. doi:10.2307/1422341

Farah, M. J., Wilson, K. D., Drain, M., & Tanaka, J. N. (1998). What is "special" about face perception? *Psychological Review, 105*, 482–498. doi:10.1037/0033-295X.105.3.482

Farkas L. G, Katic M. J., Forrest C. R., Alt K. W., Bagic I., Baltadjiev G. et al. (2005). International anthropometric study of facial morphology in various ethnic groups/races. *Journal of Craniofacial Surgery, 16*(4), 615-646.

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*(6), 1013-1027. doi:10.1037/0022-3514.69.6.1013

Freeman, J. B., Ambady, N., Rule, N. O., & Johnson, K. L. (2008). Will a category cue attract you? Motor output reveals dynamic competition across person construal. Journal of Experimental Psychology: General, 137, 673–690. doi:10.1037/a0013875

Freeman, J., Pauker, K., Apfelbaum, E., & Ambady, N. (2010). Continuous dynamics in the real-time perception of race. *Journal of Experimental Social Psychology*, doi:10.1016/j.jesp.2009.10.002

Gitter, A.G., & Satow, Y. (1969). Color and physiognomy as variables in racial misidentification among children. Proceedings of the Annual Convention of the American Psychological Association, 4 (Pt. 2), 677-678.

Goffaux, V., & Rossion, B. (2006). Faces are 'spatial'--holistic face perception is supported by low spatial frequencies. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(4), 1023-1039. doi:10.1037/0096-1523.32.4.1023.

Govorun, O, & Payne, B. K. (2006). Ego-depletion and prejudice: Separating automatic and controlled components. *Social Cognition*, *24*(2), 111-136. doi:10.1521/soco.2006.24.2.111

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. doi:10.1037/0022-3514.74.6.1464

Greenwald, A G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of*

Personality and Social Psychology, 85, 197–216. doi:10.1037/0022-3514.85.2.197

Henry, P. J., & Sears, D. O. (2002). The symbolic racism 2000 scale. *Political Psychology, 23(2),* 253-283. doi:10.1111/0162-895X.00281

Hole, G. J. (1994). Configural factors in the perception of unfamiliar faces. *Perception, 23,* 65–74. doi:10.1068/p230065

Homa, D., Haver, B., & Schwartz, T. (1976). Perceptibility of schematic face stimuli: Evidence for a perceptual Gestalt. *Memory & Cognition, 4,* 176–185.

Hugenberg, K., & Bodenhausen, G. V. (2003). Facing prejudice: Implicit prejudice and the perception of facial threat. *Psychological Science, 14*(6), 640-643. doi:10.1046/j.0956-7976.2003.psci

Hugenberg, K., & Bodenhausen, G. V. (2004). Ambiguity in social categorization: The role of prejudice and facial affect in race categorization. *Psychological Science, 15*(5), 342-345. doi:10.1111/j.0956-7976.2004.00680.x

Hutchings, P. B., & Haddock, G. (2008). Look Black in anger: The role of implicit prejudice in the categorization and perceived emotional intensity of racially ambiguous faces. *Journal of Experimental Social Psychology, 44,* 1418-1420. doi:10.1016/j.jesp.2008.05.002

Ito, T., & Urland, G. (2005). The influence of processing objectives on the perception of faces: An ERP study of race and gender perception. *Cognitive, Affective & Behavioral Neuroscience, 5*(1), 21-36. doi:10.3758/CABN.5.1.21

Ito, T., Thompson, E., & Cacioppo, J. (2004). Tracking the Timecourse of Social Perception: The Effects of Racial Cues on Event-Related Brain Potentials.

Personality and Social Psychology Bulletin, *30*(10), 1267-1280. doi:10.1177/0146167204264335.

Ito, T., Urland, G., Willadsen-Jensen, E., & Correll, J. (2006). The Social Neuroscience of Stereotyping and Prejudice: Using Event-Related Brain Potentials to Study Social Perception. *Social Neuroscience: People thinking about thinking people* (pp. 189-212). Cambridge, MA US: MIT Press.

Jablonski, N. G. (2004). The evolution of human skin and skin color. *Annual Review of Anthropology, 33,* 585-623.

Jost, J. T., Pelham, B. W., & Carvallo, M. (2002). Non-conscious forms of system justification: Cognitive, affective, and behavioral preferences for higher status groups. *Journal of Experimental Social Psychology, 38,* 586–602. doi:10.1016/S0022-1031(02)00505-X

Kubota, J., & Ito, T. (2007). Multiple cues in social perception: The time course of processing race and facial expression. *Journal of Experimental Social Psychology*, *43*(5), 738-752. doi:10.1016/j.jesp.2006.10.023

Kunda, Z. (1999). *Social cognition: Making sense of people*, Cambridge, MA: The MIT press.

Lamberts, K. (1995).Categorization under time pressure. *Journal of Experimental Psychology: General, 124,* 161-180. doi:10.1037/0096-3445.124.2.161

Lamberts, K. (2000). Information-accumulation theory of speeded categorization. *Psychological Review, 107,* 227-260. doi:10.1037/0033-295X.107.2.227

Lebrecht, S., Pierce, L. J., Tarr, M. J., & Tanaka, J. W. (2009). Perceptual other-race training reduces implicit racial bias. *PLoS ONE, 4*(1): e4215. doi:10.1371/journal.pone.0004215

Levin, D. T., & Banaji, M. R. (2006). Distortions in the perceived lightness of faces: The role of race categories. *Journal of Experimental Psychology: General, 135 (*4), 501-512. doi:10.1037/0096-3445.135.4.501

Leyens, J., & Yzerbyt (1992). The ingroup overexclusion effect: Impact of valence and confirmation on stereotypical information search. *European Journal of Social Psychology, 22,* 549-569. doi:10.1002/ejsp.2420220604

Livingston, R. W. (2002). The role of perceived negativity in the moderation of African–Americans' implicit and explicit racial attitudes, *Journal of Experimental Social Psychology, 38,* 405–413. doi:10.1016/S0022-1031(02)00002-1

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*(1), 5-18. doi:10.1037/0022-3514.82.1.5

Lowery, B. S., Hardin, C. D., & Sinclair, S. (2001). Social influence effects on automatic racial prejudice. *Journal of Personality and Social Psychology, 81*(5), 842-855. doi:10.1037/0022-3514.81.5.842

MacLin, O. H., & Malpass, R. S. (2001). Racial categorization of faces. The ambiguous race face effect. *Psychology, Public Policy, and Law, 7,* 98-118. doi:10.1037/1076-8971.7.1.98

MacLin, O. H., & Malpass, R. S. (2003). The ambiguous-race face illusion. *Perception, 32,* 249-252. doi:10.1068/p5046

MacLin, O. H., MacLin, M. K., & Malpass, R. S. (2001). Race, arousal, attention, exposure, and delay. An examination of factors moderating face recognition. *Psychology, Public Policy, and Law, 7,* 134-152. doi:10.1037/1076-8971.7.1.134

Maddox, K. B. (2004). Perspectives on racial phenotypicality bias. *Personality and Social Psychology Review, 8* (4), 383-401. doi:10.1207/s15327957pspr0804

Maddox, K.B., & Dukes, K. N. (2008). Social categorization and beyond: How facial features impact social judgments. In N. Ambady & J. J. Skowronski (Eds.), *First impressions* (pp. 205-233).New York: Guilford.

Maddox, K. B., & Gray, S. A. (2002). Cognitive representations of Black Americans: Reexploring the role of skin tone. *Personality and Social Psychology Bulletin, 28*(2), 250-259. doi:10.1177/0146167202282010

McConahay, J. B. (1986). Modern racism, ambivalence, and the modern racism scale. In J. Dovidio & S. L. Gaertner (Eds.), *Prejudice, discrimination, and racism: Theory and research* (pp.91-125). New York: Academic Press.

McConahay, J.B., Hardee, B.B., & Batts, V. (1981). Has racism declined in America? It depends on who is asking and what is asked. *Journal of Conflict Resolution, 25,* 563-579.

Medin, D. L. (1989). Concepts and conceptual structure. *American Psychologist, 44 (12),* 1469-1481. doi:10.1037/0003-066X.44.12.1469

Medin, D. L., & Smith, E. E. (1984). Concepts and concept formation. *Annual Review of Psychology, 35,* 113-138. doi:10.1146/annurev.ps.35.020184.000553

Meissner, C. A., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology, Public Policy, and Law. 7,* 3-35. doi:10.1037/1076-8971.7.1.3

Michel, C., Caldara, R., & Rossion, B. (2006). Same-race faces are perceived more holistically than other-race faces. *Visual Cognition*, *14*(1), 55-73. doi:10.1080/13506280500158761.

Michel, C., Rossion, B., Han, J., Chung, C., & Caldara, R. (2006). Holistic Processing Is Finely Tuned for Faces of One's Own Race. *Psychological Science, 17*(7), 608-615. doi:10.1111/j.1467-9280.2006.01752.x.

Migetz, D. (2004). Reassessing the Modern Racism Scale in modern times. *Dissertation Abstracts International, 65* (4-B) (UMI No. 766013101)

Nosek, B. A., Banaji, M., & Greenwald, A. G. (2002). Harvesting implicit group attitudes and beliefs from a demonstration web site. *Group Dynamics, 6,* 101–115. doi:10.1037/1089-2699.6.1.101

Olson, M. A., & Fazio, R. H. (2003). Relations between implicit measures of prejudice: What are we measuring? *Psychological Science*, *14*, 636-639. doi:10.1046/j.0956-7976.2003.psci_1477.x

Olson, M., Crawford, M., & Devlin, W. (2009). Evidence for the underestimation of implicit in-group favoritism among low-status groups. *Journal of Experimental Social Psychology*, *45*(5), 1111-1116. doi:10.1016/j.jesp.2009.06

Parra, E. J. (2007). Human pigmentation variation: Evolution, genetic basis, and implications for public health. *Yearbook of Physical Anthropology, 50,* 85–105. doi: 10.1002/ajpa.20727

Payne, B. K. (2001). Prejudice and perception: The role of automatic and controlled processes in misperceiving a weapon. *Journal of Personality and Social Psychology*, *81*, 181-192. doi:10.1037/0022-3514.81.2.181

Peery, D., & Bodenhausen, G. (2008). Black + White = Black: Hypodescent in reflexive categorization of racially ambiguous faces. *Psychological Science*, *19*(10), 973-977. doi:10.1111/j.1467-9280.2008.02185.x

Pettigrew, T. F., Allport, G. W., Barnett, E. O. (1958). Binocular resolution and perception of race in South Africa. *British Journal of Psychology, 49,* 265-278.

Plant, E. A., Peruche, B. M., & Butz, D. A. (2005). Eliminating automatic racial bias: Making race non-diagnostic for responses to criminal suspects. *Journal of Experimental Social Psychology, 41,* 141-156. doi:10.1016/j.jesp.2004.07.004

Pratto, F., Sidanius, J., Stallworth, L., & Malle, B. (1994). Social dominance orientation: A personality variable predicting social and political attitudes. *Journal of Personality & Social Psychology, 67(4),* 741-763. doi:10.1037/0022-3514.67.4.741

Richeson, J., Baird, A., Gordon, H., Heatherton, T., Wyland, C., Trawalter, S., et al. (2003). An fMRI investigation of the impact of interracial contact on executive function. *Nature Neuroscience*, *6*(12), 1323-1328. doi:10.1038/nn1156

Richeson, J., Trawalter, S., & Shelton, J. (2005). African Americans' implicit racial attitudes and the depletion of executive function after interracial interactions. *Social Cognition*, *23*(4), 336-352. doi:10.1521/soco.2005.23.4.336.

Ronquillo, J., Denson, T., Lickel, B., Lu, Z., Nandy, A., & Maddox, K. (2007). The effects of skin tone on race-related amygdala activity: An fMRI investigation.

Social Cognitive and Affective Neuroscience, *2*(1), 39-44. doi:10.1093/scan/nsl043.

Russell, R., & Sinha, P. (2007). Real-world face recognition: The importance of surface reflectance properties. *Perception, 36*(9), 1368–1374. doi:10.1068/p5779

Russell, R., Sinha, P., Biederman, I., & Nederhouser, M. (2006). Is pigmentation important for face recognition? Evidence from contrast negation. *Perception, 35,* 749–759. doi:10.1068/p5490

Sergent, J. (1984). An investigation into component and configural processes underlying face perception. *British Journal of Psychology, 75*, 221–242.

Smith, E. E, & Medin, D. L. (1981). *Categories and concepts.* Cambridge, MA: Harvard University Press.

Smith, E. R., & Zarate, M. A. (1992). Exemplar-based model of social judgment, *Psychological Review* 99 (1992), pp. 3–21. doi:10.1037/0033-295X.99.1.3

Smith-McLallen, A., Johnson, B., Dovidio, J., & Pearson, A. (2006). Black and white: The role of color bias in implicit race bias. *Social Cognition, 24*(1), 46-73. doi: 10.1521/soco.2006.24.1.46

Stepanova, E. V., & Strube, M. J (2009). Making of a face: Role of facial physiognomy, skin tone and color presentation mode in evaluations of racial typicality. *Journal of Social Psychology, 149,* 65-80. doi: 10.3200/SOCP.149.1.66-81

Stepanova, E., Strube, M., Limes, D., & Wolkoff, L. (2009). Pictorial race activation in priming measures: A literature review. Manuscript submitted for publication.

Stepanova, E., Strube, M., Yablonsky, G., Pehrson, K., & Shuman, S. (2008). Studying explicit and implicit ethnic attitudes and ethnic categorization effects in the Russian Federation [Abstract]. *International Journal of Psychology, 43,* p. 810.

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology,46*, 225–245.

Tarman, C., & Sears, D. (2005). The conceptualization and measurement of symbolic racism. *Journal of Politics*, *67*(3), 731-761. doi:10.1111/j.1468-2508.2005.00337.x

Van Hiel, A., & Mervielde, I. (2005). Authoritarianism and social dominance orientation: Relationships with various forms of racism. *Journal of Applied Social Psychology*, *35*(11), 2323-2344. Retrieved August 18, 2008, doi:10.1111/j.1559-1816.2005.tb02105.x

Wilcox, C., Seligman, L, & Cook, E. (1989). Some like it hot: Individual differences in responses to group feeling thermometers. *Public Opinion Quarterly, 53,* 246-257. doi:10.1086/269505

Willadsen-Jensen, E., & Ito, T. (2006). Ambiguity and the timecourse of racial perception. *Social Cognition, 24*(5), 580-606. doi:10.1521/soco.2006.24.5.580.

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. doi:10.1037/0022-3514.81.5.815

Wittgenstein, L. (1953). *Psychological investigations.* New York: Macmillan.

Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configurational information in face perception. *Perception, 16*, 747–759. doi:10.1068/p160747

Appendix A

Symbolic Racism Scale

Which of the following objects or statements do you have a positive or negative feeling towards? Choose the number that best represents the degree of your positive or negative feeling. Press the appropriate number key on the keyboard. The next question will appear automatically.

1. It's really a matter of some people not trying hard enough; if blacks would only try harder they could be just as well as whites.

Strongly Agree Somewhat Agree Somewhat Disagree Strongly Disagree

2. Irish, Italian, Jewish, and many other minorities overcame prejudice and worked their way up. Blacks should do the same.

Strongly Agree Somewhat Agree Somewhat Disagree Strongly Disagree

3. Some say that black leaders have been trying to push too fast. Others feel that they haven't pushed fast enough. What do you think?

Trying To Push Too Fast Going Too Slowly Moving At About Right Speed

4. How much of the racial tension that exists in the United States today do you think blacks are responsible for creating?

All Of It Most Some Not Much At All

5. How much discrimination against blacks do you feel there is in the United States today, limiting their chances to get ahead?

A Lot Some Just A Little None At All

6. Generations of slavery and discrimination have created conditions that make it difficult for blacks to work their way out of the lower class.

Strongly Agree Somewhat Agree Somewhat Disagree Strongly Disagree

7. Over the past few years, blacks have gotten less than they deserve.

Strongly Agree Somewhat Agree Somewhat Disagree Strongly Disagree

8. Over the past few years, blacks have gotten more economically than they deserve.

Strongly Agree Somewhat Agree Somewhat Disagree Strongly Disagree

Appendix B

Social Dominance Scale

Which of the following objects or statements do you have a positive or negative feeling towards? Choose the number that best represents the degree of your positive or negative feeling. Press the appropriate key on the keyboard. The next question will appear automatically.

Very Negative	Negative	Slightly Negative	Neither Positive Nor Negative	Slightly Positive	Positive	Very Positive
1	2	3	4	5	6	7

1. Some groups of people are simply not the equals of others.

2. Some people are just more worthy than others.

3. This country would be better if we cared less about how equal people are.

4. Some people are just more deserving than others.

5. It is not a problem if some people have more of a chance in life than others.

6. Some people are just inferior to others.

7. To get ahead in life, it is sometimes necessary to step on others.

8. Increased economic equality.

9. Increased social equality.

10. Equality.

11. If people were treated more equally, we would have fewer problems in this country.

12. In an ideal world, all nations would be equal.

13. We should try to treat one another as equals as much as possible. (All humans should be treated equally.)

14. It is important that we treat other countries as equals.

Appendix C

Modern Racism Scale

Read each statement carefully and indicate your degree of agreement or disagreement by pressing the appropriate number on the keyboard. The next statement will appear automatically.

Strongly Disagree	Disagree	Slightly Disagree	Neither Agree Nor Disagree	Slightly Agree	Agree	Strongly Agree
1	2	3	4	5	6	7

1. Over the past few years, the government and news media have shown more respect to blacks than they deserve.
2. It is easy to understand the anger of black people in America.
3. Discrimination against blacks is no longer a problem in the United States.
4. Over the past few years, blacks have gotten more economically than they deserve.
5. Blacks have more influence upon school desegregation plans than they ought to have.
6. Blacks are getting too demanding in their push for equal rights.
7. Blacks should not push themselves where they are not wanted.

Appendix D

Feeling Thermometers

For the next task, you will be asked to indicate how you feel about 9 different groups of people. For each group, decide how warm and favorable, or, cold and unfavorable you feel about most members of that group. Then choose the number that best represents your overall feeling and press the appropriate key on keyboard. Press 'Continue' when you are ready to begin.

1	2	3	4	5	6	7
"very cold or unfavorable"						"very warm or favorable"

1. Whites
2. Women
3. Asians
4. Blacks
5. Arabs
6. Hispanics
7. Liberals
8. Men
9. Conservatives

Appendix E

Demographic Questions

Answer the following questions and then press 'Next' to continue.

1. Age (fill in the blank)
2. Gender (Male; Female)
3. Ethnicity (European American; African American, Hispanic American, Asian American; Other)
4. Political Ideology (Very Liberal; Moderately Liberal; Slightly Liberal; Neither Liberal nor Conservative; Slightly Conservative; Moderately Conservative; Very Conservative)
5. Political Affiliation (Strongly Democratic; Moderately Democratic; Slightly Democratic; Neither Democratic nor Republican; Slightly Republican; Moderately Republican; Strongly Republican).

Footnotes

[1] Note that by using "racially typical" or "racial typicality" terminology in this paper I do not refer to any sort of anthropological or biological notion of racial or ethnic typicality. Rather, I refer to what people *perceive* as typical facial phenotypic appearance of different ethnic groups (e.g., African American and European American).

[2] Note that from here on, facial physiognomy refers to a set of features (e.g., width of nose, size of lips and eyes, etc.) rather than one specific facial feature.

[3] Both Stepanova and Strube (2009) and Stepanova et al (2008) used one-dimensional typicality scales—where a *more* African American rating signifies a *less* European American rating—which will be also used in the current research. To determine the validity of the underlying unidimensionality assumption, Stepanova and Strube (2009) had participants rate the facial stimuli for their racial typicality on two separate scales: from 1 (*Not at all African American*) to 7(*Very African American*) and from 1 (*Not at all European American*) to 7 (*Very European American*). The key finding here was that the high physiognomy (HP) Afrocentric faces were rated as less European American than the low physiognomy (LP) Afrocentric faces, which in turn, were rated as less European American than the Eurocentric faces. By contrast, the African-American rating scale showed the opposite pattern: HP Afrocentric faces were followed by LP Afrocentric faces and Eurocentric faces. Furthermore dark faces were rated as less European American than the light faces. This pattern was reversed for the African American rating scale: dark faces were followed up by light faces. We also calculated the between-participants correlations between the two ratings (European American vs. African American) for each of the six faces, controlling for color presentation mode. Each of the correlations was highly significant and ranged from -.55 to -.73 (mean r= -.66) indicating that when participants rated a given face as more African American, they also rated it as less European American. Perhaps most importantly, we also calculated within-participants correlations for the two types of ratings across the six faces. These correlations show, at the individual level, how participants used the two scales when judging the collection of faces. These correlations ranged from r=-1.00 to r=-.46, with a mean of r= -.89. In other words, most participants used the scales as if they were parallel but inverse measures of the same construct, providing strong supportive evidence for a one-dimensional typicality rating scale.

[4] Testing continuous variables' interactions requires a two-step procedure. In the first step, the continuous predictors, but not their product, are tested for significance. This provides a test of continuous variables main effects. On the second step, their product is entered to test the interaction. At the second step, the main effects are no longer interpretable, because they have the product partialled—only the product is of interest. This approach was taken for all continuous variable predictors in all analyses.

[5] When a predictor was continuous, the regression equation was used to generate predicted responses in this and all subsequent figures. "Low" and "high" values for the continuous predictor were defined as one standard deviation below (-1 *SD*) and one standard deviation above (+1 *SD*) the continuous variable mean.

[6]The present findings are also interesting in relation to recent neuroscience findings employing event-related potentials (ERPs) (Balas & Nelson, 2010). This work indicates that skin color, independently of facial features, modulates the N170 and N250 components implicated in facial processing, providing evidence that skin color might independently affect facial categorization. Note though that these studies (see also Bar-Haim, Saidel, & Yovel, 2009; Gitter & Satow, 1969; Stepanova & Strube, 2009) used only a very few levels of skin color and facial physiognomy, usually only two or three, whereas the current research employed a more sensitive manipulation of skin color and facial physiognomy. Additional supportive evidence is provided by Ronquillo, Denson, Lickel, Lu, Nandy, and Maddox (2007), who reported an interactive influence of skin color and race on amygdala activity in a fMRI study on face perception.

[7] Note that moderation by participant's ethnicity was somewhat identical to moderation by implicit and explicit racial attitudes, as African American participants had more positive implicit and explicit racial attitudes, and European American participants had less positive implicit and explicit racial attitudes.