7-27-2016

Pictorial Race Activiation In Priming Measures

Elena V. Stepanova
University of Southern Mississippi, elena.stepanova@usm.edu

Michael J. Strube
Washington University

Laura E. Clote
Washington University in St. Louis

Daniel Limes
Washington University in St. Louis

Pictorial Race Activation in Priming Measures

Elena V. Stepanova

The University of Southern Mississippi

Michael J Strube, Laura E. Clote, and Daniel Limes*

Washington University in Saint Louis

Author Note

*Contributions of the third and fourth authors are equal, and their names are arranged in alphabetical order.

We would like to thank the following individuals who kindly sent us their facial stimuli: Nazar Akrami, Jamie Barden, J.-C. Croizet, Dee Lisa Ann Cothran, Joshua Correll, Patricia Devine, Travis L. Dixon, Jack Dovidio, Jennifer L. Eberhardt, Bo Ekehammar, Susan Fiske, Anthony Greenwald, Curtis Hardin, Kurt Hugenberg, Charles Judd, Kerry Kawakami, Susan Klonis, Alan Lambert, Robert Livingston, William Maddux, Nancy Geyelin Margie, Heidi McGlothlin, Gordon Moskowitz, Brian Nosek, Keith B. Payne, Ashby E. Plant, Gillian Rhodes, Laura Scherer, Stacey Sinclair, Steven Spencer, Tracie Stewart, and Natalie Wyer. We also would like to thank those who responded to one of our listserve requests: Julian Degner, Jeremy Heider, Gregory Maio and Aaron Smith-McLallen.

We also would like to extend our thanks to the following research assistants: Christobal Cruz (The University of Southern Mississippi); Jeanellys Ramos, Brittany Chase and Katherine Harris (Florida Gulf Coast University); and Daniel Brick and Olivia Rehberger (University of Missouri – Columbia).

Correspondence concerning this article should be addressed to Elena Stepanova, Department of Psychology, The University of Southern Mississippi, Owings-McQuagge Hall,

118 College Drive, #5025, Hattiesburg, MS 39406, USA. Phone: 601-266-4342. Fax: 601-266-5580. Email: Elena.Stepanova@USM.edu

Abstract

This review explores characteristics of facial primes employed in priming studies of racial prejudice and stereotyping. It addresses the role of perceptual, cue-based processing of visual stimuli characteristics in altering racial typicality, and the effects of different moderators. The authors document the nature of variability in primes and moderators used in priming studies (N = 96) up to 2009. Methodological and conceptual implications are discussed, along with gaps in the field. Better control over facial primes employed, more accuracy in reporting and open access to procedural information are suggested in an effort to improve the state of racial priming research.

Keywords: priming; pictorial primes; implicit and automatic measures; racial prejudice and stereotyping

Pictorial Race Activation in Priming Measures

In studies of implicit racial prejudice and stereotyping, participants are frequently presented with Black and White faces in a variety of experimental priming paradigms. These faces are commonly used to activate a racial concept (e.g., "Black" or "White") in order to examine conscious and unconscious processing of racial stimuli and their effects on affect, attitudes, and behavior. However, pictorial primes vary along a number of dimensions such as racial typicality, size of the stimuli presented, mode of color presentation, degree of schematic presentation, mode of stimuli creation, and whether or not pre-testing of stimuli was performed on any dimensions. The purpose of this paper is to review research on racial priming in studies of implicit prejudice and stereotyping specifically toward Black individuals, with a focus on *pictorial primes*, that is, pictures of Black and/or White faces. We will address the potential importance of facial stimuli characteristics. The major goals of the review are (a) to describe relevant theories that address why primes are important to study, (b) to document variability in primes and characteristics of studies by tabulating these potential moderators, and (c) to identify implications of these methodological choices for current and future research in this area.

Implicit Measures of Stereotyping and Prejudice

Racial prejudice and discrimination are critical social problems, but given social desirability concerns, people are not often willing to admit to prejudicial attitudes. To address this problem, *implicit* methods have been developed (see Blair, 2001; Fazio, 2001; Fazio & Olson, 2003 for review). Implicit measures avoid asking participants about their attitudes or stereotypes explicitly. Instead they indirectly access attitudes and stereotypes. Their measurement outcomes reflect attitudes and stereotypes (a) that people might not be aware are being measured, (b) to which people might have only limited access, and (c) over which people might not be able to exert control (see De Houwer, 2006, Fazio & Olson, 2003). Research

employing implicit measures generally finds substantial evidence of negative evaluations of Blacks (e.g., Dovidio, Kawakami, Johnson, Johnson & Howard, 1997; Fazio, Jackson, Dunton, & Williams, 1995; Payne, 2001). The most commonly employed procedures are the Implicit Associations Test (IAT) (Greenwald, McGee, & Schwartz, 1998) and priming tasks (see Wittenbrink, 2007, De Houwer, 2006 for detailed overviews of the priming tasks).

Generally, in priming tasks that investigate racial biases a participant is presented with a racial stimulus (prime) followed by a target. In an evaluative judgment task the target is a positively or negatively valenced word. In a lexical decision task the target is a racial stereotypic trait, a word unrelated to race/racial stereotypes or a non-word. A participant makes a "good" versus "bad" judgment (in an evaluative judgment task) or a "word" versus "non-word" judgment (in a lexical decision task) about the target by selecting a corresponding key on a keyboard or by making a verbal response. There are variations of the evaluative judgment task, for example, when targets are positively and negatively valenced words, and additional target words (e.g., describing houses) and the judgment is whether the target word is a descriptor of a person or a house (Dovidio et al., 1997 type of task). There are some other tasks that use pictorial targets such as the "guns-tools"/Weapons Identification Task (Payne, 2001) in which the target of judgment is a gun or a tool and the type of decision is likewise "gun" or "tool" or the affect misattribution procedure (Payne, Cheng, Govorun, & Stewart, 2005) in which the target of judgment is a neutral Chinese ideograph and the decision is whether the ideograph is more or less pleasant than the average Chinese ideograph. Note that our descriptions of the tasks are abbreviated and capture the gist of priming procedures, without specific details (e.g., presence of masking before and after prime presentation, positions of targets and primes, etc.) and we refer readers to original studies for detailed procedures.

Some experimental paradigms are not based on the task types described (e.g., Bargh, Chen & Burrows, 1996; Brown et al., 2003; Chen and Bargh, 1997, Colcombe, 2000; Massey, 2003; Spencer et al., 1998; Phelps, O'Connor, Cunningham, Funayama, Getenby, Gore, & Banaji, 2002). We will refer to them as "non-classical" priming tasks, as they do not necessarily have a target of judgment, as opposed to the "classical" priming tasks that always do have a prime, a target and a judgment required.

This variety of tasks can have important conceptual implications. De Houwer (2009) argues that in affective priming tasks, such as an evaluative judgment task, there is an irrelevant stimulus-response (in)compatibility. As we previously described, a prime (i.e., stimulus) is followed by a target that needs to be classified by valence. The valence of the prime is compatible with the valence of the response on congruent trials (e.g., when Black faces are followed by negatively valenced words and the correct response is "bad") and not compatible on incongruent trials (e.g., when Black faces are followed by positively valenced words and the correct response is "good). The congruency of the task-*irrelevant* feature of a stimulus (e.g., prime valence) to a response valence varies. The most important implication is that the concept (e.g., White or Black) we are measuring attitudes toward is "implemented at the level of the irrelevant stimulus feature" (De Houwer, 2009, p. 379), that is, at the level of facial primes. De Houwer posits that due to different characteristics of facial primes, noise can be introduced into the measure of attitudes: Participants might attend to various features of stimuli, other than a category or concept under investigation. Essentially, *category/concept salience* is reduced in that type of procedure (De Houwer, 2003; De Houwer, 2009) in comparison to paradigms such as the Implicit Association Test (IAT, Greenwald, McGhee & Schwartz, 1998), outcomes of which depend on the valence of the categories under investigation (e.g., Black or White) rather than the properties of the exemplars (e.g., faces of Black and White individuals).

Nature of Primes and Racial Typicality

Most implicit racial bias research is based on the assumption that the *nature of the prime* such as racial characteristics (e.g., a Black face versus a White face) will influence participants' responses to the target (e.g., positive and negative words). Previous research has explored the extent to which the *degree* of racial typicality [1] in the primes might affect implicit and explicit judgments or evaluations (e.g., Blair, Judd, Sadler, & Jenkins, 2002; Blair, Judd, & Chapleau, 2004; Blair, Judd, & Fallman, 2004; Blair, Chapleau, & Judd, 2005; Dasgupta, Banaji, & Abelson, 1999; Dixxon & Maddox, 2005; Eberhardt, Davies, Purdie-Vaughns, & Johnson, 2006; Kahn & Davies, 2010; Livingston & Brewer, 2002; Ma & Correll, 2011; Maddox & Gray, 2002; Oliver, Jackson, Moses, & Dangerfield, 2004). These researchers manipulate racial categories by varying the extent to which a particular face is representative of a specific racial group, for example, by manipulating darkness of skin color, width and shape of nose, eyes and lips, eyebrow height and separation, and hair texture and quantity. Faces are constructed or selected to vary in their "Afrocentric" physiognomy, prototypicality, or racial phenotypicality. Even though the concept of racial typicality is very loosely defined and represented by different combinations of various features from one study to the next, images with more Afrocentric features generally elicit more negative evaluations. This holds for both traditional explicit and implicit measures of attitudes (for review, see Maddox, 2004; Maddox & Dukes, 2008).

The importance of this work lies in several key features. First, it demonstrates that when racial typicality is manipulated, it influences implicit and explicit racial attitudes. Second, the stimulus features that can influence racial typicality are apparently quite varied. Third, researchers often do not manipulate racial typicality on purpose but instead use a set of stimuli that have typicality implications arising from the various features of the stimuli.

Racial Typicality and Cue-Based Perceptual Processing

The impact of racial typicality on priming effects was originally studied by Livingston and Brewer (2002). Livingston and Brewer defined prototypicality as a set of features including hair structure, skin color, shape and size of the nose/lips/chin, and shape and size of the eyes. African American primes were divided into two groups: low prototypicality (LP) and high prototypicality (HP). In an evaluative priming task, automatic evaluations of HP African Americans were more negative than those of LP African Americans; LP African American primes received a pattern of facilitation similar to that of White primes. However, when participants were asked to perform a racial categorization task prior to performing an evaluative judgment task, both LP and HP primes received similar negative automatic evaluations. Livingston and Brewer suggested that facial features per se elicit affective responses—indicative of *perceptual cue-based* processing—that sometimes are independent of category-based processing (based on the associations of racial categories and valenced evaluations) altogether.

Further support was provided by recent studies employing affective priming tasks (Hagiwara, Kashy, & Cesario, 2012; Stepanova & Strube, 2012). In these studies the authors attempted to disentangle the effects of several facial features defining racial typicality that were confounded in previous work (i.e., skin color and facial physiognomy). This was done by manipulating skin tone and other facial features of facial primes independently. Stepanova and Strube employed six faces with two levels of skin color (dark vs. light) and three levels of facial features (high Afrocentric vs. low Afrocentric vs. Eurocentric) as primes; while Hagiwara et al. (2012) selected faces with either very dark or very light skin color and then manipulated their lip thickness and nose width simultaneously to create two levels of facial features (high vs. low Afrocentric). These studies showed that people responded more negatively toward faces with dark skin color than faces with light skin color. Additionally, above and beyond the skin color effect, people responded more negatively toward primes with high Afrocentric facial features

than primes with low Afrocentric facial features. Skin color and facial features independently affected participants' responses in affective priming tasks. These findings indicate that effects of facial cues are independent in racial priming studies. More importantly, they suggest that faces used for activation of racial categories might not necessarily tap into a general racial category "Black" and facial features (and possibly other visual cues) can evoke other types of processing. This route was described by a model of racial phenotypicality (Maddox, 2004; Maddox & Dukes, 2008) as a feature-based route of racial typicality, as opposed to a category-based route. This model suggests that differential levels of responding in priming tasks to primes of various racial typicality might be due to (a) activation of a superordinate category and subsequent activation of meaningful subcategories and thoughts and feelings associated with them; or (b) a direct association between facial cues and cognitions and feelings associated with them independent of the category activation. Note that multiple views of mental representations can be proposed explaining how less or more racially typical facial primes produce racial bias effects in priming tasks (e.g., associate network models, schema models, connectionist models among others, see a recent literature review on models of mental representations in implicit social cognition research by Payne and Cameron, 2012). Although the complexity, structure and mechanisms of these models are beyond the scope of this review, multiple models suggest that certain mental representations (e.g., dark skin tone) can be linked to evaluative/trait associations directly. If a certain marker (e.g., dark skin color) is always associated with a category (Black) evoking negative affect or a certain trait, eventually the marker will get associated with negative affect or the trait, as implicit attitudes accumulate slowly and incrementally through repeated associations of mental representations (e.g., Rydell & McConnell, 2006; Rydell, McConnell, Mackie, & Strain, 2006).

The use of more or less typical faces might influence the effects obtained in priming tasks; resulting in higher or lower estimates of racial bias correspondingly. We believe that *cue-based processing* is especially likely to occur in priming tasks, because the concept (e.g., White or Black) we are measuring attitudes toward is implemented at the level of facial primes and primes are commonly presented for a very brief period of time, as short as 13 ms. In sum, facial stimuli are not processed categorically and their nature (e.g., unique visual characteristics) may induce quite different cue-based processing depending on the features present.

Perceptual Cue-Based Processing: Typicality and Beyond

Besides features varying on racial typicality, other *visual* characteristics of stimuli might also be important. For example, Stepanova and Strube (2009) found that whether stimuli were presented in gray-scale or color affected perceivers' conscious judgments of racial typicality. 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 than in grayscale. Analogously, facial cues can have differential impact on priming effects. Stepanova and Strube (2012) found that only in the color mode was there bias toward dark skin tone primes in an affective priming task. Gray-scale mode might downplay the racial effects usually obtained: when skin tone is presented in the gray-scale mode, there is less negativity toward dark skin primes than in the color mode.

When cue-based processing is very broadly defined and encompasses a wide variety of visual cues, it suggests that quite a number of visual characteristics can alter the racial typicality of facial primes employed, thus leading researchers to under or overestimate racial biases. We discussed mode of color presentation. Other visual characteristics that may have this impact are (but are not limited to) size of the stimuli presentation (cropped vs. non-cropped), degree of

schematic presentation, and mode of stimuli creation (actual photograph, drawing, etc.). For example, when cropped images are presented, it is possible that certain facial features will be weighted differently than when a full head shot is used, producing primes with different implications for racial typicality. Analogously, when a face presented is very schematic, it might be evaluated as more racially typical than a less schematic facial stimulus (i.e., an actual photograph) because racial features are especially prominent or salient. Also, if primes are not matched on likeability, attractiveness, affective expression, or aggression, stronger racial effects can be expected potentially due to researchers' selection biases. These speculations await empirical tests but highlight that "irrelevant" aspects of primes may not be so irrelevant after all.

There are several important caveats to consider in our application of cue-based perceptual processing. First, there is the possibility that visual characteristics can change something else (e.g., perceived threat, attractiveness, age, or other characteristics) about a prime—not racial typicality—that then could lead to differential priming effects. Second, racial typicality as well as attractiveness, age, threat, etc., can potentially vary as a function of contextual manipulations in the various procedures employed. Currently, there is not enough empirical evidence for typicality effects emerging under various contextual manipulations (but see Livingston & Brewer, 2002, Experiment 4). Third, we have focused on visual cues that might affect the typicality of Black primes. Note that racial priming effects are obtained as a result of comparing participants' responses to targets when primed with Black *versus* White primes.[2] Accordingly, the impact of visual characteristics must be viewed in the context of the *difference between* responses to Black and White stimuli. Some characteristics might make White faces appear as more White AND Black faces as more Black, OR White faces as less White and Black faces as less Black, OR produce several other possible combinations of changes.

Additional Moderators

A priming task is very often just a part of a larger experimental procedure, and frequently other tasks and measures are also included, as well as additional manipulations within the priming task itself (*within*-study moderators). Hence, there is a multitude of moderators in any given study that could attenuate priming effects.

There are also several types of moderators that vary *between* studies and address various detailed aspects of procedure (*how* and *what* type of moderators), that is, *general study characteristics*. These include characteristics of participants (e.g., number, age, gender, student status, ethnic and racial make-up, attrition rate), types of priming tasks (e.g., shoot-don't shoot/First Person Shooter task, crime-related object identification, etc.), if a priming manipulation is a between or within-subjects manipulation, and types of dependent variables. Some of these (e.g., stimuli mode of presentation and information on pre-testing of stimuli), of course, have important conceptual implications as we described previously. All of them are potentially important in comparing priming effects across studies. Many of these moderators have been given attention in the priming literature, but quite a number have been ignored. Indeed, it is largely unknown how much variability exists in the priming literature regarding the vast majority of these moderators.

Inventory of Moderators in Priming Research

The purpose of this qualitative and quantitative component is to describe the representation and variability of pictorial primes and some moderators in racial priming studies published up to 2009. We reviewed the available literature, and identified and tabulated the presence of many potential moderators. We focused on documenting the number and nature of the primes and some moderators. This descriptive summary highlights conceptual and methodological implications for current and future research in the area.

Method

Research articles and theses were searched using PsycInfo and ProQuest to locate relevant studies published prior to the year of 2007. Papers in press at the time were published later in 2007-2009 and are included as well. We used various combinations of the following key words: *priming, Black, African American, race* or *racial*, *implicit, automatic, prejudice, attitudes, bias*. Also, Science Citation Index, Social Sciences Citation Index and Arts and Humanities Citation Index (using Fazio et al. 1995; and Dovidio et al. 1997, earliest priming studies, for cited reference search) and cross-referencing were used. The initial list of studies had very broad inclusion criteria, and some studies were subsequently excluded upon further examination. If selected studies did not provide the pictorial stimuli that were used, the authors of the manuscripts were contacted by mail to provide their original stimuli and any data on their norming. If stimuli were not received, we followed up with email, then again with regular mail and email once more. To obtain unpublished studies, we sent an email message to professional electronic mailing lists of the Society for Personality and Social Psychology, the Society for the Psychological Study of Social Issues, the Society of Experimental Social Psychology and posted a request at the Professional Discussion Forum at the Social Psychology Network. Original methodology, results, racial pictorial primes and any pre-testing data were requested in any form.

Exclusion Criteria

We excluded studies based on their procedural characteristics: when faces were not primes, but targets of judgment, when researchers used IAT, SC-IAT or Go/No-go Tasks as implicit measurement tasks, and when the Black primes were not used. [3] Also, we excluded studies in which the target of judgment and prime were confounded. Note that we did not exclude studies employing tasks such as shoot-don't shoot/First Person Shooter task where participants are presented with individuals holding either guns or innocuous objects and have to make a "to shoot" or "not to shoot" decision (e.g., Corell et al., 2002; Corell et al., 2006;

Greenwald et al., 2003), even though the primes and the targets of judgment were confounded. These studies were included because they tap into more automatic racial attitudes and stereotypic associations. When participants are instructed to respond to a target of judgment within a limited timeframe or as soon as possible, their cognitive resources are limited and self-control strategies/self-presentation concerns get disrupted (Bodenhausen, 1990; Govorun & Payne, 2006; Richeson et al., 2003).

Coding Guide

The coding guide included (a) general study characteristics, (b) characteristics of participants, (c) procedural aspects of the studies, (d) contextual variables, (e) characteristics of primes, and (f) information on pre-testing of facial stimuli. Two independent coders, the third and the fourth author, performed the coding of studies. A pilot coding was conducted on 23 randomly selected studies. Inter-rater agreement (Cohen kappa, κ) was calculated for each coded variable and if it was lower than .80, inconsistencies were discussed with the coders and the coding guide was revised. After the revision, coders coded the remaining studies and inter-rater agreement was calculated again. Whenever two coders disagreed on any given decision, the first author served as a tie-breaker. When new information was uncovered, the coding guide was modified to accommodate it. Final analysis revealed that for most of the variables [4], kappas were acceptable ($\kappa \geq .60$) with only a very few ($N = 6$) below .41, which is considered below moderate agreement, but fair (Landis & Koch, 1977).

Results

The final sample size was 96[5]. Not all of the variables that were coded are presented here. Instead, we focus on those relevant to the issues discussed in this review [6].

Stimuli Received

We originally contacted 46 researchers with a request to provide stimuli and any information related to pre-testing. Five more researchers—usually, second, third, or fourth authors—were contacted later per request of original first authors/academic advisors or their lack of response. From these requests, we received 29 stimulus sets or samples (although we did not include all of these studies in the final analysis, see *exclusion criteria*). Some of these stimuli were used for more than one study. Therefore, our return rate was 63% of the original 46 requests, 57% of all 51 researchers' requests and 56% of all 52 requests.

Accessibility of Stimuli

Most of the studies did not provide any of the priming stimuli materials (68.8%), some provided only samples of primes (29.2%), and only 2.1% provided all stimuli within a manuscript. We had requested all stimuli from 84 studies that were not fully available for review (if no stimuli were reproduced or only samples were given in a paper and the authors were not using the publically available Nosek et al., 2002 stimuli). We received stimuli for 64.3% of those. Most of the received stimuli were complete sets (72.2%) with the remainder being samples (24.1 %) of stimuli. Even though the return rate was high, there were stimuli that were neither provided in the original sources nor could be obtained from the authors.

Stimuli Authorship and Origination

Stimuli authorship revealed that although some of the studies used the Nosek et al. (12.5%), Fazio et al. (12.5%) and Dovidio et al. (6.2%) stimuli (see Figure 1 for examples), most employed their own stimuli (68.8%). See sources of stimuli displayed in Figure 2.

Stimuli Gender and Ethnic Composition

All studies included in this review used Black European or African American primes because the focus of this review is pictorial priming assessing stereotypes and attitudes toward Blacks. Almost all studies included White primes as well (99%) with only one study comparing

African American and Asian American primes (Colcombe, 2003, Study 4). Additionally, some studies included Asian Americans (23.4%), Hispanic Americans (16%) and Alaskan Native/Native American/Pacific Islander (1.1%) primes. Most of the studies employed male primes only (64.6%) and only 31.2% employed both female and male primes. The average number of unique male primes per study was 20.97 (*SD* = 32.32) and the average number of unique female primes was 9.56 (*SD* = 15.63). The average number of unique primes per study was 26.80 (*SD* = 32.41).

Racial Typicality of Primes

Most of the studies did not mention the status of racial typicality/prior racial categorization of their primes (85.4%), but some (9.4%) mentioned that it was controlled, and only a few (5.2%) manipulated it. Most studies (92.7%) did not mention typicality of their non-White primes, although 2.1% mentioned the same degree of typicality for all non-White primes and 5.2% reported different degrees of typicality for all non-White primes.

Various Visual Properties of Stimuli

Below we present results for different properties of the stimuli, based on coder judgments. Most of the studies (55.2%) had realistic looking stimuli, 10.4% of the studies had schematic looking stimuli, 32.3% of the studies used stimuli that were not available and 2.1% of the studies used stimuli that were not accessible due to technical difficulties. Color presentation of primes also varied: 30.2% of the studies used stimuli in grayscale, 35.4% used stimuli in color, and in 34.4% this judgment could not be made due to a lack of information and technical difficulties. There was variability in size of the presentation as well (see Figure 3).

Affective expression of primes was manipulated in 3.1% of studies, was controlled to be constant (most commonly, neutral) in 45.8%, and in 51% it was not mentioned. In 14.6% of the sets, some or all primes had a smiling expression, in 62.5% of the sets neutral expressions were

present, and in 4.2% of the sets some or all primes had negative expressions (in all the remaining cases, primes either did not have the affective expression in question or it was impossible to tell).

Backgrounds

Sometimes researchers used a plain neutral background one normally would see behind a headshot. In 50.7% of the studies that background was controlled, in 1.4% the background was manipulated and in 47.9% the background was not mentioned. The background was the same for all primes in about a third of the studies (29.4%), the background was different in about a quarter of the studies (23.5%), and was not specified in the remaining studies (47.1%). The color distribution of backgrounds was the following: colored polychromatic (11.8%), grayscale or black/white (41.2%), and in 47.1% of studies information was not available. As for the background lightness, in 40.3% of studies Black primes were presented on light background, whereas in 37.5% of the studies White primes were presented on light background, analogously, in 9.7% of the studies Black primes were presented on dark background whereas in 12.5% of studies White primes were presented on a dark background. We tried to identify if authors used more than one color when they used polychromatic backgrounds. In 20.8% of the studies both Black and White primes were presented on a background of the same color, and in 23.3% of studies they were presented on backgrounds of multiple colors.

Pre-Testing of Stimuli

Figure 4 presents results for the pre-testing. The pre-testing on several properties was mentioned only in two studies. It was most commonly performed by independent participants (71.4%) or research participants themselves after the main study (5.7%). In 22.9% of studies there was no information on who performed pre-testing. Even if pre-testing was performed, numerical data are rarely reported. At least some kind of numerical data was reported in 31.6%

of those studies that mentioned pre-testing in text; 41.7% of reported data were in the form of means, 25% reported significance tests and 33.3% reported both means and significance testing.

Tasks and Targets of Judgment

The tasks employed in priming studies also show substantial variability (see Figure 5). If a task employed was a "classical" priming task, most of the studies used lexical targets (63.3%), and some-pictorial (35.4%), while only very few employed both (1.3%).

Variability in Participants

We recorded various demographics before and after any data were discarded by the authors. Here we report data for the final samples only (see Table 1 for demographics data before and after any data were discarded). Note that even though we provide original means, means within each of the subcategories often are not correct reflections of the data because authors did not always report precisely who was recruited or dropped from the final analysis. Therefore, there are many missing cells and this is not an accurate portrayal of recruitment and drop-out rates. The only numbers that were completely available were the original total number of participants and a final total number of participants. As Table 1 shows, the majority of participants were college students, female and White, with a mean age of 19.15. Final samples (after some participants were discarded) had the following ethnicity distribution. White participants were represented in 69.8% of the studies, not present in 1% of the studies, and in 29.2% studies it was impossible to tell. Only 13.5% of studies employed Black participants in final samples, whereas 76% did not, and in 10.4% it was impossible to tell. Some studies employed Asian Americans (15.6%), but most did not (53.1%), and it could not be determined in some cases (31.2%). Some studies employed Hispanic Americans (15.6%), but most did not (54.2%), and in 30.2% of studies it could not be determined. Only 2.1% of studies reported Alaskan Native/Native American/Pacific Islander participants, 66.7% did not, and it could not be

determined in 31.2% of studies. One percent of studies reported international participants, whereas 68.8% did not, and it was not possible to determine in 30.2% of studies. As for participants whose ethnicity was not listed, 28.1% of studies reported this fact, but 60.4% did not, and it was not possible to determine in 11.5% of studies. For participants' gender, only 1% of studies employed male participants only, 5.2% of studies used female participants only, 70.8% used participants of both genders, and in 22.9% a gender break-down was not reported.

Other Moderators

In most of the studies, primes were presented within-subjects (87.5%). In 2.1% of studies purpose disclosure (e.g., whether or not participants knew the purpose of studies) was manipulated between subjects and in 2.1% studies it was manipulated within subjects. In the majority of studies it was not manipulated (95.8%). In 2.1% of studies participants were instructed to attend to race or race was activated before the priming task and in 2.1% of the studies they were instructed to attend to race or race concept was activated during the priming task. Participants' attitudes were manipulated prior to the critical task (5.2% studies), and sometimes through prior conditioning (3.1%). These manipulations were largely employed to test whether categorical information (i.e., activation of the race concept) moderates priming effects.

Discussion

In this section, we will (a) discuss the variability of stimuli and some moderators and the accuracy of reporting in priming studies, (b) relate these to important conceptual and methodological issues, and (c) discuss implications for current and future research.

Variability in Stimuli

The return rate for stimuli was high, and it is very encouraging that researchers were willing to share their stimuli. However, in 44% of all requests, we were not able to get the

stimuli and make any inferences about them. It is unfortunate that some stimuli could not be obtained, especially when they accounted for a large set of studies (e.g., Fazio et al., 1995 stimuli). Most of the studies employing pictorial primes did not provide any of their stimuli for readers (68.8%) —neither a full set of stimuli nor samples, and only a few studies provided either all stimuli (2.1%) or samples (29.2%). This is a serious limitation that may have multiple reasons. Some researchers probably did not think that the visual variations in their stimuli were of interest to others. Others probably were not aware that visual characteristics might affect the nature of processing and the results obtained. Perhaps, an exclusion of stimuli was an editorial decision driven by limited journal space. Regardless of the reasons, we argue that all stimuli should be provided in the original manuscripts or made easily available by other means.

Almost all of the studies used White and Black primes, with other groups rarely included (the two largest "other" groups were Asians and Hispanics). This largely limits conclusions about racial bias to that by Whites toward Blacks and leaves open much interesting work on other forms of racial prejudice. We specifically concentrated on White-Black racial effects, and only included corresponding studies. However, inclusion of other primes might provide opportunities to refine understanding of the processes and moderators of racial bias.

Most of the studies employed male primes, likewise limiting conclusions about racial biases. Black male and female primes might evoke different stereotypes and affective associations because there are gender-based subtypes for Black stereotypes (see Schneider, 2004). Most of the participants were young White female college students. So, current racial priming studies tap into one distinctive type of bias—that experienced by young White college women towards Black males. This type of processing might not be representative of the most common forms of real life prejudice, stereotyping, and most importantly, discrimination. Researchers (e.g., Pratto & Sidanius 2006; Sidanius & Pratto, 1999) argue that discrimination

based on arbitrary social distinctions not related to either gender or age, such as race, nationality, ethnicity, class, clan, or caste is mostly male-on-male, in which perpetrators of violence and discrimination are dominant males and targets are subordinate males. Although most of the research described here does employ male targets, the participants in these studies are not the most common perpetrators of intergroup discrimination. Currently, these middle and upper-class young college women's views of Black men are commonly presented as reflective of the general population and male segments of population, which may not be accurate. As this review indicated, samples that include male participants comprise 71.8% of studies (with 22% of studies not reporting gender of participants), and when male participants were included, they comprised about 37.6% of an average sample. These results raise questions about the generalizability of findings and illustrate that a gap exists between the forms of prejudice studied and the forms that are most frequently exhibited in everyday life.

The particular primes that have been used in research have likewise limited the inferences that can be drawn. Even though many studies employed the Nosek et al., Fazio et al., and to a lesser extent, the Dovidio et al. stimuli, many more studies (68.8%) used their own idiosyncratic stimuli (Figure 1 provides examples of Nosek et al., Dovidio et al. and other stimuli used in racial priming research). The likely impact on replication of major results is clear, but perhaps more important is the likely inability to replicate subtle findings that speak to important process differences. They may be hidden among the variability induced by primes that vary along many different dimensions besides skin tone or physiognomy. Consider the difficulties encountered by researchers trying to replicate or extend a study using their own stimuli (and not the original stimuli employed in the study they are trying to replicate), especially given the recent controversy over lack of replication in social psychology (for review, see Earp & Trafimow, 2015). The replication could be hindered if the stimuli employed in the new study differed on

one or more important dimensions (e.g., full facial versus cropped stimuli, mode of color presentation, color versus grey-scale background, etc.) compared to the stimuli used in the original study. The failure to replicate effects from the original study, or, finding new effects in the follow-up study, would then be difficult to interpret. In one sense, the studies are not directly comparable. The stimuli in the original study and the extension study may vary on a dimension that is crucial to the replication or to the new effect. As research moves beyond the mere demonstration of priming effects to explore the moderators of these effects, the ability to compare studies becomes more crucial. For example, different primes might interact with other experimental variables, and such interactions would distort the theoretical conclusions drawn from experiments and make comparisons across studies difficult.

We also showed that most of the studies used actual pictures as primes, but many other studies also employed primes based on morphed images, yearbook pictures, magazine cutouts, magazine cutouts AND actual pictures, combining for a total of 84.3% photograph-based primes. This type of presentation is more realistic than other means of stimulus presentation (e.g., computer generated sketches or animations), and probably exaggerates racial typicality less. Nonetheless, we concluded that only 55.2% of the studies had realistic-looking stimuli (note that it was impossible to view stimuli in 34.4% of studies). It is possible that processing in studies with schematic stimuli was more category-based, and perceptual processing (based on facial features of stimuli) either did not play a role or was diminished. As we previously noted, less realistic/more schematic stimuli can potentially lead to exaggeration of racial typicality, as they do not provide as much individuating information, making it easier to process stimuli categorically, and leading to stronger racial bias effects.

Primes also varied in whether they were presented in color or not. Some studies used stimuli in grayscale (30.2%) and some (35.4%) in color. However, there is evidence that color

and gray-scale stimuli produce different outcomes in implicit racial evaluations. Smith-McLallen et al. (2006), for example, found that when gray-scale cropped pictures were compared with color non-cropped photographs in two forms of the Implicit Association Test (IAT, Greenwald et al., 1998), the IAT results were not significantly correlated, even though they both produced expected race effects. Recent research has indicated that other properties of stimuli might be perceived differently (Stepanova & Strube, 2009) and differentially impact implicit racial evaluations (Stepanova & Strube, 2012) depending upon the mode of color presentation. Thus, different modes of color presentation can be sources of additional variability in racial priming studies, and racial biases are underestimated when gray-scale primes are employed. To address ecological validity concerns, we suggest use of color photographs in future research.

There was considerable variability in the size of primes. Size can potentially amplify racial typicality of racial primes, with increased size providing details that might ease or speed up racial categorization and racial processing. Most of the studies employed head shots, followed by cropped faces (mouth not shown), followed by full figures (face clearly seen), and cropped faces (mouth shown) and torsos, face clearly shown. For example, 19.8% of studies use faces that were severely cropped (mostly Nozek et al. stimuli). Such cropped stimuli that provide only limited information might have implications for perception of racial typicality, accentuating some features that full headshots do not, and potentially overestimating racial bias.

Similar variability was found with other facial characteristics. For example, affective expression was controlled in less than half of the cases. Often there were smiling and neutral primes that were used in the same set, and there was no control for affective expression. Effects of affective expressions on racial stereotyping and prejudice have been documented (Cothran, 2005; Hutchings & Haddock, 2008; Hugenberg & Bodenhausen, 2003; 2004) and so this feature of primes should be important for researchers to control in their stimuli.

Other examples of features not often controlled are the background behind the primes: Color distribution of the background was quite varied. Indeed, very often it was different for all primes (without any control), and more Black primes than White primes appeared on light background and more White primes than Black primes appeared on a dark background. Based on simultaneous lightness contrast, where perception of the lightness of an object depends upon it immediate surroundings, if a dark face is presented on a light background, it will be perceived as a darker one than if it is presented on a dark background; likewise, if a light face is presented on a dark background, it will be perceived as a lighter one than when it is presented on a light background. If color of backgrounds is not very well controlled, pure simultaneous lightness contrast effects might occur. This can alter color perception of primes, perhaps exaggerating or attenuating their "Blackness" or "Whiteness" and leading to stronger racial biases.

Most of the studies did not mention the status of their primes in terms of racial typicality or racial categorization (85.4%) and pre-testing for racial typicality/racial categorization was very low (13.5%). Indeed, it is largely unclear how researchers dealt with issues of racial typicality or if they are even aware of the need to control for that characteristic. It appears that authors do not consider the issue of racial typicality and the implications that follow if some of their primes are more typical than others. Similar conclusions hold for other dimensions. Only about half of the studies mentioned any kind of pre-testing. This inattention to pre-testing applies to a large number of potentially important characteristics (e.g., attractiveness, age, likeability, friendliness, aggression/violence/hostility, intelligence). Moreover, even if pre-testing is performed, numerical data are rarely reported.

The larger conclusion is that researchers do not attend adequately to stimulus variability. Systematic investigation of various stimulus variables awaits future empirical research. Yet, the various sources of that variability may have important implications for racial typicality (and

potentially attractiveness, threat, etc.) that could affect racial effects. Note that we do not dispute previously found robust evidence of racial biases in priming research, yet one could argue that researchers select stimuli that helped them to find these effects (e.g., by selecting stimuli that exaggerate racial typicality of primes).

De Houwer (2009) argues that to tap into attitudes towards categories, researchers should employ tasks that use category labels rather than facial primes. Even though that approach might produce racial effects that are devoid of any extraneous features of stimuli in question, it might reduce generalizability of the findings. We are routinely presented with exemplars of racial categories in real life, and elimination of pictorial priming tasks might limit the field to a somewhat narrow conceptualization of implicit attitudes. We are also aware of the dangers of too much standardization; that is, employing a single set of stimuli and one or two procedures across many studies. Rather than abandoning pictorial priming or employing one set of stimuli, we recommend that authors make their stimuli more available, describe them more thoroughly, engage in more extensive pre-testing of them and share that with the wider scientific community. Researchers should consider the potential implications when selecting faces for priming studies.

Variability in Potential Moderators

Most of the studies had the crucial racial manipulation as a within-subjects factor, which made them more powerful than those employing between–subjects manipulations. Knowledge of the statistical advantage of within-subject racial manipulations might deter researchers from exploring and developing other types of tasks, especially those that might better reflect the way that racial judgments are made outside the laboratory. More specifically, racial priming research largely investigates comparative judgments in which participants are encouraged to compare or contrast primes from different racial groups, relying mainly on a within-subjects racial manipulation. Outside the laboratory, however, judgments are often not comparative—a

member of a racial group is encountered without any reference to another group. It might be argued that absolute judgments are turned into comparative judgments by perceivers, but that has not been demonstrated in research. If it could be demonstrated, then the current task focus is less problematic, but if absolute judgments are fundamentally different, then current research largely does not speak to an important class of real-world racial decision making situations.

When categorical information is activated prior to or during priming tasks (e.g., Amodio, 2003; Amodio et al., 2004; Olson & Fazio, 2003; Payne et al., 2005), primes are either processed categorically, cue-based processing is overridden, and pictorial characteristics of stimuli do not matter OR processing primes categorically might cause features of those primes to be exaggerated to fit with the category prototype (Corneille, Huart, Becquart, & Brédart, 2004). However, only a few studies applied categorical activation before or during the procedure (less than 10%). Therefore, it is important to address how different characteristics of primes affect the racial typicality, threat, affect, attractiveness of racial stimuli.

Accuracy of Reporting

We have addressed some methodological and conceptual issues related to variability in primes and some moderators in racial priming studies. Although the information coded for this review allowed us to draw conclusions about the state of the field, almost as important was the information that we could not gather. Information pertaining to stimuli, tasks, participants, etc., was missing in much of this literature. Most of the time, it was not possible to determine an exact composition of the original samples (e.g., age, race, gender). Very often limited information was given on gender and race of participants who were dropped from studies and, consequently, the race and gender compositions of the final samples. Statistical information was often missing and authors frequently were non-compliant with the APA standards (American Psychological Association, 2001) in reporting their data. Lack of both statistical and descriptive

information raises some questions about transparency of research, not only in the field of implicit prejudice and stereotyping, but more generally for psychology as a scientific field.

Current Directions and Final Thoughts

Researchers continue to rely upon pictorial priming tasks to measure racial biases (e.g., Glaser & Knowles, 2008; Guinote, Willis, & Martellotta, 2010; Sadler, Correll, Park, & Judd, 2012; Schlauch, Lang, Plant, Christensen, & Donohue, 2009; Smith, Dijksterhuis, & Chaiken, 2008). There is still a great variability in the primes (and methods). That variability reflects a multitude of real-life features and phenomena and suggests considerable generalizability for racial priming effects. Nonetheless, this variability only distantly resembles the variability outside the laboratory. Accordingly, a thorough understanding of it is essential. In that respect, this review should not be taken as a condemnation of the variability described. It is a healthy feature of this research. What is lacking, however, is the clear and thorough description of the sources of the variability that will enable researchers to capitalize on it and use it to advance our field. Given the recent attention to replicability in psychology in general (Plasher & Wagenmakers, 2012) and of priming effects in particular (Sherman, 2014), emergence of normed face databases of individuals of different races (e.g., Chicago Face Database; Ma, Correll, & Wittenbrink, 2015), and movement toward Open Access (Reis, 2007) and more stringent journal reporting standards (American Psychological Association, 2008) supported by multiple journals and open access research platforms (e.g., Open Science Framework), the field has a potential to move forward in the right direction.

References

*Studies from these manuscripts are included in the analysis.

American Psychological Association. (2001). *Publication Manual of the American Psychological Association* (5th ed.). Washington, DC: Author.

American Psychological Association Publication and Communication Board Working Group on Journal Article Reporting Standards (2008). Reporting standards for research in psychology: Why do we need them? What might they be? *American Psychologist, 63,* 839-851.

*Amodio, D. M. (2003). Neural signals for the detection of unintentional race bias. *Dissertation Abstracts International, 64* (08), 4106B. (UMI No. 3101260)

*Amodio, D. M., Harmon-Jones, E., & Devine, P. G. (2003). Individual differences in the activation and control of affective race bias as assessed by startle eye-blink response and self-report. *Journal of Personality and Social Psychology, 84*(4), 738-753. doi: 10.1037/0022-3514.84.4.738

*Amodio, D. M., Harmon-Jones, E., Devine, P. G., Curtin, J. J., Hartley, S. L., & Covery, A. E. (2004). Neural signals for the detection of unintentional race bias. *Psychological Science, 15,* 88-93. doi: 10.1111/j.0963-7214.2004.01502003.x

*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

*Barden, J., Maddux, W. W., Petty, R. E., & Brewer, M. B. (2004). Contextual moderation of racial bias: The impact of social roles on controlled and automatically activated attitudes. *Journal of Personality and Social Psychology, 87*(1), 5-22. doi: 10.1037/0022-3514.87.1.5

Blair, I. V. (2001). Implicit stereotypes and prejudice. In G. B. Moskowitz (Ed.), *Cognitive social psychology. The Princeton symposium on the legacy and future of social cognition* (pp.359-374). Mahwah, New Jersey: Erlbaum.

Blair, I. V., Chapleau, K. M., & Judd, C. M. (2005). The use of Afrocentric features as cues for judgment in the presence of diagnostic information. *European Journal of Social* Psychology, 35, 59-68. doi: 10.1002/ejsp.232

Blair, I. V., Judd, C. M., & Chapleau, K. M. (2004). The influence of Afrocentric facial features in criminal sentencing. *Psychological Science, 15 (10),* 674-679. doi: 10.1111/j.0956-7976.2004.00739.x

Blair, I. V., Judd, C. M., & Fallman J. F. (2004). The automaticity of race and Afrocentric facial features in social judgments. *Journal of Personality and Social Psychology, 87*, 763-778. doi: 10.1037/0022-3514.87.6.763

Blair, I. V., Judd, C. M., Sadler, & Jenkins (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

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

*Boyd, B. J. (2000). Implicit racist attitudes towards African Americans and American Indians among caucasian university students. *Dissertation Abstracts International, 65* (03), 1594B. (UMI No. 3127148)

*Brown, R., Croizet, J., Bohner, G., Fournet, M., & Payne, A. (2003). Automatic category activation and social behavior: The moderating role of prejudiced beliefs. *Social Cognition, 21*(3), 167-193. doi: 10.1521/soco.21.3.167.25339

*Chen, M. (1999). Racial category activation as a precursor to nonconscious interpersonal motivation: A reconceptualization of stereotype priming effects. *Dissertation Abstracts International, 60* (09), 4960B. (UMI No. 9945264)

*Chen, M., & Bargh, J. A. (1997). Nonconscious behavioral confirmation processes: The self-fulfilling consequences of automatic stereotype activation. *Journal of Experimental Social Psychology, 33,* 541-560. doi: 10.1006/jesp.1997.1329

*Colcombe, S. J. (2000). Stereotype-consistent and inconsistent behavioral changes after image-based stereotype priming. *Dissertation Abstracts International, 61* (10), 5619B. (UMI No. 9989966).

Corneille, O., Huart, J., Becquart, E., & Brédart, S. (2004). When memory shifts toward more typical category exemplars: Accentuation effects in the recollection of ethnically ambiguous faces. *Journal of Personality and Social Psychology*, 86(2), 236-250. doi:10.1037/0022-3514.86.2.236

*Correll, J., Park, B., Judd, C. M., & Wittenbrink, B. (2002). The police officer's dilemma: Using ethnicity to disambiguate potentially threatening individuals. *Journal of Personality and Social Psychology, 83*(6), 1314-1329. doi:10.1037/0022-3514.83.6.1314

*Correll, J., Urland, G. R., & Ito, T. A. (2006). Event-related potentials and the decision to shoot: The role of threat perception and cognitive control. *Journal of Experimental Social Psychology, 42,* 120-128. doi: 10.1016/j.jesp.2005.02.006

*Cothran, D. L. (2005). The role of facial affect in the perception of threat posed by black faces. *Dissertation Abstracts International, 66* (10), 5730B. (UMI No. 3191280)

*Cunningham, W. A., Preacher, K. J., & Banaji, M. R. (2001). Implicit attitude measures: Consistency, stability, and convergent validity. *Psychological Science, 12*(2), 163-170.

doi: 10.1111/1467-9280.00328

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. (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. (2006). What are implicit measures and why are we using them. In R. W. Wiers & A. W. Stacy (Eds.), *The handbook of implicit cognition and addiction* (pp. 11-28). Thousand Oaks, CA: Sage Publishers.

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.*

*Devine, P. G., Plant, E. A., Amodio, D. M., Harmon-Jones, E., & Vance, S. L. (2002). The regulation of explicit and implicit race bias: The role of motivations to respond without prejudice. *Journal of Personality and Social Psychology, 82*(5), 835-848. doi: 10.1037/0022-3514.82.5.835

*Dixon, T. L., & Maddox, K. B. (2005). Skin tone, crime news, and social reality judgments: Priming the stereotype of the dark and dangerous black criminal. *Journal of Applied Social Psychology, 35*(8), 1555-1570. doi: 10.1111/j.1559-1816.2005.tb02184.x

*Dovidio, J. F., Kawakami, K., & Gaertner, S. L. (2002). Implicit and explicit prejudice and interracial interaction. *Journal of Personality and Social Psychology, 82*(1), 62-68. doi:10.1037/0022-3514.82.1.62 0022-3514.82.1.62

*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

*Dunton, B. C., & Fazio, R. H. (1997). An individual difference measure of motivation to control prejudiced reactions. *Personality and Social Psychology Bulletin, 23*(3), 316-326. doi: 10.1177/0146167297233009

Earp, B. D., & Trafimow, D. (2015). Replication, falsification, and the crisis of confidence in social psychology. *Frontiers in Psychology, 6,* 621. doi:10.3389/fpsyg.2015.00621.

Eberhardt, J. L., Davies, P. G., Purdie-Vaughns, V., Johnson, S. L. (2006). Looking deathworthy. Perceived stereotypicality of Black defendants predicts capital-sentencing outcomes. *Psychological Science, 17,* 383-392. doi: 10.1111/j.1467-9280.2006.01716.x

*Eberhardt, J. L., Goff, P. A., Purdie, V. J., & Davies, P. G. (2004). Seeing black: Race, crime, and visual processing. *Journal of Personality and Social Psychology, 87*(6), 876-893. doi: 10.1037/0022-3514.87.6.876

Fazio, R. H. (2001). On the automatic activation of associated evaluations: An overview. *Cognition and Emotion. 15,* 115-141. doi: 10.1080/0269993004200024

*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

*Fazio, R. H., & Dunton, B. C. (1997). Categorization by race: The impact of automatic and controlled components of racial prejudice. *Journal of Experimental Social Psychology, 33,* 451-470. doi: 10.1006/jesp.1997.1330

*Fazio, R. H., & Hilden, L. E. (2001). Emotional reactions to a seemingly prejudiced response: The role of automatically activated racial attitudes and motivation to control prejudiced reactions. *Psychology and Social Psychology Bulletin, 27*(5), 538-549. doi: 10.1177/0146167201275003

Fazio, R. H., & Olson, M. A. (2003). Implicit measures in social cognition research: Their meaning and use. *Annual Review of Psychology, 54,* 297-327. doi:10.1146/annurev.psych.54.101601.145225

Glaser, J., & Knowles, E. (2008). Implicit motivation to control prejudice. *Journal of Experimental Social Psychology, 44*(1), 164-172. doi: 10.1016/j.jesp.2007.01.002

*Govorun, O. & Payne, B. K. (2006). Ego-depletion and prejudice: Separating automatic and controlled components. *Social Cognition, 24,* 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., Oakes, M. A., & Hoffman, H. G. (2003). Targets of discrimination: Effects of race on responses to weapons holders. *Journal of Experimental Social Psychology, 39,* 399-405. doi: 10.1016/S0022-1031(03)00020-9

Guinote, A., Willis, G., & Martellotta, C. (2010). Social power increases implicit prejudice. *Journal of Experimental Social Psychology, 46*(2), 299-307. doi: 10.1016/j.jesp.2009.11.012

Hagiwara, N., Kashy, D. A., & Cesario, J. (2012). The independent effects of skin tone and facial features on Whites' affective reactions to Blacks. *Journal of Experimental Social Psychology, 48*(4), 892-898.

*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

*Jackson, J. (1998). Automatically activated racial attitudes. *Dissertation Abstracts International, 58* (08), 4524B. (UMI No. 9805416)

*Judd, C. M., Blair, I. V., & Chapleau, K. M. (2004). Automatic stereotypes vs. automatic prejudice: Sorting out the possibilities in the Payne (2001) weapon paradigm. *Journal of Experimental Social Psychology, 40,* 75-81. doi: 10.1016/S0022-1031(03)00063-5

Kahn, K. B., & Davies, P. G. (2010). Differentially dangerous? Phenotypic racial stereotypicality increases implicit bias among ingroup and outgroup members. Group *Processes & Intergroup Relations, 14,* 569-580. doi:10.1177/1368430210374609

*Kawakami, K., & Dovidio, J. F. (2001). The reliability of implicit stereotyping. *Psychology and Social Psychology Bulletin, 27*(2), 212-225. doi: 10.1177/0146167201272007

*Lambert, A. J., Payne, B. K., Jacoby, L. L., Shaffer, L. M., Chasteen, A. L., & Khan, S. R.

(2003). Stereotypes as dominant responses: On the "social facilitation of prejudice in anticipated public contexts. *Journal of Personality and Social Psychology, 84*(2), 277-295. doi: 10.1037/0022-3514.84.2.277

*Lambert, A. J., Payne, B. K., Ramsey, S., & Shaffer, L. M. (2005). On the predictive validity of implicit attitude measures: The moderating effect of perceived group variability. *Journal of Experimental Social Psychology, 41,* 114-128. doi: 10.1016/j.jesp.2004.06.006

Landis, J. R., Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33,* 159-174.

*Larsen, R. J., Lambert, A. J., & Chan, P. Y. (2006). *Perceptual consequences of threat and prejudice: Misperceiving weapons and other dangerous objects.* Unpublished manuscript.

*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. (2002). *A relational approach to the malleability of automatic prejudice.* Unpublished doctoral dissertation, University of California, Los Angeles.

*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

Ma, D. S., & Correll, J. (2011). Target prototypicality moderates racial bias in the decision to

shoot. *Journal of Experimental Social Psychology*, 47, 391-396. doi:10.1016/j.jesp.2010.11.002

Ma, Correll, & Wittenbrink (2015). The Chicago Face Database: A Free Stimulus Set of Faces and Norming Data. *Behavior Research Methods, 47,* 1122-1135.

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

*Maddux, W. W., Barden, J., Brewer, M. B., & Petty, R. E. (2005). Saying no to negativity: The effects of context and motivation to control prejudice on automatic evaluative responses. *Journal of Experimental Social Psychology, 41,* 19-35. doi: 10.1016/j.jesp.2004.05.002

*Massey, B. D. (2003). Is justice colorblind? Automatic behavioral effects of the African-American stereotype of hostility on the use of deadly force in simulated police-citizen encounter. *Dissertation Abstracts International, 64* (03), 1552B. (UMI No. 3086779)

*Moskowitz, G. B., Salomon, A. R., & Taylor, C. M. (2000). Preconsciously controlling stereotyping: Implicitly activated egalitarian goals prevent the activation of stereotypes. *Social Cognition, 18,* 151-177.

Nosek, B. A. (2008). Stimulus sets. Race faces. *Stimulus Materials for Social and Behavioral Research-Prepared by Project Implicit and Brian Nosek.* Retrieved March 14, 2008 from http://projectimplicit.net/nosek/stimuli/

Nosek, B. A., & Banaji, M. R. (2001). The go/no-go association task. *Social Cognition, 19*(6), 161-176.

Nosek, B., Banaji, M., & Greenwald, A. (2002). Harvesting implicit group attitudes and beliefs from a demonstration web site. *Group Dynamics: Theory, Research, and Practice*, *6*(1), 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. A., & Fazio, R. H. (2004). Inferences as a function of automatically activated racial attitudes and motivation to control prejudiced reactions. *Basic and Applied Social Psychology, 26*, 1-11. doi: 10.1207/s15324834basp2601_1

Pashler, H. & Wagenmakers, E.-J. (Eds.). (2012). Special Section on Replicability in Psychological Science: A Crisis of Confidence? *Perspectives on Psychological Science, 7* (6).

*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

Payne, B. K. (2003). *Cognitive control in automatic attitude measurement.* Doctoral dissertation, Washington University.

*Payne, B. K. (2005). Conceptualizing control in social cognition: How executive functioning modulates the expression of automatic stereotyping. *Journal of Personality and Social Psychology, 89,* 488-503. doi: 10.1037/0022-3514.89.4.488

Payne, B.K., & Cameron, C.D. (2012). Implicit social cognition and mental representation. In D. Carlston (Ed.), Oxford handbook of social cognition. Oxford University Press.

*Payne, B. K., Cheng, C. M., Govorun, O., & Stewart, B. D. (2005). An inkblot for attitudes: Affect misattribution as implicit measurement. *Journal of Personality and Social Psychology, 89,* 277-293. doi: 10.1037/0022-3514.89.3.277

*Payne, B. K., Lambert, A. J., & Jacoby, L. L. (2002). Best laid plans: Effects of goals on accessibility bias and cognitive control in race-based misperceptions of weapons. *Journal of Experimental Social Psychology, 38,* 384-396. doi: 10.1016/S0022-1031(02)00006-9

*Persson, A. V. (2003). The Development of racial attitudes and discriminatory behaviors from preschool age to young adulthood. *Dissertation Abstracts International*, *64*(6), 2961. (UMI No. 3095322)

*Phelps, E. A., O'Connor, K. J., Cunningham, W. A., Funayama, E. S., Gatenby, J. C., Gore, J. C., & Banaji, M. R. (2000). Performance on indirect measures of race evaluation predicts amygdala activation. *Journal of Cognitive Neuroscience, 12,* 729-738. doi: 10.1162/089892900562552

*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

*Plant, E. A., & Peruche, B. M. (2005). The consequences of race for police officers' responses to criminal suspects. *Psychological Science, 16,* 180-183. doi: 10.1111/j.0956-7976.2005.00800.x

Pratto, F., Sidanius, J., & Levin S. (2006). Social dominance theory and the dynamics of intergroup relations: Taking stock and looking forward. *European Review of Social Psychology, 17,* 271-320. doi: 10.1080/10463280601055772

Reis, H. (2007). Not your grand-advisor's journal. *Dialogue*, 22 (2), 20-21.

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

*Rodríguez Torres, R. (1998). La influencia de la inducción afectiva sobre el prejuicio racial. *Revista de Psicología Social, 13*, 279-289.

Rydell, R., & McConnell, A. (2006). Understanding implicit and explicit attitude change: A systems of reasoning analysis. *Journal of Personality and Social Psychology, 91*(6), 995-1008. doi:10.1037/0022-3514.91.6.995.

Rydell, R., McConnell, A., Mackie, D., & Strain, L. (2006). Of two minds: Forming and changing valence-inconsistent implicit and explicit attitudes. *Psychological Science, 17*(11), 954-958. doi:10.1111/j.1467-9280.2006.01811.x.

Sadler, M. S., Correll, J., Park, B., & Judd, C. M. (2012). The world is not black and white: Racial bias in the decision to shoot in a multiethnic context. *Journal of Social Issues, 68*(2), 286-313. doi: 10.1111/j.1540-4560.2012.01749.x

*Scherer, L. & Lambert, A. J. (2009). Counterstereotypic exemplars in context: Evidence for intracategory differentiation using implicit measures. *Social Cognition, 27,* 522-549. doi: 10.1521/soco.2009.27.4.522

Schneider, D. J. (2004). The psychology of stereotyping. New York: The Guilford Press.

Sherman, J. W. (Ed.). (2014). Understanding Priming Effects in Social Psychology [special issue]. *Social Cognition, 32 (*Supplement).

*Sinclair, S., Lowery, B. S., Hardin, C. D., & Colangelo, A. (2005). Social tuning of automatic racial attitudes: The role of affiliative motivation. *Journal of Personality and Social Psychology, 89,* 583-592. doi: 10.1037/0022-3514.89.4.583

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

Schlauch, R., Lang, A., Plant, E., Christensen, R., & Donohue, K. (2009). Effect of alcohol on race-biased responding: The moderating role of internal and external motivations to respond without prejudice. *Journal of Studies on Alcohol and Drugs, 70*(3), 328-336.

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

Smith, P., Dijksterhuis, A., & Chaiken, S. (2008). Subliminal exposure to faces and racial attitudes: Exposure to whites makes whites like blacks less. *Journal of Experimental Social Psychology, 44,* (1), 50-64. doi: 10.1016/j.jesp.2007.01.006

*Spencer, S. J., Fein, S., Wolfe, C. T., Fong, C., & Dunn, M. A. (1998). Automatic activation of stereotypes: The role of self-image threat. *Personality and Social Psychology Bulletin, 24,* 1139-1152. doi: 10.1177/01461672982411001

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. V., & Strube, M. J (2012). What's in a face? The role of skin tone, facial physiognomy, and color presentation mode of facial primes in affective priming effects. *The Journal of Social Psychology. 152* (2), 212-227. doi**:**10.1080/00224545.2011.597797.

*Stewart, T. L., Weeks, M., & Lupfer, M. B. (2003). Spontaneous stereotyping: A matter of prejudice? *Social Cognition, 21,* 263-298. doi: 10.1521/soco.21.4.263.27003

*Towles-Schwen, T., & Fazio, R. H. (2003). Choosing social situations: The relation between automatically activated racial attitudes and anticipated comfort interacting with African Americans. *Personality and Social Psychology Bulletin, 29,* 170-182. doi: 10.1177/0146167202239042

*Wheeler, M. E. & Fiske, S. T. (2005). Controlling racial prejudice: Social-cognitive goals affect amygdala and stereotype activation. *Psychological Science, 16,* 56-63. doi: 10.1111/j.0956-7976.2005.00780.x

Wittenbrink, B. (2007). Measuring attitudes through priming. In B. Wittenbrink & N. Schwarz (Eds.), *Implicit Measures of Attitudes* (pp. 17-58). New York: Guilford Press.

*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

Footnotes

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

[2] There are several different computational approaches used to detect if there are general racial negativity or stereotypic associations, largely determined by a specific procedure and design involved.

[3] 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 Single Category IAT (Karpinski & Steinman, 2006) and Go/No-go Task (Nosek & Banaji, 2001) were excluded on the same grounds (i.e., the effects reflect attitudes towards the categories *Black* and *White* rather than the faces used to activate those categories).

[4] Data for interrater agreement (Cohen's kappa, κ) for all variables collected and reported in this manuscript is available from authors upon request.

[5] Most of the studies included in this review were published articles (84.4%), with smaller numbers of dissertations (14.6%) and manuscripts under review (1%). The earliest publication date was 1995, with the majority of studies published in 2003 (19.6%), 2005 (17.5%) and 2002 (13.4%).

[6] The results of analyses, including some additional analyses, data set, and the final version of the coding guide used for this review are available from authors at https://www.dropbox.com/sh/5lvyzt9lp0wz4bk/AABbAoI6eXuxlY0sYougTXzla?dl=0

Table 1.
Mean Numbers of Participants in Original and Final Samples per Study Grouped by Demographic Characteristics and Means for Participants' Age

	Original Samples, Mean *N* (N of samples included)	Final Samples, Mean *N* (N of samples included)	M_{age}
Age group/profession			
College students	69.09 (90)	64.31 (90)	19.58
Children*	80 (3)	80 (3)	
Law enforcement officers	50 (1)	48 (1)	37
Other adults	52 (1)	46 (1)	
Ethnicity/Race			
African-American/Black	13.17 (24)	20.92 (12)	
White	59.21(68)	55.08 (61)	
Hispanic American	5.06 (17)	5.92 (12)	
Asian American	8.79 (19)	11.67 (12)	
Alaskan Native/Native American/ Pacific Islander	1.00 (4)	1.00 (1)	
International	3.67 (3)	1.00 (1)	
Ethnicity not reported	34.23 (35)	50.60 (25)	
Gender			
Males	25.77 (52)	24.35 (55)	
Females	43.88 (56)	39.46 (59)	
Total	69.91 (96)	64.81(96)	19.15

*Note: both young children (M_{age} = 5.31) and older children (M_{age} = 9.24) are included. Number of studies on which these means are based is included in parenthesis.

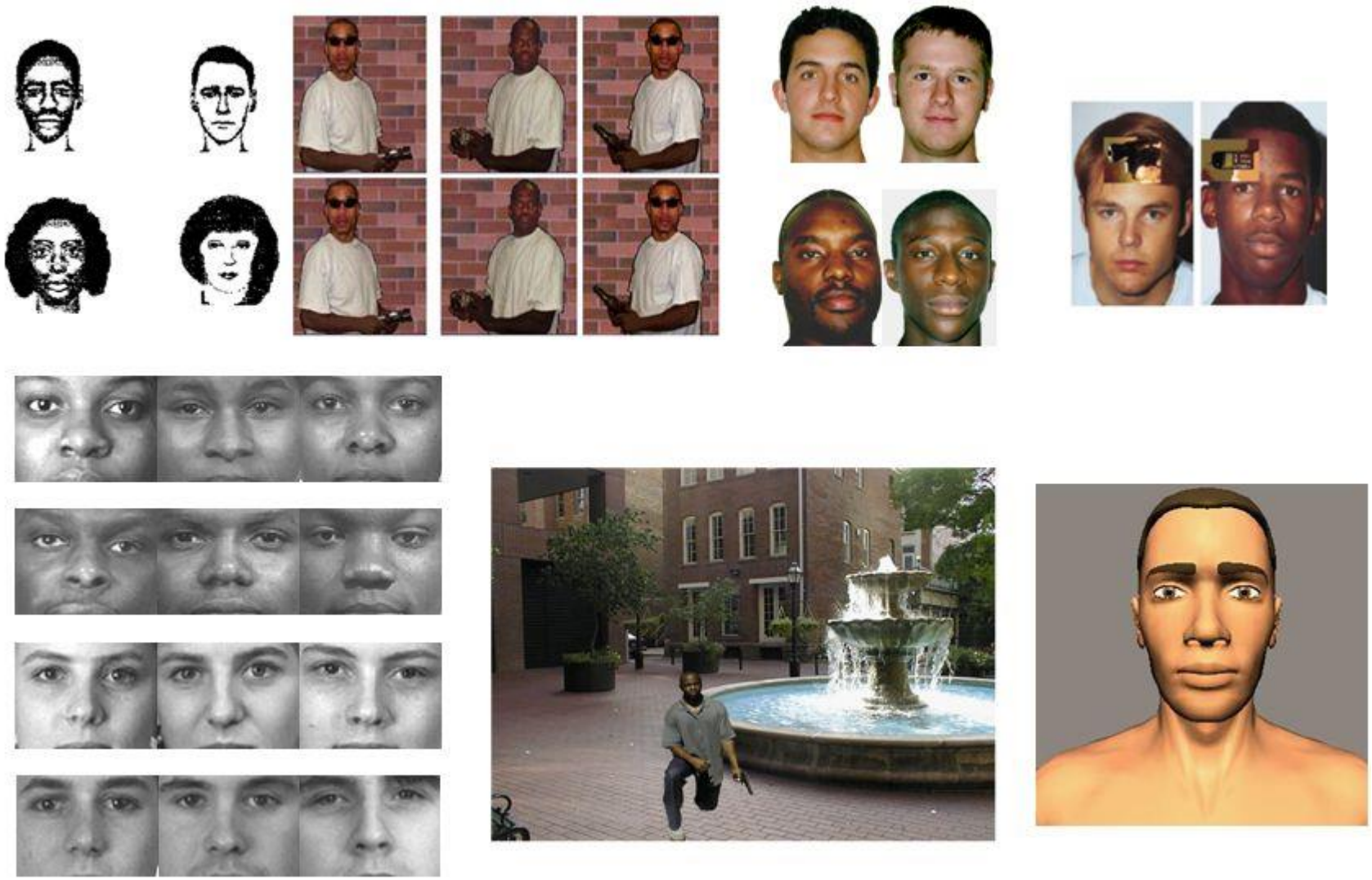

Figure 1. Sample stimuli, read clockwise from the lower left corner: (a) Nosek et al. (2002) stimuli, available at http://projectimplicit.net/nosek/stimuli/, (b) Dovidio et al. (1997) stimuli, from Dovidio et al. (1997), Figure 1. copyright 1997 by Elsevier, reprinted by permission, (c) Greenwald, Oakes, & Hoffman (2003) stimuli, received from authors and available at http://faculty.washington.edu/agg/pdf/AppendixPhotos.pdf, (d) Barden, Maddux, Petty, & Brewer, (2004) stimuli, received from authors, (e) Plant, Peruche, & Butz (2005) stimuli, from Plant et al. (2005), Appendix A, copyright 2005 by Elsevier, reprinted by permission, (f) Hugenberg & Bodenhausen (2003) stimuli, received from authors, (g) Correll, Park, Judd, & Wittebrink (2002) stimuli, received from authors.

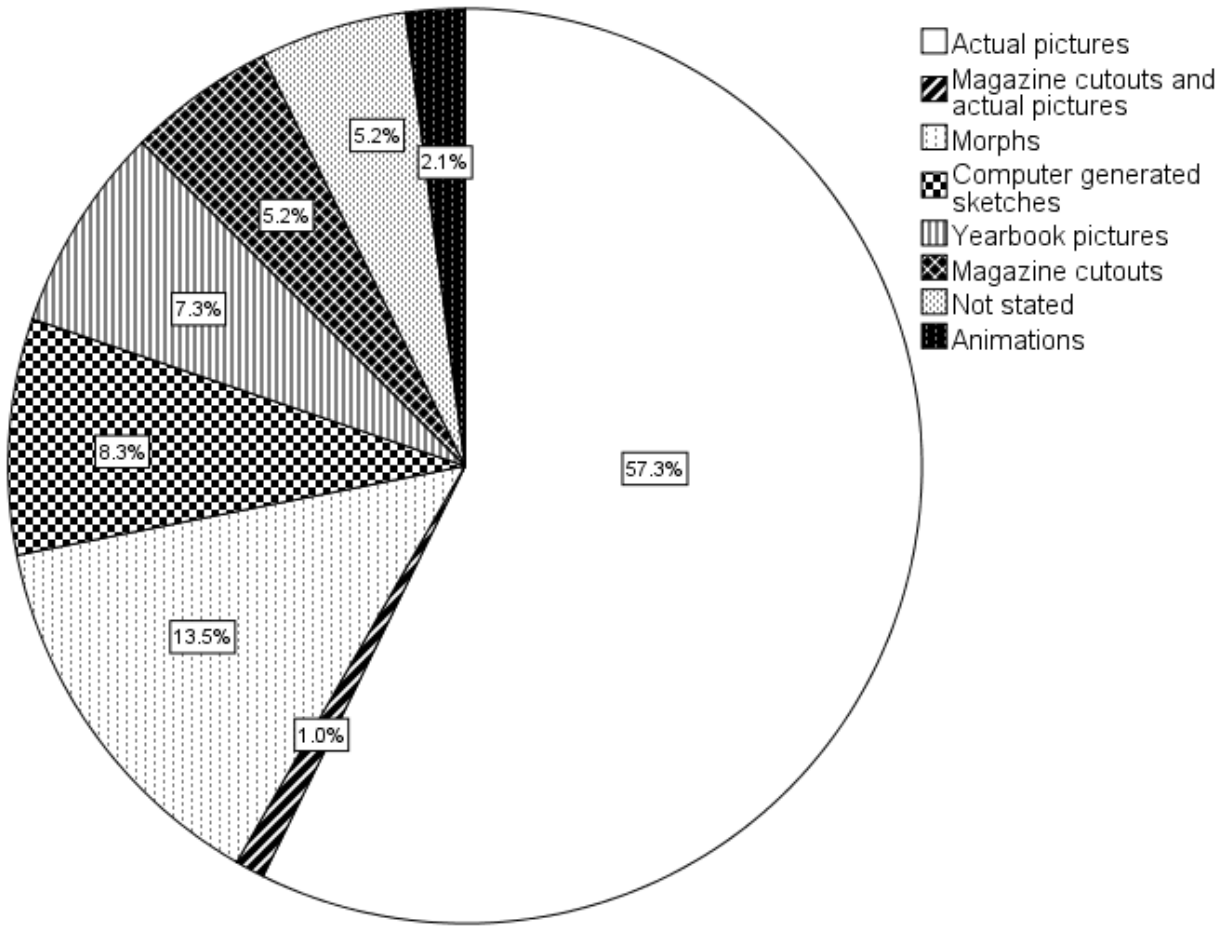

Figure 2. Stimuli origination (in percentages).

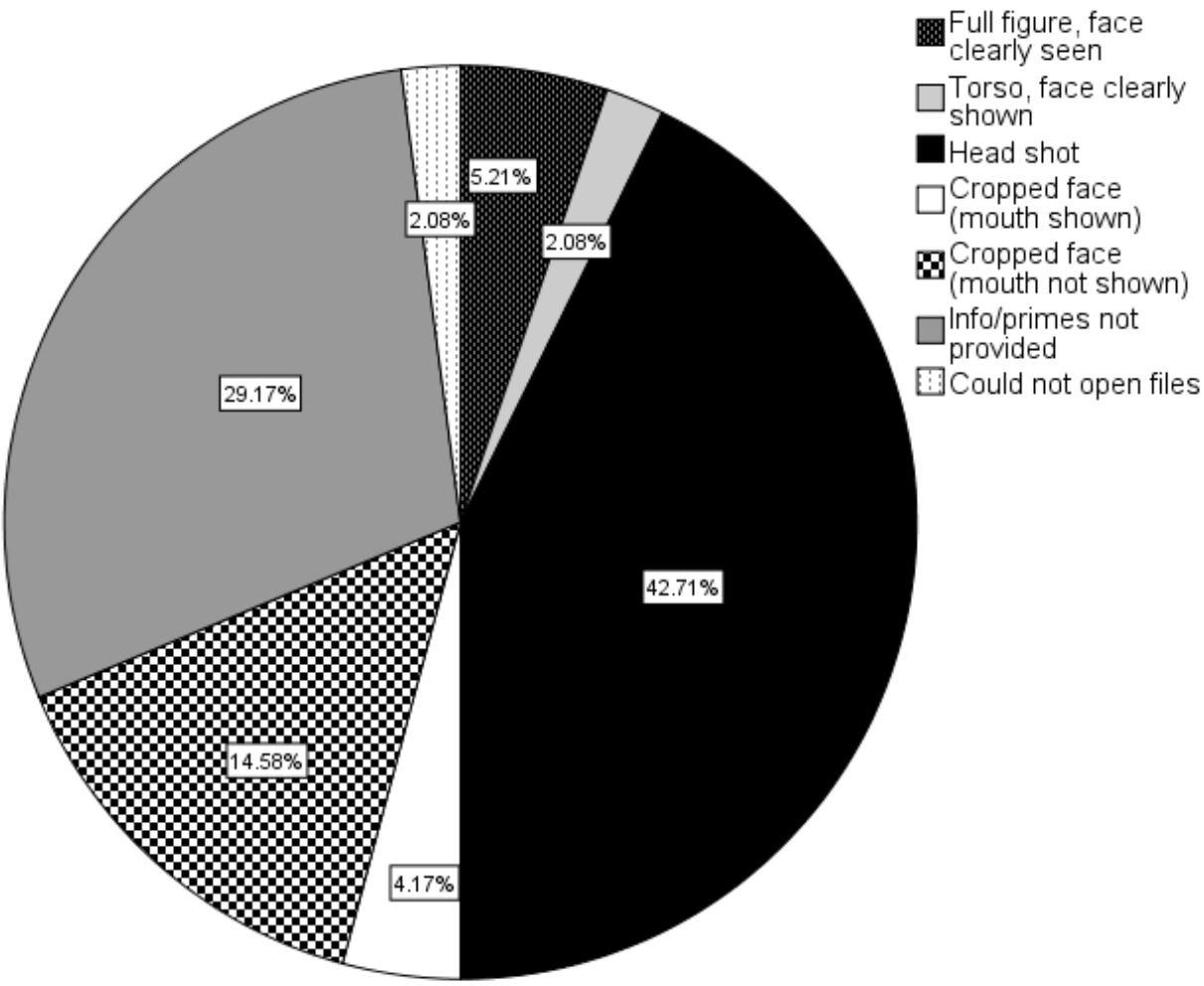

Figure 3. Size of presentation (coders' judgment) in percentages.

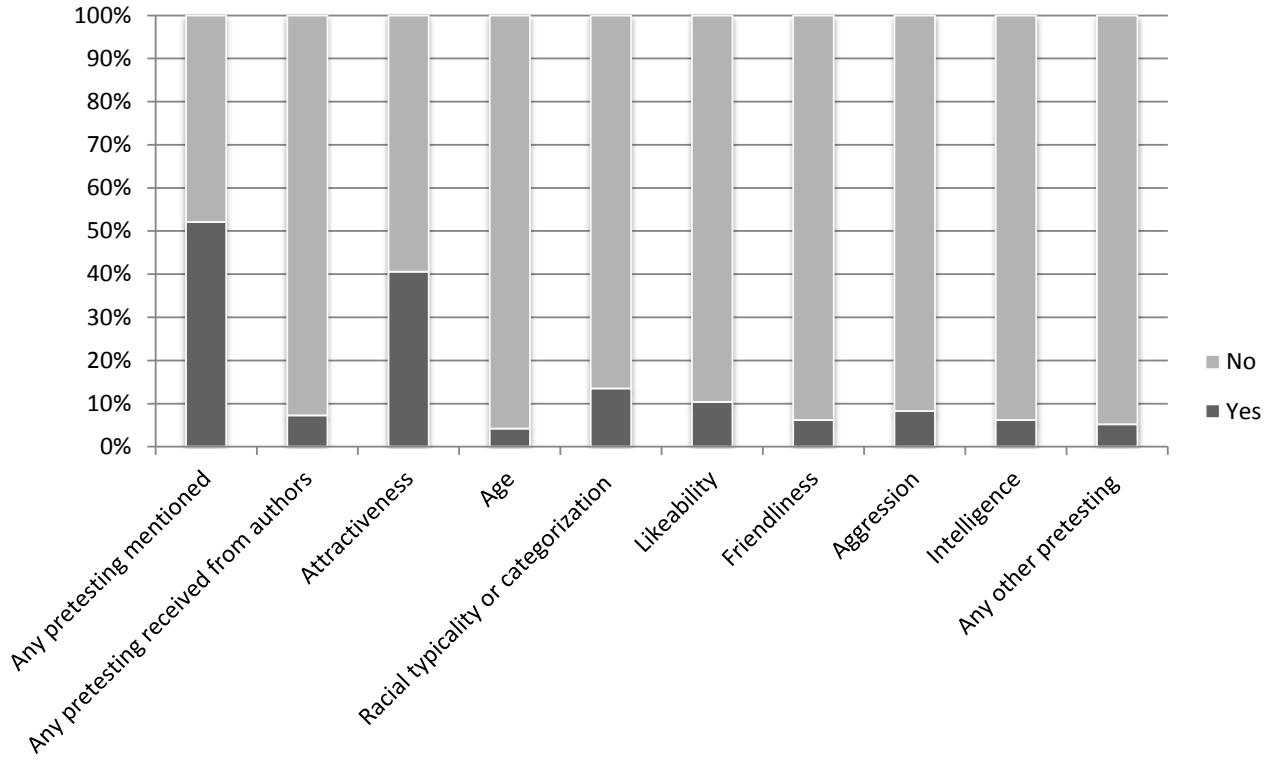

Figure 4. Information on pretesting of stimuli.

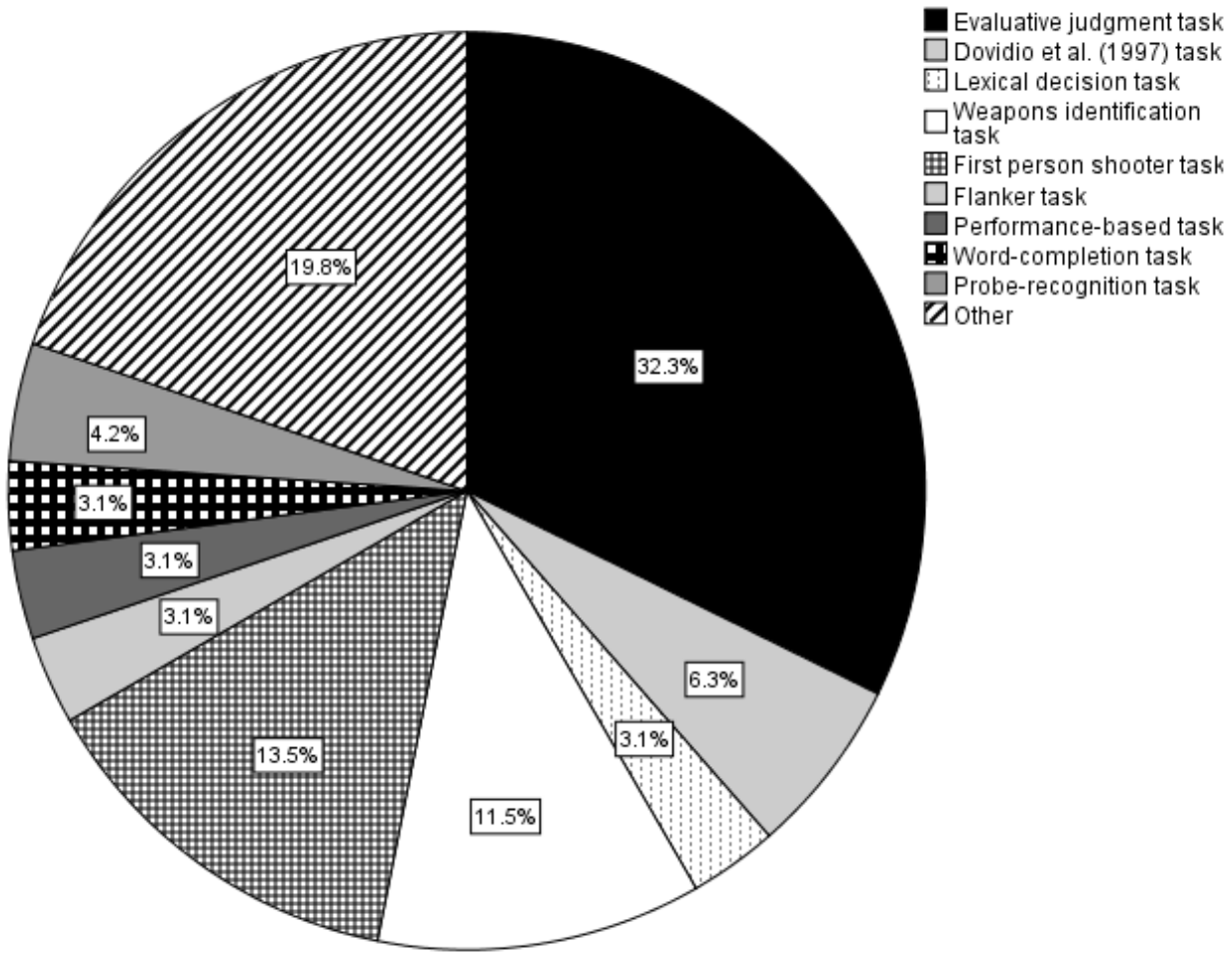

Figure 5. Types of priming tasks. Note: Flanker task is a variation of an affective priming task with an adddional manipulation of where on the screen a target word is presented. In performance type-based tasks changes in performance are measured after being presented with a racial prime, e.g. in a musical or math task, etc. In word completition tasks, participants are asked to complete incomplete words after racial prime presentation. The words are constructed in such a way that they can be completed in stereotypic/nonstereotypic manner. In probe recognition tasks, primes are paired with stereotypic or nonstereotypic behavioral descriptions and then a recognition test is completed where participants determine whether certain words (e.g., stereotypic traits) had been previously seen in behavioral descriptions.