RACIAL CATEGORIZATION OF AND DECISION-MAKING RESPONSES TO AMBIGUOUS AND UNAMBIGUOUS FACIAL STIMULI: AN EXAMINATION OF RACIAL PROFILING PROCESSES

A Thesis

Submitted to the Graduate School of the University of Notre Dame in Partial Fulfillment of the Requirements for the Degree of

Master of Arts

by

Hilary C. Cisco Reuter, B.A.

__________________________
Alexandra F. Corning, Director

Graduate Program in Psychology
Notre Dame, Indiana
July 2004
RACIAL CATEGORIZATION OF AND DECISION-MAKING RESPONSES TO AMBIGUOUS AND UNAMBIGUOUS FACIAL STIMULI: AN EXAMINATION OF RACIAL PROFILING PROCESSES

Abstract

by

Hilary C. Cisco

Categorization and decision-making were combined in a task with photorealistic faces to test for the presence of racial profiling processes. Participants were asked to make a racial categorization and a decision about whether and for which type of crime to apprehend the target face. Results indicated that participants did engage in racial profiling behaviors, and did so of African Americans and European Americans at statistically indistinguishable rates. In addition, results indicated that crime decision-making was dependent on racial categorization and participants made race categorizations before deciding crime apprehension, even when the order of judgments required a crime decision prior to a race categorization. Additional findings were that ambiguous-race faces were not racially profiled, and also that these faces were twice as likely to be categorized as African American.
DEDICATION

I thank God for everything. I have been very fortunate to have a supportive family, loving friends, and caring faculty while preparing this project. I dedicate this thesis to Dave and my perfect parents. I could not have been blessed anymore by your love. Thank you and I love you.
ACKNOWLEDGMENTS

I must acknowledge several amazing people that contributed to this project developing into a master’s thesis. First and foremost is my advisor, Alexandra F. Corning. Alexandra, I cannot begin to thank you enough for your unwavering support and incredible mentoring. You helped me develop all the skills necessary at the master’s level. The thesis has extraordinarily grown under your guidance. Secondly, Angelina Copeland for making many seemingly impossible software issues become realities, giving me advice and a supportive friendship, and helping me check my every move for accuracy throughout the project. Thirdly, Johnnathan Velandia, my research assistant who morphed many pictures and was dedicated to collecting data, and the research assistants from the Social Cognition lab, Kathryn Bethea, Kristen Henkel, and Carrie Graf who also contributed to the project. Fourth, I would like to thank my committee members, Drs. Anre Venter and Tom Merluzzi for your insights, suggestions, and directions for further pursuit. Lastly, I would like to thank many Notre Dame graduate students that contributed the success of this project: Mike Lau, Angie Krumm, Jarren Gonzales, Stacey Tiberio, Joe Rausch, Mignon Montpetit, Eric Covey, and Drs. David Copeland and Christoph Rasche. I am grateful to each and every one of you for making this research project a master’s thesis and want to acknowledge your contribution.
INTRODUCTION

Racial discrimination remains a persisting problem in the United States (Karlsen & Nazroo, 2002), with a recent survey indicating that 80% of Americans report having experienced racial discrimination at some point in their lives (Karlsen & Nazroo). More alarming perhaps is that 98% of African Americans report having experienced racial discrimination (Brown, Jackson, Neighbors, Torres, Sellers, & Brown, 2000). Furthermore, one of the most ubiquitous forms of discrimination is the stereotypical association of African Americans with criminal behaviors (Ruffins, 2002).

In contemporary American society, race and crime are conflated. This may be a result of unequal and stereotypical treatment in the media (Hurwitz & Peffley, 1997). African Americans comprise 12-13% of the U.S. population, yet they are arrested for more than one half of all robberies and murders (Hurwitz et al., 1997; Meier, 1994). It is thought that the rate of arrest for African Americans contributes to American’s stereotypes of African Americans as aggressive, violent, and associated with crime (Hurwitz et al., 1997). These stereotypes are maintained by the disproportionately negative portrayal of African Americans in television media (Hurwitz et al.). For example, although African Americans are marginally featured on the local news, the majority of that time is devoted to coverage of their involvement in violent crime (Entman, 1992). In addition, Entman has found that the coverage of alleged African American criminals’ omits their names, shows them already in police or correction’s
custody and in handcuffs, and does not show them speaking for themselves. It seems plausible, then, that the association of violence and crime with ethnic group members, particularly African Americans, across various media may contribute to racial profiling attitudes held by the media-viewing public.

The purpose of this study was to examine individuals’ propensities, within a law-enforcement context, to engage in racial-profiling behaviors when presented with African American, European American, and ambiguous-race targets. In this paper, I will describe the current climate of racial profiling in the U.S., discuss psychological theories relating to racial profiling behaviors, and review the relevant findings from cognitive psychology that offer explanations about how the construct of race is perceived and processed. As such, I will propose connections between racial profiling and psychological theory and already well-substantiated psychological phenomena.

Defining Racial Profiling

Racial profiling is the targeting of individuals for criminal investigation on the basis of their race or on the basis of their race in combination with other aspects of criminality (Ramirez, McDevitt & Farrel, 2000). Several definitions of racial profiling have been forwarded by a range of bodies including the U.S. government, civil liberty groups, academic experts including historians and law professionals, politicians, and persons indicating they have been recipients of repeated exposure to racial profiling. Of note is the absence of definition of this phenomenon by psychologists. Although the American Psychological Association (APA) has posted a position-oriented web page endorsing legislation that protects citizens from police-initiated racial profiling practices
APA, 2003), formed committees that support anti-racial profiling practices, and made reference to the practice of racial profiling in APA-published journals (e.g., O’Conner, 2001; Tyler, 2001), it has yet to offer an encompassing and straightforward definition of racial profiling.

Some definitions, such as the one offered by the U.S. Department of Justice, have had more extensive impact on the study of racial profiling. The U.S. Department of Justice asked researchers at Northeastern University to develop a racial profiling study design guide. The design guide forwards this definition:

Racial profiling is any police-initiated action that relies on the race, ethnicity or national origin of an individual rather than the behavior of an individual or information that leads the police to a particular individual who has been identified as being, or having been, engaged in criminal activity (Ramirez, McDevitt & Farrel, 2000, p.3).

The Justice Department clearly defines and delineates racial profiling without political editorial. However, it is intentionally legal and social, rather than psychological in nature, and thus does not provide sound grounding for the study of racial profiling from a psychological perspective.

The Justice Department definition and other lay definitions may encounter additional problems in their applications. Of particular importance is the lack of attention to the concept of base rates as a function of race. Base rates are “values that indicate how often an item occurs in the population” (Kahneman & Tversky, 1973, p. 238). For example, according to the 2000 U.S. Census, there are approximately 36 million African Americans in the U.S. population of 281 million, so the base rate for African Americans...
is the U.S. is about 13 percent. Kahneman and Tversky described a principle they coined the base-rate fallacy. The base-rate fallacy involves overly emphasizing representativeness, and thus underemphasizing important information about the base rate of the population. Representativeness refers to similarity on an important characteristic between a sample and its corresponding population. Representativeness can also occur in the form of a racial stereotype; for instance, the stereotype that African Americans are more criminal than are European Americans. Therefore, people are most likely to think that there are more African Americans committing crime than European Americans even though there are far more European Americans in the United States than there are African Americans.

In tests of the base-rate fallacy, Tversky and Kahneman (1974) conducted an experiment in which subjects were given highly stereotypic descriptions of a person from a professional career (i.e., engineers), and given the base rate for the prevalence of that professional career (30%) as well as the base rate for another professional career (i.e., lawyers, 70%). Their results indicated that subjects tended to ignore the base rate information that would lead them to conclude that there was a greater likelihood that the person was from the non-stereotypic profession (i.e., lawyers) due to the sheer number of people in that given field, and instead focused on their representative heuristic or stereotype (e.g., details indicating that the person was shy, withdrawn, had very little interest in people, and had a passion for detail) to guide their decisions.

The relationship between use of base rate information and an individual’s prejudices requires empirical attention. Although, in contemporary society, people, on average, may not be willing to admit they endorse racial stereotypes, they may continue
to ignore base rate information regarding the likelihood of African Americans committing crimes in favor of their stereotypes.

*Explicit and Implicit Prejudice*

The prejudice literature offers an abundance of evidence suggesting that people are motivated to respond in ways that do not reveal that they are racially prejudiced (Brauer, Wasel, & Niedenthal, 2000; Wittenbrink, Judd, & Park, 1998). Therefore, researchers interested in better understanding prejudice have worked to develop more covert ways of measuring individuals’ real (or implicit) biases. Implicit bias is the unconscious or automatic operation of stereotypes that are unavailable for direct measure (Bargh, 1994; Devine, 1989; Dovidio, Evans, & Tyler, 1986; Fiske, 1989; Geis, 1993; Gilbert & Hixon, 1991; Greenwald & Banaji, 1995) and explicit bias is the conscious, prejudice cognition that is available for direct measure (Greenwald et al., 1995). Researchers do not agree, however, on the precise definitions of explicit and implicit prejudice (Brauer et al., 2000). Specifically, scholars continue to debate whether prejudice is one construct ranging on a continuum from explicit to implicit prejudice (see Brauer et al., 2000; Kawakami, Dion, & Dovidio, 1998; Lepore & Brown, 1997; Locke, MacLeod, & Walker, 1994; Neumann, 1998; Wittenbrink, et al., 1997), or if explicit and implicit prejudice are two separate constructs sharing similarities and differences which make them conceptually alike but operationally distinct (see Devine 1989; Dovidio, Gaertner, Validzic, Matoka, Johnson, & Frazier, 1997).

There is substantial evidence to suggest that both the single-continuum and the separate-constructs models of implicit and explicit prejudice describe important
components of racial prejudice. Both contend that social stereotypes play a role in the automatic component of prejudice and that learning can moderate the manifestation of explicit prejudice. However, it is important to investigate whether a social stereotype alone is enough to motivate prejudicial attitudes and discriminatory behavior.

Correll, Park, Judd, and Wittenbrink (2002) tested the effect of target race, using African Americans and European Americans targets, on “shoot/don’t shoot” decisions. European American college student participants performing in this shooter bias videogame were shown slides of armed and unarmed target and were instructed to shoot only armed targets. Participants more quickly made the decision to appropriately shoot armed targets if the target was African American and more quickly decided to not shoot unarmed European American targets. The researchers further tested whether prejudice may account for this effect over and above bias. Importantly, prejudice in this study included measures of explicit prejudice against African Americans and dislike for African Americans. Their prediction was that the stereotype of African Americans as violent would be strong enough of an association to account for the biased responses. None of the explicit measures of prejudice, however, were significantly correlated with the shooter bias videogame experiment results.

In a separate community-based sample, Correll, et al. (2002) replicated these results. In addition, the researchers used a community sample of both African Americans and European Americans and discovered that both groups favored accurately shooting the armed African American targets and accurately failed to shoot the unarmed European American targets. In other words, for these participants, it was easier to correctly shoot the armed target if the target was African American and it was easier to “not shoot” the
unarmed targets if the target was European American. The researchers concluded that the African American participants’ desire to not be prejudiced had no effect on racial bias.

Although social psychological research has demonstrated across multi-method research the stability of these findings, it does not suggest that these processes are unyielding. It remains unclear the extent to which these implicit stereotypes account for so much of the behavioral variance that explicit prejudice adds little. In addition, it is also unclear to what extent these automatic stereotypes are ingrained, necessary, or unfortunately optimal to the perception of people. Cognitive psychology offers a great deal of insight into how individuals perceive, categorize, and decide how to respond to new faces and new people.

**Contributions from Cognitive Psychology**

When examining stereotypes, it is important to know what motivates an individual to create a stereotype and to pose the question of whether that learned response has become automatic. If the stereotype has become automatic, it would be key to understand the chain of cognitive events that precede the racial stereotyping behavior, specifically, the relationship among three aspects of racial discrimination behavior: stereotypes, the racial categorization, and the decision-making processes.

In order to make a categorization, a person first must have access to the stimulus and the facilities to perceive the stimulus. If the person has sufficient information to categorize the stimulus, she or he can then react to the stimulus. The decision to “interact with a stimulus” (of either another person or a representation of another person, i.e., a picture of a face) has been referred to as an interaction decision (Townsend, Silva,
Spencer-Smith, & Wenger, 2000). When a categorization contains necessary or sufficient information to form and follow through on an interaction decision, the decision involves performing a specifiable type of response to another person. For example, when a woman is walking alone on a dark street and she sees a tall humanlike figure approaching her at a rapid speed, she can create a strategy to secure her safety based on how she categorizes the figure. If she uses race as the necessary and or sufficient information, she is using a stereotype to guide her interaction decision. She has then used a social categorization, race, to determine her decision to flee, fight, or be calm.

**Purpose and Hypotheses**

The purpose of this study was to test for the presence of racial profiling by examining categorical and decision-making processes in an adult sample. The operational definition of racial profiling applied in this study is disproportionate decision making to arrest a person of a given race (African American or European American) based on the presence of the target’s race and the racial categorization. Racial profiling is examined in this study by observing the extent to which participants overestimate the number of people to apprehend for a crime that they are informed is stereotypic of their race.

Five hypotheses were tested. The first two hypotheses test for the presence of racial profiling and the remaining three examine the extent to which categorization and decision-making differ between unambiguous and ambiguous faces. The first hypothesis is that participants will racially profile the targets. That is, they will apprehend targets for the stereotypical crime at a rate significantly greater than the base rates for those crimes.
Second, in every meaningful comparison among race and crime, decision-making about crime will be dependent on racial categorization; however, racial categorization will not be dependent on the crime decision. Third, participants will be likely to categorize ambiguous photographs as African American more than they categorized ambiguous photographs as European American. Fourth, participants will choose “do not apprehend” when presented with ambiguous race stimuli more than when presented with unambiguous race stimuli. Finally, participants will choose “do not apprehend” when presented with persons with ambiguous faces that they had categorized as European American more than the persons with ambiguous faces they had categorized as African American.
METHOD

Stimuli Development Study

Participants

Seventy-eight participants were recruited from the Notre Dame undergraduate subject pool to rate a series of African American and European American photographs altered by Gryphon's CD-Morph software with respect to race. Of the 78 participants, 59 identified as European American, 10 as Hispanic, 8 as Asian, and 2 as African American. Participants were required to have normal or corrected-to-normal vision and unencumbered use of their hands. They were compensated with psychology class extra-credit points or a required research participation class credit.

Materials

Sixty images were used as the basis for the creation of a set of morphed images, designed to range from unambiguously African American to unambiguously European American, with racially ambiguous images within this range. To create the morphed images, 30 African American images were randomly paired with 30 European American images. A total of 200 anchor points were placed at prominent locations on each pair of images in identical anatomical locations (e.g., at the rightmost corner of the mouth in both photographs). Image morphing software (i.e., CD-Morph) was used to create a
sequence of 100 images, involving a transformation of one image in the pair (e.g., the African American image) into the second image in the pair (e.g., the European American image). From this sequence of 100 images, a total of 21 images (including the two original images) were selected. This was accomplished by starting with the first image in the pair, then choosing images at five-image intervals, and ending with the second image in the pair. The resulting set of 21 images generated for each of the original pair of images can be thought of as composing a physical transformation of an unambiguously African American face into an unambiguously European American face. Because the physical (computational) transformation cannot be unambiguously mapped onto a human observer's psychological judgment (see Busey & Tunnicliff, 1999 and MacLin et al., 2001), the final set of 630 images was presented to observers to obtain human ratings of perceived race.

All images were presented to participants on 40.64 cm monitors. Presentation and timing of the images, and recording and timing of all responses (both accurate to \(\pm 1\) ms), was done on PC-compatible microcomputers.

**Design and Procedure**

Participants rated all of the photographs on a 9-point Likert-type scale ranging from 1 (*European American*) to 9 (*African American*). Participants were tested in groups of one to four in a dimly lit room. Each trial consisted of the following events. First, a fixation cross was displayed in the center of the screen for between 500 and 1000 ms, with the actual duration of each trial determined on the basis of a uniform distribution over the integer values in this range. Next, the image to be rated was displayed and
remained on the screen either until a rating response had been provided or 2000 ms elapsed. Participants were instructed to give a response indicating the extent to which they perceived the face shown to be either African American or European American, using the scale ranging from 1 (European American) to 9 (African American). Participants were instructed to respond as quickly as they could, while being as accurate as they could, with their racial judgments of each face. If a participant required more than 3000 ms to respond, the trial was logged as a "no response" trial. Time pressure was applied during these ratings to minimize the use of any strategies and evaluative heuristics. No feedback was provided for responses, and the inter-trial interval was 1000 ms. Individual stimuli response times (RTs) were recorded but not analyzed.

**Stimuli Development Results**

A frequency analysis was conducted on every photograph and results of these analyses were used to categorize photographs as either European American, African American, or racially ambiguous. Specifically, using the frequency analysis resulting from participants' ratings of the faces, I categorized faces as having European American, African American, or racially-ambiguous membership. The European American photographs were selected based on having the highest frequency ratings of ones and twos. The African American photographs were selected based on having the highest frequency ratings of eights and nines. The ambiguous photos were selected based on high frequency ratings of fours, fives, and sixes. This preliminary analysis resulted in selecting 90 photographs out of the original 630. The 90 photographs were then subjected to chi-square ($\chi^2$) analyses to test that the selected photographs were
categorically distinct from each other. The null test of the $\chi^2$ tested whether all the photographs belonged to a uniform distribution; this was violated ($\chi^2(2, N=78), p < .05$). The alternative hypotheses tested each racial category of photographs (i.e., unambiguous European American, unambiguous African American, and ambiguous race) separately. Results indicated that every $\chi^2$ was violated ($\chi^2(2, N=78), p < .05$) in the corresponding and expected direction; thus all ninety photographs were retained for use in the racial profiling experiment, which is discussed next.

*Racial Profiling Experiment*

*Participants*

A total of 189 participants were recruited from the Notre Dame undergraduate psychology subject pool to participate in this experiment. Of the 189 participants, 148 identified as European American, 18 as Hispanic, 13 as Asian, 7 as African American, 1 as Native American, and 2 as "other." All participants were compensated with one psychology class extra credit point, and none had participated in the stimuli development study.

*Stimuli*

The 90 photographs generated in the stimuli development study were the stimuli that were used in this experiment. All photographs were grey scale, 196 x 196 pixel photographs of male faces without hair. Faces were absent of hair, a salient racial feature, to minimize the categorization of African American on the ambiguous photographs (see
MacLin, 2001). The stimulus set, then, contained 30 unambiguous European American male faces, 30 unambiguous African American male faces, and 30 ambiguous-race male faces.

**Apparatus**

The face stimuli were presented on PC-compatible computers with 43 cm (diagonal) color monitors at 800 \( \times \) 600 pixel resolution. All participants viewed monitors from an unconstrained viewing distance of approximately 60 cm. A numeric keypad of the standard PC keyboard was used for participant responses. All display durations and responses were timed by the PC to ± 1 ms.

**Design and Procedure**

Participants were asked to imagine that they were police officers assigned to patrol a crime-ridden area of a city. Two contrived, artificial crimes, *nolvined* and *hovrict*, were created by the experimenter. Nolvined, participants were told, involves the wrecking and devastation of individual possessions and the imminent intent to inflict harm on someone (see Appendix A). Hovrict, participants were told, involves damage and destruction to personal property and expressed threats to personal safety (see Appendix A). These two contrived crimes were devised to be similar yet distinguishable. In the onscreen instructions to participants, which were also read to participants, the experimenter assigned nolvined to be highly associated with European Americans, and hovrict to be highly associated with African Americans; that is, participants were told that
the former was committed most often by European Americans and the latter most often by African Americans (see Appendix A).

The instructions also informed participants that most people were guilty of a crime, but that a small number of people who have not committed any crime live in this section of the city because they cannot afford to live elsewhere. This allowed for the participants to indicate “do not apprehend” if they did not want to make a crime decision in the absence of evidence of criminal behavior.

Participants were instructed that they would be asked to make two types of judgments each time they were presented with a photograph: a racial categorization and a crime decision. They were told they would be asked to judge whether the target photograph was a European American or an African American, and whether they would arrest the target for the crime of nolvined or hovrict, or would choose “do not apprehend.” (In addition to deciding to arrest for nolvined or hovrict, participants could alternatively indicate that they did not want to arrest for either crime, and thus would select "do not apprehend.").

Participants were given 1 block of 180 trials that displayed 180 faces for 2000 ms each. Participants indicated their judgment about race by pushing a button on the numeric keypad. The keypad technique was counterbalanced such that half the participants indicated 1 for European American and 3 for African American whereas the remaining half indicated 3 for European American and 1 for African American. Participants indicated their crime decision for each face stimulus by pressing either 4, 5, or 6 on the numeric keypad. A judgment of 4 indicated that the participant wished to apprehend for nolvined, 5 indicated that the participant wished to choose “do not
apprehend,” and 6 indicated that the participant wished to apprehend for hovrict. The stimulus presentation block lasted approximately 15 minutes.

All photographs and questions appeared in random order and the participant was always shown the face prior to being asked to make either type of judgment (i.e., about race or arrest). For each participant, in half of the trials, the race-categorization question preceded the crime-decision question (i.e., race-to-crime), for the other half of the trials, the crime-decision question preceded the race-categorization question (i.e., crime-to-race). The computer program was developed so as to randomly ask race-to-crime set of questions and crime-to-race set of questions each, one time per stimulus. That is for each of the 90 faces, participants indicated judgments for both race and crime in the race-to-crime ordering and again in the crime-to-race ordering. However, the ordering of the two types of judgments was randomized for all participants, so that race-to-crime or crime-to-race was not always asked first for any stimulus or for any participant.

Base Rate Creation

To obtain objective comparisons between participants' tendencies to arrest for a stereotypic crime and what I will loosely term “true guilt,” base-rates were created. True guilt in this experiment refers to an assignment of guilt for either nolvined, hovrict, or “do not apprehend” (see base rates listed in Appendix C, Figure 1). Arbitrary assignment and artificial crimes were used to create the experimental design base rates. The base rate values for race and crime were subjectively, and thus somewhat arbitrarily, created. These base rates were designed to reflect strong stereotypes that European Americans commit nolvined and African Americans commit hovrict. To observe a
potential racial profiling effect, the criterion values could not be so high such that an
effect could not be observed, for instance 30 arrests for nolvined out of 30 European
American photographs. Thus, the base rates for race and crime were set as follows. The
unambiguous race face photographs (i.e., unambiguously European American and
unambiguously African American) were paired with their corresponding highly
associated crime 20 of 30 times; they were paired 5 of 30 times with the other race’s
crime (i.e., the opposite crime); and they were paired 5 of 30 times with the “do not
apprehend” choice (which was not associated with either race).

To test whether ambiguous faces would be perceived as African American or
European American, a new set of base rates for perceiving race were developed to
compare participants’ categorizations of ambiguous-race faces against. Ambiguous race
to see
categorized as European American). In condition 2, the base rates favored more European American race judgments; it was expected that 60% of the faces would be categorized as European American (and the other 40% of the faces would be categorized as African American).
RESULTS

Five major hypotheses were tested. In addition to the first two specific hypotheses about racial profiling, the Markov property was tested as well. The Markov property states that the future state of the system relies only on the present state of the system, not any previous states. This can also be understood in terms of “conditional independence of the stimulus” which says that once the categorization process has been initiated and concluded there is no further information relative to the succeeding decision response that is associated with the stimulus. In this study, the Markov model would predict that the face photograph itself would not contribute to the decision to apprehend once the participant has categorized the photograph as belonging to one race or the other. The Markov property states that categorization should predict decision making, and that decision making should predict categorization. One aspect of the Markov model is that no temporal ordering effect should occur; that is, it should not matter whether the participants made a race categorization first or a crime-apprehension decision first if the Markov property holds.

Of the 189 participants, five were excluded from all analyses because they provided at least one response after 750 ms had elapsed. The 750 ms criterion was set because it was 2 standard deviations from the mean of response times for all participants. This resulted in 184 participants’ responses for analyses.
The first hypothesis was that evidence would emerge indicating that participants were using racial profiling behaviors, and this indeed was observed and it was observed regardless of target race. I tested for the presence of racial profiling by examining the race-categorization and arrest decision-making processes of participants. As discussed earlier, racial profiling behavior would be evidenced by participant judgments that favor indicating that they are most likely to arrest someone for a highly stereotypical crime based on the racial category they believe the individual belongs to, above and beyond the base rates.

To test hypothesis 1, several chi-square ($\chi^2$) tests were conducted. Specifically, for all 184 participants, and only using the responses to the 60 unambiguous faces, 11,040 $\chi^2$ tests comparing observed participant responses to the expected values given by the rates were conducted. If participants could perceive all aspects of the design “correctly,” namely the correct racial categorizations and the correct crime apprehensions set forth by the experimental design in the base rates, the null would be satisfied as the participants would have been acting in accordance with the base rates (see Figure 1).

Of the 184 participants, 135 violated the null test for base rates in the unambiguous trials. Given an alpha level of .05, approximately 9 violations are expected by chance. Therefore, this indicates that participants apprehended for a stereotypical crime at a significantly greater frequency than expected by base rates.

The second hypothesis stated that decision-making about crime apprehension will be dependent on racial categorization and racial categorization will not be dependent on the crime decision. Indeed, the race-then-crime temporal order best fit all the data, including, importantly, the data obtained when participants were asked to indicate crime.
before race. This finding violates the Markov property which states that there is no
temporal order between judgments.

The second set of $\chi^2$ tested whether participants violated independence in a given
direction. Following the example of Townsend et al. (2000), probabilities were assigned
such that for each European American stimulus there is a probability, $P_1$, that the
participant will classify the face (correctly) as European American (see Figure 2), and a
probability, $1-P_1$, that she or he will categorize the stimulus (incorrectly) as African
American; similarly, given an African American stimulus, there is a probability, $P_2$, of
the participant categorizing the face as European American and a probability, $1-P_2$, of
categorizing the stimulus as African American. The participant is given no feedback
about the accuracy of any of their race judgments.

$\chi^2$ statistics for each participant were created to test the predicted probabilities in
both the trials that asked the racial categorization first followed by the decision to
apprehend (race-then-crime) and in the trials that asked the decision to apprehend
followed by the racial categorization (crime-then-race). Specifically, the equations that
compute the probabilities are as follows: $P_1 = P(\text{categorization} = \text{EA race} \mid \text{EA photo}); 1-
P_1 = P(\text{categorization} = \text{AA race} \mid \text{EA photo}); P_2 = P(\text{categorization} = \text{AA race} \mid \text{EA
photo}); 1-P_2 = (\text{categorization} = \text{AA race} \mid \text{AA photo}); Q_1 = P(\text{decision} = \text{nolvined} \mid
\text{categorization} = \text{EA race}); 1-Q_1 = P(\text{decision} = \text{hovrict} \mid \text{categorization} = \text{EA
race}); Q_2 = P(\text{decision} = \text{nolvined} \mid \text{categorization} = \text{AA race}); 1-Q_2 = P(\text{decision} = \text{hovrict} \mid
\text{categorization} = \text{AA race}); Q_3 = P(\text{decision} = \text{no apprehend} \mid \text{categorization} = \text{EA
race}); 1-Q_3 = P(\text{decision} = \text{no apprehend} \mid \text{categorization} = \text{AA race}).
To control the Type I error rate, a Bonferoni adjustment for multiple tests was used to insure that the overall alpha remained at .05 for this family of tests; this brought the individual-test alpha level to .016 per test. The race-then-crime order was never violated across all three racial categories (i.e., European American, African American and ambiguous race). Crime-then-race, however, was violated 47 times for European American faces, 31 times for African American faces, and 0 times for the ambiguous faces. This result reveals that racial profiling occurred (i.e. participants relied on race to make a decision to apprehend someone for a stereotypic crime) and refuted that participants were just associating race and crime together, because crime-then-race was not upheld; just race-then-crime which is necessary temporal order to exhibit the occurrence of racial profiling when in conjunction with the base rate violations in hypothesis 1.

The third hypothesis was that participants would be more likely to categorize ambiguous photographs as African American than European American. This result, in fact, was observed; participants categorized ambiguous photographs as African American twice as often as they categorized the ambiguous faces as European American. In particular, a simple frequency count of racial judgments of the ambiguous faces was conducted. Of the 60 times that ambiguous faces were shown to participants, there were 2400 judgments that the race was African American and 1200 judgments that the face was European American and further analyses indicated that this difference was statistically significant ($F(1, 59) = 2.716, p = .009$).

The fourth hypothesis stated that participants would be more likely to choose “do not apprehend” when presented with ambiguous race stimuli than when presented with
unambiguous race stimuli. This hypothesis was also supported. Participants chose “do not apprehend” significantly more often for ambiguous faces than for unambiguous faces ($F(1, 182) = 246.75, p = .000$).

The fifth hypothesis stated that participants would select “do not apprehend” for the ambiguous faces they categorized as European American more than for the ambiguous faces they categorized as African American. This hypothesis was not supported. In fact, a significant effect in the opposite direction emerged: Participants were more likely to indicate “do not apprehend” for the ambiguous faces judged as African American than for the ambiguous faces judged as European American ($F(1, 163) = 14.52, p = .000$). (To control the Type I error rate, a Bonferroni adjustment for multiple tests was used to regulate alpha at .05 for this family of tests, bringing the individual test alpha level to .008).
DISCUSSION

The main purpose of this study was to test for the presence of racial profiling by examining categorical and decision-making processes in an adult sample. The first two hypotheses were, taken together, tests for the presence of racial profiling behavior. Three additional hypotheses were tested that examine the extent to which categorization and decision-making differ between unambiguous and ambiguous faces.

The results of this study indicate that racial profiling behaviors were evident among participants. At the same time, participants’ rates of racial profiling did not significantly vary across European American and African American targets. In addition, I tested the Markov property (i.e., that the future state of the system relies only on the present state of the system, not any previous states) using a new set of parameters -- racial categorization and crime decision-making -- and results suggested that the Markov property does not hold under these particular parameters. The decision to interact with a representation of another person (i.e., the photograph of an African American face) is the interaction decision that I tested. Specifically, results indicated that participants create quite different interaction decisions in race-then-crime judgments than they do in crime-then-race judgments.

Five major findings emerged from this study. The first is that participants did racially profile the targets. That is, they apprehended targets for the stereotypical crime (i.e., the race-matched crime) at a rate significantly higher than the base rates for those
crimes. Interestingly, participants tended somewhat to over-apprehend -- or racially profile -- European Americans more so than African Americans, though this difference did not reach statistical significance. This supports previous research from the stereotype literature (e.g., Devine, Monteith, Zuwerink & Elliott, 1991; Katz & Hass, 1986; Kinder & Sears, 1981; McConahay, 1986; Monteith, 1996; Sears & Kinder, 1971) that people have become much more savvy about prejudice and avoid associating themselves with racial bias; it is also consistent with findings from prior base rate studies (Kahneman et al., 1973) in that participants tended to overestimate the amount of arrests for both populations given the base rates, and it supports the interaction decisions literature (Townsend et al., 2000) in that people relied heavily on categorical information to make their interaction decision -- that is, their apprehend decision -- with unambiguous stimuli. At the same time, these results are not consistent with research pertaining to the shooter bias (Correll et al., 2002) and other well-documented patterns indicating that African Americans are targeted more than European Americans, in that, participants did not arrest African Americans for their assigned stereotypic crime more than they arrested European Americans for their assigned stereotypic crime.

Second, results indicated that in every meaningful comparison among race and crime, decision-making about crime was dependent on the racial categorization decision; however, racial categorizing was not dependent on the crime apprehension decision. This finding is consistent with the findings of Townsend, Silva, Spencer-Smith, and Wenger (2000), and at a general level replicates their findings. That is, Townsend et al. demonstrated that the interaction decision to be friendly or hostile to an alien was dependent on what type of alien their participants had categorized the stimulus to be. The
results of the present study replicate this finding, and do so using both more stringent base rates and a genuine social categorization (i.e., race), as opposed to two contrived alien categories.

The racial profiling finding, if replicated and later confirmed, has implications for law enforcement. Namely, it illustrates that people act on racial stereotypes when making very important decisions about the guilt or innocence of a person. As a society, we know this is in violation of the 1st, 4th, 5th and 14th Amendments and law enforcement officers should be aware that it violates an Executive Order titled “The End Racial Profiling Act” that bans the use of racial profiling in any U.S. law enforcement setting and is enforced by the Justice Department. The End Racial Profiling Act defines racial profiling as using racial or ethnic stereotypes to target people for law enforcement attention. Racial profiling also contributes to potentially guilty people not being apprehended for a crime they committed if they don’t belong to the race with the racial stereotype for that crime. Racial profiling is bad police work that allows innocent targets to have their civil rights violated and criminals that do not belong to the stereotypic group to go unattended.

Third, participants were twice as likely to categorize ambiguous photographs as African American than European American. This runs contrary to findings from previous research (e.g., MacLin et al., 2001). In the absence of salient racial features such as hair, participants were expected to have difficulty categorizing racially ambiguous photographs as African Americans – which would be evidenced by participants’ about equally categorizing ambiguous faces as African American or European American (see MacLin, 2001). Although these face photographs were developed to be, and tested as, ambiguous in terms of race, and as such were absent of
hair (a salient racial feature), participants showed apparent sensitivity to other classical African American racial features. It may be the case that we live with the decades-old legacy of racism and "one-drop" rules that perhaps have facilitated oversensitivity to African American facial features (see Beutler et al., 1996).

Fourth, participants were more likely to decide to “not apprehend” when presented with ambiguous race stimuli than when presented with unambiguous race stimuli. This finding is not surprising given the representativeness of the unambiguous stimuli in their racial category. More specifically, the photographs selected as African American and European American are much closer to the prototype for those races than are the ambiguous photographs. This result has not previously been documented using racially-ambiguous stimuli. One possible explanation for this finding is that participants did not want to racially profile targets for a stereotypic crime if they were unsure that the target belonged to that race.

Finally, when presented with ambiguous faces they had categorized as African American, participants tended to choose “do not apprehend” more than when they had categorized ambiguous faces as European American. This finding was not consistent with hypothesis 4 and it is possible that this might be inconsistent with racial profiling African Americans. One would expect that participants presented with an ambiguous face would show stronger reliance on the well-learned social stereotype that African Americans are more criminal (Correll et al., 2002; Cosby, 1994; Devine & Elliot, 1995; Gerbner, Gross, Morgan, & Signorielli, 1986). However, participants may have questioned whether the ambiguous faces could truly belong to the African American race.
and this might explain why there was a weaker association between race and crime in the ambiguous faces that were judged to belong to the African American race.

These finding lend further support that people are using race to make decisions about who to apprehend. It may also be the case that ethnic group members that attempt to distinguish themselves from the prototypicalities of their race may be able to side step the menacing effects of racial profiling or possibly stereotyping if they are successful at making their appearance ambiguous with respect to race. Based on result 4 alone, African Americans in particular would benefit from privileges that ambiguating race allows.

The results of this study provide a number of new findings, which at the same time, reflect a theme found throughout much prior research. Specifically, the present results are consonant with studies finding that the detection of racial prejudice can be highly varied and contextual (Biernat, 2003; Biernat & Fuegen, 2001; Biernat & Kobrynowicz, 1997; Dunton & Fazio, 1997; Fazio, Jackson, Dunton, & Williams, 1995; Devine, Monteith, Zuwerink, & Elliott, 1991; Katz & Hass, 1986; Katz, 1981). The more unique contributions of this study have to do with its focus on assessing for the presence of racial profiling. Indeed, racial profiling behavior was observed among study participants and there were no differences in the rates at which unambiguous African American and unambiguous European American targets were racially profiled. Additionally, two important findings emerged that add significantly to the cognitive and pragmatics literature on the Markov model. Again, the Markov model proposes that the order of categorizations and decision-making responses should be of no consequence, however, this was rejected. Specifically, there was strong evidence that categorization
precedes decision-making even when people are forced to give the crime-apprehension
decision-making response prior to the racial categorization.

The ambiguous race allowed a test of categorical representativeness and the effect
on decision-making. Recall that representativeness refers to similarity on an important
characteristic between a sample and its corresponding population. When the
representativeness was decreased in the ambiguous stimuli, participants chose the “do not
apprehend” choice at a rate that was far below that which they had been instructed to do.
In other words, if people judge a target to be African American but the target does not
seem to fully belong to that racial category, people will be less likely to arrest the target
for the stereotypical crime. This particular finding may be quite important to researchers
interested in the construct of phenotype as well as the construct of privilege of passing as
European American (see Betancourt & Lopez, 1993; Beutler, Brown, Crothers, Booker,
Seabrook, 1996; Bianchi, Zea, Belgrave, & Echeverry, 2002; Helms, 1994; Helms &
Talleyrand, 1997; Root, 1992).

**Limitations and Future Directions**

Although this study represents an advance in psychologists' understanding of
racial profiling behavior, several limitations warrant discussion. First, during debriefing
(see Appendix B), participants were not asked to provide their estimations of the study's
purpose. Doing so would have generated a clearer picture of how motivated participants
were to respond in ways that did not endorse racial profiling. Others (e.g., Devine, Plant,
Amodio, Harmon-Jones, & Vance, 2002), for example, demonstrated that people are
motivated to self-regulate race bias, and the more implicit and elusive the task, the more
racial bias was revealed by their participants who were motivated to respond without prejudice.

Second, and related to the first limitation, is that perhaps some of the unexpected findings that favored African American targets and more readily penalized European American targets could be explained by social desirability. The use of social desirability covariates probably would have been informative in better understanding this study’s results, and inclusion of this control is recommended for future research in this area. Recall that it is thought that social desirability and self-presentational styles affect responses to explicit measures prejudice and, in fact, weaken the correlation between explicit and implicit measures of prejudice (Brauer et al., 2000; Wittenbrink, Judd, & Park, 1997), it is possible that social desirability may have weakened the sensitivity to detect racial bias and racial profiling in this study much as it is thought to have weakened the association of explicit and implicit prejudice in the studies cited above.

Related, there may have been a “Hawthorne effect” that confounded social desirability and participant performance. That is, participants may have been acutely aware of being observed and therefore may have tried to optimally perform. They may not have known what “optimal” responses were, however, and therefore easily could have acted on their beliefs of what was optimal. Any strategy that they may have used, whether based on a correct or incorrect guess, has the potential to obscure the results of the study, this is especially important when noting the absence of differences in the tendencies to racially profile African Americans and European Americans as observed in the results for hypothesis 1.
Third, the study does not have differing level of the dependent variable in each crime. Differing levels of the dependent variable would have allowed an examination of the threshold needed for participants to make zero sum and non-zero sum decisions for each of the races (see Biernat, 2003). The actions of the participants in this study amounted to non-zero sum behavior. That is, without rewards, punishment, or other feedback, participants were not forced to make difficult decisions. Theoretically, participants could demonstrate political correctness and feel good about their lack of bias favoring African Americans without having to take away any resources (i.e. “do not apprehends”) from the European Americans. To be more clear, participants were at no time given a finite number of arrests to make or a limit on the number of “do not apprehends” that they were allowed to make that may have unearthed some stereotypic bias if they had been subjected to a zero-sum choice. The limitless amount of non-zero sum behaviors at the participants’ disposal might even lead to a sense of “moral credentials” that they could use to justify that they do not act biased, and yet it may make them more likely to discriminate in a zero-sum way at a subsequent time (see Beirnat, 2003; Monin & Miller 2001).

Fourth, the external validity of this study is limited. Because the crimes were arbitrarily contrived and had been intentionally designed to do little to spark existing racial stereotypes, little can be said about how real social stereotypes would affect judgments of criminal apprehension. In addition, this study was designed to examine criminal apprehension behaviors with known base rates. Objective crime base rates, however, are not known in the real world, for example, by law enforcement personnel. In fact, base rates are very difficult to obtain in terms of criminality because of the before
mention confound of the U.S. Justice systems. It is possible that law enforcement may be behaving in accordance with their own subjective sense of the base rates. Particularly, base rates may vary depending on location. For example, one police officer may consider the entire U.S. her population whereas another officer may think of a two-mile radius in a particular neighborhood as her population to assume is the base rate for African Americans and African Americans who commit drug offenses. Both the arbitrariness of the crime and base rates should be taken into consideration with respect to the study’s external validity.

Finally, the participants elected to be in the sample and were not randomly sampled, which both create limitations in terms of generalizability. Of course even given random sampling procedures, generalizability is not assured and threats to external validity persist. The best way to examine the extent to which these finding would generalize to U.S. adult populations would be to replicate or heteromethod-replicate this study with other samples (i.e., Biernat & Eidelman, 2003; Biernat & Kobrynowicz, 1997; Correll et al., 2002). As always, prior to such replication, it is important recognize that these findings may represent sets of spurious results and sample artifacts that coincidentally converge on a hypothesis.

This is the first study to date that has examined questions pertaining to racial profiling in an experimental context. Overwhelmingly, across stimuli classes it was demonstrated that behaviors consistent with racial profiling were used when making decisions to apprehend targets. However, as previously mentioned, there was very little “real world” representation in the experiment. Future research would be enhanced by layering in behavioral choices that may alter the manifestation of racial profiling. By
giving participants the opportunity to make decisions that include behaviors in addition to just categorizing the race and making the decision about crime apprehension, a more substantial conclusion about racial profiling processes could be made. For instance, following the lead of the shooter bias study (Correll et al., 2002), subsequent researchers might vary weapon possession in the photographs or vary the criminal record of the target.

Related, caution should be used in attempting to extrapolate these findings to behavior in law enforcement settings and other civic institutions. I set high, but arbitrary, baselines on contrived stereotypes. Unlike the complex tasks that police officers perform, in this study I created base rates that were arbitrarily mapped onto the targets. In the real world, law enforcement personnel never objectively know the guilt or innocence of their suspects, or the corresponding base rates for guilt or innocence of their racial group. Even after an acquittal, conviction, or even a confession of guilt, there is no certainty about true guilt because of the lingering effects of racism and other cognitive and social-psychological phenomena which exert influences in the U.S. criminal justice system.
APPENDIX A

Participant Instructions

Imagine that it is the near future, and you are a law enforcement officer assigned to patrol the streets of one section of a major city. This part of the city is known to be home to a large number of criminals, who are primarily involved in one of two types of violent, destructive crimes. Almost every one of the residents of this section of the city is guilty of one of these two crimes, although there are a small number of people who have not committed any crime and have to live in this section of the city because they cannot afford to live elsewhere. It is your job to bring to justice everyone who is responsible for these crimes. The population of this area of the city is racially balanced, with about half the residents being Black and half being white. But the crimes these two racial groups commit are very different and distinct.

The first crime is called Hovrict. This crime involves damage and destruction to personal property and expressed threats to personal safety. The majority of the times, those responsible for Hovrict are Black; still, in a small number of cases, those responsible for Hovrict are White.

The second crime is called Nolvined. This crime involves the wrecking and devastation of individual possessions and the imminent intent to inflict harm on someone. The majority of the time, those responsible for Nolvined are White; still, in a small number of cases, those responsible for Nolvined are Black. Imagine now that you
are walking your beat. You will meet a number of people, and you will need to make two
decisions. First, you will need to decide if the individual should be apprehended for
Hovriect, should be apprehended for Nolvined, or should not be apprehended at all. In
addition, you will need to decide if the individual is Black or White. You have very little
time to make these decisions: if you are wrong about the decision as to whether you
should apprehend someone, there is a high likelihood that person may go on to commit
another crime. You also need to be as accurate as you can be with the limited time that
you have.

During this part of the experiment, you will be making both decisions. Sometimes
you will need to make the decision about apprehending the person first, other times you
will need to make the decision about the person's race first. You will be using the
numeric keypad to make your decisions.

To make your decision about apprehending, you will use these three keys:

4 = apprehend for Hovriect
5 = do not apprehend
6 = apprehend for Nolvined

To make your decision about race, you will use these two keys:

1 = Black
3 = White
APPENDIX B

Playing Detective Debriefing Statement
Spring 2004

This is a study of perception and decision making for visual information. Our goal was to answer important questions about how visual information is perceived, stored, and remembered in various settings. Specifically we wanted to see if racial profiling can occur in a mock law enforcement setting. We were not interested in individual differences or individual scores, rather, we want to see if as a group there is evidence that racial profiling occurs.

Your participation involved seeing human faces and making simple judgments about what category the face belongs in and for what crime you would arrest that person.

Your individual responses to the study are private and will remain anonymous and confidential. Any published summaries of this study will refer only to the grouped, not individual data.

We must ask you to be sure to refrain from discussing your experiences with and responses to this survey with others. We are interested in each individual’s personal opinion. Sharing information about the study with potential participants will bias their expectations of and responses to the study.

If you have any questions, please ask the experimenter now. If you have any questions later, please contact either Dr. Michael Wenger, Department of Psychology, University of Notre Dame (574-631-9429, mwenger1@nd.edu) or Hilary Cisco (631-4486, hcisco@nd.edu). You will be given a copy of this form to keep.
Figure 1. Diagram of the base rates of crime assigned to each stimulus class.
Figure 2. Diagram of probabilities for each race and crime in all stimulus classes.

P1 = P (categorization = EA race | EA photo)
1-P1 = P (categorization = AA race | EA photo)
P2 = P (categorization = AA race | EA photo)
1-P2 = (categorization = AA race | AA photo)
Q1 = P (decision = nolvined | categorization = EA race)
1-Q1 = P (decision = hovrict | categorization = EA race)
Q2 = P (decision = nolvined | categorization = AA race)
1-Q2 = P (decision = hovrict | categorization = AA race)
Q3 = P (decision = no apprehend | categorization = EA race)
1-Q3 = P (decision = no apprehend | categorization = AA race)
Figure 3. Diagram of the conditional probabilities.
REFERENCES


