UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

GREEBLES AND BRONCOS:

A TEST OF EXPERTISE IN THE CROSS-RACE EFFECT

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

DEBRA CONLY TOWER
Norman, Oklahoma
2012

GREEBLES AND BRONCOS:
A TEST OF EXPERTISE IN THE CROSS-RACE EFFECT

A DISSERTATION APPROVED FOR THE
DEPARTMENT OF PSYCHOLOGY

BY

Dr. Scott Gronlund, Chair

Dr. Teresa DeBacker

Dr. Sowon Hahn

Dr. Lara Mayeux

Dr. Joe Rodgers

This work is dedicated to my mother, my best friend.

Table of Contents

List of Tables

List of Figures

Abstract

People are more accurate at recognizing faces of individuals of their own race than individuals of another race; this well-replicated phenomenon is known as the cross-race effect (CRE). The reasons underlying the CRE are not clear; research indicates that expertise with own- versus cross-race faces may play a role, but there is also evidence that social categorization is a factor. Two experiments examine the roles of expertise and categorization by comparing the effects of cross-race faces to novel stimuli (Experiment 1) and rival university members (Experiment 2). Experiment 1 used inversion as a test of expertise and holistic processing to compare White faces, Black faces, and Greebles. In addition, the EZ Diffusion model was utilized to gather insight into the components underlying the differences in accuracy and reaction time. Results indicate that, for most measures, Black faces shared more similarities with Greebles than White faces, pointing to a strong component of expertise in the CRE. In Experiment 2, upright and inverted images of White and Black males were identified as members of the participants' own or rival university. Again, the EZ Diffusion model was applied to the results. A cross-race effect was found, but contrary to work by other researchers, no cross-university effect was detected, again indicating that expertise, rather than categorization, is a critical factor in the CRE. The results are discussed within the context of models that incorporate both perceptual expertise and social categorization, as well as how factors such as saliency might have affected the results.

Introduction

In 1987, a Black man entered a Tulsa laundromat and attacked an employee, Cyndi Doe[1], a twenty-year-old White woman. She was tied up and locked in the restroom. Eventually, the man returned to the restroom, carried Ms. Doe to a car, then drove to another location and raped her.

Four months later, Cyndi Doe was presented with a five-man photographic array and asked if her attacker was among the images. Cyndi Doe identified Mr. Arvin McGee, who was subsequently prosecuted. During the trial, it was revealed that Ms. Doe had been inconsistent in her description of the man that attacked her, and, in an earlier photographic lineup, had previously picked out a different man. In addition, the defense presented evidence that Arvin McGee was injured at the time the crime was committed and would have been physically unable to carry a person from a building to a car. In spite of these discrepancies, Mr. McGee was eventually convicted and sentenced to 298 years.[2] He served almost thirteen of those years before DNA evidence proved his innocence and identified the true perpetrator of the crime, Oklahoma inmate Edward Alberty ("Know the cases: Arvin McGee," n.d.). Attempts in 2002 to prosecute Mr. Alberty for this crime failed; the statute of limitations had passed. Edward Alberty was released from prison in 2004. He was arrested again in 2007 for theft, and sentenced to 15 years after pleading guilty to five felony charges (Braun, 2007). He is up for parole in April of 2013 ("Oklahoma Department of Corrections: Offender lookup," n.d.).

---

[1] Cyndi Doe is not the victim's real name.

[2] A mistrial was declared for Alvin McGee's first trial, and the second trial resulted in a hung jury. At his

[2] A mistrial was declared for Alvin McGee's first trial, and the second trial resulted in a hung jury. At his third trial, he was found guilty of robbery, kidnapping, forcible sodomy, and rape. The final sentence of 298 years was a reduction of his original sentence.

Cyndi Doe made a mistake. She looked at several photographs and pointed at the one that best matched her memory of the man that attacked her, and Arvin McGee paid for this error with 13 years of his life. The case of Arvin McGee is far from unique; experts have long known that eyewitnesses are often mistaken and that memories are imperfect and easily altered (Münsterberg, 1909; Borchard & Lutz, 1932; Cutler & Penrod, 1995).

The importance of the identification and testimony of mistaken eyewitnesses has increasingly gained attention, due in large part to The Innocence Project (www.innocenceproject.org), a non-profit legal clinic based in New York. This organization, affiliated with the Benjamin Cardozo School of Law at Yeshiva University, uses DNA evidence to exonerate wrongly convicted individuals. As of May 2012, there have been 289 post-conviction DNA exonerations in the United States, 222 of which have been since 2000. On average, those exonerated had served 13.5 years, and 17 had been sentenced to death before their innocence was established and they were released.

While there are a number of important contributing factors to wrongful conviction, such as improper forensics, false confessions, government and/or forensic science misconduct, poor legal representation, and the testimony of dishonest informants, eyewitness misidentification is, by far, the greatest factor in wrongful convictions. Faulty eyewitness identification was a factor in the conviction of over 75% of the individuals that were subsequently exonerated through DNA evidence ("Understand the causes: Eyewitness misidentification," n.d.).

As demonstrated by the case of Mr. McGee, the mistaken eyewitness identification presents a serious dilemma in the legal world (Brigham, 1989; Leippe, 1995; Chance & Goldstein, 1996; Brigham, Wasserman, & Meissner, 1999; Lindsay, Brigham, Elizabeth Brimacombe, & Wells, 2004). Eyewitness identification is a critical component of finding, arresting, and convicting alleged perpetrators of a crime (Devenport, Penrod, & Cutler, 1997; Lindsay, et al., 2004). In spite of the well-established fallibility of eyewitness memory (Brigham, Maass, Snyder, & Spaulding, 1982; Penrod, Loftus, & Winkler, 1982; Cutler & Penrod, 1995), the testimony given in court by eyewitnesses tends to have great influence. Eyewitnesses generally believe their recollection to be accurate, as do the members of the jury (Brigham & Bothwell, 1983; Cutler, Penrod, & Stuve, 1988; Lindsay, Wells, & O'Connor, 1989; Cutler, Penrod, & Dexter, 1990; Sporer, 2001a; Flowe, Mehta, & Ebbesen, 2011).

In the criminal justice system, numerous factors have been identified as contributing to mistaken eyewitness identification. Researchers often categorize the factors as either *system variables*, those factors that are within the control of the criminal justice system, or *estimator variables*, which are beyond the influence or control of the individuals or processes of the criminal justice system (Wells, 1978).

System variables include such factors as the interval between crime and identification, interrogation style, and composition and presentation of the lineup. The manner of questioning (i.e., the use of suggestive questioning, question bias, and open-ended or unstructured interviews vs. leading or structured interview), whether the lineup photographs are presented simultaneously or sequentially, the selection of ‘fillers’ (members of a photo array that are not considered suspects), instructions to the eyewitness, and the reaction by the lineup administrator(s) to the witness’s selection may affect the identification made by the eyewitness, as well as the witness’s confidence in the selection (Wells, 1978; Shapiro & Penrod, 1986). The memory of the witness can be altered, whereby the image of the person s/he selected from the lineup essentially replaces the face of the actual perpetrator in the witness's mind. When the latter occurs, the witness may have complete confidence in the memory; with a very clear mental image of the culprit, when, in fact, the image is of the person seen in a lineup, rather than the actual perpetrator of the crime (Wells & Loftus, 2007).

In part, Mr. McGee's wrongful conviction was due to variables that were under the control of the criminal justice system. According to a letter from R. Thomas Seymour, Ms. Doe's attorney, to Alan Jackere, the prosecuting attorney in Mr. McGee's case (2006), Cyndi Doe was presented with a lineup of five individuals, four of whom lacked any resemblance to her description of her attacker. The city of Tulsa had no policy on assembling a fair lineup, and the men included in the array varied greatly in terms of complexion, age, build, and facial hair. Only one man included in the array matched Ms. Doe's memory: Alvin McGee. In addition, the detective in charge of the case admitted that Cyndi Doe was led to believe that her attacker was included in the photos in presented to her. Also, Cyndi Doe asked repeatedly to hear the voice of the suspect; that memory was very strong. Her request was refused[3]. Whether done intentionally or unintentionally, several decisions and actions made by police officers influenced Cyndi Doe to identify Alvin McGee, and, as the lead detective in the case admitted during the trial, this identification was the only direct evidence the prosecution had against Mr. McGee.

[3] When, some 15 years later, DNA evidence identified Edward Alberty as her assailant, Ms. Doe claimed that upon hearing his voice, she immediately recognized it as the voice of the man who had kidnapped and raped her.

The conduct of the police may have distorted Cyndi Doe's memory of her attacker's face, but it is possible that it was not all that clear in the first place. Research indicates that recognition accuracy decreases under conditions of negative arousal (Deffenbacher, 1984; Shapiro & Penrod, 1986; MacLin, MacLin, & Malpass, 2001). The trauma of the experience could have negatively impacted her ability to remember her attacker. This potential factor is an example of an estimator variable, a variable that is not under the control of the criminal justice system. Estimator variables are often related to the circumstances of the crime, such as the presence of a weapon (referred to as *weapon focus*) (Loftus, 1979; Steblay, 1992), and factors that affect visibility, such as lighting or distance, which can affect how well the witness is able to encode, store, and retrieve the memory of the perpetrator (Wells, 1978; Cutler & Penrod, 1995; Wells, Wright, & Bradfield, 1999).

According to the Innocence Project, one of the most important variables in eyewitness misidentification is race: the race of the victim and of the perpetrator, and whether they are the same or different. Of the 289 DNA exonerees, 62% are African American, Caucasians comprise 28% of the cases, and about 7% are Latino. Cross-race identifications constitute at least 40% of the cases in which mistaken eyewitness identification played a role. Two-thirds of exoneration cases for the Innocence Project are cross-race cases of sexual assault; in contrast, cross-race sexual assaults comprise only 15% of all rapes in United States ("Facts on post-conviction DNA exonerations," n.d.). According to Peter Neufeld, the co-director of the Innocence Project, several factors play a role in this large discrepancy, including, "a long history of the American criminal justice system treating the rape of a White woman by a Black man as a particularly vile crime. One consequence of treating such crimes with particular zeal is that people of color will be wrongfully convicted more frequently," (Patton, 2009). In addition, the socioeconomic status of many of the accused may impede their ability to hire adequate legal assistance.

Systemic racial bias may have played a role in the case of Arvin McGee. In 2006, Arvin McGee Jr. was awarded $12.25 million in a wrongful conviction lawsuit, which claimed that the Tulsa police acted with "deliberate indifference" in the manner in which they conducted the photographic lineup. It was established that the lineup was biased; under cross-examination, the lead detective, a White male, admitted that Edward Alberty and Arvin McGee were so similar that they could be brothers ("Tulsa to Settle Wrongful Conviction Suit," 2006), though mug shots of Alberty and McGee, taken in 2002 and 2001, respectively, do not appear to show an uncanny resemblance between the two men (see Figure 1). If Cyndi Doe and her attacker were the same race, would she have been able to see beyond the similarity in age and facial hair and recognize that her attacker was not included in the array, or at least felt sufficient doubt as to declare that she did not know? This question is unanswerable, but research demonstrates that an eyewitness to a crime will tend to be less accurate in identifying the perpetrator of a crime if the witness and the perpetrator are not the same race; this phenomenon is part of what is known as the cross-race effect. In this sense, race is an estimator variable; the witness's memory is impacted by the difference in race. Like many of the men that have been exonerated by the Innocence Project, the cross-race effect appears to have been an important factor in Alvin McGee's wrongful conviction.

Figure 1. Oklahoma Department of Corrections: Edward Alberty, 2002 (left) and Arvin McGee, 2001 (right).

The cross-race effect (CRE, also known as own-race-bias) refers to the ability of people to more readily distinguish and accurately recognize faces of their own race than faces of individuals from another race. The CRE is remarkably robust (Yarmey & Jones, 1983; Chance & Goldstein, 1996), having been replicated numerous times and across a variety of conditions (Bothwell, Brigham, & Malpass, 1989; Meissner & Brigham, 2001; Brigham, 2008). While the majority of the studies in this area have examined the cross-race effect for Black and White participants and were conducted in North America (Malpass & Kravitz, 1969; Cross, Cross, & Daly, 1971; Chance, Goldstein, & McBride, 1975), the CRE has been replicated in several countries and across different racial and ethnic groups, including Asian and White (Ng & Lindsay, 1994), Germans viewing Turkish faces, (Sporer, 1999), Mexican American, White, and Black (Platz & Hosch, 1988), Hispanic and Black (MacLin, et al., 2001), Canadian First Nation and White (Jackiw, Arbuthnott, Pfeifer, Marcon, & Meissner, 2008), Whites and Blacks from South

Africa and England (Wright, Boyd, & Tredoux, 2001; Chiroro, Tredoux, Radaelli, & Meissner, 2008), Jewish and Arab (Weimann, Fishman, & Rattner, 1986; Rattner & Fishman, 1998), and the Tswana and Pedi of South Africa (Coetzee, Greeff, Barrett, & Henzi, 2009).

Ideally, cross-race studies include participants of both races tested on both sets of stimuli (for example, Black and White participants presented with Black and White faces), rather than just one group tested on both sets of stimuli (e.g., White participants presented with Black and White faces (Bothwell, et al., 1989; Slone, Brigham, & Meissner, 2000; Meissner & Brigham, 2001; Sporer, 2001b). It is only under these circumstances that a complete cross-over interaction (whereby performance by White participants is worse for Black faces and performance of Black participants is worse for White faces), can possibly be identified and fully examined (Lindsay & Wells, 1983; Sporer, 2001b, 2001a). However, the availability of minority participants is often an issue, as it was for the research presented here, and many studies include only one racially or ethnically homogeneous group of participants. In studies that include both White and Black participants, a cross-over interaction is reliably found, though there has been some inconsistency in the results of meta-analyses as to whether the magnitude of the CRE tends to be the same for Black and White participants (Sporer, 2001b). Bothwell, Brigham, and Malpass (1989) found that the CRE was approximately the same across groups, but other meta-analyses found that White participants tended to exhibit a greater CRE than Black participants; while Black participants were less accurate in identifying White faces versus Black faces, White participants were even worse at identifying Black faces versus White faces, (Anthony, Copper, & Mullen, 1992; Meissner & Brigham, 2001).

A meta-analysis of 30 years of research on this phenomenon (Meissner & Brigham, 2001) found that own-race identifications were 1.4 times more likely to be correct than cross-race identifications. In addition, mistaken identifications were 1.56 times more likely when the stimulus face was of another race than when it was the same race as the participant. Comparing the studies across time, the researchers found that measures of overall cross-race identification accuracy (*A', d'*), as well as the proportion of correct positive identifications, or 'hits,' had increased in more recent years. However, the proportion of 'false alarms,' or incorrect identifications when the face was not present, had also increased. This trend may be due to a shifting response criterion; participants with a liberal response criterion tend to respond that a face has been seen before, whereas those with a conservative response criterion will expend the time and effort to arrive at a greater level of internal certainty before indicating that a face had been presented previously. The meta-analysis found a significant cross race effect for response criterion in that participants were more likely to have a liberal response criterion with cross-race faces than when the face was of someone of their own race. There are several possible contributing factors to a liberal response criterion, including an overall lower concern regarding accuracy and/or a sense that extra effort will not necessarily help in terms of achieving greater accuracy. However, the tendency to default to making a positive identification rather than a negative identification (declaring that the face had not been seen previously) has serious repercussions when considering the real-life implications of an eyewitness who is too willing to identify someone in spite of uncertainty or a poor memory (Wells & Loftus, 2007).

The consequences of the CRE can range from devastating, as exemplified by the case of Arvin McGee, to socially awkward. An individual that wrongly identifies or fails to recognize a person of a different race may feel embarrassed or judged by others as prejudiced. The person that is misidentified could interpret the lack of recognition as a racially induced slight or insult, and may feel discounted and stereotyped (Hugenberg, Young, Bernstein, & Sacco, 2010). Future interactions could well be negatively impacted by incidents of cross-race misidentification; perhaps even the intuitive knowledge that we tend to recognize individuals of a different race less well inhibits interaction in the first place (Brigham, 2008). Studies have demonstrated that individuals with developmental prosopagnosia (or 'face blindness', a condition characterized by severe difficulty in recognizing familiar faces) often prefer to avoid social interaction due to feelings of embarrassment and anxiety about offending others due the inability to recognize people (Kress & Daum, 2003; Yardley, McDermott, Pisarski, Duchaine, & Nakayama, 2008). Even mild prosopagnosia can hamper social interaction and may contribute to higher levels of shyness or social phobia: in one study, compared to non-shy peers, shy children were less accurate in facial discrimination, but only when faces differed in the spacing of features; there were no significant differences for faces that had different sets of eyes and mouths, or for faces that differed in the shape of external contour (Anzures, Quinn, Pascalis, Slater, & Kang, 2009). If it can be said that we tend to suffer from some degree of prosopagnosia for cross-race faces, when we are also acutely aware of the importance of the face and interpersonal recognition in terms of facilitating social interaction (Ellis, 1990), the cross-race effect may be, in part, a self-perpetuating deficit.

In spite of decades of research and relatively consistent findings for a CRE, there is not a definitive answer to why it occurs (Teitelbaum & Geiselman, 1997). In general, there are two main theoretical explanations that underlie the proposed factors of the CRE; the perceptual expertise model or a model based on social-cognitive influences. Various mechanisms have been proposed, including group disparities in facial similarity (Goldstein & Chance, 1976, 1978; Chiroro, et al., 2008), categorization of cross-race members (Levin, 1996, 2000; MacLin & Malpass, 2001), the effect of attitudes and level of contact (Brigham & Malpass, 1985; Wright, Boyd, & Tredoux, 2003), a lack of expertise for cross-race faces (Carey, 1992; Gauthier, Skudlarski, Gore, & Anderson, 2000; Rhodes, Hayward, & Winkler, 2006; Stahl, Wiese, & Schweinberger, 2008; Heron-Delaney et al., 2011) and the use of different cognitive processes (Rhodes, Brake, Taylor, & Tan, 1989; Rhodes, et al., 2006; Greenberg & Goshen-Gottstein, 2009; Michel, Corneille, & Rossion, 2010; Richler, Cheung, & Gauthier, 2011). Although many researchers may tend to favor one model over another, these proposed bases for the CRE are not considered mutually exclusive; recent models (Sporer, 2001b; Susa, Meissner, & de Heer, 2010) combine perceptual as well as social-cognitive components of the cross-race effect.

One of the earliest proposals, that there are between-group differences in physiognomic homogeneity, has not been supported (Chance, et al., 1975; Goldstein & Chance, 1976, 1978, 1979). Japanese faces, for example, have not been found to be quantifiably more similar to each other in terms of facial features than White faces, and there were no differences in homogeneity for Black faces when compared to White or Japanese faces (Goldstein, 1979). Also, given that the CRE commonly occurs, at least to some extent, for both groups of participants (i.e., both groups perform better with same race faces), and that the CRE has been found even when the categories of race were artificially projected on ambiguous race faces (MacLin & Malpass, 2001; Corneille, Huart, Becquart, & Brédart, 2004) differences in racial physiognomic variance do not provide a sufficient explanation.

A participant's classification of a face into racial category, and not the face itself, can lead to errors in memory that contribute to the CRE. If a person holds beliefs or attitudes about a group of people, the categorization of a face into that group may interfere with processing of that face as an individual person. There is evidence that memory for cross-race faces may be affected by the viewer's facial schema of the racial or ethnic category. Using prototypical faces from two distinct racial categories, Corneille, Huart, Becquart, and Brédart (2004) created several continua of mixed race faces. For example, an Asian face and a Caucasian face were used to create one set of morphed faces, and a North African and a Caucasian face were used to create another. At encoding, participants viewed faces from various locations on the continuum (for example, a face derived 50% from the Asian face and 50% from the Caucasian face). Subsequent testing, in which the target face was presented with foils from other locations of the same continuum, found that moderately ambiguous faces (30% and 70%) were often remembered as being closer to the racial source face than they actually were. This phenomenon has also been demonstrated for racially ambiguous Caucasian/Hispanic faces (MacLin & Malpass, 2001). Such a distortion in memory towards a racial prototype points to a tendency to endow faces with features deemed to be typical of a category. This type of shift in memory is regarded as one possible factor in the CRE.

On the other hand, a portion of CRE research involves simple in-group/out-group categorization for which there are no identifiable facial prototypes. When participants were presented with photos of students designated either as from the participant's or a rival university, 'fellow' students were more accurately recognized at test than the 'rivals' (Bernstein, Young, & Hugenberg, 2007). In reality, the photos were of people from neither university, and the presentation of a face as belonging to the in-group or out-group varied between participants. In this case, faces of out-group members were not recognized as well as faces of in-group members, even when the stimuli were arbitrarily assigned to one group or the other. As such, distortions of memory towards a prototype cannot explain the differences in memory accuracy between the facial categories; rather, in-group membership may enhance memorability, whereas out-group membership may weaken it.

Possible explanations include attentional differences at encoding and/or test: participants may pay more attention to the faces of their perceived allies because the ability to recognize them at a later time is somehow advantageous. Moreover, future interactions with in-group members are considered to be more likely than with out-group members, thereby increasing the chance of an opportunity or need for facial recognition (Harvey, Yarkin, Lightner, & Town, 1980). Researchers have proposed that strong negative attitudes towards out-group members could lead to a greater CRE, but this has not borne out; in general, there is not a significant relationship between attitude about a group and facial recognition ability for members of that group (Lavrakas, Buri, & Mayzner, 1976; Brigham & Barkowitz, 1978; Yarmey, 1979).

Alternatively, but leading to similar results, is the idea that people have limited cognitive resources, and the tendency to discount or prejudge an out-group member is simply one of the many cognitive short-cuts that people utilize. Out-group members are categorized, whereas in-group members are individuated. This is referred to as cognitive disregard (Rodin, 1987), and has been demonstrated for age (Wright & Stroud, 2002; Anastasi & Rhodes, 2006), gender (Wright & Sladden, 2003; Rehnman & Herlitz, 2006; Zhao & Bentin, 2008), and attractiveness (Cross, et al., 1971). This tendency is well studied in personality and social psychology; research consistently demonstrates two strong effects in the study of intergroup perception. First, in-group members are perceived more positively than out-group members, and second, in terms of beliefs and attitudes, out-group members are judged to be more homogeneous than in-group members. The latter is referred to as the out-group homogeneity effect, and it occurs whether group membership is real or artificially created, though the effect is diminished for laboratory-created groups (Mullen & Hu, 1989).

Disregard cues depend on the situation, not on the category per se. For example, Rodin (1987) found that while young people are better at recognizing same-age faces than middle-aged or elderly faces, elderly participants did not show a similar lack of ability to recognize young faces. According to the theory, this interaction is due to a difference in cognitive disregard; young people are likely to discount the middle-aged and elderly as unimportant, whereas older people do not demonstrate the same cognitive disregard for younger people.

Consistent with the idea of categorization and cognitive disregard, Levin (2000) found that the White participants who were quickest to categorize the photos of African Americans as Black were found to be less accurate in subsequently recognizing Black faces. Cognitive disregard can also potentially explain the typically asymmetrical interaction of many cross-race studies: while it is true that same-race identification is better than cross-race, White participants are typically worse at recognizing or identifying members of the other race/ethnicity (Asian, African American, Hispanic) than the minority groups are at recognizing Caucasian faces. As the majority population in the United States, White people may be generally associated with power, opportunity, and the possibility of a beneficial interaction. As such, it may be prove to be more useful for minority members to expend the energy to individuate, rather than categorize, White people (Shriver & Hugenberg, 2010). It is plausible, too, that, individual members of minority populations will have, on average, more interactions with White people than a White person will have with a specific minority. For example, an African American, across her lifetime, may interact more frequently with White people compared to a White woman's interactions with African Americans (Harvey, et al., 1980).

A recent test of predictions based on ideas underlying the concept of cognitive disregard produced interesting results: the CRE for White participants viewing Black faces was eliminated when the photos depicted angry Black faces (Ackerman et al., 2006). Furthermore, when attentional resources were limited, the White participants actually exhibited greater accuracy for angry Black faces than angry White faces. The finding can be interpreted as a switching off of cognitive disregard when the situation alters the relative importance of the pictured individual. However, the authors note that categorization and individuation are not mutually exclusive; both can occur, resulting in a curious paradox of increased stereotypical assessments about members of an out-group along with improved recognition ability (Judd, Park, Yzerbyt, Gordijn, & Muller, 2005; Ackerman, et al., 2006)

Some researchers hypothesize that the CRE is due, in part, to a deficit in perceptual memory expertise for cross-race faces as compared to same-race faces (Chance, et al., 1975; Brigham & Malpass, 1985; Ng & Lindsay, 1994; Chance & Goldstein, 1996). The relatively constant exposure to same-race faces leads to an increased ability to focus on the features that are most useful for distinguishing between individual faces within that group, but these features differ between different racial and ethnic groups. This direction of thought is supported by findings that facial features that are generally used to describe or identify a face vary between racial groups; a study by Ellis, Deregowski, and Shepherd (1975) compared the facial descriptions provided by Black Africans and White Scottish participants. While the White participants were more likely to mention hair color, hair texture, and eye color, the Black participants included descriptions of ears, eye size, hairstyle, chin, face shape, eyebrows, and the nose/nostrils.

Chance, Turner, & Goldstein (1982) posit that children develop a face schemata as a result of frequent contact with (same-race) faces, and this allows for fast and accurate processing. Because the relative amount of exposure to other-race faces is often far less, expertise does not develop to the same degree, and the face schemata that facilitates processing for same-race faces does not function well for cross-race faces. Same-race facial recognition in children improves at a greater rate than cross-race facial recognition: in a cross-sectional study on White participants ranging in age from six to 20, found that for the youngest participants, there were no significant differences in recognition accuracy of a White versus and Asian faces, but that the older participants did exhibit the cross-race effect.

If the CRE is due to expertise that develops as a result of the differential between day-to-day exposure to same- versus other-race faces, it follows that if a person *does* have a great deal of cross-race interaction, that individual should become an expert in cross-race faces as well, and the cross-race effect should decrease. The contact hypothesis, which states that increased contact with the other group will improve cross-race recognition ability, has some support in the literature, though it is far from universal (Brigham & Barkowitz, 1978). It was found that children and adolescents from integrated neighborhoods exhibited better cross-race facial recognition than children and adolescents from segregated neighborhoods (Cross, et al., 1971; Feinman & Entwisle, 1976). Some research has indicated the reverse; increased contact is associated with worse cross-race facial recognition ability: in a study by Lavarkas, Buri, & Mayzner (1976), White students that attended a predominately Black high school tended to demonstrate a greater cross-race deficit than White students attending a predominately White high school. Other research found no significant effects: an early study by Brigham and Barkowitz (1978) did not find a relationship between level of interracial contact and recognition ability, nor did Malpass & Kravitz (1969).

In order to test the effect of differential contact on three, rather than two, racial groups, Chance, Goldstein, and McBride (1975) presented Black and White participants with three sets of faces: White, Black and Japanese. The authors pointed out that, in one respect at least, Black faces are more physically homogenous than White faces: Black people have dark hair. The Japanese faces, like the Black faces, all had dark hair; if variety in hair color was boosting the recognition accuracy of White participants for White faces, then Japanese faces and Black faces should be recognized at about the same level. Also, the authors suggested that, due to the dark skin of Black faces, details are less easy to distinguish. If dark versus light skin was a major factor in the CRE, Japanese faces, similar in tone to White faces, should be recognized with higher accuracy than Black faces for White participants. However, given the lack of contact for both groups of participants with Japanese people, if differential experience was responsible for the CRE, recognition accuracy would be the lowest for the Japanese faces for both Black and White participants. The researchers found a CRE for both groups of participants: White performed better with White faces, and Blacks performed better with Black faces. Compared to the other sets of stimuli, recognition accuracy for the Japanese faces was the lowest for both White and Black participants.

Some studies show that the quality, not quantity, of cross–race contact is the more critical factor in reducing the CRE (Brigham & Malpass, 1985). Rather than mere level of contact, some researchers have examined the idea that it is the need to differentiate between people of another race that leads to a decrease in the cross-race effect (Valentine, Chiroro, & Dixon, 1995). Based on the idea that, in order to follow professional basketball, White basketball fans would need to differentiate between players, 75% of whom are Black, Dunning, Li, and Malpass (1998) compared White basketball fans to White participants that did not watch NBA basketball ("novices"), and found that the White fans did exhibit better facial recognition for Black faces than the White "novices". However, a later study on CRE of basketball fans versus novices by

MacLin, MacLin, Van Sickler, and Li (2004) failed to replicate this aspect of the Dunning et al. (1998) study.

Following the idea that differentiation leads to expertise, which leads to a better cross-race facial recognition, researchers have examined whether or not prejudicial/negative beliefs about members of another race are associated with a higher CRE. Early studies seemed to indicate that there was some relationship between memory performance for cross-race faces and level of prejudice or attitudes about individuals of another race (Allport & Kramer, 1946), but more recent studies have not found any relationship (Lavrakas, et al., 1976; Brigham & Barkowitz, 1978; Slone, et al., 2000). However, prejudicial attitudes are related to interracial contact (see Meissner & Brigham, 2001, for a review).

When discussing how faces are processed, researchers in the field of facial recognition tend to distinguish between featural processing and configural processing (Carey & Diamond, 1977; Farah, Tanaka, & Drain, 1995; Greenberg & Goshen-Gottstein, 2009). Featural processing involves the use of individual facial features or parts, such as the nose or eyes. Each feature is treated separately, and not considered in relation to the other parts of the face. Configural or holistic processing utilizes the face as a whole, its size and shape, and the angles and distances between the features. In this type of processing, many characteristics of the faces are incorporated at once. It bears noting that there is not a true consensus as to what is meant by configural processing, and not all researchers equate the terms “holistic” and “configural,” (Mondloch et al., 2010), but in general, and for this paper, the terms are used interchangeably in contrast with featural processing. Studies in the development of facial recognition indicate that as a child develops expertise in facial recognition, the process is increasingly governed by a holistic, rather than featural, process (Carey, Diamond, & Woods, 1980). However, both processes are generally considered to be important in accurate facial recognition (Diamond & Carey, 1986).

Yin (1969) developed an inversion paradigm for examining the process of facial recognition. In the original inversion task, a mixed array of upright and inverted faces or objects is presented to the viewer. At test, the stimuli are presented either right-side up or upside-down (in Yin's method, the encoding and presentation orientation were the same), and the participant is asked whether the stimulus was included in the previous presentation (old) or not (new). The accuracy or latency difference in performance between the upright and inverted test stimuli is the inversion deficit. In this two-choice recognition task, Yin found that inverted faces suffered a greater recognition deficit than other inverted objects, such as houses. According to Yin, this indicated that faces are processed primarily holistically, and that inversion disturbs this. On the other hand, specific features help us to recognize a house, and such features can still be identified when the stimulus is presented upside-down. Inversion at test (or at presentation and test (Valentine, 1988) eliminates the recognition advantage that faces carry and leads to processing that is not as efficient or accurate (i.e., carried out piecemeal or featurally). While Yin and other researchers (Kanwisher, McDermott, & Chun, 1997) contend that these and other findings indicate that faces are special and processed in a way unlike any other stimuli, others (Diamond & Carey, 1986; Gauthier & Tarr, 1997; Farah, Wilson, Drain, & Tanaka, 1998; Gauthier, Williams, Tarr, & Tanaka, 1998) have found evidence that expertise in other types of visual stimuli (for example, birds for a birdwatcher) evoke the same sort of holistic processing, which tends to be highly accurate and efficient.

Inversion experiments used to examine the CRE indicate that cross-race faces are not processed holistically, as they do not suffer the same inversion effect (decrease in recognition accuracy) as same race faces (Valentine, 1988; Golby, Gabrieli, Chiao, & Eberhardt, 2001; MacLin, et al., 2004; Rhodes, et al., 2006; Hancock & Rhodes, 2008; Michel, et al., 2010). For example, White European participants suffered a greater inversion deficit for White faces than Chinese faces, whereas Chinese participants had a larger inversion deficit for Chinese faces. Same-race faces exhibited greater susceptibility to inversion in terms of accuracy and reaction time (Rhodes, et al., 1989). This indicates that people are experts in recognizing same-race faces, but this expertise is reduced for cross-race faces.

Neuroimaging studies indicate that an area in the fusifrom gyrus (often referred to as the fusiform face area or FFA) is particularly important in facial recognition (Kanwisher, et al., 1997; Rhodes, Byatt, Michie, & Puce, 2004). Golby, Gabrieli, Chiao, & Eberhardt (2001) found that activation in this area was not as strong for participants viewing cross-race faces as for same-race faces, but the difference in activation varied between participants. In a subsequent recognition task, the CRE was bigger for those participants that had previously demonstrated the greatest difference in brain activation for same- versus cross-race faces.

If expertise is responsible for differences in own- versus cross-race facial recognition, the use of a non-expert object can provide a control condition against which to examine accuracy and response latency. Greebles (Gauthier & Tarr, 1997) have been used in numerous studies involving facial recognition, including assessment of brain activity. Greebles are computer-generated “creatures” that were designed to be able to compare the processing of faces to a set of novel object, and to examine how the processing changes as people became “Greeble experts.” An example of four Greebles is depicted in Figure 2. Greebles have some characteristics in common with faces: they are bilaterally symmetrical and, though they differ somewhat in shape and detail, the features of Greebles are always in the same configuration, as are facial features. Greebles are very similar to one another, but have a gender, belong to a family, and with expertise, can be individuated. Greebles have names; both a family name and an individual name. For participants without previous exposure to Greebles, recognition attempts are difficult and depend largely on featural processing. As with other non-face objects, participants tend to focus on and try to remember the separate parts of the Greebles rather than using holistic processing. However, experiments have demonstrated that, with training, Greebles are processed like faces; as participants become Greeble experts, recognition is fast and easy, and depends largely on configural processing. (Gauthier & Tarr, 1997; Gauthier, et al., 1998).

Neuroimaging studies show that Greeble experts utilize the part of the brain that is often designated as specialized for facial processing, the fusiform face area or FFA (Gauthier, Tarr, Anderson, Skudlarski, & Gore, 1999). Research utilizing other non-facial stimuli, such as birds and cars, have found similar FFA activation by experts with those objects. It is hypothesized that the FFA, rather than specialized for facial recognition, is utilized when fine distinctions need to be made between familiar objects; essentially, those things we are expert in distinguishing (Gauthier, et al., 2000; Rhodes, et al., 2004).

This research supports the idea the CRE may be due to a difference in expertise with same- versus other-race faces: Without training, Greebles, like cross-race faces, are processed featurally, or piece-meal: participants will attempt to remember aspects of the individual parts rather than the configural characteristics of the Greeble as a whole. However, with experience, Greebles are processed more holistically and in the same part of the brain that is generally associated with facial processing (Gauthier & Tarr, 1997; Gauthier, et al., 1998).

Figure 2. Four Greebles

In inversion tasks, reaction time often is used as a second measure of differences in task difficulty or participant skill. For example, the increase in reaction time for inverted faces (as compared to upright faces) is examined along with a drop in accuracy to assess the effect of the manipulation on the participant's ability to complete the task (i.e., the effect of preventing holistic processing). However, this use of the data fails to consider the contributions of task difficulty and response criterion. This is especially critical given that the application of a signal detection analysis to a CRE two-choice facial recognition task often reveals a lower response criterion for cross-race than own-race faces (Meissner & Brigham, 2001). This is exhibited in, for example, a higher false alarm rate for Black faces that for White faces when the participant is White.

Given the problems inherent with eyewitness identification, an increased tendency to simply pick a person if a lineup is cross-race, and to show greater caution if the lineup is own-race, is a matter of real concern. The difference in accuracy of White participants/witnesses in recognizing White faces over Black faces may be due to a relative inability to distinguish between faces of another race, a tendency to be more cautious when making a selection from a same race lineup, or both. Models, such as the signal detection model (McNichol, 1972; Wickens, 2002) and the diffusion model (Ratcliff, 1978), can help parse whether differences in cross and same race facial recognition tasks are due primarily to discrimination ability (memory) or response criterion (decision), and to quantify the contribution of each.

Reaction time and accuracy exist in a trade-off relationship that can confound the researcher's ability to determine level of skill; while one participant may answer quickly but have low accuracy, another participant may answer more carefully, thereby increasing both reaction time and accuracy. Without examining both variables, there is no way to separate response conservativeness from subject ability, and, even examined together, the researcher cannot readily determine the relative contributions of skill (or task difficulty) and response conservativeness to a participant's performance.

Models allow the researcher to estimate the latent variables that affect performance. The Ratcliff diffusion model (Ratcliff, 1978) utilizes the reaction time for accurate and inaccurate responses to estimate seven parameters, including the rate of information accumulation (quality of the memory), the amount of information required to make a decision (the criterion), and the nondecision (encoding, retrieval, and response) components of processing. While the Ratcliff diffusion model has been found to fit data well from a variety of two-choice reaction time tasks (Ratcliff & Rouder, 1998), the fitting procedure is rather complex, which may discourage researchers from attempting to apply it to their data. It also requires entire reaction time distributions for incorrect decisions, so that in paradigms with a low error rate (such as many facial recognition tasks), the model will not function very well (Wagenmakers, van der Maas, & Grasman, 2007).

The EZ-Diffusion model (Wagenmakers, et al., 2007) is a simplified version of Ratcliff's (1978) diffusion model. It estimates only three parameters: *v*, the drift rate (quality of information), *a*, boundary separation (response conservativeness), and T_{er}, non-decision time. The drift rate and boundary separation determine the decision time, while other processes such as stimulus encoding, memory access and retrieval, and response, etc. are combined into T_{er}. Although the cost of this simplification is less information, the remaining parameters are of those of primary interest. Applying the EZ diffusion model to the results of cross and same races facial inversion tasks allows for separate estimates of discrimination ability, response criterion, and the contributions of encoding, retrieval, and response processes across conditions of race and photo orientation.

In sum, the use of the EZ Diffusion model can provide insight into the components underlying the differences in accuracy and reaction time for the six conditions created by the combinations of stimuli (White faces, Black faces, or Greebles) and photo orientation (upright or inverted at encoding). The inversion effect reveals differences in holistic versus featural processing; if the CRE is due to a lack of expertise in cross-race faces, leading to featural instead of holistic processing, cross-race faces and Greebles should share characteristics in terms of latent variables underlying performance. If differential expertise is not a factor in the CRE, the parameters that summarize the Greeble data should be relatively distinct from same and cross-race faces.

Experiment 1

Method

Participants. A total of 103 University of Oklahoma students enrolled in an undergraduate psychology course participated in study. Participants earned research participation credit for the course in which they were enrolled, and were recruited through an online experiment management system that listed the available studies in need of subjects. Seventy-five of the participants identified themselves as White, non-Hispanic, six identified as Black, non-Hispanic, four participants were White-Hispanic, four identified as Asian, two as multi-racial, one as Black-Hispanic, and eleven as Other. Although no participants were excluded on the basis of race or ethnicity, due to the constraints of the experiment and the demographic make-up of the student population at the University of Oklahoma, in which 75.3% of the students are White non-Hispanic, analysis of the data was restricted to White participants only. Due to computer malfunction, data from eight of the White participants were lost.

Materials.

Photographic Stimuli. During the experiment, participants viewed 80 computer images of photographs of college-age White and Black males of average weight (stimulus images courtesy of Dr. Chris Meissner, Investigative Interviewing Research Laboratory, University of Texas at El Paso, http://iilab.utep.edu/stimuli.htm). In order to eliminate any potential advantage due to a difference in physiognomic homogeneity (Chance, et al., 1975), we selected only images of individuals with dark hair and brown eyes for both the Black and White stimulus sets.

For many years in the field of facial recognition, the standard procedure used identical photographs of each individual at study and at test (Shapiro & Penrod, 1986). Although this consistency serves as a control and allows researchers to explore how experimental variables affect recognition performance, it has been criticized for lacking ecological validity (Laughery, Alexander, & Lane, 1971). In reality, when asked to identify someone, a witness will not see the identical image of the person they saw previously; there may be a multitude of differences, including lighting, expression, age, clothing, angle, facial hair, and hair style. Recognition of one image of a person's face is not necessarily the same as recognizing the person's face (Ellis & Deregowski, 1981; Hay & Young, 1982). Research has shown that simply changing the pose or expression between study and test results in decreased recognition ability (Laughery, et al., 1971; Krouse, 1981). For example, Bruce (1982) found that participants did less well in recognizing faces that had changed in either angle or expression (compared to no change) between study and test; when both angle and expression changed, performance declined even more. Due to our interest in examining facial recognition rather than image recognition, we utilized different images at study and at test.

The stimuli included two images of each individual, one with the individual wearing street clothes and smiling, and the other with a neutral expression, wearing a gray sweatshirt. The images were cropped and digitally altered to depict only face and shoulders against a neutral background at a size of 270 x 360 pixels. No photographs of individuals with long hair, eyeglasses, jewelry, or distinctive facial hair were included. Any facial blemishes or marks were digitally removed. At encoding, participants saw the smiling image; at test, the image with the neutral expression and gray sweatshirt was presented.

The photos had been rated previously by ten individuals for distinctiveness. Research has demonstrated that the uniqueness of a face positively impacts the participants' recognition ability for that face (Shepherd, Gibling, & Ellis, 1991; Valentine, 1991). Distinctiveness impacts both hit and false alarm rate; participants tend to recognize previously seen distinctive faces better than typical faces, and, for distractor faces (that is, faces that are not presented during the encoding phase, but only at test), participants are less likely to mistakenly respond that the face was presented earlier if the face is distinctive rather than typical (Valentine, 1991; Wickham, Morris, & Fritz, 2000).

Participants in this phase of the experiment saw each image and indicated how distinctive or memorable they found the face to be, on a scale of zero (very distinctive; "very easy to remember or pick out of a crowd") to six (very typical; "very difficult to remember or pick out of a crowd"). No photo received a rating of zero or six, and the typicality ratings for the White faces (M = 3.25, SD = .57) and Black faces (M = 3.37, SD = .72) were not found to be statistically different, paired $t(9) = -.40$, $p = .696$, two-tailed.

In addition to the photographs of Black and White males, participants saw images of eighty Greebles (stimulus images courtesy of Dr. Michael J. Tarr, Center for the Neural Basis of Cognition and Department of Psychology, Carnegie Mellon University, http://www.tarrlab.org/). The images were resized so that the Greebles were approximately the same size as the head and shoulder images of the males, and cropped to the same dimensions (250 x 360). In an attempt to prevent a floor effect in terms of accuracy of Greeble recognition, the participants viewed an image of ten Greebles (one of each gender from all five 'families') at the beginning of the experiment. The images were labeled according to family and gender. Participants were instructed to examine the image of the different Greebles for 1 min.

Equipment and software. The photographs and the questionnaire were presented via computer, utilizing the experimental software E-Prime 2.0.1.68 (Psychology Software Tools), allowing for randomization of the order of conditions presented (upright or inverted; Black, White, or Greeble), as well as the specific stimuli presented with each condition.

Procedure. After reading the informed consent form, participants were asked to read brief instructions about the stimuli, as well as the subsequent recognition test. Participants acknowledged readiness by pressing the space bar to begin the experiment.

All participants completed four blocks. Each block consisted of two randomly ordered trials that differed as to the orientation of the encoding stimuli (inverted or upright). The test stimuli were always upright. Each of the four blocks included all three stimuli: faces of White males, faces of Black males, and images of Greebles. The encoding phase consisted of 15 randomly presented images, five each of Black males, White males, and Greebles, presented for 2 s each, with a 1 s fixation screen between each image.

Following the encoding phase of each block, participants worked on a word search puzzle for 2 min. Following this task, participants viewed the instructions regarding the upcoming old-new recognition test. Participants could begin the test once they had completed reading the instructions. During the test phase, 30 images, all in an upright orientation, were presented sequentially. Ten images of each stimulus category (Black, White, Greeble) were presented; five had been presented during the encoding phase, five had not. Participants were instructed to press the letter ‘e’ if the object depicted in the image was old (one they had seen during the encoding phase), or to press the letter ‘i’ if the image was new (not presented during encoding). The order of blocks, inclusion of a particular image in any block, whether an image was presented at encoding and at test or at test only, as well as the order of presentation of images within a block, was completely randomized via E-Prime.

Results and Discussion

Data were analyzed using repeated-measures ANOVA and paired-samples *t*-tests (with Bonferroni adjustment). The design was a 3 (stimulus: White vs. Black vs. Greeble) x 2 (orientation: upright vs. inverted) within-subjects factorial. Seven dependent variables were examined: hit rate, hit rate reaction time, false alarm rate, false alarm rate reaction time, as well as the three parameters of the EZ Diffusion model, v (drift rate), a (boundary separation), and T_{er} (time needed to encode and respond at test).

For hit rate, there was a main effect for orientation of stimulus, $F(1, 63) = 7.891$, $p = .006$, partial $\eta^2 = .112$, a main effect for race, $F(2, 62) = 5.124$, $p = .009$, partial $\eta^2 = .142$, and a significant interaction effect of orientation and race, $F(2, 62) = 3.255$, $p = .045$, partial $\eta^2 = .095$. The interaction effect prevents interpretation of the main effects, and because identifying the specific differences in the effects of inversion on the each set of stimuli is of great interest in this experiment, planned comparisons were conducted to further our understanding of the results.

In the upright condition, there were no significant differences between the hit rate for White faces (M=.722, SD = .224) versus Black faces (M=.668, SD = .229), $t(63) = 1.5$, p = *ns*. This is consistent with existing research which indicates that the CR effect is due, in large part, not to fewer CR correct identification, but rather to a higher false alarm rate (false positive identification) for cross-race faces (Meissner & Brigham, 2001).The hit rate for Greebles (M=.584, SD = .260), was significantly lower than for White faces, $t(63) = 3.521$, $p = .001$. Although participants' ability to correctly identify a previously seen Black face tended to be higher than the hit rate for Greebles, this difference did not reach the necessary level of significance, $t(63) = 2.021$, $p = .047$ [4].

In the inverted condition, there were no significant differences between White faces (M=.573, SD = .205), Black faces (M=.628, SD = .228), or Greebles (M=.556, SD = .238), $ps > .077$.

The inversion effect is examined by looking at the drop in hit rate between the upright versus inverted condition. This is an indicator of whether the stimulus is processed holistically. A significant inversion effect indicates that holistic processing

[4] Bonferroni adjustment based on nine planned comparisons requires a p-value less than $\alpha/9 = .006$ to be significant.

occurs in the upright condition, but that this is disrupted in the inverted condition. In examining the difference in hit rate for White faces, Black faces, and Greebles, only White faces showed a significant decline between the upright and inverted conditions, $t(63) = 3.669$, $p = .001$. Neither Black faces or Greebles had a significant inversion effect, *ps* >.318.

The lack of an inversion effect for Black faces and Greebles indicates that a featural approach may have been used for both sets of stimuli, regardless of orientation, in contrast to the holistic process used for upright White faces. However, in the upright condition, White and Black faces did not differ in hit-rate, indicating that Black faces, while not processed with the same expertise as White faces, share some characteristics with White faces as well as with Greebles. Figure 3 presents a graphical depiction of hit rate and hit reaction time in the upright and inverted conditions for White faces, Black faces, and Greebles.

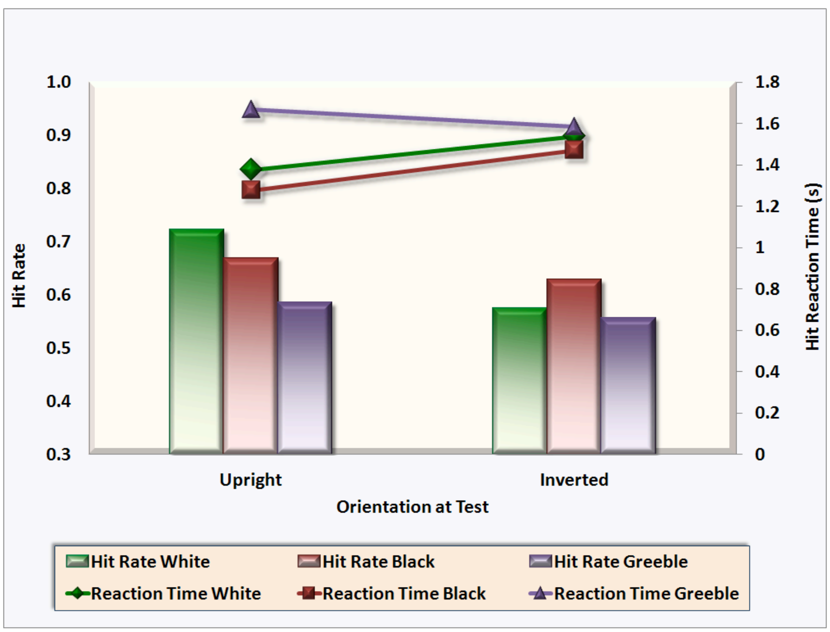

Figure 3. Mean hit rate and hit reaction time (s) as a function of target stimuli and orientation at test in Experiment 1.

Turning to hit reaction time (RT), there was a main effect for stimulus $F(2, 52) = 5.679$, $p = .006$, partial $\eta^2 = .179$. Pairwise comparisons based on the estimated marginal means indicated that the hit RT was slower for Greebles ($M = 1.630$, $SE = .084$) than for Black faces ($M = 1.376$, $SE = .054$), $p = .005$. The hit RT for White faces ($M = 1.63$, $SE = .084$) did not differ significantly from either Greebles or Black faces, $ps < .193$. Orientation did not show a main effect, $F(1, 53) = 1.507$, $p = .225$, partial $\eta^2 = .028$, and there was no significant interaction, $F(2, 52) = 1.088$, $p = .344$, partial $\eta^2 = .040$.

Planned comparisons were conducted in order to compare the stimuli by orientation and to explore the effect of orientation on each set of stimuli. In the upright condition, there were no significant differences in hit RT between Black ($M = 1.28$, $SD = .400$) and White faces ($M = 1.38$, $SE = .486$), $t(60) = 1.156$, $p = .252$, and the difference in hit RT for White faces versus Greebles ($M = 1.67$, $SD = .836$) did not reach significance, $t(59) = -2.432$, $p = .018$. Participants tended to take more time to make a correct decision regarding a previously seen Greeble versus a Black face, $t(58) = -3.34$, $p = .001$.

In the inverted condition, there were no significant differences in hit RT among the three sets of stimuli (White: $M = 1.54$, $SD = .653$, Black: $M = 1.47$, $SD = .653$, Greeble ($M = 1.59$, $SD = .882$), $ps > .505$.

Although there appeared to be a slight inversion effects for both Black ($t(59) = -2.479$, $p = .016$) and White ($t(61) = -2.117$, $p = .038$) faces, the differences in hit rate RT between upright versus inverted orientation were not statistically significant for either category of stimulus, and there was no inversion effect for Greebles, $t(59) = .282$, $p = .779$.

Overall, White and Black faces were similar in terms of hit RT. The longer reaction time for correct positive identifications of Greebles indicates that extra processing was necessary to make an accurate decision, and emphasizes the difficulty of the task.

A false alarm occurs when a participant signals that the face was presented during the encoding phase when it was not. Figure 4 depicts the false alarm rates for each condition and the corresponding false alarm RTs. For false alarm rate, a main effect for orientation was found: the false alarm rate was significantly higher when the stimuli were inverted at encoding ($M = .367$, $SE = .019$) versus upright ($M = .285$, $SE = .017$), $F(1,63) = 13.962$, $p =.0004$, partial $\eta^2 = .181$. There was also a significant main effect for stimulus $F(2,62) = 14.832$, $p = .000005$, partial $\eta^2 = .324$. Pairwise comparisons on the estimated marginal means (Bonferonni adjusted) revealed that White faces ($M = .247$, $SE = .021$) had a marginally lower false alarm rate than Black ($M = .367$, $SE = .019$) faces, $p = .069$, and that the false alarm rate for Greebles ($M = .420$, $SE = .022$) was significantly higher than for White faces, $p = .0000004$, or Black faces, $p =.0004$. There was no interaction effect between orientation and stimuli $F(2,62) = 1.106$, $p = .337$, partial $\eta^2 = .034$.

Due to the interest in the inversion effect, we proceeded with the planned comparisons. In the upright condition, the false alarm rate for White faces ($M = .191$, $SD = .180$) was significantly lower than for Black faces ($M = .2875$, $SD = .228$), $t(63) = -2.906$, $p = .005$ and Greebles ($M = .378$, $SD = .233$), $t(63) = -5.036$, $p =.000004$.

Although the false alarm rate for Black faces trended towards being lower than the false alarm rate for Greebles, the difference did not reach significance, $t(63) = -2.305$, $p = .024$.

In the inverted condition, the false alarm rate was significantly higher for Greebles ($M = .463$, $SD = .216$), than for White faces ($M = .303$, $SD = .242$), $t(63) = -3.832$, $p < .001$, and Black faces ($M = .334$, $SD = .255$), $t(63) = -3.358$, $p = .001$, but there were no significant differences in false alarm rate between White faces and Black faces , $t(63) = -.837$, $p = .406$.

In terms of an inversion effect, only White faces suffered a significant increase in false alarms when inverted as compared to upright presentation, $t(63) = -3.320$, $p = .002$. Although Greebles approached an inversion effect, it was not statistically significant, $t(63) = -2.353$, $p = .022$, and there was not a significant difference between the upright and inverted conditions for Black faces, $t(63) = -1.269$, $p = .209$.

The finding that White faces differed from both Black faces and Greebles in the upright condition, and that the inversion effect again occurred only for White faces supports the idea that holistic encoding drives the same-race advantage. However, although White faces appear to be processed in a way that is distinct from Black faces and Greebles, Black faces do not appear to be processed in exactly the same way, or with similar accuracy, as Greebles. In the inverted condition, when the possibility of holistic encoding is disrupted, the false alarm rate for Greebles is significantly higher than both Black and White faces. If participants rely on features to make identifications in the inverted condition, it is possible that inverted Greeble features are not as salient or distinguishable as facial features, and thus harder to encode and/or retrieve.

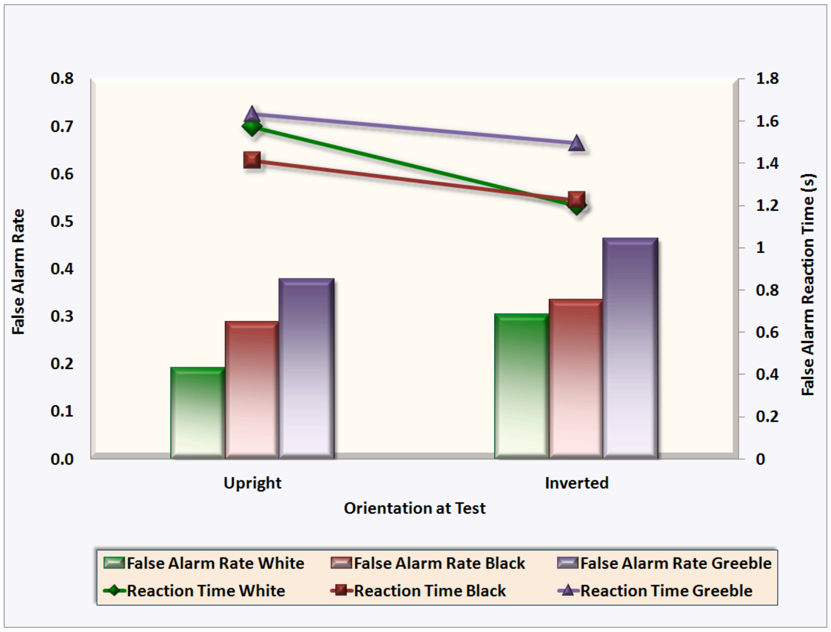

Figure 4. Mean false alarm rate and false alarm reaction time (s) as a function of target stimuli and orientation at test in Experiment 1.

For false alarm RT, orientation led to a marginally significant main effect $F(1, 20) = 4.034$, $p = .058$, partial $\eta^2 = .168$, but stimulus did not, $F(2, 19) = 1.668$, $p = .215$, partial $\eta^2 = .149$, and there was not an interaction effect $F(2, 19) = .344$, $p = .713$, partial $\eta^2 = .035$. The reaction time of false alarms was significantly longer in the inverted ($M = 1.307$, $SE = .097$) versus upright ($M = 1.541$, $SE = .127$) orientation. The planned comparisons yielded no significant findings.

Longer response latency gives us some insight into the participants' effort at a task; in the inverted condition, the need to process faces featurally rather than holistically takes more time and is less accurate. There were no significant differences between the different types of stimuli in either orientation; if White faces are processed holistically, whereas Black faces and Greebles are not, a main effect of stimulus (i.e., shorter RT for White faces) might be anticipated. However, the this measure assesses the response latency for an incorrect response, and reaction time for accurate facial recognition has consistently been found to be faster than for incorrect responses (Sporer, 1993) . Dunning & Perretta (2002) argue that the potential accuracy of eyewitness identifications can be judged, in part, on the speed with which the decision is made. Accurate selections tend to be automatic, whereas incorrect identifications are characterized by a slow reaction time and effortful attempts to remember. Thus, it is possible that the participants, without feeling a sense of recognition at seeing a face, attempted to utilize more cognitive resources in an attempt to supply a correct response, but failed due to lack of memory strength. False alarm responses necessarily occur when the previously viewed face is not present at test; this is a harder task than correctly realizing that a face has been seen. Essentially, a "target absent" condition, whether for a Black face, a White face, or a Greeble, may elicit an examination of features, since there is no true holistic match to memory or chance of recognition.

Hit and false alarm rate are subject to influences from response criterion placement and do not allow for a full understanding of a person's memory strength or skill. For example, if a participant responds "old" (previously seen) to every image presented at test, the resulting hit rate is 100%. This seems impressive, unless one notes that the false alarm rate is also 100%, since each new test face will also have been identified as having been seen during encoding. For a clearer understanding, discrimination ability must be decoupled from criterion. As discussed previously, the EZ-Diffusion model measures these two components of the decision process with v (drift rate) and a (criterion), along with the non-decision component, T_{er}, which are depicted in Figure 5.

In looking at the parameters of the EZ diffusion model, *v* showed a significant interaction effect, $F(2,62) = 4.741$, $p = .012$, partial $\eta^2 = .133$, as well as a main effect of orientation, $F(1, 63) = 31.030$, $p = .000001$ partial $\eta^2 = .330$, and a main effect for stimulus, $F(2, 126) = 23.199$, $p < .001$, partial $\eta^2 = .395$. Due to the significant interaction effect, the main effects cannot be interpreted. The planned comparisons were conducted to examine the different conditions.

In the upright orientation, there was a marginally significant difference between White ($M = .085$, $SD = .055$) and Black faces ($M = .060$, $SD = .052$), $t(63) = 2.653$, $p = .010$, a significant difference between White faces and Greebles ($M = .028$, $SD = .054$), $t(63) = 6.110$, $p < .001$, and a significant difference between Black faces and Greebles, $t(63) = 3.340$, $p < .001$.

In the inverted orientation, the drift rate for Black faces ($M = .043$, $SD = .052$) was significantly higher than for Greebles ($M = .006$, $SD = .073$), $t(63) = 3.335$, $p < .001$, but the difference between White faces ($M = .032$, $SD = .047$) and Greebles did not reach significance, $t(63)=2.458$, $p = .017$.

For White faces, the drift rate was significantly higher in the upright condition than in the inverted condition, $t(63) = 5.669$, $p = .001$, but orientation did not result in significant difference in the drift rate for Black faces or for Greebles, *ps* > .062. Panel 1 of Figure 3 shows that while inversion had a dramatic effect on the drift rate for White faces, neither Black faces nor Greebles were likewise affected. Overall, the task of accurately judging whether or not a Greeble had been seen previously was difficult for participants, and even if it is the case that lack of expertise for cross-race faces is responsible for the cross-race effect, participants would have more lifetime experience in recognizing Black faces than for recognizing Greebles.

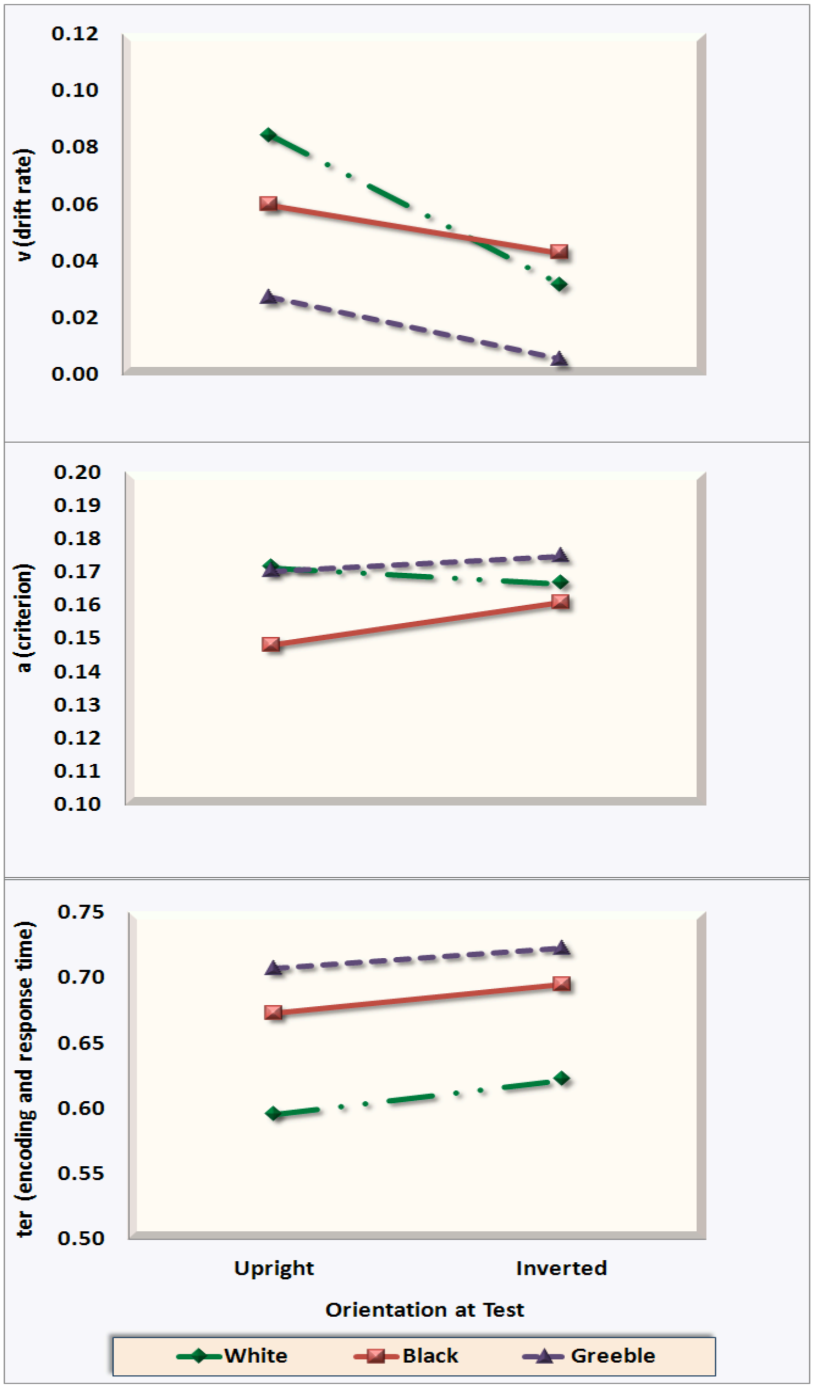

Figure 5. EZ diffusion model parameters (v, a, and ter) as a function of target stimuli and orientation at test in Experiment 1.

The parameter that assesses criterion placement is *a*. In *a*, the EZ diffusion model calculates how much information the participant requires before making a decision. If a participant has a liberal criterion, they will tend to reach their decision boundary quickly, potentially sacrificing accuracy for speed. If a participant is motivated to answer correctly, the decision boundaries are farther apart; the participant takes time to gather and assess information before responding.

Boundary separation, or *a*, showed no main effect for orientation, $F(1,63) = .377$, $p = .542$, partial $\eta^2 = .006$, or interaction effect, $F(2,62) = 1.374$, $p = .261$, partial $\eta^2 = .042$, but there was a significant main effect for stimulus, $F(2,62) = 4.846$, $p = .011$, partial $\eta^2 = .135$. Based on the marginal means, pairwise comparisons indicated that the wider boundary for White faces ($M = .169$, $SE = .006$) versus that for Black faces ($M = .155$, $SE = .004$) was not statistically significant, $p = .084$, and that the boundary for Greebles ($M = .173$, $SE = .007$) was significantly larger than for Black faces, $p = .012$.

Planned comparisons for the stimuli in the upright condition indicated that the participants had a more conservative criterion for White faces ($M = .171$, $SD = .057$) than for Black faces ($M = .148$, $SD = .044$), $t(63) = 2.720$, $p = .008$. However, the response criterion for White faces and Greebles ($M = .171$, $SD = .071$ appeared to be very similar, $t(63) = .101$, $p = .920$, and although participants tended to use a more liberal criterion for Black faces than for Greebles, this difference was not statistically significant, $t(63) = -2.483$, $p = .016$.

In the inverted condition, there were no significant differences between the response criteria for White faces ($M = .167$, $SD = .066$), Black faces ($M = .161$, $SD =$

.046), or Greebles (*M* = .175, *SD* = .071), *ps* > .100. In addition, none of the stimuli exhibited a significant inversion effect, *ps* > .084.

Participants appear to have been motivated to respond correctly to White faces, and were willing to sacrifice a quick response in order to avoid making an error. This was less true for Black faces, and this replicates the usual findings in research of cross-race facial recognition; there appears to be less motivation to work to answer correctly for other- versus own-race faces. It is possible that this is an aspect of in-group/out-group behavior; the criterion is more liberal for those individuals of a different category or group than one's own because inaccuracy is less important. The conservative response criterion for Greebles was not expected and cannot be well-explained by applying ideas about in-group/out-group behavior or cognitive disregard. It seems unlikely that participants shared any sense of affiliation with Greebles or intuited that Greebles held a position of power that might be of benefit to the participant in the future. Perhaps due to the novelty of the stimuli, as well as the similarity of Greebles to each other (at least, to the Greeble non-expert), participants realized that, in order to have a chance of responding correctly, care and time were necessary before making a decision.

T_{er}, the time taken to encode the stimuli at test and to respond, did not demonstrate a main effect of orientation, $F(1,63) = .417$, $p = .521$, partial $\eta^2 = .007$, or an interaction effect, $F(2,62) = .011$, $p = .898$, partial $\eta^2 = .000$. There was a main effect of stimulus, $F(1,62) = 4.013$, $p = .023$, partial $\eta^2 = .115$. Pairwise comparisons on the estimated marginal means (Bonferroni adjusted) indicated that T_{er} was significantly shorter for White faces (*M* = .609, *SE* = .030) than for Black faces (*M* = .684, *SE* = .029), *p* = .048 or Greebles (*M* = .715, *SE* = .036), *p* = .036, but there were no significant differences between Black faces and Greebles, $p = 1.00$. The planned comparisons revealed no significant differences, $ps > .074$.

T_{er} was significantly shorter for White faces than for Black faces or Greebles, but Black faces and Greebles were not significantly different in terms of the non-decision time as measured by T_{er}. This similarity may reflect a lack of expertise in processing both types of stimuli, although given the disparity for most individuals in lifetime exposure to Greebles versus Black faces, a measure that reflected expertise or experience might be expected to show some difference between the two sets of stimuli. Contrary to our expectations, there was no inversion effect for White faces; if holistic processing facilitates encoding, the time taken to encode the photos of White faces should have increased in the inverted condition, when primarily configural, rather than holistic, processing is used. It could also be hypothesized that T_{er} reflects categorization of out-group members. In Sporer's model, categorization occurs first in the encoding process and may interfere with the thoroughness and ease of the non-decision components of response time. In this scenario, it is possible that both Black faces and Greebles were, upon presentation, categorized rather than individuated, leading to a similar "hiccup" in response. The lack of an inversion effect supports T_{er} as capturing a "categorization effect" rather than expertise. However, there is no way to establish that the similarity in T_{er} for Greebles and Black faces is due to the same process. For example, categorization may increase T_{er} for Black faces, while lack of experience is responsible for the increase in Greeble T_{er}.

Overall, the cross race effect was found: For upright orientation, participants had a lower false alarm rate, higher accuracy (as measured by v), a higher response criterion (*a*), and a faster latency for components of encoding, retrieval, and response (T_{er}) for White faces than for Black faces. Recognizing Greebles, even in the upright condition, was fairly difficult for most participants. Performance, as measured in false alarm rate and *v*, reflected poor accuracy for Greebles. Participants did try to answer correctly, as evidenced by the conservative response criterion as measured by *a*.

Inversion has a disproportionate effect on White faces over both Black faces and Greebles in terms of hit rate and *v* (drift rate). A lack of inversion effect is hypothesized to be related to relatively greater featural, instead of configural, processing; inversion disrupts this process for White faces, but, because upright Black faces and Greebles are not processed in this manner, inversion does not lead to a marked drop in performance for either one. In terms of *v*, as well as false alarm rate, Black faces were more similar to Greebles than to White faces, though the graphs indicate that Black faces may occupy a middle ground between the other two stimuli.

If expertise in same- versus cross-race faces is responsible for the cross-race effect, this would be reflected in a similarity in the response variables between Black faces and Greebles as compared to White faces. The findings of Experiment 1 provide some support for the idea that expertise is a factor in the disparity between own-race and cross-race recognition. However, it is possible that, while lack of expertise in recognizing Greebles affected the response variables in terms of the measures of accuracy and RT, Black faces might be affected in similar ways by either lack of expertise or increased categorization, or both. Experiment 1 could not be used to examine the contributions of categorization or cognitive disregard to the cross-race effect, nor did it reveal information about how different levels of expertise might affect cross-race facial recognition.

A second experiment, incorporating both race and categorization, was developed. In addition, we sought to measure participants' interracial contact, as well as implicit attitudes about Black people, to examine the influence of these factors on cross-race facial recognition ability.

Experiment 2

In order to examine the role of in-group/out-group categorization in facial recognition, we utilized a paradigm similar to that of Bernstein, Young, and Hugenberg (2007), in which facial recognition for own- versus rival university students was tested. By identifying the images as members of the participants' university or students at another university, Bernstein et al. (2007) established that social categorization alone could lead to differences in facial recognition, but used images of white males only. In Experiment 2, the inclusion of both cross-race (Black and White faces) and cross-university (the faces identified as students of the University of Oklahoma or of a rival university) allowed for a comparison of the cross-race effect to the cross-category effect.

In order to further test the expertise hypothesis, level of cross-race interaction was assessed via a self-report. If the own race advantage develops as a result of contact with same-race individuals, the difference between own and cross-race identification should diminish with increased interracial contact. Strong underlying negative stereotypes about Black people may correspond to increased levels of depersonalization of African Americans. According to Sporer's (2001b) theory, out-group members are less well recognized because they are first categorized, while in-group members are individuated. Attitudes about individuals of a different race and cross-race interaction may well be intertwined: negative beliefs lead to limited and superficial contact; increased interaction diminishes the tendency to stereotype (Shepherd, 1971).

Previous research in this area has relied primarily on self-reports, which are vulnerable to measurement reactivity, particularly when the questions involve potentially sensitive topics such as prejudicial beliefs. It is preferable, though often not possible, to use a measure that is not entirely within the participants' control (Greenwald, McGhee, & Schwartz, 1998). Instead of asking questions about beliefs and stereotypes, the Implicit Association Test (IAT) utilizes response time. Developed by Greenwald, McGhee, and Schwartz (1998), the IAT has been used to examine the positive and negative associations of many different target concepts. Though somewhat controversial (Blanton, Jaccard, Gonzales, & Christie, 2006; van Ravenzwaaij, van der Maas, & Wagenmakers, 2011), the IAT was of interest due to its purported ability to assess attitudes about race that might not be revealed in a self-report, whether due to a participant's desire to present him/herself in a more positive light (i.e., without prejudiced attitudes) or simply because s/he is unaware of some of the negative associations s/he may have about people of a different race.

There are two tasks in the IAT, a concept task and an attribute task. For the concept task, the participant views exemplars of two distinct target concepts ("insects" and "flowers") and categorizes them into the appropriate group. For example, a participant may be told to press the letter 'e' if an image of a flower appears and press 'i' if the image depicts an insect. During the second phase, the attribute task, the participant classifies attributes into categories ("positive" or "negative"). As an example, if instructed to press the letter 'e' for all positive attributes and the letter 'i' for negative attributes, upon reading a word with a pleasant meaning or connotation (e.g., “happy”, “peace”) the participant should press ‘e’, and press ‘i’ if the word means something unpleasant (“rotten, “war”). During the test phase of the IAT, the concept and attribute tasks are performed in alternation, using the same response keys. Thus, the same response key (for example, the letter ‘e’) may be needed to classify the picture of a daisy into the category “flower” and to identify the word “beautiful” as “positive,” whereas ‘i’ is mapped onto both the category “insect” and the attribute “negative.” Alternatively, the category “flower” could be paired with the attribute “negative,” and “insect” with positive. After the first test block, the response key mapping and category/attribute pair are re-paired (if, in test block 1, “flower” was paired with “negative,” in test block 2, flower will be paired with “positive”), and a second test block begins. For most people, the task will be easier if the same response key is mapped onto “flower” and “positive” than “insect” and “positive” because there is a stronger association between the former category and attribute.

The IAT uses response latencies and error rates to examine the differential association of two concepts with the attributes; compatible mapping (e.g., flowers and positive) will result in faster and more accurate responses than incompatible mapping. Simply put, the mean reaction time of combined incompatible trials is compared to the combined mean reaction time of the compatible trials; this is the “IAT effect.” Theoretically, the more incompatible a concept is with an attribute, the slower and more error prone the responses. The individual that struggles more when images of White people are paired with negative words and images of Black people are paired with positive words is hypothesized to have stronger negative stereotypes about Black people than a person with a smaller difference in reaction time between the different pairings of race to attribute.

Method

Participants. Participants were students (N=80) at the University of Oklahoma who were enrolled in an undergraduate psychology class. The participants received research participation credit for the class in which they were enrolled. Sixty-three of the participants identified themselves as White, non-Hispanic, two identified as Black, non-Hispanic, five participants were White, Hispanic, four identified as Asian, and six as Other. As in Experiment 1, only the data from the White participants were included in the analysis.

Materials.

Photographic Stimuli. The computer images of White and Black males used in Experiment 1 were used in Experiment 2. However, instead of a white background, the faces were displayed against either a maroon background with the caption "University of Oklahoma" below the faces, or against a dark blue background with "Boise State University" as the caption.[5] As for Experiment 1, the faces used in the encoding session bore a happy expression and wore street clothes; in the test component of the experiment, the males had neutral expressions and wore gray sweatshirts. The background color and caption were consistent across encoding and test.

Social Experiences Questionnaire (SEQ). The SEQ, developed by John Brigham (1993), is a 67-item questionnaire and was used to assess the frequency and the quality of cross-race interactions of the participants. Different types (personal, business, social,

[5] In the football season preceding the experiment, the University of Oklahoma lost to Boise State University at the Fiesta Bowl.

public) and outcomes (unpleasant/pleasant; costly/beneficial) of cross-race contact can be measured with the SEQ, including past and present levels of interaction. Participants that self-identified as Black were asked about their quantity and quality of contact with White people. All other participants (White, Asian, Hispanic, Other) answered questions about their interactions with Black people. For any participant that identified as bi-racial (Black and White), the version of the SEQ to be administered was decided by a coin-toss. Only the results from White participants regarding their level and type of interaction with Black people are included.

Implicit Association Test (IAT). We administered the Racial (Black/White) IAT, developed by Greenwald, McGhee, and Schwartz (1998). This version of the IAT consists of ten black and white photos each of White and Black males and females. During a trial, participants viewed either an image of a White person or a Black person, or one of 16 words, eight of which were positive or pleasant ("love," "joy"), and eight of which were negative or unpleasant ("hate," "horrible"). Critical trials included compatible/congruent blocks[6] in which White images were paired with positive words and Black images were paired with negative words. Participants were instructed to press a letter if the stimulus was an image of a White person *or* a positive word, and a different letter if the presented image was a Black person *or* a negative word. In the incompatible/incongruent blocks, White images were paired with negative words and Black images were paired with positive words. The order of these critical blocks and response key mappings were counterbalanced between subjects. For each task, stimuli appeared one at a time in the center of the computer screen, and participants pressed the

[6] The terms compatible/congruent and incompatible/incongruent reflect the theoretical basis of the IAT and are derived from the use of the terms by the developers of the IAT. The terms were never used during at any point during the administration of the IAT to the participants.

‘e’ key or ‘i’ key on the computer keyboard to classify the stimuli as quickly and accurately as possible into the appropriate categories. All pictures were placed against a black background and sized to 300 x 450 pixels.

Equipment and software. The photographs and the questionnaire were presented via computer, utilizing the experimental software E-Prime 2.0.1.68 (Psychology Software Tools), which allows for randomization of the order of conditions presented (upright and upside-down; Black and White), as well as the photographs presented with each condition.

Procedure. Participants first read the informed consent form. After reading some brief instructions about the facial images they were about to see and the subsequent recognition test, the first encoding sequence began. All participants completed 2 blocks consisting of two randomly ordered trials each. The trials varied by the orientation of the encoding stimuli (inverted or upright), while the test stimuli was always upright.

During encoding, participants saw 12 facial images, six each of White and Black males. For each racial category, half were depicted as being affiliated with OU, while the other three were affiliated with BSU. University affiliation was randomly assigned via computer for each participant. The images displayed on the computer monitor for 2 s each, with a 1 s fixation screen before each image. Immediately following the encoding session, participants were instructed regarding the upcoming recognition test, and, as in Experiment 1, were asked to press ‘e’ if they recognized the face as one that had been presented during encoding, or ‘i’ if it had not been displayed previously. Participants were informed that a change in clothing or facial expression was forthcoming, and that their recognition decision should be based on the individual, not on the characteristics of the images themselves. Each test session consisted of 24 faces, half of which had been present previously and half of which were new. The new faces were divided evenly into the categories of Black and White, OU and BSU, so that, in total, participants saw 12 Black faces, 6 OU affiliated, 6 BSU affiliated, and 12 White faces, 6 each affiliated with OU or BSU. If a face had been presented during the encoding session, university affiliation was the same at encoding and at test.

Results and Discussion

As in Experiment 1, repeated-measures ANOVA and paired comparisons (with Bonferroni adjustment) were used to test for significant differences. The design was a 2 (stimulus: White males vs. Black males) x 2 (orientation at encoding: upright vs. inverted) x 2 (university affiliation: OU vs. BSU) within-subjects factorial. In total, 7 dependent variables were examined: hit rate, false alarm rate, hit rate RT, false alarm RT, as well as the EZ diffusion parameters of v (drift rate), a (boundary separation) and T_{er} (time for encoding and response).

Orientation of stimuli led to a main effect for hit rate, $F(1, 62) = 20.990$, $p < .001$, partial $\eta^2 = .253$, demonstrating an overall inversion effect: participants were less accurate at correctly identifying previously presented faces when the test stimuli were inverted ($M = .589$, $SD = .183$) than when the stimuli were presented upright at test ($M = .691$, $SD = .151$). Stimuli at encoding were always presented upright. No significant effects were found for race or university, and there were no interaction effects. However, in order to gain a better understanding of potential trends in the different categories of stimuli, we proceeded with the planned comparisons. Eight paired t-tests were conducted: in the upright condition, performance for White OU-affiliated versus White BSU-affiliated images, Black OU versus Black BSU affiliated faces, White OU versus Black OU images, and White BSU versus Black BSU faces. In addition, the inversion effect was examined through paired t-tests for the upright versus inverted condition of each race/university category. A significant inversion effect was found for White OU-affiliated images; the hit rate for upright White OU faces ($M = .737$, $SD = .241$) was significantly larger than the hit rate for inverted White OU faces ($M = .587$, $SD = .319$), $t(62) = 3.297$, $p = .002$. The inversion effect for White BSU faces approached, but did not achieve, significance (Upright: $M = .686$, $SD = .232$; Inverted: $M = .578$, $SD = .274$), $t(62) = 2.739$, $p = .008$[7]. Inversion disrupted the holistic processing that occurred for White faces, particularly White OU-affiliated faces. As in Experiment 1, Black faces did not show a significant inversion effect, indicating that holistic processing is not used for upright Black faces. Figure 6 depicts the hit rate and hit reaction time.

For hit reaction time, orientation at encoding had a significant effect, $F(1, 52) = 11.320$, $p = .001$, partial $\eta^2 = .179$. Accurate positive identifications were made more quickly in the upright condition ($M = 1.266$, $SD = .444$) than when the stimuli were inverted ($M = 1.451$, $SD = .532$). No other significant effects were detected. In general, hits in the upright condition were achieved relatively quickly, whereas, in the inverted condition, a more effortful and time-consuming process was needed. This points to the use of a featural, rather than holistic, approach for inverted images. Paired comparisons were again conducted; the hit reaction time for Black OU-affiliated faces was shorter in the upright condition ($M = 1.255$, $SD = .539$) versus the inverted condition ($M = 1.523$, $SD = .566$), $t(61) = 3.539$, $p = .001$. For Black BSU-affiliated faces, the difference in

[7] Bonferroni adjustment based on eight planned comparisons requires a *p*-value less than $\alpha/8 = .00625$ to be significant.

reaction time between the upright ($M = 1.281$, $SD = .558$) versus inverted ($M = 1.513$, $SD = .592$) conditions indicated a possible trend that participants took less time in the upright condition, but this did not reach the necessary level of significance, $t(58) = 2.691$, $p = .009$.

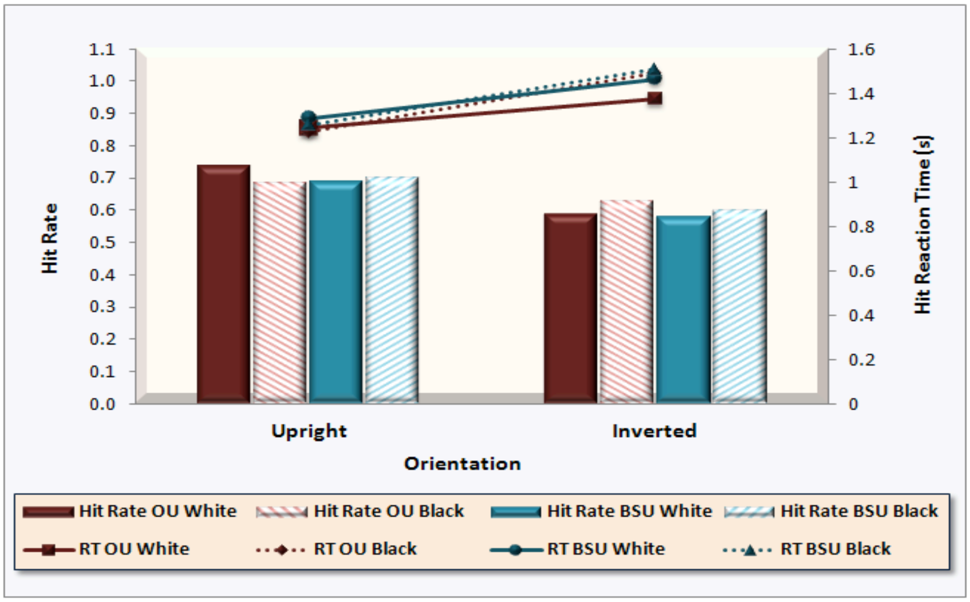

Figure 6. Mean hit rate and hit reaction time (s) as a function of target stimuli and orientation at test in Experiment 2.

For false alarm rate, there was a main effect of orientation, $F(1, 62) = 47.020$, $p < .001$, partial $\eta^2 = .431$. Orientation and race showed a significant interaction effect, $F(1, 62) = 4.512$, $p = .038$, partial $\eta^2 = .068$. Comparisons of White versus Black faces within each orientation revealed that Black faces ($M = .365$, $SD = .222$) had a higher false alarm rate than White faces ($M = .265$, $SD = .151$) in the upright orientation only. Contrary to expectations, there were no significant differences due to university affiliation. The planned comparisons showed a significant inversion effect for White OU-affiliated faces (Upright: $M = .248$, $SD = .205$, Inverted: $M = .470$, $SD = .305$), $t(62) = 5.116$, $p < .001$ and White BSU-affiliated faces, (Upright: $M = .283$, $SD = .211$, Inverted: $M = .460$, $SD = .270$), $t(62) = 4.209$, $p < .001$, but Black faces did not suffer an inversion effect in terms of false alarm rate. As in Experiment 1, when compared to the false alarm rate for White faces, participants were more likely to incorrectly indicate that a Black face had been seen previously. However, the recognition advantage for White faces disappeared in the inverted condition, supporting the hypothesis that, in part, the cross-race effect is due to holistic processing for White faces.

Examining false alarm reaction time (RT), there was a main effect for race $F(1, 25) = 8.307$, $p = .008$, partial $\eta^2 = .249$. The false alarm response latency for Black faces ($M = 1.701$, $SD = .510$) was longer than the false alarm response latency for White faces ($M = 1.454$, $SD = .480$). Neither orientation nor university affiliation showed a main effect, and there were no significant interactions. None of the paired comparisons yielded significant differences. A graphical display of false alarm rate and false alarm response time is presented in Figure 7.

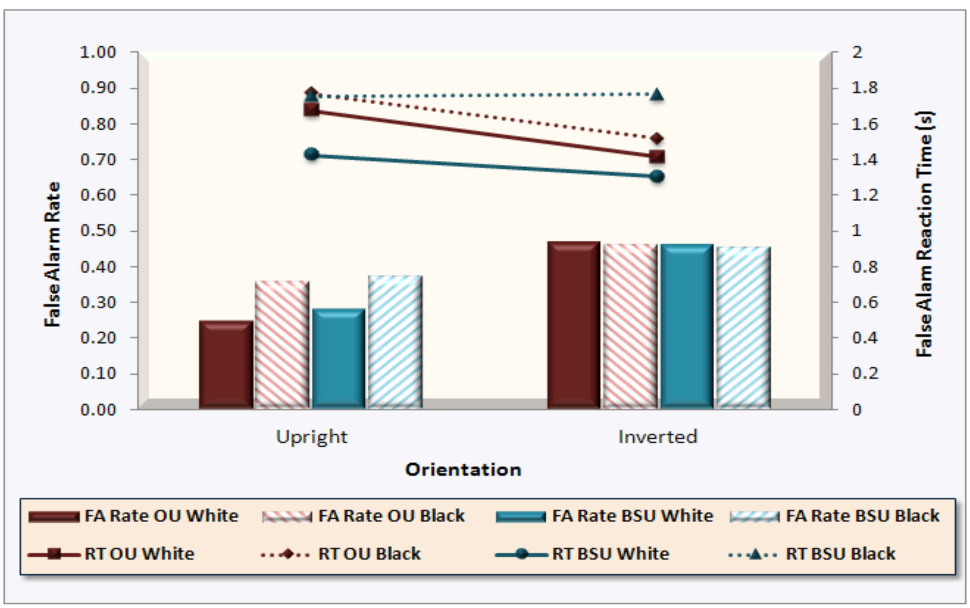

Figure 7. Mean hit rate and hit reaction time (s) as a function of target stimuli and orientation at test in Experiment 2.

There was a main effect of orientation for *v*, $F(1, 62) = 37.642$, $p < .001$, partial $\eta^2 = .378$. The interaction between orientation and race was significant, $F(1,62) = 13.336$, $p = .001$, partial $\eta^2 = .177$. Based on the estimated marginal means, the drift rate for White faces ($M = .067$, $SD = .040$) was higher than for Black faces ($M = .044$, $SD = .040$) in the upright, but not inverted, condition. University affiliation showed no main effect, nor was there an interaction effect for university and orientation. The planned follow-up comparisons indicated that, in the upright condition, participants performed significantly better when the image was of a White OU-affiliated face ($M = .076$, $SD = .052$) versus a Black OU-affiliated face ($M = .046$, $SD = .055$), $t(62) = 3.433$, $p = .001$. Significant inversion effects were found for White OU faces (Upright: $M = .076$, $SD = .052$, Inverted: $M = .020$, $SD = .070$), $t(62) = 5.70$, $p < .001$, and White BSU faces (Upright: $M = .059$, $SD = .053$, Inverted: $M = .023$, $SD = .044$), $t(62) = 3.835$, $p < .001$.

The EZ diffusion measure of criterion, *a*, showed a main effect of orientation, $F(1,62) = 4.196$, $p = .045$, partial $\eta^2 = .063$, and an interaction effect of orientation and race, $F(1,62) = 5.407$, $p = .023$, partial $\eta^2 = .080$. Participants had a higher criterion for White faces ($M = .175$, $SD = .063$) than for Black faces ($M = .158$, $SD = .048$) in the upright condition, but there was not a significant difference in the inverted condition. There were no other significant differences, contrary to our prediction that participants would have a lower criterion for BSU-affiliated faces than for OU-affiliated faces.

For the parameter T_{er}, there were no significant main or interaction effects, and none of the planned comparisons were significant. Unlike Experiment 1, race did not have a significant impact on T_{er}. Likewise, university affiliation did not appear to influence the time needed by participants to encode and respond to the stimuli. This was contrary to our hypothesis that, as a potential measure of categorization, T_{er} would be larger for Black faces and for BSU-affiliated images.

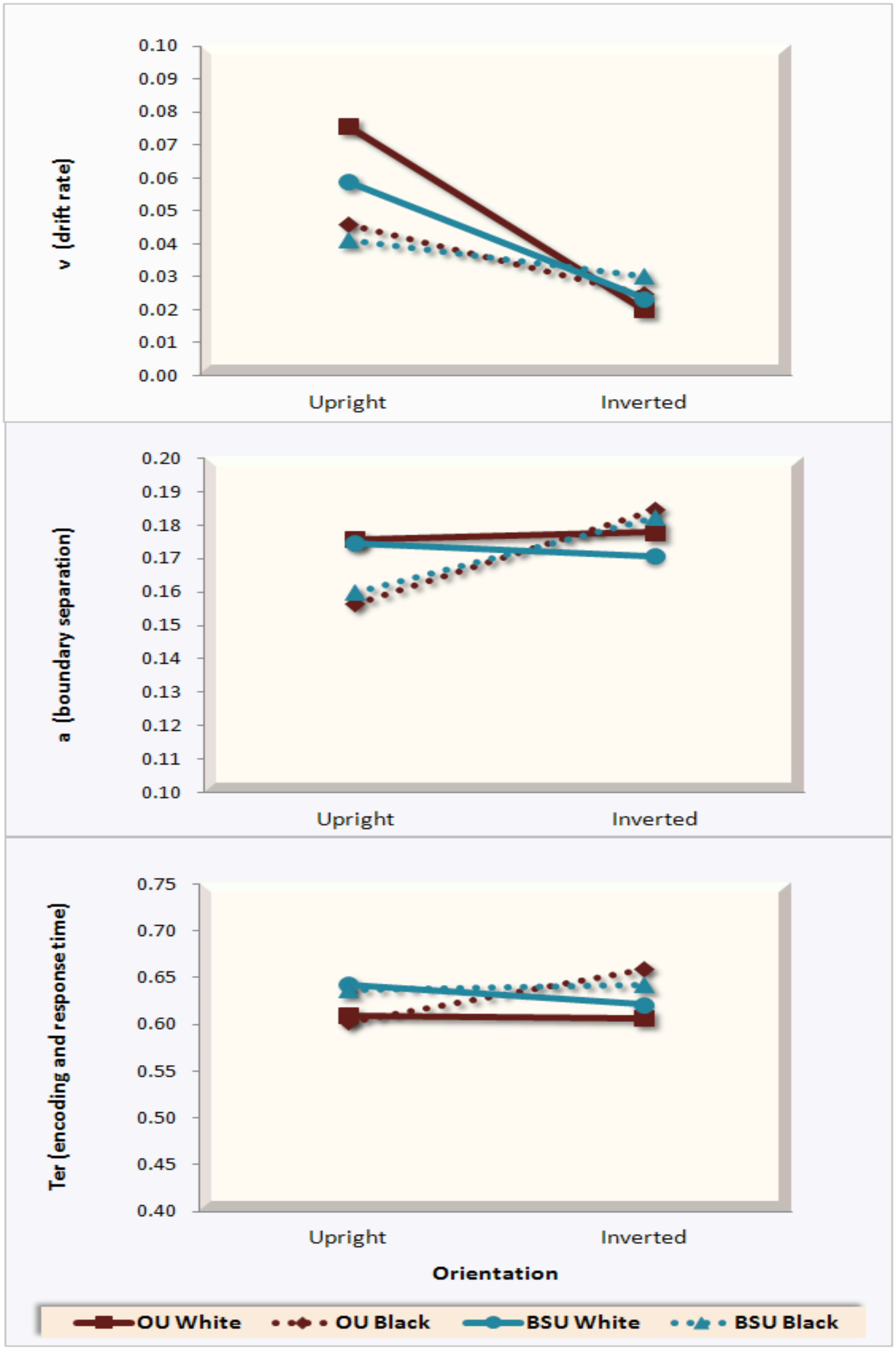

Figure 8. EZ diffusion model parameters (v, a, and ter) as a function of target stimuli and orientation at test in Experiment 2.

To determine if cross-race contact and prejudice had an effect on cross-race recognition performance we examined the relationship between the items on the Social Experience Questionnaire, the results of the Implicit Association Test, and the cross-race differences for v, a, and T_{er}. In order to calculate the cross-race effects, the mean responses for the Black stimuli were subtracted from the mean responses for the White stimuli for each participant. Due to the number of interaction effects between race and orientation, only the upright condition was used in the calculations. Thus, the three variables included in the analysis were $v_{\text{upright\_White-upright\_Black}}$, $a_{\text{upright\_White-upright\_Black}}$, and $T_{er\ \text{upright\_White- upright\_Black}}$.

The 56 SEQ items were subjected to an orthogonal factor analysis (varimax rotation) that yielded six meaningful factors. The first factor, labeled public/social, consisted of contact with Blacks in public or social non-intimate settings. It contained 8 items (factor loadings of .40 or above) regarding how pleasant and how beneficial past interactions were with Blacks in public settings and personal– nonintimate settings, as well as expectations for future interactions in these settings (eigenvalue = 5.98, 14.2% of total variance). The second factor, past contact, consisted of 8 items that inquired as to the level of contact with Blacks in elementary, middle, and high school, the percentage of Black people in the neighborhood in which one grew up, as well as level of past interaction overall, in public settings, and in personal, non-intimate settings (eigenvalue = 4.94, 11.8% of total variance accounted for). The third factor, which we labeled present contact, included 9 items that asked about the extent of present contact with Blacks on campus, in the dorms, at work, number of Black friends (eigenvalue = 3.58, 8.2% of variance). The fourth factor, labeled personal/intimate, contained five questions regarding the pleasantness and benefit of past interactions, as well as expected pleasantness and benefit of future interactions, in intimate settings (eigenvalue = 3.45, 38.2% of variance). The fifth factor, business, consisted of four items about the pleasantness and benefit of past interactions with Blacks in business settings (eigenvalue = 2.43, 5.8% of variance). The sixth factor, past friends, contained three items about the number of Black friends in elementary, middle and high schools (eigenvalue=1.94, 4.6% of variance).

The IAT scores were derived following the D scoring algorithm as recommended by Greenwald, Nosek, and Banaji (2003). Response latencies were calculated from the onset of the trial until the participant made a correct response. Any response latencies less than 400 ms or greater than 10,000 ms were removed. IAT scores can range from -2.00 to +2.00. According to Greenwald et al. (1998), a score of -2.00 (a result of high accuracy and quick response time when the pairings are Black/positive and White/negative, as well as low accuracy and long reaction time for Black/negative and White/positive pairings) indicates a very strong preference for Black people and suggests a negative attitude, conscious or unconscious, about White people. A score of +2.00 corresponds to a strong preference for white people and a negative attitude about Black people.

Table 1 displays the correlations between the dimensions of the SEQ and the IAT. There were three significant correlations: interracial contact and experience in public setting correlated positively with cross-race contact in business settings, a positive correlation was found between level of cross-race experience the personal domain and past interracial contact, and level of present cross-race contact was positively correlated with the items involving past interracial friendship. Contrary to our hypothesis, level of bias or prejudice, as measured by the IAT, did not correlate significantly with any of the interracial experience dimensions from the SEQ. This finding mirrors other studies that have failed to demonstrate a relationship between cross-race experience and prejudice. The effect has been shown to be small, so it is possible that the IAT is not sensitive enough to capture the differences in levels of prejudice, as the majority of White participants tend to score relatively high on the IAT.

Table 1

Correlations Between Measures of Cross-Race Experience and the Implicit Association Test for Race

	1	2	3	4	5	6	7
1. Personal		0.24	0.14	0.30*	0.10	-0.03	-0.09
2. Public			0.55**	-0.02	0.04	0.10	0.13
3. Business				-0.13	-0.07	-0.11	0.04
4. Past					0.22	-0.15	-0.17
5. Present						0.39**	-0.26
6. Past Friends							0.44
7. IAT							

Note. ns = 46-64. IAT = Implicit Association Test. Personal, public, business, past, present, and past friends represent six factors derived from the Social Experience Questionnaire (SEQ).
* $p < .05$. ** $p<.01$

Correlational analyses were performed on the measures of cross-race recognition (v, a, and T_{er}) and the results of the SEQ and IAT. The results are displayed in Table 2.

Table 2

Correlations Between Measures of Cross-Race Experience and Cross-Race Recognition Performance

	Personal	Public	Business	Past	Present	Past Friends	IAT
ter_{UW-UB}	-0.12	-0.29*	0.08	0.12	-0.05	-0.11	-0.14
a_{UW-UB}	0.88	0.07	-0.08	-0.02	0.22	0.11	0.06
V_{UW-UB}	0.08	-0.03	-0.14	0.20	0.44**	0.62**	0.08

Note. ns = 46-64. IAT = Implicit Association Test. Personal, public, business, past, present, and past friends represent six factors derived from the Social Experience Questionnaire (SEQ).
* $p < .05$. ** $p<.01$

A negative correlation was found between the cross-race experience in the public domain and T_{er}, the time taken to encode and respond to the stimuli at test. The participants that had a large cross-race difference in T_{er} tended to have a low score in terms interracial contact in public settings. This was an unexpected finding; although increased interracial contact was expected to potentially lower T_{er} due to increased ease of processing, we did not expect that this would occur for the public domain only. The measure of discrimination, or *v*, correlated positively with the SEQ dimensions of present interracial contact, and cross-race friendships in the past. This is contrary to our prediction that increased cross-race experience would correspond to a smaller cross-race effect in recognition ability. Participants that reported high levels of interracial contact in the present and in past friendships actually tended to have a higher cross-race effect, indicating that they were able to accurately identify White faces much better than Black faces. Contrary to our hypothesis, the measure of criterion, *a*, did not show any significant correlations with the SEQ factors. Level of cross-race contact did not correspond to a difference in caution in terms of deciding whether or not a White versus a Black face had been seen before.

As in Experiment 1, the cross-race effect was found. Recognition performance was better for White faces versus Black faces for false alarm rate, v, and a. In general, inversion led to disruption of the processing of White faces, and not Black faces, as indicated in the response variables of hit rate, false alarm rate, v and a. Unlike the study by Bernstein et al. (2007), in which a significant cross-university effect was found, university affiliation had very little impact on accuracy or response time in Experiment 2.

The findings support the contribution of expertise rather than categorization in the CRE; the cross-category manipulation failed to yield significant results, allowing for little comparison between the CRE and the out-group effect. This could be because race is a more salient characteristic and therefore supersedes other categories. Thus, when presented with Black and White faces for each university affiliation, participants categorized the Black faces and individuated the White faces.

Other researchers have found evidence for this idea. Shriver, Young, Hugenberg, Bernstein, and Lanter (2008) found that facial recognition was worse for White faces presented in impoverished versus wealthy contexts. In a series of experiments, they found support for a social cognitive rather than a perceptual expertise model of the CRE. However, although out-group categorization did result in a decline in recognition for White faces, Black faces were poorly recognized regardless of socioeconomic status. Likewise, in an experiment similar to the second experiment presented in this paper, images were affiliated with either the same or other-university; recognition dropped for White faces, but not for Black faces. Thus, while this research appears to make a strong case for the contribution of categorization on the CRE, it does not, as the researchers admit, preclude the possibility that expertise also plays a role in the CRE, although Shriver et al. (2008) suggest that, instead, the tendency to categorize Black people as out-group members was too strong to be ameliorated by in-group status in terms of wealth or university affiliation. Their stance is supported by research in social cognition that demonstrates the power of certain characteristics such as sex, race, and age to spontaneously elicit categorization and to persist as the dominant category across contexts (Brewer, 1988; Stangor, Lynch, Duan, & Glas, 1992). However, such research can still coexist logically with the argument that a lack of expertise contributes to the CRE; indeed, they are complementary. Out-group categorization on the basis of race appears to be, to some extent, spontaneous and involuntary. This tendency, then, may further impede the development of expertise, even with greater contact, and this lack of expertise may only compound the strength of the tendency to categorize (Judd & Park, 1988).

General Discussion

Taken together, the results of Experiments 1 and 2 support the importance of expertise in the CRE. Black faces were more similar to Greebles than to White faces on many of the measures assessed in Experiment 1. A significant cross-race effect, but not a cross-category effect, was found in Experiment 2. Inversion affected the processing of White faces, but did not significantly impact the identification accuracy of Black faces or Greebles, and the inversion effect was not impacted by university affiliation.

However, these results are not inconsistent with the possibility that categorization of out-group members interferes with facial recognition and contributes to the CRE, particularly in light of current research in this field that illuminates the importance of category saliency. For example, Hehman, Mania, and Gaertner (2009) induced categorization by either race or university affiliation by spatially organizing the presentation of eight faces. When race was made salient, a significant cross-race effect was found, and there was no cross-university effect. However, when university affiliation was the salient category, participants were significantly more accurate at identifying own- versus other-university faces, and there was no cross-race effect. While these intriguing results bolster arguments for the role of social categorization rather than perceptual expertise in the CRE, they still cannot discount the contribution of the latter. Scrutiny of the different manipulations reveals that the CRE was eliminated primarily through a reduction in own-race recognition in the university manipulation rather than due to an increase in ability to recognize cross-race faces. Thus, it appears there is a complex interplay between expertise and categorization.

Susa, Meissner, and de Heer (2010) used path analysis to examine the influences of cross-race contact, racial attitudes, perceptual discrimination, racial categorization, and recollection on cross-race facial recognition. They suggest a cascade of direct and indirect effects, beginning with the social psychological influences (i.e., level of interracial contact and racial attitudes), to aspects of perception and working memory (perceptual discrimination and categorization), and the contributions recollection and recognition, which are higher-level cognitive processes associated with long-term memory. Susa, et al. (2010) assessed participants on four different tasks: level of interracial experience and racial attitudes; speed and accuracy of racial categorization; ability to discriminate a previously seen face from a set of 2, 4, 6, or 8 distracter faces; and, pitting the participants' capacity to use recollection rather than relying on a sense of familiarity, correct identification of faces presented during the study phase when randomly presented in a sequential array of both novel faces and faces that had previously been presented in the test phase.

Their findings were consistent with previous research: participants categorized other-race faces more accurately and faster than own-race faces, perceptual discrimination and recognition memory were better for own race faces, and participants were more likely to commit a "repetition error" for cross-race faces, erroneously identifying faces presented previously in the test phase as faces that had been presented in the study phase.

Separate path models were estimated for own and cross-race faces. The results indicated that while facial recognition performance for own-race faces was due primarily to the higher-level cognitive processes of familiarity and recognition, racial categorization and perceptual discrimination accounted for cross-race recognition performance. Also, in the cross-race model, greater interracial contact was associated with a lower occurrence of the repetition error and longer categorization latency was linked with better discrimination and recognition abilities. Perceptual discrimination scores predicted recognition ability and repetition errors. A comparison of the two models revealed two significant differences: participants were more likely to commit the “repetition error” for cross-race faces versus own-race faces, and a more positive association between quick categorization and lower accuracy was found for cross-race faces.

These findings indicate that higher level processes account for differences in recognition performance for own-race faces more than for cross-race faces, ostensibly because the more conscious, effortful process of recollection relies on a strong and accurate memory, which is not as available for cross-race faces. Also, the results emphasize the role of categorization in the CRE. The easier it was for a participant to categorize, as indicated by a shorter latency, the lower the recognition accuracy, which suggests that the process of out-group categorization interferes with the encoding process, perhaps because individuating features are ignored. Future research in this field will need to be attentive to such multifaceted contributions to the CRE.

Level of interracial contact and implicit attitudes about Black people were not well established as influential factors in the CRE in Experiment 2. Previous findings about increased contact and racial attitudes are often inconsistent and somewhat contradictory, and it is possible that the measurements used were not sensitive enough to capture real differences. It may be injudicious to expect that attending school with a few African American children or having positive experiences with Black people in public settings would begin to approach the level of opportunity for developing the expertise for White faces, a complex proficiency garnered over a lifetime with innumerable faces and types of contact. In addition, relative lack of variety within the sample regarding levels of contact and implicit attitudes weakened our ability to detect the influence of these factors on cross race facial recognition. A more diverse population, with greater variance in levels of interaction and beliefs, would greatly enhance this sort of study, as well as allow the inclusion of other races.

The inversion effect was found for White faces, but not for own-university faces, lending credence to the idea that expertise leads to holistic processing, which facilitates recognition. However, due to the issue of race saliency, the results cannot disprove or support the possibility that out-group categorization would disrupt holistic processing. An experiment that examined the effect of inversion on in-group versus out-group members might yield findings that could address this question.

Another avenue for future research is the potential change in response patterns as one gains expertise. Training in Greebles appears to lead to an inversion effect (Rossion, Gauthier, Goffaux, Tarr, & Crommelick, 2002); researchers speculate that this is evidence of the holistic processing that comes into play with increased skill in recognizing and individuating members of a group. In addition, neurological studies demonstrate that, with sufficient training, Greeble recognition takes place in the FFA, the area known to be active during facial recognition (Rossion, Gauthier, Goffaux, Tarr, & Crommelinck, 2002). By testing participants on facial (White and Black) and Greeble recognition before, during, and after engaging in training in Greeble recognition, an assessment of the response variables and EZ diffusion parameters could reveal information about the development of expertise.

The study of cross-race facial recognition is increasingly becoming a complicated and fascinating endeavor, one that holds important ramifications for the legal field, but also informs us about the impact of group membership, motivation, and context, as well as the development of expertise. Determining how the factors coact requires sophisticated experiments and analysis; through careful inquiry, researchers may be able to better understand the CRE, and, perhaps, work to reduce it.

References

Ackerman, J. M., Shapiro, J. R., Neuberg, S. L., Kenrick, D. T., Becker, D. V., Griskevicius, V., Maner, J. K., & Schaller, M. (2006). They all look the same to me (unless they're angry): from out-group homogeneity to out-group heterogeneity. *Psychological Science, 17*(10), 836-840.

Allport, G. W., & Kramer, B. M. (1946). Some roots of prejudice. *Journal of Psychology, 22*, 9-39.

Anastasi, J. S., & Rhodes, M. G. (2006). Evidence for an own-age bias in face recognition. *North American Journal of Psychology, 8*(2), 237-252.

Anthony, T., Copper, C., & Mullen, B. (1992). Cross-racial facial identification: A social cognitive integration. *Personality & Social Psychology Bulletin, 18*, 296-301.

Anzures, G., Quinn, P. C., Pascalis, O., Slater, A. M., & Kang, L. (2009). Categorization, categorical perception, and asymmetry in infants' representation of face race. *Developmental Science, 13*(4), 553-564.

Bernstein, M. J., Young, S. G., & Hugenberg, K. (2007). The cross-category effect: Mere social categorization is sufficient to elicit an own-group bias in face recognition. *Psychological Science, 18*(8), 706-712.

Blanton, H., Jaccard, J., Gonzales, P., & Christie, C. (2006). Decoding the implicit association test: Implications for criterion prediction. *Journal of Experimental Social Psychology, 42*(2), 192-212.

Borchard, E. M., & Lutz, E. R. (1932). *Convicting the innocent : Sixty-five errors of criminal justice* London: Oxford University Press, 1932.

Bothwell, R. K. K., Brigham, J. C., & Malpass, R. S. (1989). Cross-racial identification. *Personality & Social Psychology Bulletin, 15*(1), 19-25.

Braun, B. (2007). Rape suspect sentenced in theft, *Tulsa World*. Retrieved from http://www.tulsaworld.com/news/article.aspx?articleID=070828_1_A7_spanc63544

Brewer, M. B. (1988). A dual process model of impression formation. In T. K. Srull & R. S. J. Wyer (Eds.), *Advances in Social Cognition* (Vol. 1, pp. 1-36). Hillsdale, NJ: Lawrence Erlbaum.

Brigham, J., C. (1993). College students' racial attitudes. *Journal of Applied Social Psychology, 23*, 1933-1967.

Brigham, J., C. (2008). The role of race and racial prejudice in recognizing other people. In C. Willis-Esqueda (Ed.), *Motivational Aspects of Prejudice and Racism* (pp. 68-110). New York: Springer.

Brigham, J., C., Maass, A., Snyder, L. D., & Spaulding, K. (1982). Accuracy of eyewitness identifications in a field setting. *Journal of Personality and Social Psychology, 42*, 673-681.

Brigham, J., C., Wasserman, A. W., & Meissner, C. A. (1999). Disputed eyewitness identification evidence: Important legal and scientific issues. *Court Review, 36*, 40902.

Brigham, J. C. (1989). Disputed eyewitness identifications: can experts help? *The Champion, 8*(5), 10-18.

Brigham, J. C., & Barkowitz, P. (1978). Do "They all look alike?" The Effect of Race, Sex, Experience, and Attitudes on the Ability to Recognize Faces'. *Journal of Applied Social Psychology, 8*(306-318).

Brigham, J. C., & Bothwell, R. K. (1983). The ability of prospective jurors to estimate accuracy of eyewitness identifications. *Law and Human Behavior, 7*(1), 19-30.

Brigham, J. C., & Malpass, R. S. (1985). The role of experience and contact in the recognition of faces of own- and other-race persons. *Journal of Social Issues, 41*(3), 139-155.

Bruce, V. (1982). Changing faces: Visual and non-visual processes in face recognition. *British Journal of Psychology, 73*, 105-116.

Carey, S. (1992). Becoming a face expert. In V. Bruce, A. Cowey, A. W. Ellis & D. I. Perrett (Eds.), *Processing the facial image.* (pp. 95-103). New York, NY US: Clarendon Press/Oxford University Press.

Carey, S., & Diamond, R. (1977). From piecemeal to configurational representation of faces. *Science Magazine, 195,* 312-314.

Carey, S., Diamond, R., & Woods, B. (1980). Development of face recognition: A maturational component? *Developmental Psychology, 16*(4), 257-269.

Chance, J. E., & Goldstein, A. G. (1996). The other-race effect and eyewitness identification. In S. L. Sporer, R. S. Malpass & G. Koehnken (Eds.), *Psychological issues in eyewitness identification* (pp. 153-176). Hillsdale, NJ England: Lawrence Erlbaum Associates, Inc.

Chance, J. E., Goldstein, A. G., & McBride, L. (1975). Differential experience and recognition of memory for faces. *Journal of Social Psychology, 97*(2), 243-254.

Chance, J. E., Turner, A. L., & Goldstein, A. G. (1982). Development of differential recognition for own- and other-race faces. *Journal of Psychology, 112*(1), 29-37.

Chiroro, P. M., Tredoux, C. G., Radaelli, S., & Meissner, C. A. (2008). Recognizing faces across continents: The effect of within-race variations on the own-race bias in face recognition. *Psychonomic Bulletin & Review, 15*(6), 1089-1092.

Coetzee, V., Greeff, J. M., Barrett, L., & Henzi, S. P. (2009). Facial-based ethnic recognition: insights from two closely related but ethnically distinct groups. *South African Journal of Science, 105*(11/12), 464-466.

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.

Cross, J., Cross, J., & Daly, J. (1971). Sex, race, age, and beauty as factors in recognition of faces. *Attention, Perception, & Psychophysics, 10*(6), 393-396.

Cutler, B. L., & Penrod, S. D. (1995). *Mistaken identification: The eyewitness, psychology, and the law*. Cambridge, England: Cambridge University Press.

Cutler, B. L., Penrod, S. D., & Dexter, H. R. (1990). Juror sensitivity to eyewitness identification evidence. *Law and Human Behavior, 14*(2), 185-191.

Cutler, B. L., Penrod, S. D., & Stuve, T. E. (1988). Juror decision making in eyewitness identification cases. *Law and Human Behavior, 12*(1), 41-55.

Deffenbacher, K. A. (1984). The influence of arousal on reliability of testimony. In S. Lloyd-Bostock & B. R. Clifford (Eds.), *Evaluating witness evidence: Recent psychological research and new perspectives* (pp. 235-251). Chichester, England: Wiley.

Devenport, J. L., Penrod, S. D., & Cutler, B. L. (1997). Eyewitness identification evidence: Evaluating commonsense evaluations. *Psychology, Public Policy, and Law, 3*(2), 338-361.

Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General, 115*(2), 107-117.

Dunning, D., Li, J., & Malpass, R. S. (1998). *Basketball fandom and cross-race identification among European Americans: Another look at the contact hypothesis.* Paper presented at the Paper presented at the biennial conference of American Psychological-Law Society, Redondo Beach, CA.

Dunning, D., & Perretta, S. (2002). Automaticity and eyewitness accuracy: A 10- to 12-second rule for distinguishing accurate from inaccurate positive identifications. *Journal of Applied Psychology, 87*(5), 951-962.

Ellis, H. D. (1990). Developmental trends in face recognition. *The Psychologist: British Bulleting of the Psychological Society, 3*, 114-119.

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

Ellis, H. D., Deregowski, J. B., & Shepherd, J. W. (1975). Descriptions of White and Black faces by White and Black subjects. *International Journal of Psychology, 10*(2), 119-123.

Facts on post-conviction DNA exonerations. (n.d.). *The Innocence Project* Retrieved April 10, 2011, 2011

Farah, M. J., Tanaka, J. W., & Drain, H. M. (1995). What Causes the Face Inversion Effect? *Journal of Experimental Psychology: Human Perception and Performance, 21*(3), 628-634.

Farah, M. J., Wilson, K. D., Drain, H. M., & Tanaka, J. W. (1998). What Is "Special" About Face Perception? *Psychological Review, 105*(3), 482-498.

Feinman, S., & Entwisle, D. R. (1976). Children's ability to recognize other children's faces. *Child Development, 47*, 506-510.

Flowe, H. D., Mehta, A., & Ebbesen, E. B. (2011). The role of eyewitness identification evidence in felony case dispositions. *Psychology, Public Policy, and Law, 17*(1), 140-159.

Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience, 3*(2), 191-197.

Gauthier, I., & Tarr, M. J. (1997). Becoming a 'Greeble' expert: Exploring mechanisms for face recognition. *Vision Research, 37*(12), 1673-1682.

Gauthier, I., Tarr, M. J., Anderson, A. W., Skudlarski, P., & Gore, J. C. (1999). Activation of the middle fusiform 'face area' increases with expertise in recognizing novel objects. *Nature Neuroscience, 2*(6), 568-573.

Gauthier, I., Williams, P., Tarr, M. J., & Tanaka, J. (1998). Training 'greeble' experts: A framework for studying expert object recognition processes. *Vision Research, 38*(15-16), 2401-2428.

Golby, A. J., Gabrieli, J. D. E., Chiao, J. Y., & Eberhardt, J. L. (2001). Differential responses in the fusiform region to same-race and other-race faces. *Nature Neuroscience, 4*(8), 845.

Goldstein, A. G. (1979). Race-related variation of facial features: Anthropometric data. *Bulletin of the Psychonomic Society, 13*, 187-190.

Goldstein, A. G., & Chance, J. E. (1976). Measuring psychological similarity of faces. *Bulletin of the Psychonomic Society, 7*(4), 407-408.

Goldstein, A. G., & Chance, J. E. (1978). Judging Face Similarity in Own and Other Races. *The Journal Of Psychology, 98*, 185-193.

Goldstein, A. G., & Chance, J. E. (1979). Do 'foreign' faces really look alike? *Bulletin of the Psychonomic Society, 13*(2), 111-113.

Greenberg, S. N., & Goshen-Gottstein, Y. (2009). Not all faces are processed equally: Evidence for featural rather than holistic processing of one's own face in a face-imaging task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*(2), 499-508.

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*(6), 1464-1480.

Hancock, K. J., & Rhodes, G. (2008). Contact, configural coding and the other-race effect in face recognition. *British Journal of Psychology, 99*(1), 45-56.

Harvey, I., Yarkin, K., Lightner, J., & Town, J. (1980). Unsolicited interpretation and recall of interpersonal events. *Journal of Personality and Social Psychology, 38*, 551-568.

Hay, D. C., & Young, A. W. (1982). The human face. In I. A. W. E. (Ed.) (Ed.), *Normality and pathology in cognitive functions* (pp. 173-202). New York: Academic Press.

Hehman, E., Mania, E. W., & Gaertner, S. L. (2009). Where the division lies: Common ingroup identity moderates the cross-race facial-recognition effect. *Journal of Experimental Social Psychology, 46*(2), 445-448.

Heron-Delaney, M., Anzures, G., Herbert, J. S., Quinn, P. C., Slater, A. M., Tanaka, J. W., Lee, K., & Pascalis, O. (2011). Perceptual training prevents the emergence of the other race effect during infancy. *Plos One, 6*(5), e19858.

Hugenberg, K., Young, S. G., Bernstein, M. J., & Sacco, D. F. (2010). The categorization-individuation model: An integrative account of the other-race recognition deficit. *Psychological Review, 117*(4), 1168-1187.

Jackiw, L. B., Arbuthnott, K. D., Pfeifer, J. E., Marcon, J. L., & Meissner, C. A. (2008). Examining the cross-race effect in lineup identification using Caucasian and First Nations samples. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement, 40*(1), 52-57.

Judd, C. M., & Park, B. (1988). Out-Group Homogeneity: Judgments of Variability at the Individual and Group Levels. *Journal of Personality & Social Psychology, 54*(5), 778-788.

Judd, C. M., Park, B., Yzerbyt, V., Gordijn, E. H., & Muller, D. (2005). Attributions of intergroup bias and outgroup homogeneity to ingroup and outgroup others. *European Journal of Social Psychology, 35*(6), 677-704.

Kanwisher, N., McDermott, J., & Chun, M. (1997). The Fusiform Face Area: A Module in Human Extrastriate Cortex Specialized for the Perception of Faces. *Journal of Neuroscience, 17*, 4302-4311.

Know The Cases: Arvin Mcgee. (n.d.). *The Innocence Project* Retrieved March 1, 2010, 2010, from http://www.innocenceproject.org/Content/Arvin_McGee.php

Kress, T., & Daum, I. (2003). Developmental prosopagnosia: A review. *Behavioural Neurology, 14*, 109-121.

Krouse, F. L. (1981). Effects of pose, pose change, and delay on face recognition performance. *Journal of Applied Psychology, 66*(651-654).

Laughery, K. R., Alexander, J. F., & Lane, A. B. (1971). Recognition of human faces: Effects of target exposure time, target position, pose position, and type of photograph. *Journal of Applied Psychology, 55*, 477-483.

Lavrakas, P. J., Buri, J. R., & Mayzner, M. S. (1976). A perspective of the recognition of other race faces. *Perception and Psychophysics, 20*, 475-481.

Leippe, M. R. (1995). The case for expert testimony about eyewitness memory. *Psychology, Public Policy, and Law, 1*, 909-959.

Levin, D. T. (1996). Classifying faces by race: The structure of face categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*(5), 1364-1382.

Levin, D. T. (2000). Race as a Visual Feature: Using Visual Search and Perceptual Discrimination Tasks to Understand Face Categories and the Cross-Race Recognition Deficit. *Journal of Experimental Psychology: General, 129*(4), 559-574.

Lindsay, R. C., Brigham, J. C., Elizabeth Brimacombe, C., & Wells, G. L. (2004). Eyewitness Research Taking Psychology and Law into the Twenty-First Century (Vol. 14, pp. 199-223): Springer US.

Lindsay, R. C., & Wells, G. L. (1983). What do we really know about cross-race eyewitness identification? In B. R. C. E. In S. M. A. Lloyd-Boystock (Ed.), *Evaluating witness evidence*. Chichester: Wiley.

Lindsay, R. C., Wells, G. L., & O'Connor, F. J. (1989). Mock-juror belief of accurate and inaccurate eyewitnesses: A replication and extension. *Law & Human Behavior, 13*, 333-339.

Loftus, E. (1979). *Eyewitness Testimony*. Cambridge, MA: Harvard University Press.

MacLin, O. H., MacLin, M. K., & Malpass, R. S. (2001). Race, arousal, attention, exposure, and delay - An examination of factors moderating face recognition. *Psychology Public Policy and Law, 7*(1), 134-152.

MacLin, O. H., & Malpass, R. S. (2001). Racial categorization of faces: The ambiguous race face effect. *Psychology, Public Policy, and Law, 7*(1), 98-118.

MacLin, O. H., Van Sickler, B. R., MacLin, M. K., & Li, A. (2004). A Re-examination of the Cross-race Effect: The Role of Race, Inversion, and Basketball Trivia. *North American Journal of Psychology, 6*(2), 189-203.

Malpass, R. S., & Kravitz, J. (1969). Recognition for faces of own and other race. *Journal of Personality & Social Psychology, 13*, 330-334.

McNichol, D. (1972). A primer of signal detection theory *A primer of signal detection theory B2 - A primer of signal detection theory*. London: George Alien & Unwin Publishers.

Meissner, C. A., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology, Public Policy, and Law, 7*(1), 3-35.

Michel, C., Corneille, O., & Rossion, B. (2010). Holistic face encoding is modulated by perceived face race: Evidence from perceptual adaptation. *Visual Cognition, 18*(3), 434-455.

Mondloch, C. J., Elms, N., Maurer, D., Rhodes, G., Hayward, W. G., Tanaka, J. W., & Zhou, G. (2010). Processes underlying the cross-race effect: an investigation of holistic, featural, and relational processing of own-race versus other-race faces. *Perception, 39*(8), 1065-1085.

Mullen, B., & Hu, L. (1989). Perceptions of ingroup and outgroup variability: A meta-analytic integration. *Basic and Applied Psychology, 10*, 233-252.

Münsterberg, H. (1909). *On the witness stand : Essays on psychology and crime*: New York : Doubleday, Page, 1909.

Ng, W. J., & Lindsay, R. C. L. (1994). Cross-race facial recognition: Failure of the contact hypothesis. *Journal of Cross-Cultural Psychology, 25*(2), 217-232.

Patton, S. (2009, June 26, 2009). Innocence Denied [Interview with Peter Neufeld, the co-director of the Innocence Project]. Retrieved from http://www.thedefendersonline.com/2009/06/26/innocence-denied/

Penrod, S., Loftus, E., & Winkler, J. (1982). Reliability of Eyewitness Testimony - A Psychological Perspective. In N. L. Kerr & R. M. Bray (Eds.), *The Psychology of the Courtroom* (pp. 119-168). Orlando, FL: Academic Press.

Platz, S. J., & Hosch, H. M. (1988). Cross-Racial/Ethnic Eyewitness Identification: A Field Study. *Journal of Applied Social Psychology, 18*(11), 972-984.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85*(2), 59-108.

Ratcliff, R., & Rouder, J. (1998). Modeling Response Times For Two-Choice Decisions. *Psychological Science, 9*(5), 347-356.

Rattner, A., & Fishman, G. (1998). *Justice for all? : Jews and Arabs in the Israeli criminal system*. Westport, CT: Praeger Publishers.

Rehnman, J., & Herlitz, A. (2006). Higher face recognition ability in girls: Magnified by own-sex and own-ethnicity bias. *Memory, 14*(3), 289-296.

Rhodes, G., Brake, S., Taylor, K., & Tan, S. (1989). Expertise and configural coding in face recognition. *British Journal of Psychology, 80*(3), 313-331.

Rhodes, G., Byatt, G., Michie, P. T., & Puce, A. (2004). Is the Fusiform Face Area Specialized for Faces, Individuation, or Expert Individuation? *Journal of Cognitive Neuroscience, 16*(2), 189-203.

Rhodes, G., Hayward, W. G., & Winkler, C. (2006). Expert face coding: configural and component coding of own-race and other-race faces. *Psychonomic Bulletin & Review, 13*(3), 499-505.

Richler, J. J., Cheung, O. S., & Gauthier, I. (2011). Holistic processing predicts face recognition. *Psychological Science, 22*(4), 464-471.

Rodin, M. J. (1987). Who is memorable to whom: A study of cognitive disregard. *Social Cognition, 5*(2), 144-165.

Rossion, B., Gauthier, I., Goffaux, V., Tarr, M. J., & Crommelinck, M. (2002). Expertise Training with Novel Objects Leads to Left-Lateralized Facelike Electrophysiological Responses. *Psycholocial Science, 13*(3), 250-257.

Seymour, R. T. (2006). [Letter from R. Thomas Seymour, attorney for Arvin McGee, to Alan Jackere, City Attorney].

Shapiro, P. N., & Penrod, S. (1986). Meta-analysis of facial identification studies. *Psychological Bulletin, 100*(2), 139-156.

Shepherd, J. W., Gibling, F., & Ellis, H. D. (1991). The effects of distinctiveness, presentation time and delay on face recognition. *European Journal of Cognitive Psychology, 3*, 137-145.

Shriver, E. R., & Hugenberg, K. (2010). Power, individuation, and the cross-race recognition deficit. *Journal of Experimental Social Psychology, 46*(5), 767-774.

Shriver, E. R., Young, S., G., Hugenberg, K., Bernstein, M. J., & Lanter, J. R. (2008). Class, Race, and the Face: Social Context Modulates the Cross-Race Effect in Face Recognition. *Personality & Social Psychology Bulletin, 34*(2), 260-274.

Slone, A. E., Brigham, J. C., & Meissner, C. A. (2000). Social and Cognitive Factors Affecting the Own-Race Bias in Whites. *Basic and Applied Social Psychology, 22*(2), 71-84.

Sporer, S. L. (1993). Eyewitness identification accuracy, confidence, and decision times in simultaneous and sequential lineups. *Journal of Applied Psychology, Crime & Law, 78*, 22-33.

Sporer, S. L. (1999). *The own-race bias in Germany: Testing the contact hypothesis with Turks and Germans*. Paper presented at the 4th European Conference on Psychology and Law of the American Psychology-Law Association and the European Psychology-Law Association, Dublin, Ireland.

Sporer, S. L. (2001a). The cross-race effect: Beyond recognition of faces in the laboratory. *Psychology, Public Policy, and Law, 7*(1), 170-200.

Sporer, S. L. (2001b). Recognizing faces of other ethnic groups: An integration of theories. *Psychology, Public Policy, and Law, 7*(1), 36-97.

Stahl, J., Wiese, H., & Schweinberger, S. R. (2008). Expertise and own-race bias in face processing: an event-related potential study. *Neuroreport, 19*(5), 583-587.

Stangor, C., Lynch, L., Duan, C., & Glas, B. (1992). Categorization of individuals on the basis of multiple social features. *Journal of Personality and Social Psychology Bulletin, 62*, 207-218.

Steblay, N. (1992). A meta-analytic review of the weapon focus effect. *Law and Human Behavior, 16*(4), 413-424.

Susa, K. J., Meissner, C., & de Heer, H. (2010). Modeling The Role Of Social-Cognitive Processes In The Recognition Of Ownand Other Race Faces. *Social Cognition, 28*(4), 523-537.

Teitelbaum, S., & Geiselman, R. E. (1997). Observer mood and cross-racial cognition of faces. *Journal of Cross-Cultural Psychology, 28*(1), 93-106.

Tulsa to Settle Wrongful Conviction Suit. (2006, 6/3/2006). *Houston Chronicle*. Retrieved from http://www.chron.com/disp/story.mpl/ap/nation/3924453.html

Understand the causes: Eyewitness misidentification. (n.d.). *The Innocence Project* Retrieved May 15, 2012, 2012, from http://www.innocenceproject.org/understand/Eyewitness-Misidentification.php

Valentine, T. (1988). Upside-down faces: A review of the effect of inversion upon face. *British Journal of Psychology, 79*(4), 471.

Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion and race on face recognition. *Quarterly Journal of Experimental Psychology, 43A*, 161- 204.

Valentine, T., Chiroro, P., & Dixon, R. (1995). An account of the other race effect and the contact hypothesis based on a 'face space' model of face recognition. In I. T. V. (Ed.), *Cognitive and Computational Aspects of Face Recognition: Explorations in Face Space* (pp. 69-94). London: Routledge.

van Ravenzwaaij, D., van der Maas, H. L. J., & Wagenmakers, E. J. (2011). Does the Name-Race Implicit Association Test Measure Racial Prejudice? *Experimental Psychology, 58*(4), 271-277.

Wagenmakers, E. J., van der Maas, H. L. J., & Grasman, R. P. P. P. (2007). An EZ-diffusion model for response time and accuracy. *Psychonomic Bulletin And Review, 14*(1), 3-22.

Weimann, G., Fishman, G., & Rattner, A. (1986). Social Distance and Misidentification. *International Journal of Comparative Sociology (Brill Academic Publishers), 27*(3/4), 217-225.

Wells, G. L. (1978). Applied eyewitness-testimony research: System variables and estimator variables. *Journal of Personality & Social Psychology, 36*(12), 1546-1557.

Wells, G. L., & Loftus, E. (2007). Eyewitness memory for people and events. In R. C. L. Lindsay (Ed.), *Memory for People*. Mahwah, NJ: Lawrence Erlbaum Association.

Wells, G. L., Wright, E. F., & Bradfield, A. L. (1999). Witnesses to crime: Social and cognitive factors governing the validity of people's reports. In S. D. H. In R. Roesch, & J. R. P. Ogloff (Eds.) (Ed.), *Psychology and law: The state of the discipline* (pp. 53-87). New York: Kluwer Academic/Plenum.

Wickens, T. D. (2002). *Elementary signal detection theory*. New York, NY US: Oxford University Press.

Wickham, L. H. V., Morris, P. E., & Fritz, C. O. (2000). Facial distinctiveness: Its measurement, distribution and influence on immediate and delayed. *British Journal of Psychology, 91*(1), 99-123.

Wright, D. B., Boyd, C. E., & Tredoux, C. G. (2001). A field study of own-race bias in South Africa and England. *Psychology, Public Policy, and Law, 7*(1), 119-133.

Wright, D. B., Boyd, C. E., & Tredoux, C. G. (2003). Inter-racial contact and the own-race bias for face recognition in South Africa and England. *Applied Cognitive Psychology, 17*(3), 365-373.

Wright, D. B., & Sladden, B. (2003). An own gender bias and the importance of hair in face recognition. *Acta Psychologica, 114*(1), 101-114.

Wright, D. B., & Stroud, J. N. (2002). Age differences in lineup identification accuracy: People are better with their own age. *Law & Human Behavior, 26*, 641-654.

Yardley, L., McDermott, L., Pisarski, S., Duchaine, B., & Nakayama, K. (2008). Psychosocial consequences of developmental prosopagnosia: A problem of recognition. *Journal of Psychosomatic Research, 65*(5), 445-451.

Yarmey, A. D. (1979). *The psychology of eyewitness testimony*. New York: Free Press.

Yarmey, A. D., & Jones, H. T. (1983). Is the psychology of eyewitness identification a matter of common sense? In B. R. C. E. In S. Lloyd-Bostock (Ed.), *Recent psychological research and new perspectives* (pp. 13-40). Chichester, England: Wiley.

Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology, 81*, 141-145.

Zhao, L., & Bentin, S. (2008). Own- and other-race categorization of faces by race, gender, and age. *Psychonomic Bulletin & Review, 15*(6), 1093-1099.