5-4-2020

An Investigation of Criminal Face Bias in the Innocence Project Cases

Amanda Clevinger

Recommended Citation
Clevinger, Amanda, "An Investigation of Criminal Face Bias in the Innocence Project Cases." Dissertation, Georgia State University, 2020.
doi: https://doi.org/10.57709/17545317

AN INVESTIGATION OF CRIMINAL FACE BIAS IN THE INNOCENCE PROJECT CASES

by

AMANDA M. CLEVINGER

Under the Direction of Heather Kleider-Offutt, Ph.D.

ABSTRACT

Eyewitness memory error is the most frequent cause of wrongful conviction in the United States (The Innocence Project, 2008). Many studies have investigated causes of these misidentifications such as type of questioning, lineup presentation, and witness instructions, all which can be controlled by police procedure or laws (Wells, Memon, & Penrod, 2006). Other factors outside the control of the legal system, such as crime scene context, duration of the crime, and individual differences among individuals involved may also have effects on sentencing.

Research shows that risk for misidentification is high when an innocent suspect in a lineup looks more salient to an eyewitness than do other lineup members (Flowe & Humphries, 2011). One way in which a suspect may appear more salient in a lineup is if they possess facial features that reflect the criminal stereotype: scars, pockmarks, tattoos, long, dark, shaggy hair, prominent jaw (MacLin & Herrera, 2006; Reed & Reed, 1973). Research shows that facial ap-

pearance is not an accurate reflection of character, yet individuals continue to make character judgments based on facial features and that these judgments influence decision making (Brandt, 1980; Fiske, 1998). In the current study, I investigate the use of the criminal face stereotype in a sample of misidentified and convicted persons from the court cases indexed in the Innocence Project on-line database.

Research shows that face perception is not only a function of the face being viewed, but also of the individual perceiving the face (Hehman et al., 2017). In the current study, I test whether there are systematic differences in the perception of trait characteristics criminality, trustworthiness, and dominance as a function of conviction status (perpetrator, exonerate, filler) and features of the target face. Another factor that may affect perception is the cross-race effect (Malpass & Kravitz, 1969). The current study examines differences in facial ratings for cross-race versus same-race judgments.

Further, how perceiver characteristics relate to character judgments of others have not been fully explored. The current study will investigate whether characteristics of the perceiver, including interracial anxiety, previous victimization, and fear of future victimization are associated with facial ratings of trait impressions. These characteristics of the perceiver may be related to face judgments that pertain to an eyewitness context.

INDEX WORDS: Eyewitness identification, Criminal face stereotype, Cross-race effect

AN INVESTIGATION OF CRIMINAL FACE BIAS IN THE INNOCENCE PROJECT CASES

by

AMANDA M. CLEVINGER

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

in the College of Arts and Sciences

Georgia State University

2020

AN INVESTIGATION OF CRIMINAL FACE BIAS IN THE INNOCENCE PROJECT CASES

by

AMANDA M. CLEVINGER

Committee Chair: Heather Kleider-Offutt

Committee: Lee Branum-Martin

Chris Henrich

Erin Tone

Electronic Version Approved:

Office of Graduate Services

College of Arts and Sciences

Georgia State University

May 2020

DEDICATION

For Dr. John Mace. For many years of inspiration and guidance, I thank you.

ACKNOWLEDGEMENTS

Many thanks to Dr. Heather Kleider-Offutt for your mentorship over the years. Also, I give warm thanks to my colleagues and dear friends Samantha Emerson, Megan Capodanno, and Alesha Bond who have supported me and cheered me on. I want to express my deep appreciation for my stellar research assistants Elaina Gebhardt and Ja'Keila Scott. Also, heartfelt thanks to my committee members for giving me time you didn't have to spare. Your input on this project has been invaluable. I am most grateful to all of you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 INTRODUCTION

Eyewitnesses to a crime are often called upon to assist police in the identification of possible perpetrators. Research shows that factors such as poor viewing conditions at the crime, weak memory for the perpetrator's face, and racial differences between the witness and the perpetrator all impact memory performance with the result of making accurate lineup identification difficult (Brewer & Wells, 2011). Other factors can bias witnesses towards choosing individuals whom the police suspect to be perpetrators. For example, when the physical appearance of the suspect differs from that of other lineup members, witnesses' attention is drawn to the suspect, making a misidentification more likely (Flowe & Humphries, 2011). This difference in appearance can put a person at a greater risk for misidentification, as that individual "stands out" from the other lineup members. Such bias is of particular concern if the suspect in the lineup is actually innocent (Brewer & Wells, 2011).

A large body of research has investigated how individual differences in appearance among people in a lineup can influence identification accuracy. For instance, evidence confirms the importance of matching lineup members for such characteristics as gender, race, height, and weight (Charman, Wells, & Joy, 2011). However, other potential effects of appearance, such as the possibility that specific facial features may increase an individual's prominence within the lineup or their risk for being selected out of a lineup, have not been rigorously studied. Of particular concern to lineup decisions are feature-based judgments related to purported criminal appearance. Given that people commonly stereotype some faces as "criminal-looking" (Bull, 1992; MacLin & Herrera, 2006; Shoemaker, South, & Lowe, 1973), it is important that we understand whether and how the use of this facial stereotype may affect eyewitness identification.

Furthermore, research on person perception suggests that judgments made about other individuals are not only a function of target characteristics, but are also highly influenced by characteristics of the perceiver (Hehman et al., 2017). Thus, individuals may vary in their likelihood of using the criminal face stereotype as a function of personal characteristics or experience, such as whether they have been a victim of a violent crime in the past, have anxiety about violent crime, or have anxiety about interracial contact. Published studies to date, however, have not fully explored whether and how perceiver characteristics relate to eyewitness identification decisions in the context of police lineups. Research aimed at clarifying ways in which individual differences among eyewitnesses themselves may be associated with how they perceive faces and subsequently make decisions about them would enhance our current knowledge of eyewitness memory and decision-making in a legal context.

1.1 Purpose of the Study

In the current paper, I evaluate the strength of association between ratings of criminality and possession of facial characteristics stereotypically associated with criminality. In addition, I examine, using archival data from the Innocence Project, if rated criminality varies according to whether the target was an actual perpetrator; was convicted of a crime, but later exonerated; or has no known history of involvement with the criminal justice system. Also, I examine whether rated criminality varies as a function of the race match between the target face and the perceiver (rater). I also examine whether rater characteristics, including past exposure to violent crime, anxiety about violent crime, and anxiety about interracial contact, influence perceived criminality. This work will inform several gaps in the literature. Specifically, the bulk of previous research on face perception has been conducted using photos from actor databases, volunteers, or computer generated faces (e.g., Funk, Walker, & Todorov, 2017; Hehman, Flake, & Freeman,

2015; Oosterhof & Todorov, 2008). The current study uses images of actual convicted persons as stimuli. Further, this is the only study known to the author to directly compare exonerates and perpetrators. Also, this is one of only a few studies to examine face perception as a function of both the target and perceiver. To support my investigation, I review information on eyewitness memory and lineup procedure, first impressions, stereotypes and their consequences, the cross-race effect, and how perception may be altered by experience.

1.2 Eyewitness Memory

The reliability of eyewitness memory has been questioned as far back as 1908, in Hugo Munsterberg's seminal book, *On the Witness Stand: Essays on Psychology and Crime*. To corroborate Munsterberg's suspicions, law professor Edwin Borchard documented 65 cases of eyewitness error and discussed them in his 1932 book, *Convicting the Innocent*. Further, in the 1967 court case of *United States v. Wade*, the U.S. Supreme Court noted the shortcomings of eyewitness memory and suggested that cases of mistaken identification were quite common.

Despite these early warnings, psychological research on eyewitness memory and identification did not start in earnest until the 1970s. This body of more recent work has yielded support for earlier claims that eyewitness memory is fallible and can be shaped by factors such as eyewitness biases, stereotypes, emotion, and experience (Loftus, 2005; Wixted, 2004). However, it was only with the advent of DNA testing in the 1990s, and, as a result, DNA exonerations, that the criminal justice system began to take this work on eyewitness identification seriously (Wells & Olson, 2003).

Although eyewitness memory has been repeatedly shown to be unreliable, eyewitnesses often provide the only evidence available in a case (Lindsay, Wells, & Rumpel, 1981). In fact, it is estimated that eyewitness evidence plays a role in over 75,000 cases per year in North Amer-

ica (Goldstein, Chance, & Schneller, 1989). Despite its demonstrated unreliability, eyewitness testimony is powerful in court and is often convincing to judges and jurors, especially when the witness appears highly confident (Magnussen et al., 2010; Wells, Ferguson, & Lindsay, 1981; Wells, Lindsay, & Ferguson, 1979).

Most commonly, investigations into eyewitness misidentification distinguish between two sources of potential error (Wells, 1978). Some of these potential errors constitute system variables, in that they can be controlled by police procedure or laws and include type and number of lineup presentations (Steblay et al., 2003), witness instructions (Wells, 1978; Wells, Memon, & Penrod, 2006), and how physically similar the fillers (known innocent lineup members) are to the suspect (Charman, Wells, & Joy, 2011; Fitzgerald et al., 2013; Lindsay & Wells, 1980; Wells, Rydell, & Seelau, 1993).

Other sources of error lie outside the control of the legal system; these estimator variables include factors that are specific to the circumstances of the crime, such as viewing conditions at the crime, the duration of the crime, type of crime, whether a weapon was used, and eyewitness stress level (Wells, 1978). They also encompass individual characteristics of the people involved, like perpetrator and witness race, age, and gender.

These two types of factors interact with one another to influence judgments during lineup procedures administered by police officers. In the United States, a traditional lineup procedure involves the simultaneous presentation of six individuals, either live or in photos. One of these six individuals is the suspect, who may be guilty or innocent; the remaining five are fillers—foils who are known to be innocent (Wixted et al., 2016). Police create lineups by selecting fillers whose general features of appearance, such as gender, race, height, and weight match the witnesses' description of the suspect (Wixted et al., 2016).

The greatest risk for misidentification comes when the real perpetrator is not in the lineup, and thus the suspect in the lineup is innocent (Well, 1984). Witnesses often assume that the perpetrator is in the lineup and will try to make an identification of someone, even when they do not recognize any faces (Brewer & Wells, 2006). This assumption increases witnesses' tendency to make relative judgments, in which they compare lineup members *to each other*, rather than to their memory of the perpetrator, and engage in various reasoning processes in order to figure out who the suspect is (Wells, 1984). This type of reasoning or comparison can be particularly problematic when the innocent suspect is salient in some way and thus may encourage their misidentification. The tendency to choose the suspect at above chance levels is known as suspect bias and occurs when the suspect is particularly salient or "sticks out" in some way (Brigham, Meissner, & Wasserman, 1999). Indeed, there are many ways in which a suspect may "stick out" from other lineup members (Brigham, Meissner, & Wasserman, 1999). Currently, suspect bias has not been fully investigated. It may be that certain facial features encourage bias in choosing the suspect from a lineup and so facilitates misidentification. A better understanding of what factors encourage biased identification would shed some light on who is at risk for getting misidentified.

1.3 Physiognomy & First Impressions

During the 18th century, Johann Kaspar Lavater, a prominent Swiss writer with an interest in physiognomy, argued that one's facial features provide information about one's character (Lavater, 1797). For example, he held that the chin was an indicator of willpower and endurance; the forehead an indicator of broadmindedness and intelligence (Lavater, 1797). During the 19th century, Cesare Lombroso attempted to apply physiognomy to the prediction of criminal behavior (Lombroso, 2006; first English translation of the 1876 original). Similarly, Francis Gal-

ton sought to identify different "types" of people, including criminals, based on their facial features, using his novel technique of facial composites (Galton, 1892).

Not everyone during this time agreed with the practice and applications of physiognomy, but it was not until the 20th century that systematic experiments were conducted to test the validity of physiognomy as a science. In one such study (Litterer, 1933), participants viewed pictures of men and women from magazines and matched each picture with the "correct" social type (e.g., college president, newspaper editor, politician, university professor, actress). With only facial features to rely on, participants achieved accuracy rates that were little better than chance, and so this result failed to support the theory of physiognomy. Interestingly, the author argued that although participants were not accurate in their judgments, they were highly consistent. That is, many times participants chose (although incorrectly) the same person for the same role, which suggests consensus regarding the facial appearances of people in various social roles. Similar findings regarding accuracy emerged in a study that examined how accurately personnel managers and social workers could judge the intelligence of students from their photographs (Cook, 1939). Collectively, these studies were assumed to provide evidence that the physiognomist's claim that personal character or social role can be inferred from facial features was inaccurate.

In the 1950s, psychologist Paul Secord changed the focus of face judgment research from inferred character (i.e., physiognomy) to the *perception* of character, and so began the study of first impressions (Todorov, 2017). Across his body of work, Secord found strong consensus among participants on facial impressions for personality traits such as conscientiousness, friendliness, honesty, and intelligence (Secord, Bevan, & Dukes, 1953). Additionally, he found these impressions to generalize across cultures (Secord & Bevan, 1956). Importantly, Secord found

that faces that were rated similarly on personality traits such as trustworthiness, kindness, honesty, and friendliness, tended to have similar facial features (e.g., light eyebrows and complexion), and that participants had little or no idea regarding what facial features were causing their impressions (Bevan, Secord, & Richards, 1956; Secord, Dukes, & Bevan, 1954). This body of work suggests that people readily form impressions based on facial appearance, and these impressions are generally agreed upon within and across cultures.

More recent research replicates Secord's seminal work and also finds that there is strong agreement on what those possessing traits such as trustworthiness, aggressiveness, and emotional stability should look like (Oosterhof & Todorov, 2008; Willis & Todorov, 2006). Extending Secord's work, modern research has found that individuals make character inferences almost instantly and often without conscious awareness (Hassin & Trope, 2000; Willis & Todorov, 2006). In one such study, participants were shown pictures of females and males with neutral facial expressions for 100 ms, 500 ms, and 1,000 ms and were asked to rate each target face on trustworthiness, attractiveness, aggressiveness, and competence. Results showed that participants made consistent judgments in 100 ms or less (Willis & Todorov, 2006). Having more viewing time (500 ms and 1000 ms) did not change participants' impressions; instead, it served only to increase participants' confidence in their judgments. Similarly, Bar, Neta, and Linz (2006) found that individuals need only 39 ms—which is barely enough time to process the face—to form a first impression. Consistent with previous work, they found that impressions did not change with a presentation rate of greater than 200 ms. These findings suggest that first impressions based on facial features are nearly automatic.

1.4 Stereotypes

Feature-based judgments about character are rooted in stereotypes (Lippman, 1922). Secord and Backman (1964) define stereotyping as a process involving categorization of individuals or groups according to some set of attributes, consensus as to what those attributes are, and discrepancies between the supposed and actual traits of the person or group involved. Individuals use these stereotypes in efforts to extract more information about someone when information is limited. Despite a lack of support for the physiognomist's claims (Cook, 1939; Litterer, 1933), there is ample evidence that people still use these stereotypes today and that they influence decision making (Brandt, 1980; Fiske, 1998). In fact, studies have shown that reliance on face stereotypes can have serious real-world consequences (see Olivola, Funk, & Todorov, 2014; Todorov et al., 2015 for reviews). For example, studies show that businesses hire CEOs rated as having a more competent appearance and pay them more than they do less competent-looking applicants (Stoker, Garretsen, & Spreeuwers, 2016). Also, those perceived to have a highly dominant facial appearance are more likely to get promoted in the military than are those whose faces are less dominant (Mueller & Mazur, 1996). In an interesting study, Todorov and colleagues (2005) found that candidates for the U.S. senate whom undergraduate participants judged as having a more competent facial appearance were more likely to win elections in the years 2000 and 2002 than were those whose faces were perceived as less competent. This result has been replicated using participants from different cultures (Poutvaara, Jordahl, & Berggren, 2009; Sussman, Petkova, & Todorov, 2013) and actual voters registered in the United States (Lawson & Lenz, 2011).

Facial features associated with innocence also appear to influence judgments. In a seminal study, Zebrowitz and McDonald (1991) found that when a defendant in a court case had a

babyfaced appearance and the case was about intentional harm, the defendant was less likely to lose the case; however, a babyfaced defendant was more likely to lose when the case was about accidental harm. In contrast, plaintiffs who were babyfaced were granted more rewards (if they won), particularly when the defendant had a more mature looking appearance.

Facial features that are linked with trustworthiness also appear to influence decisions. In a lab-based study, Rezlescu and colleagues (2012) found that undergraduate participants were more willing to invest money with a trustworthy-looking partner than one who appeared less trustworthy during an investing game, even when participants were provided with each partner's past investment history (Rezlescu et al., 2012). A study conducted using data from the money-lending website, Prosper, replicated these findings. Among potential borrowers who uploaded pictures of themselves with their applications, those whose pictures looked trustworthy were more likely to get a loan and to get one with lower interest rates than were those applicants who looked less trustworthy, even though applicants also provided their credit history, debt, income, and employment information (Duarte, Siegel, & Young, 2012). Also, Wilson and Rule (2015; 2016) compared trustworthiness ratings for individuals convicted of first degree murder who were sentenced to life in prison to those for individuals who were convicted of first degree murder and were sentenced to death. Results showed that those convicted of first degree murder and subsequently sentenced to death were perceived as less trustworthy based on their facial appearance than were those who received life in prison.

The use of stereotypes regarding facial appearance has also been found to affect judgments regarding one's likelihood to behave violently or aggressively. For example, Afrocentric facial features (some combination of darker skin, wider nose, and fuller lips) have been linked to judgments of violence, aggressiveness, or likelihood to engage in criminal behavior (e.g.,

Eberhardt et al., 2004; Knuycky, Kleider, & Cavrak, 2014). In one lab-based study, for instance, participants made predictions about the likelihood that different Black men, who varied in the Afrocentricity of their features would behave aggressively in the future. Participants were also given information regarding each depicted person's past aggressive or nonaggressive behavior. Results revealed that participants predicted that men with Afrocentric facial features would engage in more aggressive behaviors in the future than would men with less Afrocentric features, regardless of their past histories of aggression (Eberhardt et al., 2004).

In another recent study, participants were tasked with selecting the perpetrator in each of two mock-crime carjacking scenarios; they were presented with perpetrator-present or perpetrator-absent lineups. Accurate identification rates were high, regardless of the perpetrator face type [high Afrocentric (87%), low Afrocentric (100%)] in the perpetrator-present lineup. In the perpetrator-absent lineup, however, innocent foils were frequently identified as the perpetrator of the crime, with higher misidentification rates for lineup members with high Afrocentric features (60%) than for those with low Afrocentric features (24%) (Knuycky, Kleider, & Cavrak, 2014).

Applied studies also indicate that Afrocentric facial features influence the way in which individuals are perceived. For example, research shows that Black defendants with more-Afrocentric facial features receive more convictions and harsher sentences (Blair, Judd, & Chapleau, 2004) and more death penalty recommendations (Eberhardt et al., 2006) than do defendants with less-Afrocentric facial features. Exonerates from the Innocence Project who were convicted on the basis of eyewitness misidentification were rated as having a more Afrocentric face type compared to those who were convicted for other reasons (Kleider-Offut et al., 2017).

Extant research consistently demonstrates that a wide range of facial stereotypes are instantly activated and used to make decisions in a variety of contexts, sometimes with severe con-

sequences. One face stereotype that is not well understood or studied in the context of lineup decisions is the criminal face stereotype. Surprisingly, little work has been done in this area despite its relevance to the criminal justice system.

1.5 Criminal face stereotype

Evaluations of "criminality" may be activated by a person's facial features. Here, the term criminality refers to the extent to which "a person's appearance activates stereotypes about criminals" rather than about any actual criminal behavior (MacLin, Malpass, & Herrera, 2001). Evidence suggests that there is widespread agreement regarding what constitutes a criminal appearance (Flowe & Humphries, 2011; MacLin & MacLin, 2004). Depending on the study, typical criminals have been described as males (Reed & Reed, 1973) with shaggy long dark hair, scars, pockmarks, facial hair, tattoos, beady eyes (MacLin & Herrera, 2006), a prominent chin, lowered eyebrows, and darker pigmentation (Funk, Walker, & Todorov, 2017). In addition, research shows that physically unattractive people are rated as more likely to commit a crime than attractive people (MacLin & MacLin, 2004).

Funk, Walker, and Todorov (2017) posit that evaluations of criminality are largely driven by perceived untrustworthiness—an indicator of intent to harm—and dominance—an indicator of physical ability to act on that intent (see also Oosterhof & Todorov, 2008; Todorov et al., 2008). Relationships among criminality, (un)trustworthiness, and dominance have been suggested in the literature. For example, Flowe (2012) presented mugshots from a police database to participants and had them make various ratings of each person pictured. Results showed that faces rated high on criminality were also rated as less trustworthy and more dominant. Tendencies to judge a face's trustworthiness and dominance may have adaptive value in that they can guide quick decisions regarding whom to approach and whom to avoid (Oosterhof & Todorov,

2008). However, Todorov (2017) argues that these tendencies can also be problematic, because people may draw inaccurate conclusions regarding others' intentions and capabilities, especially when they have very little other information available to guide their decisions, if those conclusions are based on culturally shared stereotypes, emotional expressions, and their own personal experience.

Like other facial judgments such as trustworthiness, aggressiveness, competence, and attractiveness, judgments of the criminality of a face occur almost instantly, in 100 ms or less; longer exposure time only serves to increase confidence in these judgments (Klatt et al., 2016). Further, research suggests that criminality ratings do not decline over multiple presentations of the same face (Klatt et al., 2016) and that faces rated high on perceived criminality are more likely to be remembered (MacLin & MacLin, 2004; Yarmey, 1993).

Shoemaker, South, and Lowe (1973) conducted the first lab study to demonstrate the criminal face stereotype. In this study, undergraduate participants predicted how likely middle-aged Caucasian men were to be found guilty of crimes described in contrived scenarios. In a similar classic study, Goldstein, Chance, and Gilbert (1984) asked participants to select the mass murderer, armed robber, rapist, medical doctor, clergyman, and engineer from a set of middle-aged Caucasian men whose photos had been taken from an acting directory. Participants were very consistent in whom they chose for each role, suggesting that they shared similar expectations about social roles and facial characteristics. Further, faces selected as criminals were distinct from those selected as professionals. This finding has been replicated many times, in studies that have used pictures of researchers' friends as stimuli and in samples of both undergraduate participants and police officers (see Dumas & Teste, 2006; Macrae, 1989; Macrae & Shepherd, 1989). Surprisingly, results were highly similar across undergraduate participants and po-

lice officers (Bull & Greene, 1980). This shows that individuals have an expectation (i.e., stereotype) about criminal appearance and that police may use these face stereotypes as well as ordinary citizens. There may be serious consequences if police officers use these stereotypes while looking for suspects that meet a witness description as well as during lineup construction.

A few studies have also demonstrated the criminal face stereotype in actual court cases and lineups. Flowe and Humphries (2011) investigated the relationship between criminality and mock witness identification and the decision processes underlying those choices. They used six person photo lineups that had been used in real cases as well as the actual eyewitness descriptions from those cases. Participants either received descriptive information on the perpetrator (from the eyewitness report) or no description. Participants were to determine for each lineup which person they thought was the suspect. Results showed that mock witnesses were more likely to report that they had used criminal appearance to make a lineup choice when no description of the perpetrator was provided. This suggests that witnesses may use criminal appearance as an alternative decision strategy when they have no information or memory of the perpetrator's appearance, but are presented with a lineup (study 1). In study 2, authors investigated whether the faces chosen in study 1 because of criminal appearance were, in fact, more criminal in appearance than other lineup members. Indeed, results showed that faces rated as high in criminality were chosen more often from lineups when no description was provided. This result suggests that participants used a criminal face stereotype to infer guilt. Mock witnesses did not appear to use criminal face bias when they were provided with a suspect description. Authors suggest that the criminal face stereotype may be most utilized when witness memory is weak or when the witness is highly suggestible. In that case, witnesses may choose the lineup member who looks the most criminal compared to other lineup members. Additional analyses showed that these

lineups were biased in that police had chosen foils that were less criminal in appearance compared to the suspect, making the suspect "pop out" in the lineup, highlighting that suspect bias with regard to criminal appearance occurs in actual cases.

Additional studies have also demonstrated that criminal appearance can render eyewitness lineups biased (Flowe, Klatt, & Colloff, 2014; McQuiston & Malpass, 2002). Similarly, using a mock witness paradigm, MacLin and colleagues (2001) created lineups from actual mugshots. Lineups were created by matching members to a physical suspect description, as is common police practice. The lineups were shown to participants, who received either crime information, physical appearance information, or both. Results revealed that when participants were given physical information or a combination of physical and crime information, the lineup member who best fit the physical description was chosen most often. However, when witnesses were provided with crime information only, the most criminal looking lineup member was chosen most often. Thus, the use of the criminal stereotype may have guided participants' decision process regarding who to pick from the lineup. The results of this study are supported by the finding that levels of suspect bias, the tendency to choose the suspect at above chance levels, varied depending on whether or not mock witnesses had a description of the culprit (Brigham, Meissner, & Wasserman, 1999). Further, Flowe and Humphries (2011) found that, on average, actual witnesses report an average of six features (five physical descriptors and one other type) when describing a perpetrator.

Given that witnesses often report very few details of the perpetrator, it could be the case that mock witnesses, as well as actual witnesses, rely on stereotypes to make decisions and identifications when information or memory is lacking. In support of this reasoning, research has shown that individuals are likely to rely on heuristic processing when information is lacking or

memory is weak (Kleider et al., 2008). Additionally, studies have demonstrated that individuals are more likely to fill in informational gaps with schema consistent information, regardless of its accuracy (Hastie & Park, 1986; Stangor & McMillan, 1992). Further, individuals are more likely to use stereotypes when cognitive resources are limited, (Macrae et al., 2002), such as when doing a difficult task (Bodenhausen & Lichtenstein, 1987).

It is clear from the literature that perceived criminal appearance can create suspect bias and that those with a criminal appearance are more likely to be chosen from a lineup given a lack of actual evidence (Flowe & Humphries, 2011). I examine this in the current study by using a sample of actual misidentification cases to compare the criminal appearance of those that were misidentified (exonerates), verified perpetrators, and those outside the criminal justice system. Given that many actual witnesses are able to provide only minimal physical descriptors of a perpetrator, it is possible that the criminal face stereotype was used when making lineup decisions in these cases. In misidentification cases, witnesses were presented with at least one identification procedure (e.g., lineup, showup) that contained the innocent suspect (i.e., exonerate) and the witness *should have* responded that the perpetrator was not present. However, the innocent suspect was mistakenly identified as the perpetrator. It is possible that there are systematic errors related to decision-making strategies that are committed during these misidentifications. One potential strategy that witnesses may use is the criminal face stereotype, which would systematically influence who gets selected from a lineup. A working hypothesis is that exonerates "looked criminal" and this supported the misidentification of the exonerate. This study aims to replicate previous research on the use of the criminal face stereotype in actual court cases and to extend previous work on face-type bias as a factor in actual trial outcomes by comparing the criminal appearance of exonerates (proven innocent), perpetrators (proven guilty), and fillers (no known connec-

tion with the criminal justice system). Also, criminal features have not yet been investigated in a sample of convicted persons.

Additionally, perceptions of criminality may be influenced by characteristics and background of the perceiver. For instance, it may be the case that factors such as one's own race, anxiety regarding interracial contact, previous exposure to violent crime, and anxiety about experiencing violent crime in the future may impact one's perception of criminality.

1.6 Perceiver race, the Cross-race effect, and perception

The Innocence Project reports that of its eyewitness misidentification cases, 42% involved cross-racial identifications (Innocence Project, 2018). This is alarming in light of the cross-race effect, also called the own-race bias, which is the finding that people more accurately identify people of one's own race compared to those of different races (Malpass & Kravitz, 1969). This effect has been replicated across different racial groups (Chiroro et al., 2008), different ages (Pezdek, Blandon-Gitlin, & Moore, 2003), and using different memory tests (Meissner & Brigham, 2001; Young et al., 2012). In one novel demonstration of the cross-race effect, researchers utilized unaware store cashiers as participants. Researchers had individuals of different races go through the cashiers' lines to make a purchase. Two hours later the cashiers' memory for customers of different races was tested. Results were in line with the cross-race effect, showing that cashiers had poorer memory for cross-race customers compared to same-race customers. When researchers removed data from participants who guessed at identification, the accuracy rate for cross race identification was 34% (Brigham et al., 1982).

The cross-race effect has been shown to affect misidentification rates. In fact, in a meta-analysis, Meissner and Brigham (2001) showed that individuals were 1.4 times more likely to correctly identify a face if it was of their own race. Conversely, incorrect identifications were

1.56 times more likely for cross-race compared to same-race faces. Research also shows that, in general, the greater time between the crime and identification procedure, the less accurate identification becomes (Deffenbacher et al., 2008). However, Marcon and colleagues (2010) found that the reduction in identification accuracy occurs even more quickly for cross-race compared to same-race identifications. This is important given that the interval between crime and lineup procedure can span weeks, months, or even years. Additionally, cross-race differences have been demonstrated in lineup construction. Brigham and Ready (1985) had Black and White participants create lineups using both Black and White suspects. Results showed that both Black and White participants were less selective regarding lineup foils when constructing cross-race lineups rather than same-race lineups. When constructing cross-race lineups, both Black and White participants chose lineup foils that were easy to reject, thus creating a biased lineup in which the suspect would have a greater likelihood of being chosen (suspect bias). Lineups tended to be fair when foils were the same race as the participant.

Many theories for the cross-race effect have been proposed, but causes for this effect are generally broken down into two categories. One set of theories posits that the decreased accuracy in recognizing cross-race faces comes from a lack of expertise with cross-race faces. This differential experience may have its roots in historical racial segregation laws and in present day de facto racial segregation (Goldsmith & Blakely, 2010). This segregation causes individuals to have more contact with people of their own race and so individuals become experts at processing faces of their own race and are better at recognizing people of their own race. By this theory, more experience with cross-race faces should yield better recognition accuracy for cross-race faces. Research has shown mixed support for this theory with some studies failing to find a relationship between experience with cross-race individuals and the cross-race effect (e.g., Malpass

& Kravitz, 1969). However, other work has demonstrated that practice distinguishing cross-race faces from one another has worked over a short period of time (Malpass, Lavigueur, & Weldon, 1973). Additionally, other studies have shown longer-term results. For example, Koreans who had been adopted by White European families showed a reversed cross-race effect as adults. That is, as adults, the adopted Koreans better recognized White Europeans compared to Koreans (Sangrigoli et al., 2005). A meta-analysis revealed that experience with cross-race faces only accounts for approximately 2% of the variance in face recognition (Meissner & Brigham, 2001).

The second set of theories regarding the cross-race effect draws on evidence that individuals tend to think categorically about others (Wilson, Hugenberg, & Bernstein, 2013). This categorization happens automatically and across multiple characteristics such as race, sex, and age (Ito & Urland, 2003). Sporer's (2001) account of the cross-race effect posits that the effect occurs due to differential attention allocation based on perceived in-group/out-group differences. Specifically, same-race faces cause automatic and deep encoding of facial features and attention is focused on the facial features that enable one to differentiate between same-race faces. However, when viewing other-race faces, automatic categorization occurs and encoding stops at a superficial level. Detailed facial information is not stored, thus encoding the face into the category "Asian", for example. That is, categorization makes the commonalities of out-group members more salient and the group therefore seems more homogeneous, and so individual differences are lost (Levin, 1996; 2000). Further, individuals are also more likely to judge out-group members in a negative or harsh way (Hugenberg & Sacco, 2008) as well as judge out-group members as more guilty and worthy of harsher punishments compared to in-group members (Bodenhausen & Lichtenstein, 1987; Bodenhausen & Wyer, Jr., 1985).

Research shows that people perceive less variability within categories compared to between categories, such that individuals see large differences between their in-group and their out-group; however, they see all out-group members as being almost identical (Beale & Keil, 1995). By processing individuals shallowly as a category, individuating information is missed, which results in low recognition rates. Support has been found for this theory. For example, the cross-race effect was not demonstrated if participants were informed about the cross-race effect and asked to pay attention to what differentiates faces rather than what makes them similar (Hugenberg, Miller, & Claypool, 2007). Importantly, this so called cross-race effect does not seem to be race specific, but rather specific to in- and out-groups. For example, Bernstein and colleagues (2007) found better recognition accuracy for students that went to the same university as the study participants versus a competing university, regardless of student race.

As it applies to an eyewitness context, if the witness and perpetrator are of different races, then the in-group/out-group model would predict that the face of the perpetrator was shallowly encoded by the witness due to race-based categorization. So, few, if any, individuating features of the face would be committed to memory. When that witness is then presented with a lineup, and given the fact that their memory is weak, they may rely on stereotypes such as the criminal face stereotype to aid them in making the identification. Having witnesses and perpetrators who are cross-race may exacerbate any systematic differences that lie between exonerates and perpetrators. Cross-race identifications may make differences between the in-group and out-group more salient, with the out-group seeming more homogenous and more likely to be rated in a negative (criminal) way.

However, previous work has not examined all the ways that the cross-race effect may exert its effects. Historically, the cross-race effect has been examined using various memory tasks

(e.g., Meissner & Brigham, 2001; Young et al., 2012); however it has not been examined in relation to perceptions of criminality. This study could potentially extend the cross-race effect literature by showing cross-race differences not only as measured directly by memory errors, which have been demonstrated in the past, but also on facial judgments. The current study examines whether cross-race judgments have an effect on criminality ratings. The goal is to make preliminary steps towards understanding the decision processes that may contribute to misidentification rather than the rate of misidentification that has been studied previously. Based on the theories discussed above, it may be that participants abide by the in- group/out-group model and rate out-group faces harsher (more criminal) than same-race faces so as to maximize the difference between in-group and out-groups.

Additionally, the majority of past work on the cross-race effect has used White college students as participants. The current study utilized a sample of self-identified African American women. Although the study was initially open to college students regardless of gender or race, I encountered difficulty recruiting participants who identified as male or non-African American. Consequently, I narrowed recruitment to a demographically circumscribed group that was conveniently accessible. However, there are merits to my focus on African American women in particular, because they constitute an under-studied group who merit representation in this literature. Accordingly, my study fills a gap in the literature in addition to serving as a potentially informative investigation of the cross-race effect.

Further, expertise models of the cross-race effect posit that the effect may be influenced by factors such as amount and quality of interracial contact (e.g., Freeman, Pauker, & Sanchez, 2016). The current study thus included measures of these constructs (included as subscales of the Interracial Anxiety Scale) as predictors of criminality ratings.

1.7 Perceiver background and perception

In addition to effects of race, research has demonstrated that personal life events such as trauma may affect the perception of faces. For example, abused children were compared to non-abused controls (matched for age and IQ) on a task that required them to distinguish between facial expressions of happiness, anger, fear, and sadness. Results showed that abused children judged more faces as angry compared to children with no history of abuse (Cicchetti & Curtis, 2005). Similarly, other studies have found increased attention to threat cues in children with a history of abuse (Curtis & Cicchetti, 2011, 2013). These results highlight the role of learning and experience in face perception.

Additionally, research with adult populations shows greater neural activation to threatening versus non-threatening faces (Batty & Taylor, 2003; Calvo & Beltran, 2013). Extending this work, Chu and colleagues (2016) investigated the impact of childhood interpersonal, childhood non-interpersonal, and adult trauma in a sample of healthy adults (no anxiety or depression diagnosis) with a history of childhood trauma in response to threatening and non-threatening faces. Results showed reduced differentiation between threatening faces and non-threatening or even positive faces for individuals with both adult trauma and interpersonal childhood trauma. Authors find this in keeping with the hyper-aware states of those diagnosed with post-traumatic stress disorder, responding with a high level of vigilance to both threatening and non-threatening faces.

Based on these findings that prior trauma alters face perception, the current study hopes to extend this work by investigating the relation between being the victim of a violent crime and perceptions of criminality. Consistent with the work reviewed above, it may be the case that if individuals have previously been the victim of a violent crime, they are more likely to rate faces

as highly criminal compared to those who have not been a victim of violent crime. Further, it may be the case that just worrying about becoming a victim of a violent crime in the future is enough to change face perception such that faces are rated higher in criminality compared to individuals who are less concerned or anxious about future victimization. The current study tests this.

1.8 Current Study

The current study aims to investigate systematic differences in criminality ratings for real-life exonerates, perpetrators, and fillers. A second aim is to evaluate whether rater characteristics and cross-race judgments, as well as possessing criminal facial features affect those ratings of criminality. Ultimately, this work may shed light on decision processes underpinning misidentification.

One challenge in this type of research lies in identifying a legal source through which case information and defendant photographs are publicly available. A second challenge is ensuring that exonerates are factually innocent. One option for addressing these challenges is to use images of individuals exonerated through DNA. In these cases, convictions based in part on eyewitness misidentification or other errors (e.g., false confessions, police misconduct) were later overturned based on physical evidence.

The Innocence Project maintains a publicly accessible database of DNA exonerate cases. This database provides the information needed to conduct an initial test of criminal face type bias as a factor contributing to misidentifications in real-world criminal cases. Importantly, we are assured through physical evidence that exonerates in this database are, in fact, innocent. Further, this same physical evidence that freed exonerates also implicated the true perpetrators in these cases. That is, for cases where the perpetrator has been found, their guilt has been confirmed by

DNA evidence. Although it is impossible to control for every variable that may have contributed to a conviction in a real court case, reliance on the reported case facts taken from court records provides a reasonable amount of assurance that the data are accurate. Using this database also provides an established baseline of information that may be subject to replication in other studies.

Based on the literature reviewed above, the current study investigates whether criminality ratings can be predicted from the number of criminal facial features (e.g., beady eyes (having small, glittering eyes, especially eyes that seem to gleam with evil, greed, or lust.), long and shaggy hair, scars, pock marks, etc.), conviction status of the target (exonerate, perpetrator, or filler), race match between the rater and the target face, and rater background (experience with violent crime, anxiety about violent crime, and interracial anxiety scale). To investigate, I followed the general procedure of Eberhardt and colleagues (2006), wherein, using archival murder cases, participants rated pictures of convicted persons based on a number of perceived characteristics.

Hypotheses:

1. Number of criminal features in target faces will correlate positively with criminality and dominance ratings, but negatively with trustworthiness ratings.
2. Criminality ratings will be predicted by conviction status, with exonerates being rated significantly more criminal and dominant-looking and less trustworthy than both perpetrators and fillers.
3. Criminality and dominance ratings will be significantly higher and trustworthiness ratings lower when raters evaluate faces of a different race from their own than when they evaluate faces of their own race.

4. Criminality and dominance ratings will be higher and trustworthiness ratings lower among participants who have experienced violent crime than among those who have not been the victim of a violent crime.
5. Criminality and dominance ratings will be higher and trustworthiness ratings lower among participants who are highly anxious about being a victim of violent crime in the future than among those who are less anxious about becoming a victim of a violent crime.
6. No specific hypothesis was formed for the relationship between trait ratings (criminality, dominance, trustworthiness) and interracial anxiety. Analyses were exploratory in nature.

2 EXPERIMENT

2.1 Methods

2.1.1 Participants

A power analysis revealed that a total of 200 participants would be needed for the current study. This estimate was derived from a power curve calculated in the program, Optimal Design. A power analysis assuming 20 faces per rater, 80% power, alpha = .05, ICC of faces within rater of .05 to .30, and rater characteristic R^2 of zero to .30 suggests that 120 – 200 participants would be needed to detect an effect size of between .12 and .27. While little is known about the magnitude of relations among these variables and the degree of clustering for raters versus target faces in a cross-classified model, the suggested sample size is likely to be informative, given general sample size conventions described in Maas and Hox (2005).

A total of 307 undergraduate students at Georgia State University participated in this study in exchange for course credit. Data were cleaned on a case by case basis. Five participants were excluded for being male. Seventeen participants were excluded for being a race other than African American. Fourteen participants were excluded from analyses due to completing the study in less than 20 minutes. As participants had to make 14 ratings for each of 24 faces and then complete multiple questionnaires, it is unlikely that participants who finished in less than 20 minutes answered questions in a thorough manner. Additionally, 57 participants were excluded from analyses for providing the same response to all items. Data from the resulting 214 participants were included in the analyses detailed below. All participants identified as African American and female. All participants were 18 years of age or older (range 18 – 60 years, with the majority (84.41%) between the ages of 18 to 21 years.

2.1.2 Design and Analysis

I conducted a cross-classified regression analysis to evaluate how well number of criminal facial features, conviction status (exonerate, perpetrator, or filler), race match between the rater and the target face, and rater background (experience with violent crime, anxiety about violent crime, and interracial anxiety scale) predict criminality ratings.

I conducted two additional cross-classified regression analyses using dominance and trustworthiness as outcome variables as each of these constructs have a known relationship with criminality. The predictor variables in these two analyses were the same as with the criminality analysis.

2.1.3 Materials

2.1.3.1 Stimuli

2.1.3.1.1 Innocence Project photographs.

The Innocence Project website contains 356 profiles of individuals exonerated via DNA evidence, 246 cases of which involved eyewitness identification. Profiles include information such as cause for incarceration, exonerate and eyewitness race, jurisdiction, and information on whether the actual perpetrator has been found. A companion site, available with permission from the Innocence Project, is the Innocence Record. The Innocence Record has much the same information available as on the Innocence Project site, but for many cases, also contains a wealth of original case files such as police reports, court transcripts, and news articles associated with the case. Additionally, the Innocence Record contains, for some cases, the original mugshots of both exonerates and perpetrators (for cases in which the perpetrator has been found). However, some of the case files are incomplete, as it is up to the individuals in possession of the original files to upload them to the Innocence Project/Innocence Record databases.

Out of the 246 eyewitness misidentification cases, a total of 75 mugshots of exonerates and perpetrators were available at the time of the search. However, 11 of those photos were in black and white format and so were excluded, leaving a total of 64 color mugshots. All photographs are of Black and White men: 22 Black exonerates, 18 Black perpetrators, 15 White exonerates, and 9 White perpetrators. All photographs are front-facing (i.e., no side views) and were cropped to include only features above the neck. These photos were cropped close to the chin, ears, and top of the head so that no clothes (i.e., jumpsuit) were visible and minimal background showed. Adobe Photoshop was used to resize the images to approximately 3 x 3 inches. Photos were presented to participants via Qualtrics software.

2.1.3.1.2 *Filler photographs*.

As a comparison group and in order to increase the generalizability of the study, filler faces were presented along with the exonerate and perpetrator photos. These filler faces have no known connection to the criminal justice system and were taken from an online database. I chose to use faces from The Chicago Face Database, which was developed at the University of Chicago and is made freely available for use in scientific research (Ma, Correll, & Wittenbrink, 2015). It contains high resolution photographs of male and female faces of varying ethnicities and facial expressions, ranging in age from 17 – 65 years. The database includes information on each model and includes both physical measures of each face (nose length and width, thickness of the lips, distance between the eyes) as well as subjective ratings completed by independent judges (e.g., ratings of masculinity, attractiveness, babyfacedness, happiness). Photos were standardized, maintaining the same face angle, eye level, and lighting conditions. All models wore a light gray t-shirt and were photographed against a white background. (For more information on norming data see Ma, Correll, and Wittenbrink, 2015).

All pictures chosen from this database for the current study were Black and White males: 40 White males and 24 Black males. This specific amount was chosen so as to ensure an equal number of Black and White faces overall (64 White faces and 64 Black faces).

As attractiveness has a known relation to ratings of criminality, pilot ratings were collected regarding the attractiveness of exonerate and perpetrator faces. As in the proposed study, participants (N = 41) rated the stimuli photos of perpetrators and exonerates on a 1 (not at all attractive) to 7 (very attractive) scale. Results showed an average rating of 2.24. Filler faces were matched to exonerate and perpetrator's average attractiveness (M = 2.33).

Additionally, all selected photos from the Chicago Face Database were front facing with neutral facial expressions. All photos were presented in color and were resized to approximately 3 x 3 inches so as to match perpetrator and exonerate photos. Like exonerate and perpetrator photos, filler photos were cropped close to the face, excluding clothing and minimizing background. Photos were presented to participants via Qualtrics software.

2.1.3.2 Measures

2.1.3.2.1 Criminal Facial Feature Inventory.

Previous research shows that individuals have a strong consensus as to what constitutes the "criminal" look. In order to gain an objective measure of criminal features for each target face, two independent raters indicated for each target photo (exonerates, perpetrators, and filler faces), the presence/absence of each feature (e.g., beady eyes, scars, shaggy hair). For each target photo, the number of criminal features rated as present was summed. Possible scores for each target ranged from 0 (no criminal features present) to 17 (all features present). See appendix for full inventory. Inter-rater reliability was high (Cronbach's Alpha = .94). Accordingly, ratings for each face were averaged across the two independent raters.

2.1.3.2.2. Survey of Exposure to Community Violence (SECV).

This measure is a self-report inventory that measures lifetime exposure to a wide variety of types of community violence, ranging from being chased to being raped and shot. Participants rate the frequency with which they have experienced each of these events on a 9-point scale (0 = never; 8 = almost every day). Participants also answer items for each violence type regarding where the incident took place, if the participant knew the perpetrator, and how long ago each incident occurred. This measure contains over 250 items and can be divided into subscales based

on victimization and witnessing of given crime events. The length of this scale and many of its questions (e.g., "Have you only heard about someone being chased. If so, how long ago was the person chased? Where was the person chased? Who was this person chased by?") are not well-suited to the purposes of the current study. As such, I used one of the measure's subscales called "all types of violence combined". This subscale comprises 4 items inquiring about how many times the participant has been the victim of a violent crime and the timeframe, as well as how many times the participant has witnessed violent crime and the timeframe (for the current study $\alpha = .67$). In addition to using this subscale, I added 4 items inquiring as to whether the gender and race of the perpetrator were the same as or different from those of the participant. These 4 additional items may be used in the future for another purpose. For the current study, I was particularly interested in how many times each participant has been the victim of a violent crime. I used this question to determine if each participant has been a victim of violence and coded responses as yes or no. Victim score was determined by the self-reported frequency of victimization. Participant selection of any occurrence of victimization (one time through almost every day) constitutes victimization. Participant selection of "never" constitutes no victimization. Whether a participant has been victimized in the past was entered into the analysis as a predictor variable. See appendix for modified subscale.

2.1.3.2.3 Anxiety about Violence Questionnaire (pilot version).

This is a 21 item self-report measure on which participants rate the degree to which they have a range of safety concerns (for the current study $\alpha = .90$). Most items are rated on 5-point semantic differential scales. For the current analysis, I analyzed data from the question, "How concerned are you that you will be a victim of a violent crime?" This question is rated on a 5-

point scale (1 = Not at all; 5 = Always). Level of rated concern was entered into the analysis as a predictor variable. See appendix for full measure.

2.1.3.2.4. Interracial Anxiety Scale.

This 31-item self-report measure assesses interracial anxiety along six dimensions: the quantity and quality of participants' previous contact with Black individuals, their outcome expectancies regarding interactions with Black individuals, their anxiety resulting from such interactions, their desire to avoid interactions with Black individuals, and their degree of hostility resulting from interacting with Black individuals (Plant & Devine, 2003). Overall reliability for this measure as reported by Plant and Devine (2003) is $\alpha = .71$. In this measure, participants rate items on a Likert scale from 1 (strongly disagree) to 7 (strongly agree). Some items are reverse-coded. Per measure instructions, responses were averaged to create an overall interracial anxiety score, with higher scores indicating more interracial anxiety.

For the current study, I modified the wording of the original scale slightly to reflect broader interracial anxiety rather than interracial anxiety regarding Black individuals specifically. To accomplish this, I replaced the word "Black" with the phrase "people of a different race from your own". See modified scale in appendix. Overall reliability for this modified measure was $\alpha = .78$.

2.1.3.2.5 Face Ratings

Participants rated each presented target face on the following properties based on their perception: trustworthiness, masculinity, babyfacedness, skin tone, aggressiveness, approachability, distinctiveness, Afrocentricity, White stereotypicality, attractiveness, dominance, criminality, threat, and pleasantness of facial expression. All perceptions were rated on 1 – 7 semantic differential scales where a rating of one indicated an absence of the rated property and seven indi-

cated extreme presence of the property. For example, with regard to criminality, for each face, raters made their judgments on a scale where 1 indicated "not at all criminal" and 7 indicated "extremely criminal." The exception to this convention was the skin tone scale where 1 indicated light skin tone and 7 indicated dark skin tone.

For the current study, only criminality, trustworthiness, and dominance ratings were analyzed.

2.1.4 Procedure

All instructions and stimuli were presented via computer via Qualtrics software. Participants were told that we were interested in investigating first impressions and assessing how people perceive faces. Further, participants were told:

> Your task is to rate a series of faces on different properties. Your ratings should be based on your own opinions and criteria for each property. There are no right or wrong answers.

Participants then rated 12 randomly selected White faces and 12 randomly selected Black faces (total of 24 faces). Order of presentation of White and Black faces were counterbalanced. Each photograph remained on the screen while the participant answered all questions for that face before moving on to the next face. Questions for each face were presented in random order. Participants had unlimited time in which to make their ratings.

Next, participants completed three questionnaires: Survey of Exposure to Community Violence, Anxiety about Violence Scale, and Interracial Anxiety Scale. The order of questionnaires was randomized for each participant. Lastly, participants provided demographic information, including age, gender, and race.

3 RESULTS

In person perception, there are two important sources of variability: the target face and the perceiver (Hehman et al., 2017). In order to test my hypotheses about the relative contribution of each of these sources, I conducted a univariate cross-classified multilevel regression analysis predicting criminality ratings from number of criminal facial features, conviction status of the target face (exonerate, perpetrator, filler), race match, and rater background (experience with violent crime, anxiety about violent crime, and interracial anxiety). Criminality ratings were nested in both rater and target face. The regression was fit in a sequence, adding predictors, in order to evaluate the stability of estimates (see model descriptions below).

To further evaluate the stability of my results, I conducted two additional univariate cross-classified multilevel regression analyses, one with trustworthiness as the outcome variable (same predictors), and the other with dominance as the outcome variable (same predictors). Both dominance and trustworthiness ratings were nested in both rater and target face. Both trustworthiness and dominance have been well-studied and have a known relationship to criminality, thus, it is possible to evaluate the models with criminality as the outcome variable relative to the trustworthiness and dominance models.

The hypothesized models were analyzed using the lme4 package (Bates, Maechler, & Bolker, 2012) for R software (R Core Team, 2018). Models were estimated using restricted maximum likelihood (REML). The two predictor variables anxiety about violent crime and interracial anxiety were grand mean centered. Grand mean centering was not necessary for the other predictor variables as they were dummy coded and so already had a meaningful zero.

To examine the possibility that there were preexisting differences in the distribution of criminal facial features across conviction status and target race, I conducted a 2 (target race:

black, white) x 3 (conviction status: exonerate, perpetrator, filler) ANOVA. Number of criminal facial features was the dependent variable.

Results revealed a significant main effect of conviction status on criminal facial features, $F(2, 210) = 37.30$, $p < .001$. A post hoc Bonferroni test revealed that both perpetrators and exonerates had significantly more criminal features than fillers ($p < .05$).

Results revealed a significant main effect of target race on criminal facial features, $F(1, 210) = 13.71$, $p < .001$. A post hoc Bonferroni test revealed that Black targets had significantly more criminal facial features than White targets ($p < .05$).

Results revealed a significant interaction between conviction status and target race $F(2, 208) = 5.05$, $p < .01$.A series of post hoc Bonferroni tests revealed that White exonerates had significantly more criminal facial features compared to Black exonerates. Likewise, White perpetrators had significantly more criminal facial features compared to Black perpetrators. However, Black fillers had significantly more criminal facial features than White fillers. All *p*s < .05. See limitations and future directions for more discussion on the results of this analysis.

3.1. Criminality Analyses

The mean criminality rating for target faces was 3.65 (SD = 1.90), collapsing across conviction status. Criminality rating had a strong negative correlation with rated trustworthiness (r = -.46) such that participants that gave a high criminality rating to a face was likely to give a low trustworthiness rating to that face, and vice versa. Criminality rating also show a moderate positive correlation with rated dominance (r = .29) such that participants that gave a high criminality rating to a face were likely to give a high dominance rating to that same face. Criminality rating also had a small positive correlation with number of criminal facial features (r = .16) meaning that as the number of criminal facial features a target face possessed increased, the higher the

criminality rating. Criminality rating also had a small negative correlation (r = -.16) with race match between the participant and target face such that when participant and target race matched, rated criminality tended to be lower than when there was a race mismatch. Also, there was a small positive correlation between criminality ratings and whether the target face was a perpetrator (r = .14) such that if the target face was a perpetrator, the criminality rating tended to be higher than if the target face was not a perpetrator. Lastly, there was a small positive correlation between criminality ratings and whether the target face was an exonerate (r = .15) such that if the target face was an exonerate, the criminality rating tended to be higher than if the target face was not an exonerate.

Models were fit in three stages. First, an intercept only model (baseline model), then a model containing only the face variables (perpetrator, exonerate, race match, and feature) (basic model), and, lastly, the full model containing all predictor variables (both face and rater variables). Each multilevel model can be represented with an equation. See Table 1 for an explanation of variable abbreviations.

Baseline: $Y_{trait} = I_f + I_r + e$

Basic: $Y_{trait} = B_0 + B_1Perp + B_2Exon + B_3Feat + B_4Match + I_r + e$

Full: $Y_{trait} = B_0 + B_1Perp + B_2Exon + B_3Feat + B_4Match + B_5Victim + B_6AnxVic + B_7AnxInt + e$

Number of criminal facial features was a significant predictor of criminality. The effect was similar in magnitude across the basic and full models (B = .10 and .11, respectively). See figures 1, 2, and 3 for the baseline, basic, and full models of criminality, respectively.

Conviction status was a significant predictor of criminality with both perpetrators and exonerates being rated as significantly more criminal-looking compared to filler faces. However, the results of a Wald test show no significant difference between rated criminality of perpetrators

and exonerates. The Wald statistic was calculated for both the basic and full models. $Wald_{basic}$ = 1.01 [95% CI = .68, 1.38]. $Wald_{full}$ = .90 [95% CI = .66, 1.38].

In the basic model, the average criminality rating (intercept) for a filler face was 3.02. If the face was of a perpetrator, the criminality rating increased by 1.03. If the race was of an exonerate, the criminality rating increased by .86. In the full model, the average criminality rating (intercept) for a filler face was 2.83. If the face was of a perpetrator, the criminality rating increased by 1.02. If the face was of an exonerate, the criminality rating increased by .85.

Race match was a significant predictor of criminality. Specifically, when there was a race match between rater and the face pictured, the criminality rating tended to be lower than when the rater was evaluating a face different from their race. See Table 5.

Past victimization was not a reliable predictor of rated criminality. Likewise, neither anxiety about future victimization nor interracial anxiety was found to be predictors of rated criminality.

One way to determine the separate contributions of faces and raters to the outcome variable, in this case, criminality, is to compute the intra-class correlation (ICC). This method parses out the variance in the outcome (criminality) that is due to the target face and the variance that is due to rater characteristics. By computing an ICC for each outcome, I will investigate whether the variance in faces and raters are the same across criminality, dominance, and trustworthiness. Variance estimates from the full criminality model were used to compute the intra-class correlation for both face and rater: face ICC = .22, rater ICC = .27.

3.2 Dominance Analyses

The mean dominance rating for target faces was 3.62 (SD = 1.76), collapsing across conviction status. Dominance rating also show a moderate positive correlation with rated criminali-

ty (r = .29) such that participants that gave a high dominance rating to a face were likely to give a high criminality rating to that same face. Dominance rating also had a small positive correlation with number of criminal facial features (r = .15) meaning that as the number of criminal facial features a target face possessed increased, the higher the dominance rating. Dominance rating also had a small positive correlation (r = .12) with race match between the participant and target face such that when participant and target race matched, rated dominance tended to be higher than when there was a race mismatch. Also, there was a small positive correlation between dominance ratings and whether the target face was a perpetrator (r = .13) such that if the target face was a perpetrator, the dominance rating tended to be higher than if the target face was not a perpetrator. Lastly, there was a small positive correlation between dominance ratings and whether the target face was an exonerate (r = .11) such that if the target face was an exonerate, the dominance rating tended to be higher than if the target face was not an exonerate. See Tables 2, 3, and 4 for descriptive statistics.

Like the criminality models, dominance models were also fit in 3 stages (baseline, basic, and full). The only difference between dominance and criminality models were the outcome variable. All predictor variables were the same.

Number of criminal facial features was a significant predictor of dominance and was similar in magnitude across the basic and full models (B = .07 and .08, respectively).

Conviction status was a significant predictor of dominance with both perpetrators and exonerates being rated as significantly more dominant-looking compared to filler faces. The results of a Wald test show no significant difference between rated dominance of perpetrators and exonerates. The Wald statistic was calculated for both the basic and full models. $Wald_{basic}$ = 1.58 [95% CI = .43, .97]. $Wald_{full}$ = 1.27 [95% CI = .39, .93].

In the basic model, the average dominance rating (intercept) for a filler face was 2.86. If the face was of a perpetrator, the dominance rating increased by .70. If the face was of an exonerate, the dominance rating increased by .49. In the full model, the average criminality rating (intercept) for a filler face was 2.82. If the face was of a perpetrator, the dominance rating increased by .66. If the face was of an exonerate, the dominance rating increased by .49. See Table 6.

Race match was a significant predictor of dominance. Specifically, when there was a race match between rater and the face pictured, the dominance rating tended to be higher than when the rater was evaluating a face different from their race.

Past victimization was not a reliable predictor of rated dominance. Likewise, neither anxiety about future victimization nor interracial anxiety was found to be predictors of rated dominance.

Additionally, variance estimates from the full dominance model were used to compute the ICC for both face and rater: face ICC = .12, rater ICC = .46.

3.3 Trustworthiness Analyses

The mean trustworthiness rating for target faces was 2.62 (SD = 1.53), collapsing across conviction status. Trustworthiness rating had a strong negative correlation with rated criminality ($r = -.46$) such that participants that gave a high trustworthiness rating to a face was likely to give a low criminality rating to that face, and vice versa. Trustworthiness rating also had a small positive correlation ($r = .19$) with race match between the participant and target face such that when participant and target race matched, rated trustworthiness tended to be higher than when there was a race mismatch. Trustworthiness rating also had a small negative correlation with number of criminal facial features ($r = -.09$) meaning that as the number of criminal facial fea-

tures a target face possessed increased, the lower the trustworthiness rating. Also, there was a small negative correlation between trustworthiness ratings and whether the target face was a perpetrator ($r = -.11$) such that if the target face was a perpetrator, the trustworthiness rating tended to be lower than if the target face was not a perpetrator. Lastly, there was a small negative correlation between trustworthiness ratings and whether the target face was an exonerate ($r = -.05$) such that if the target face was an exonerate, the trustworthiness rating tended to be lower than if the target face was not an exonerate.

Like the criminality and dominance models, trustworthiness models were also fit in 3 stages (baseline, basic, and full). The only difference between trustworthiness, criminality, and dominance models was the outcome variable. All predictor variables were the same.

Number of criminal facial features was a significant predictor of trustworthiness. The effect was similar in magnitude across the basic and full models ($B = -.06$ and $-.06$, respectively). See Table 7.

Conviction status was a significant predictor of trustworthiness. Specifically, both perpetrators and exonerates were rated as significantly less trustworthy compared to filler faces. Perpetrator faces were rated as significantly less trustworthy compared to exonerates. The results of a Wald test confirmed the difference between rated trustworthiness of perpetrators and exonerates. The Wald statistic was calculated for both the basic and full models. $Wald_{basic} = 2.26$ [95% CI = -.82, -.40]. $Wald_{full} = 2.21$ [95% CI = -.81, -.38].

In the basic model, the average trustworthiness rating (intercept) for a filler face was 2.85. If the face was of a perpetrator, the trustworthiness rating decreased by .61. If the face was of an exonerate, the trustworthiness rating decreased by .37. In the full model, the average trustworthiness rating (intercept) for a filler face was 2.98. If the face was of a perpetrator, the

trustworthiness rating decreased by .60. If the face was of an exonerate, the trustworthiness rating decreased by .36.

Race match was a significant predictor of trustworthiness. Specifically, when there was a race match between rater and the face pictured, the trustworthiness rating tended to be higher than when the rater was evaluating a face different from their race.

Neither past victimization nor interracial anxiety was found to be reliable predictors of rated trustworthiness. Anxiety about becoming a victim of violent crime was found to be a significant predictor of trustworthiness. Specifically, individuals who reported greater anxiety about future victimization tended to rate target faces as more trustworthy.

Additionally, variance estimates from the full trustworthiness model were used to compute the ICC for both face and rater: face ICC = .10, rater ICC = .55.

4 DISCUSSION

The goal of the current study was to make preliminary steps towards understanding the decision processes that may contribute to eyewitness misidentification. Of particular interest was the use of the criminal face stereotype and the factors that may promote its use. I examined this in the current study by using a sample of actual misidentification cases to compare the criminal appearance of those that were misidentified (exonerates), verified perpetrators, and those outside the criminal justice system. Given the relationship of perceived criminality to that of perceived dominance and trustworthiness noted in the literature (Funk, Walker, and Todorov, 2017; Oosterhof & Todorov, 2008; Todorov et al., 2008), those trait impressions were also explored here.

4.1 Criminality

4.1.1Face Variables

All four face variables were found to be significant predictors of criminality. Results of the current study show that number of criminal facial features reliably predicts ratings of criminality. Specifically, those targets that were rated as having more criminal facial features were rated as higher in criminality, which supported my hypothesis.

As predicted, results also revealed that conviction status reliably predicted rated criminality. Specifically, both perpetrators and exonerates were rated as significantly more criminal compared to filler faces. However, perpetrator faces were not significantly different compared to exonerates, which was counter to my hypothesis that exonerates would be rated a significantly more criminal than perpetrators.

The current study examined whether cross-race judgments have an effect on criminality ratings. Results showed that when there was a race match between rater and the face pictured, the criminality rating tended to be lower than when the rater was evaluating a face different from their race. That is, African American females rated Black men lower in criminality compared to white men. This finding is in line with my hypothesis.

4.1.2 Rater Variables

None of the rater variables including experience with violent crime, anxiety about future victimization, and interracial anxiety were found to be significant predictors of criminality.

4.1.3 ICC

In person perception, there are two important sources of variability: the target face and the perceiver (Hehman et al., 2017). The ICC analyses inform about the relative contribution of each of these sources of variability for impressions of criminality. The variance explained in criminality models were fairly evenly split between rater and face.

4.2 Dominance

4.2.1 Face Variables

As with criminality, all face variables were found to be significant predictors of dominance. Results of the current study show that number of criminal facial features reliably predicted ratings of dominance. Specifically, those targets that were rated as having more criminal facial features were rated as higher in dominance. This finding supports my hypothesis.

Also, results revealed that conviction status reliably predicted rated dominance ratings. Specifically, both perpetrators and exonerates were rated as significantly more dominant compared to filler faces. However, perpetrator faces were not significantly different compared to exonerates, which was counter to my hypothesis that exonerates would be rated as significantly more dominant than perpetrators.

The current study also examined whether cross-race judgments have an effect on dominance ratings. Results revealed that when there was a race match between rater and the target, the dominance rating tended to be higher than when the rater was evaluating a face different from their race. That is, African American females tended to rate Black men higher in dominance compared to white men, which was counter to my hypothesis that white men would be rated as higher in dominance than Black men.

4.2.2 Rater Variables

None of the rater variables including experience with violent crime, anxiety about future victimization, and interracial anxiety were found to be significant predictors of dominance.

4.2.3 ICC

Approximately half (46%) of variance in dominance models was explained by who the rater was. However, only 12% of variance in the dominance models was explained by characteristics of the face being rated.

4.3 Trustworthiness

4.3.1 Face Variables

As with both criminality and dominance, all four face variables were found to be significant predictors of trustworthiness. Results of the current study show that number of criminal facial features reliably predicts ratings of trustworthiness. Specifically, those targets that were rated as having more criminal facial features were rated as low on trustworthiness, which supports my hypothesis.

Also, results revealed that conviction status reliably predicted rated trustworthiness. Specifically, both perpetrators and exonerates were rated as significantly less trustworthy compared to filler faces. Further, perpetrators were rated significantly less trustworthy than exonerates, which was opposite of my prediction.

Lastly, the current study also examined whether cross-race judgments have an effect on trustworthiness ratings. Results showed that when there was a race match between rater and the face pictured, the trustworthiness rating tended to be higher than when the rater was evaluating a face different from their race. That is, African American females rated Black males as more trustworthy compared to white males, which was in line with my hypothesis.

4.3.2 Rater Variables

Neither past victimization nor interracial anxiety were significant predictors of trustworthiness ratings. However, anxiety about future victimization was found to be a significant predictor. Specifically, individuals who reported greater anxiety about future victimization tended to rate faces as *more* trustworthy. This result was unexpected, in that no other rater variables

were significant predictors in any of the other models. Additionally, this specific result was only marginally significant. This result should be further examined through replication.

4.3.3 ICC

Approximately half (55%) of variance in trustworthiness models can be explained by who the rater is. However, only 10% of variance in the trustworthiness models can be explained by characteristics of the face being rated.

4.4 General

Several face and rater variables were explored here in trying to predict trait impressions of criminality, dominance, and trustworthiness. Largely, it was only the face variables that had predictive power. Below I discuss the contributions of individual predictors and their relation to existing literature.

4.4.1 Features

Number of criminal features was predictive of all three outcomes. The result that number of criminal facial features predicts trait judgments supports previous studies finding that individuals make feature-based judgments (stereotypes) to create expectations about what certain traits should look like and that the use of these facial stereotypes influence decision making (e.g., Blair, Judd, & Chapleau, 2004; Brandt, 1980; Eberhardt et al., 2006; Fiske, 1998; Zebrowitz & McDonald, 1991). In particular, this finding supports previous research finding that there is strong agreement on what those possessing traits such as trustworthiness and dominance should look like (e.g., Oosterhof & Todorov, 2008; Willis & Todorov, 2006). The current study also adds to the accumulating evidence that suggests that there is widespread agreement regarding what constitutes a criminal appearance (criminal stereotype) (Flowe & Humphries, 2011; MacLin & MacLin, 2004). In previous studies, typical criminals have been described as males

(Reed & Reed, 1973) with shaggy long dark hair, scars, pockmarks, facial hair, tattoos, beady eyes (MacLin & Herrera, 2006), a prominent chin, lowered eyebrows, and/or darker pigmentation (Funk, Walker, & Todorov, 2017). This was the first study known to the author to investigate so-called criminal features in a sample of actual convicted persons. All of these facial features were examined in the current study and found to be related to criminal appearance, with more criminal facial features constituting a more criminal appearance. Lastly, this result supports previous work suggesting a reliable relationship between criminality, dominance, and trustworthiness, specifically, that target faces that are rated as high in criminality are also rated high in dominance, but low on trustworthiness (Flowe, 2012; Oosterhof & Todorov, 2008; Todorov, Said, Engell, Oosterhof, 2008).

4.4.2 Conviction Status

Likewise, conviction status was also predictive of all three outcomes. As discussed above, these findings are in accordance with the literature in that criminality and dominance ratings have an inverse relationship with trustworthiness (Flowe, 2012; Oosterhof & Todorov, 2008; Todorov, Said, Engell, Oosterhof, 2008).

In considering the result that perpetrators and exonerates were found to be more criminal and dominant-looking and less trustworthy compared to filler faces, there are many possible explanations. One reason lies in the conditions under which photos were taken. Todorov and Porter (2014) point out that the face is rarely stationary, but continually changes to express different emotions, some subtle and others quite obvious. A microexpression is a brief, involuntary facial expression (Ekman & Friesen, 1969). Research shows that microexpressions occur as quickly as .4 to .5 seconds (Matsumoto & Hwang, 2018) and may be signs of concealed emotions or intentions (Porter & ten Brinke, 2008). Previous research has shown that even faces perceived as neu-

tral at a macro level show emotion through microexpressions (Ekman & Friesen, 1969; Said, Haxby, & Todorov, 2011) and these emotions may be interpreted as traits (e.g., trustworthiness) (Knutson, 1996; Malatesta, Fiore, & Messina, 1987; Montepare & Dobish, 2003). It follows that photos may capture microexpressions and thus influence character judgment. That is, character judgment may then, be highly dependent upon the exact moment the picture is taken.

The microexpression literature may be especially relevant here because of the specific conditions under which pictures of the targets were taken. Individuals who posed for the filler photos did so voluntarily; those who provided mugshots presumably did so involuntarily. It seems plausible that individuals being placed under arrest would have been under stress at the time of being photographed. So, even though all stimulus pictures were taken in a similar manner – front facing, neutral expressions, background minimized, there could be subtle differences (microexpressions) that reflect the stress perpetrators and exonerates were under when their pictures were taken. So, it is possible that photos taken during the process of an arrest may differ in some fundamental ways from photos taken in a lab setting and that picture selection here may have contributed to the result that perpetrators and exonerates had higher criminality ratings compared to filler faces.

Todorov and Porter (2014) found that random variations in pictures of the same individual resulted in different trait impressions. In fact, they found that there was just as much variation in impressions of the same person across different pictures as compared to variation of different individuals. They argue that trait impressions are highly dependent on which picture is chosen. Further, Todorov and Porter (2014) challenge findings from Valla and colleagues (2011) and Porter and colleagues (2008) for their choice of control images. Valla and colleagues (2011) compared mugshots to photos of students in terms of criminality. Results showed that mugshots

were rated as more criminal than students. Similarly, Porter and colleagues (2008) compared photos of criminals featured on America's Most Wanted with photos of recipients of the Nobel Peace Prize. Results showed that photos chosen from America's Most Wanted were rated as significantly more criminal than Nobel Peace Prize winners. Todorov and Porter (2014) argue that it may be the difference in context under which these photos were taken that may be responsible for results and not actual differences in criminality or appearance.

In support of Todorov and Porter's arguments, Rule and colleagues (2013) conducted a study in which they compared criminals and non-criminals within the same context. In a multipart study, Rule and colleagues (2013) compared criminals and non-criminals within the business world and found that these two groups were indistinguishable in terms of perceptions of trustworthiness (Study 2). Again, comparing photos within the same domain, authors found no perceived differences in trustworthiness for military criminals and military heroes (Study 3). Finally, authors compared students who cheated on a test in the lab versus students who did not. Again, their faces were indistinguishable in terms of a trustworthy appearance. Considering this series of studies in combination with Todorov and Porter's (2014) work and the current study may suggest that significant differences in trait appearances may be highly dependent upon photo selection and the conditions those photos were taken under.

Another possible explanation for these results may come from the difference in photos. The lab photos from the Chicago Face Database were all quite uniform in posing. However, although mugshots taken from the Innocence Project were all front facing, not every aspect of the posing was as controlled as in the database photos. For example, research shows that head tilt has the result of influencing trait perceptions such that a downward head tilt increases perceptions of facial femininity (Burke & Sulikowski, 2010) and an upward tilt increases perceptions of

facial masculinity (Burke & Sulikowski, 2010). Faces with an upward tilt have also been perceived to have a heavier weight, which relates to increased perception of intimidation and dominance (Schneider, Hecht, & Carbon, 2012). Further, research shows that manipulating the facial width-to-height ratio (fWHR) can influence perceptions of dominance and increase perceptions of intimidation. Facial WHR relies on an individual's bone structure and is measured as the distance between the cheekbones divided by the distance between the upper lip and mid-brow (Carre, McCormick, & Mondloch, 2009). Research shows that greater fWHR is related to perceptions of aggressiveness (Carre, McCormick, & Mondloch, 2009; Geniole et al., 2012) and untrustworthiness (Stirrat & Perrett, 2010) compared to those with smaller fWHR. The fWHR is generally thought to be an unchangeable feature of the face; however, Hehman and colleagues (2013) investigated the possibility that fWHR could be manipulated to appear to be increased by tilting the head upwards or downwards. They found that both male and female targets were perceived to be more intimidating when tilting the head upwards or downwards compared to no head tilt. Authors attribute this perceived intimidation to be due to the appearance of an increased fWHR. In the current study, it could be the case that at the time mugshot photos were taken, heads were tilted in an upwards or downwards tilt as an expression of dominance or intimidation out of anger at the arresting officer, thus making mugshot photos of perpetrators and exonerates appear more dominant and criminal and less trustworthy.

Furthermore, the fWHR may be a potential explanation for the finding that perpetrators were rated as less trustworthy compared to exonerates. Not only is a greater fWHR related to perceptions of aggression and untrustworthiness, studies have also shown that greater fWHR is positively correlated with greater physical aggression (Carre & McCormick, 2008; Geniole et al., 2014, Goetz et al., 2013; Haselhuhn, Ormiston, & Wong, 2015; but see Deaner et al., 2012 and

Gomez-Valdes et al., 2012 for another view), physical strength (Hehman, Flake, & Freeman, 2015), and untrustworthy behavior (Stirrat & Perrett, 2010; but see Rule et al., 2013 for another view). So, the true perpetrators of the crimes in the Innocence Project database were factually guilty of an untrustworthy behavior (a heinous crime) and physically and mentally able to carry out these ill intentions. The criminal history of exonerates was not explored in this study. We know that exonerates were innocent of the crime they were incarcerated for, but we do not know of previous arrest records for crimes they may have committed before or after that time. Presumably the filler faces had no known connection to the criminal justice system and likely varied in their intentions and ability to carry those intentions out.

4.4.3 Race Match

Further, race match was also predictive of all three outcomes. The current results support previous work finding that individuals are more likely to judge out-group members in a negative or harsh way (Hugenberg & Sacco, 2008) as well as judge out-group members as more guilty compared to in-group members (Bodenhausen & Wyer, Jr., 1985). Here, African American women rated out-group, White men, as more criminal than Black men. The finding that participants rated their out-group as more negative is consistent with the theory of the cross-race effect that posits that individuals tend to think categorically about others (Wilson, Hugenberg, & Bernstein, 2013).

In the current study, African American women rated Black men as less criminal, but more dominant. This is an interesting finding in light of the literature. Previous studies have shown that faces rated high on criminality also tend to be rated as more dominant (Flowe, 2012; Oosterhof & Todorov, 2008; Todorov et al., 2008). Dominance is typically associated with physical ability and potentially even aggression (Oosterhof & Todorov, 2008). Framed in this

way, this finding may then support the idea that individuals see their own in-group as more positive (physically capable) and their out-group as possessing more negative traits (less physically capable). Indeed, research shows that how one views dominance may largely depend on context (Hehman et al., 2015). For example, being dominant may be viewed as a positive trait in certain kinds of competitions (e.g., sports).

4.4.4. Rater Variables

There are many possible reasons why rater variables were not significant predictors of trait impressions here. For the anxiety about becoming a victim of a crime in the future variable, perhaps it is the case that participants have become desensitized to crime and victimization due to living in a major city and hearing about the widespread stories about different crimes. Furthermore, being anxious about becoming a victim may depend of context. Participants may not be an accurate judge when making ratings from home, which is likely a safe and comfortable environment. So, making ratings about concerns over future crime in a safe context may not make for accurate ratings. Of course, it could be the case that simply worrying over future victimization is not enough to alter face perception.

Past victimization was not found to be a significant predictor of trait impressions in the current study. Previous work (Chu et al., 2016; Cicchetti & Curtis, 2005; Curtis & Cicchetti, 2011, 2013) has demonstrated that past experience, including trauma and abuse, has the ability to influence face perception. That finding was not replicated here. One thing to consider in light of the current findings is that we do not know the nature of participants' past victimization. It could be that changes in face perception occur only after trauma or abuse of a certain severity or that happens repeatedly over time.

It could be the case that interracial anxiety was not a significant predictor of trait impressions when considering that participants were all African American females and that the stimulus photos were of both Black and white men. In this case, it is possible that interracial anxiety only has an effect on trait impressions when viewing white faces. That is, it may be the case that racial match interacts with interracial anxiety to influence ratings of trait impressions. Future research should explore this possibility.

4.4.5 ICC

The variance explained in criminality models were fairly evenly split between characteristics of the rater and characteristics of the face being rated. However, dominance and trustworthiness showed a different pattern of clustering, with rater variables accounting for the bulk of the variance in the model. In examining these values, it is apparent that rater characteristics play an important role in trait impressions, especially for dominance and trustworthiness. However, none of the rater variables in the current study were found to be significant. Future work should either refine the measures used for rater variables in the current study or seek new variables that might impact trait impressions. Given the amount of variance explained for rater characteristics, this seems like a worth-while avenue for future research.

This different pattern of face and rater variance across trait impressions has been previously studied. Hehman and colleagues (2017) suggest that some impressions are driven largely by the characteristics of the perceiver while other impressions are largely driven by the target. Indeed, the current pattern of results across the three models supports this idea in that the rater and face variance was different for criminality compared to trustworthiness and dominance. Hehman and colleagues (2017) further explain that impressions of trustworthiness and dominance may vary greatly across perceivers because each perceiver may have a different idea of

what constitutes a specific trait. Authors give the example that one perceiver may think that a target with a strong build is dominant; however, another perceiver may think that a target with a confident smile is dominant. As such, it is likely that perceivers focus on different facial features in which to make their trait judgments. This is in agreement with the current study as variability due to rater was much larger for the trait impressions of dominance and trustworthiness.

4.5 Implications

Factors influencing eyewitness misidentification merit investigation as the consequences of being misperceived as criminal may be severe. Moreover, lives are potentially put in danger if the true perpetrator in an eyewitness misidentification case remains free.

The current study was the first known to directly compare ratings of the appearance of perpetrators and exonerates. The results showed that exonerates and perpetrators were perceived as more criminal and dominant, while being less trustworthy than fillers with no known connection to the criminal justice system. Even though my hypotheses regarding exonerates possibly having a more criminal look than perpetrators was not supported, it is still worthwhile to examine the result that for criminality and dominance, perpetrators and exonerates were indistinguishable. If perpetrators and exonerates are equally criminal and dominant-looking and it is the innocent suspect (future exonerate) in the lineup who looks criminal amongst other lineup members who do not look criminal or look less criminal, that may cause suspect bias and may facilitate misidentification. This possibility is potentially increased given that the fillers in lineups are often off duty police officers, interns, or taken from old drivers license photos (i.e., individuals with no known connection to the criminal justice system) (Wells & Olson, 2003). Furthermore, faces rated high on perceived criminality (e.g., exonerates) are more likely to be remembered (MacLin & MacLin, 2004; Yarmey, 1993), which is a concern given that many criminal investi-

gations conduct multiple lineups and showups in a case, which promotes an innocent suspect being misidentified as the perpetrator because they are familiar (because they have been viewed many times) rather than because they were the actual perpetrator of the crime.

The current findings on differences in criminal appearance together with past research call for criminal appearance to be taken into account, along with physical appearance, when constructing lineups. Flowe and Humphries (2011), who analyzed real police lineups, found that in over half of the lineups used, the suspect was rated as the most or second most criminal-looking member. Measures of lineup fairness should include both measures of physical similarity of lineup members and measures of criminal appearance similarity. This ensures a fair lineup in which all members have an equal chance of being selected. In creating a fair lineup, criminality could be held constant across lineup members so as to reduce the impact of criminal face bias and thus reduce the likelihood of misidentifications.

However, more research is needed before firm conclusions can be made. In order to make the conclusion about exonerates from the Innocence Project being victims of suspect bias, the original lineups and lineup members from these cases would need to be studied. However, a thorough search of the Innocence Project and the Innocence Record reveal that these photo lineups are rare to find.

Race match was found to significantly affect trait ratings of criminality, dominance, and trustworthiness in the current study. Accordingly, cross-race effects may have an effect on trait inferences at different stages in a case. Cross-race effects may occur when a witness sees a cross-race perpetrator commit a crime, cross-race effects may occur again if a police officer is creating a lineup of people that are of a different race than they are, and finally again, when the witness attempts a cross-race identification from the lineup. Police should be trained on cross-

race effects and other biases that might affect lineup creation. When possible, police officers should construct only same-race lineups to avoid inserting bias into the lineup.

Yet another application for this work may be the implementation of police training in the use of stereotypes. Research has shown that police officers are just as likely as ordinary citizens to use stereotypes (Dumas & Teste, 2006; Macrae, 1989; Macrae & Shepherd, 1989), so training them to avoid creating biased lineups may help reduce the problem as witnesses will actually have to discriminate the lineup members rather than just defaulting to the most criminal-looking member.

Another avenue for improvement may be jury instructions. Currently, at the end of a trial, many courts give jury instructions containing information about the unreliable nature of eyewitness memory (National Academy of Science, 2014). It may be possible to include additional information about face biases and cross-race effects in jury instructions. However, the utility of jury instructions on eyewitness identification is debated, with some studies finding that such instructions only serve to make jurors more suspicious of all eyewitness information (Martire & Kemp, 2009).

4.6 Limitations and Future Directions

One limitation of the current study was that pictures taken from the Innocence Project and Innocence Record vary in quality (e.g., photos taken with different cameras, different lighting, etc.), which is a drawback in using real case information. This may have been problematic, as filler photos from the Chicago Face Database were all high-quality digital images taken under standardized conditions in which models were instructed to pose with neutral facial expressions. If perpetrator and exonerate photos contained even subtle emotional expressions, this could have influenced results. Further, there was an uneven distribution of facial features categorized as

"criminal" across target race, which may have affected results. Specifically, it could be the case that White faces were rated as more criminal and less trustworthy solely because they possessed a higher number of criminal facial features and may have nothing to do with cross-race judgment. If future studies could keep constant the number of criminal facial features across race, that would allow a clearer picture to emerge regarding the contribution of race versus criminal facial features. Lastly, there were only quality photos available for 37 exonerates and 27 perpetrators. As a result, caution should be exercised in drawing strong conclusions from this study until findings can be replicated.

Another limitation pertains to the analyses. Outcomes were on a 7-point ordinal scale, but for the purposes of the current project, were treated as intervally scaled. A multilevel model for categorical outcomes (e.g., logistic) could provide validation of the current results. Also, additional analyses are needed in order to determine any possible moderating effect of racial mismatch and interracial anxiety.

Additionally, the three outcomes were fit separately. While this approach allows us to evaluate predictors' effects multiple times, it may invite Type-I error. Further, the separate analyses neglect relations across face ratings. A three-outcome multilevel SEM could provide more complete information about the joint relations across these three outcomes.

As this was a preliminary investigation, replication is needed in order to strength findings. Future studies should utilize different photo stimuli sets from other databases. Given the possible limitations of using highly standardized photos as a control set for mugshots discussed above, it will be important for future studies to either take extra steps to control for additional aspects of differences in photos or find a different control group that is more comparable to mugshots. Additionally, given the nature of proposed criminal facial features noted in the litera-

ture, future work may investigate the possibility that criminal features are confounded with socioeconomic status. Future studies may utilize different measures of rater characteristics or investigate additional rater characteristics than the ones examined in the current study. Lastly, other facial features could be coded and modeled in order to account for factors such as stress, duress, or emotional content.

REFERENCES

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*(4), 390-412. doi:10.1016/j.jml.2007.12.005

Bates, D.M., Maechler, M., & Bolker, B. (2012). lme4: Linear mixed-effects models using S4 classes. R package version 0.999999-0.

Bar, M., Neta, M., & Linz, H. (2006). Very first impressions. *Emotion, 6,* 269-278.

Batty, M. & Taylor, M. J. (2005). Early processing of six basic facial emotional expressions. *Cognitive Brain Research, 17*, 613-620.

Beale, J. M. & Keil, C. (1995). Categorical effects in the perception of faces. *Cognition, 57*, 217-239.

Bernstein, D. P. & Fink, L. (1998). Childhood Trauma Questionnaire: a retrospective self-report manual. San Antonio, TX: The Psychological Corporation.

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*, 706-712.

Bevan, W., Secord, P. F., & Richards, J. M. (1956). Personalities in faces: Personal identification and the judgment of facial characteristics. *Journal of Social Psychology, 44,* 289-291.

Blair, I. V., Judd, C. M., & Chapleau, K. M. (2004). The influence of Afrocentric facial features in criminal sentencing. *Psychological Science, 15*, 674-679.

Bodenhausen, G. V. & Lichtenstein, M. (1987). Social stereotypes and information-processing strategies: the impact of task complexity. *Journal of Personality and Social Psychology, 52*(5), 871-880.

Bodenhausen. G. V. & Wyer, R. S., Jr. (1985). Effects of stereotypes on decision making and information processing strategies. *Journal of Personality and Social Psychology, 48*, 267-282.

Borchard, E. M. (1932). Convicting the innocent: Sixty-five actual errors of criminal justice. New York: Garden City Publishing Company.

Brandt, A. (1980). Face-reading: the persistence of physiognomy. *Psychology Today*, December, 90-96.

Brewer, N. & Wells, G. L. (2011). Eyewitness identification. *Current Directions in Psychological Science, 20*(1), 24-27.

Brewer, N. & Well, G. L. (2006). The confidence-accuracy relationship in eyewitness identification: Effects of lineup instructions, foil similarity, and target-absent base rates. *Journal of Experimental Psychology: Applied, 12*, 11-30.

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

Brigham, J. C., Meissner, C. A., & Wasserman, A. W. (1999). Applied issues in the construction and expert assessment of photo lineups. *Applied Cognitive Psychology, 13*, 73-92.

Brigham, J. C. & Ready, D. J. (1985). Own-race bias in lineup construction. *Law and Human Behavior, 9*, 415-424.

Bull, R. (1992). Physical appearance and criminality. *Current Psychological Reviews, 2*, 269 – 281.

Bull, R. H. & Greene, J. (1980). The relationship between physical appearance and criminality. *Medicine, Science, and the Law, 20*, 79-83.

Burke, D. & Sulikowski, D. (2010). A new viewpoint on the evolution of sexually dimorphic human faces. *Evolutionary Psychology, 21,* 573-585.

Calvo, M. G. & Beltran, D. (2013). Recognition advantage of happy faces: Tracing the neurocognitive processes. *Neuropsychologia, 51,* 2051-2060.

Carre, J. M. & McCormick, C. M. (2008). In your face: Facial metrics predict aggressive behavior in the laboratory and in varsity and professional hockey players. *Proceedings of the Royal Society B: Biological Science, 275,* 2651-2656.

Carre, J. M., McCormick, C. M., & Mondloch, C. J. (2009). Facial structure is a reliable cue of aggressive behavior. *Psychological Science, 20,* 1194-1198.

Charman, S. D., Wells, G. L., & Joy, S. W. (2011). The dud effect: Adding highly dissimilar fillers increases confidence in lineup identifications. *Law and Human Behavior, 35,* 479-500.

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 and Review, 15,* 1089-1092.

Chu, D. A., Bryant, R. A., Gatt, J. M., Harris, A. W. (2016). Failure to differentiate between threat related and positive emotion cues in healthy adults with childhood interpersonal or adult trauma. *Journal of Psychiatric Research, 78,* 31-41.

Cicchetti, D. & Curtis, W. J. (2005). A event-related potential study of the processing of affective facial expressions in young children who experienced maltreatment during the first year of life. *Developmental Psychopathology, 17,* 641-677.

Cook, S. W. (1939). The judgment of intelligence from photographs. *Journal of Applied Psychology, 8,* 215-231.

Curtis, W. J. & Cicchetti, D. (2013). Affective facial expression processing in 15-month-old infants who have experienced maltreatment: An event-related potential study. *Child Maltreatment, 18* (3), 140 – 154.

Curtis, W. J. & Cicchetti, D. (2011). Affective facial expression processing in young children who have experienced maltreatment during the first year of life: An event-related potential study. *Development and Psychopathology, 23* (2), 373 – 395.

Deaner, R. O., Goetz, S. M., Shattuck, K., & Schnotala, T. (2012). Body weight, not facial width-to-height ratio, predicts aggression in pro hockey players. *Journal of Research in Personality, 46*, 235 – 238.

Deffenbacher, K. A., Bornstein, B. H., McGorty, E. K., & Penrod, S. D. (2008). Forgetting the once-seen face: Estimating the strength of an eyewitness's memory representation. *Journal of Experimental Psychology: Applied, 14*(2), 139-150.

Dotsch, R., Wigboldus, D. H., & van Knippenberg, A. (2011). Biased allocation of faces to social categories. *Journal of Personality and Social Psychology, 100*, 999-1014.

Duarte, J., Siegel, S., & Young, L. (2012). Trust and credit: the role of appearance in peer-to-peer lending. *Review of Financial Studies, 25*, 2455-2484.

Dumas, R. & Teste, B. (2006). The influence of criminal facial stereotypes on juridic judgments. *Swiss Journal of Psychology, 65*, 237-244.

Eberhardt, J. L., Davies, P. G., Purdie-Vaughns, V. J., & Johnson, S. (2006). Looking deathworthy: Perceived stereotypicality of Black defendants predicts capital-sentencing outcomes. *Psychological Science, 17*, 383-386.

Eberhardt, J. L., Goff, P., Purdie, V. J., & Davies, P. G. (2004). Seeing Black: Race, crime, and visual processing. *Journal of Personality and Social Psychology, 87*, 876-893.

Ekman, P. & Friesen, W. V. (1969). Nonverbal leakage and clues to deception. *Psychiatry*, *32*, 88 – 105.

Farah, M. J., Wilson, K. D., Drain, M., & Tanaka, J. N. (1998). What is "special" about face perception? *Psychological Review, 105* (3), 482 – 498.

Fiske, S. T. (1998). Stereotyping, Prejudice, and Discrimination. p. 357–411. In Daniel T. Gilbert, Susan T. Fiske and Gardner Lindzey (Eds.). *The handbook of social psychology*. Boston: McGraw-Hill.

Fitzgerald, R. J., Price, H. L., Oriet, C., & Charman, S. D. (2013). The effect of suspect-filler similarity on eyewitness identification decisions: a meta-analysis. *Psychology, Public Policy, and Law, 19*(2), 151-164.

Flowe, H. D. (2012). Do characteristics of faces that convey trustworthiness and dominance underlie perceptions of criminality? *PLoSONE, 7*(6), e37253.

Flowe, H. D. & Humphries, J. E. (2011). An examination of criminal face bias in a random sample of police lineups. *Applied Cognitive Psychology, 25*, 265-273.

Flowe, H. D., Klatt, T., & Colloff, M. F. (2014). Selecting fillers on emotional appearance improves lineup identification accuracy. *Law and Human Behavior, 38*, 509-519.

Freeman, J. B., Pauker, K., & Sanchez, D. T. (2016). A perceptual pathway to bias: Interracial exposure reduces abrupt shifts in real-time race perception that predict mixed-race bias. *Psychological Science, 27*, 502 – 517.

Funk, F., Walker, M., & Todorov, A. (2017). Modelling perceptions of criminality and remorse in faces using a data-driven computational approach. *Cognition & Emotion*, *31*(7), 1431-1443.

Galton, F. (1892). Inquiries into human faculty and its development. London: Macmillan.

Geniole, S. N., Keyes, A. E., Carre, J. M., I& McCormick, C. M. (2014). Fearless dominance mediates the relationship between the facial width-to-height ratio and willingness to cheat. *Personality and Individual Differences, 57,* 59-64.

Goetz, S. M., Shattuck, K. S., Miller, R. M., Campbell, J. A., Lozoya, E., Weisfeld, G. E., & Carre, J. M. (2013). Social status moderates the relationship between facial structure and aggression. *Psychological Science, 24,* 2329-2334.

Goldsmith, W. W. & Blakely, E. J. (2010). *Separate societies: Poverty and inequality in U. S. Cities*. Philadelphia: Temple.

Goldstein, A. G., Chance, J. E., & Gilbert, B. (1984). Facial stereotypes of good guys and bad guys: a replication and extension. *Bulletin of the Psychonomic Society 22,* 549-552.

Goldstein, A. G., Chance, J. E., & Schneller, G. R. (1989). Frequency of eyewitness identification in criminal cases: a survey of prosecutors. Bulletin of the *Psychonomic Society, 27,* 71-74.

Gomez-Valdes, J., Hunemeier, T., Quinto-Sanchez, M., Paschetta, C., Azevedo, S., Gonzalez, M. F.,….Gonzalez-Jose, R. (2013). Lack of support for the association between facial shape and aggression: a reappraisal based on a worldwide population genetics perspective. *PLoS ONE, 8*(1), doi: 10.1371/journal.pone.0052317

Haselhuhn, M. P., Ormiston, M. E., & Wong, E. M. (2015). Men's facial width-to-height ratio predicts aggression: a meta-analysis. *PloS ONE, 10,* e0122637

Hassin, R. & Trope, Y. (2000). Facing faces: Studies on the cognitive aspects of physiognomy. *Journal of Personality and Social Psychology, 778,* 837-852.

Hastie, R. & Park, B. (1986). The relationship between memory and judgment depends on whether the judgment task is memory-based or on-line. *Psychological Review, 93*(3), 258-268.

Hehman, E., Flake, J. K., & Freeman, J. B. (2015). Static and dynamic facial cues differentially affect the consistency of social evaluations. *Personality and Social Psychology Bulletin, 41*, 1123 – 1134.

Hehman, E., Leitner, J. B., Deegan, M. P., & Gaertner, S. L. (2015). Picking teams: when dominant facial structure is preferred. *Journal of Experimental Social Psychology, 59*, 51 – 59.

Hehman, E., Leitner, J. B., Gaertner, S. L. (2013). Enhancing static facial features increases intimidation. *Journal of Experimental Social Psychology, 49*, 747-754.

Hehman, E., Sutherland, C. A., Flake, J. K., & Slepian, M. L. (2017). The unique contributions of perceiver and target characteristics in person perception. *Attitudes and Social Cognition, 113* (4), 513 – 529.

Hugenberg, K., Miller, J., & Claypool, H. M. (2007). Categorization and individuation in the cross-race recognition deficit: Toward a solution to an insidious problem. *Journal of Experimental Social Psychology, 2007*, 334-340.

Hugenberg, K. & Sacco, D. F. (2008). Social categorization and stereotyping: how social categorization biases person perception and face memory. *Social and Personality Psychology Compass*, *2*(2), 1052-1072.

Innocence Project (2018). DNA Exonerations in the United States. Retrieved June 4, 2018. https://www.innocenceproject.org/dna-exonerations-in-the-united-states/

Ito, T. A. & Urland, G. R. (2003). Race and gender on the brain: Electrocortical measures of attention to the race and gender of multiply categorizable individuals. *Journal of Personality and Social Psychology, 85*, 616-626.

Klatt, T., Phelps, M., Maltby, J. J., Smailes, H. L., Ryder, H., Humphries, J. E., & Flowe, H. (2016). Looking bad: Inferring criminality after 100 ms. *Applied Psychology in Criminal Justice, 12*(2), 114-125.

Kleider, H. M., Pezdek, K., Goldinger, S. D., & Kirk, A. (2008). Schema-driven source misattribution errors: Remembering the expected from a witnessed event. *Applied Cognitive Psychology, 22*, 1-20.

Kleider-Offutt, H. M., Knuycky, L. R., Clevinger, A. M., & Capodanno, M. M. (2017). Wrongful convictions and prototypical black features: can a face-type facilitate misidentifications? *Legal and Criminological Psychology, 22*(2), 350-358.

Kline, R. B. (2005). Principles and practice of structural equation modeling (2nd ed.). New York: Guilford Press.

Knutson, B. (1996). Facial expressions of emotion influence interpersonal trait inferences. *Journal of Nonverbal Behavior, 20*, 165-181.

Knuycky, L. R., Kleider, H. M., & Cavrak, S. E. (2014). Line-up misidentifications: when being 'prototypically black' is perceived as criminal. *Applied Cognitive Psychology, 28*(1), 39-46.

Lavater, J. C. (1797). Essays on Physiognomy: Calculated to extend the knowledge and the love of mankind. London.

Lenz, G. S. & Lawson, C. (2011). Looking the part: Television leads less informed citizens to vote based on candidates' appearance. *American Journal of Political Science, 55*, 574-589.

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

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

Lindsay, R. C. & Wells, G. L. (1980). What price justice? Exploring the relationship of lineup fairness to identification accuracy. *Law and Human Behavior, 4*, 303-313.

Lindsay, R. C., Wells, G. L., & Rumpel, C. M. (1981). Can people detect eyewitness identification accuracy within and across situations? *Journal of Applied Psychology, 66*, 79-89.

Lippman, W. (1922). *Public Opinion*. New York: Harcourt, Brace.

Litterer, O. F. (1933). Stereotypes. *Journal of Social Psychology 4*, 59- 69.

Loftus, E. F. (2005). Planting misinformation in the human mind: a 30-year investigation of the malleability of memory. *Learning & Memory, 12*(4), 361-366.

Lombroso, C. (2006). *Criminal man*. Durham, NC: Duke University Press.

Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago Face Database: A free stimulus set of faces and norming data. *Behavior Research Methods, 47*, 1122-1135.

Maas, C. J. M., & Hox, J. J. (2005). Sufficient sample sizes for multilevel modeling. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences, 1*(3), 86-92. doi:10.1027/1614-2241.1.3.86

MacLin, M. K., & Herrera, V. (2006). The criminal stereotype. *North American Journal of Psychology, 8,* 197-208.

MacLin, O. H. & MacLin, K. M. (2004). The effect of criminality on face attractiveness, typicality, memorability, and recognition. *North American Journal of Psychology, 6,* 145-154.

MacLin, M. K., Malpass, R. S., & Herrera, V. T. (2001). The effect of criminality on mock witness identification. Paper presented at the Annual Meeting of the Society for Applied Research in Memory and Cognition, Kingston, Ontario, Canada.

Macrae, C. N. (1989). The good, the bad, and the ugly: Facial stereotyping and juridic judgments. *Police Journal, 62,* 195-199.

Macrae, C. N., Schloerscheidt, A. M., Bodenhausen, G. V., & Milne, A. B. (2002).Creating memory illusions: Expectancy-based processing and the generation of false memories. *Memory, 10*(1), 63-80.

Macrae, C. N. & Shepherd, J. W. (1989). Do criminal stereotypes mediate juridic judgements? *British Journal of Social Psychology, 28,* 189-191.

Magnussen, S., Melinder, A., Stridbeck, U., & Raja, A. Q. (2010). Beliefs about factors affecting the reliability of eyewitness testimony: a comparison of judges, jurors, and the general public. *Applied Cognitive Psychology, 24*(1), 122-133.

Malatesta, C. Z., Fiore, M. J., & Messina, J. J. (1987). Affect, personality, and facial expressive characteristics of older people. *Psychology and Aging, 2*, 64-69.

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

Malpass, R. S. Lavigueur, H., & Weldon, D. E. (1973). Verbal and visual training in face

recognition. *Attention, Perception, & Psychophysics, 14*, 285-292.

Marcon, J. L., Meissner, C. A., Frueh, M., Susa, K. J., & MacLin, O. H. (2010). Perceptual identification and the cross-race effect. *Visual Cognition, 18*(5), 767-779.

Martire, K. A. & Kemp, R. I. (2009). The impact of eyewitness expert evidence and judicial instruction on juror ability to evaluate eyewitness testimony. *Law and Human Behavior, 33*(3), 225-236.

Matsumoto, D. & Hwang, H. C. (2018). Microexpressions differentiate truths from lies about future malicious intent. *Frontiers in Psychology, 9*, 1 – 11. doi: 10.3389/fpsyg.2018.02545

McQuiston, D. E. & Malpass, R. S. (2002). Validity of the mockwitness paradigm: Testing the assumptions. *Law and Human Behavior, 26*, 439-453.

Mehta, P. D. (2013). n-level structural equation modeling. In Y. Petscher, C. Schatschneider, & D. L. Compton (Eds.), *Applied quantitative analysis in the social sciences* (pp. 329-362). New York: Routledge.

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

Montepare, J. M. & Dobish, H. (2003). The contribution of emotion perceptions and their overgeneralizations to trait impressions. *Journal of Nonverbal Behavior, 27*, 237-254.

Mueller, U. & Mazur, A. (1996). Facial dominance of West Point cadets as a predictor of later military rank. *Social Forces, 74*, 823-850.

Munsterberg, H. (1908). *On the witness stand: essays on psychology and crime*. Garden City, NY: Clark, Boardman.

National Academy of Sciences Report on Assessing Eyewitness Identification (2014). *Identifying the culprit: Assessing eyewitness identification*. Washington, DC: National Academies Press.

Olivola, C. Y., Funk, F., & Todorov, A. (2014). Social attributions from faces bias human choices. *Trends in Cognitive Sciences, 18,* 566-570.

Oosterhof, N. N. & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences of the United States of America, 105,* 11087-11092.

Pezdek, K., Blandon-Gitlin, I., & Moore, C. (2003). Children's face recognition memory: More evidence for the CRE. *Journal of Applied Psychology, 88,* 760-763.

Plant, E. A. & Devine, P. G. (2003). The antecedents and implications of interracial anxiety. *Personality and Social Psychology Bulletin, 29*(6), 790-801.

Pollack, S. D. & Kistler, D. J. (2002). Early experience is associated with the development of categorical representations for facial expressions of emotion. *Proceedings of the National Academy of Sciences of the United States, 99*(13), 9072-9076.

Porter, S., England, L., Juodis, M., ten Brinke, L., & Wilson, K. (2008). Is the face the window to the soul?: Investigation of the accuracy of intuitive judgments of the trustworthiness of human faces. *Canadian Journal of Behavioural Science, 40,* 171 – 177.

Porter, S. & ten Brinke, L. (2008). Reading between the lies: Identifying concealed and falsified emotions in universal facial expressions. *Psychological Science, 19* (5), 508 -514.

Poutvaara, P., Jordahl, H., & Berggren, N. (2009). Faces of politicians: Babyfacedness predicts inferred competence but not electoral success. *Journal of Experimental Social Psychology, 45,* 1132-1135.

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Reed, J. P. & Reed, R.S. (1973). Status, images and consequence: Once a criminal always a criminal. *Sociology and Social Research, 57*(4), 460-472.

Rezlescu, C., Duchaine, B., Olivola, C. Y., & Chater, N. (2012). Unfakeable facial configurations affect strategic choices in trust games with or without information about past behavior. *PLoSONE, 7*(3), e34293.

Rule, N. O., Krendl, A. C., Ivcevic, Z., & Ambady, N. (2013). Accuracy and consensus in judgments of trustworthiness from faces: Behavioral and neural correlates. *Journal of Personality and Social Psychology, 104,* 409 – 426.

Said, C. P., Haxby, J. V., & Todorov, A. (2011). Brain systems for assessing the affective value of faces. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 366,* 1660-1670.

Sangrigoli, S., Pallier, C., Argenti, A. M., Ventureyra, V. A., & de Schonen, S. (2005). Reversibility of the other-race effect in face recognition during childhood. *Psychological Science, 16,* 440-444.

Schneider, T. M., Hecht, H., & Carbon, C. C. (2012). Judging body-weight from faces: the height-weight illusion. *Perception, 41,* 121-124.

Secord, P. F. & Backman, C. W. (1964). *Social Psychology*. New York: McGraw-Hill.

Secord, P. F. & Bevan, W. (1956) Personalities in faces: a cross-cultural comparison of impressions of physiognomy and personality in faces. *Journal of Social Psychology 43,* 283-288.

Secord, P. F., Bevan, W., & Dukes, W. F. (1953). Occupational and physiognomic stereotypes in the perception of photographs. *Journal of Social Psychology, 37,* 261-270.

Secord, P. F., Dukes, W. F., & Bevan, W. (1954). Personalities in faces: an experiment in social perceiving. *Genetic Psychology Monographs, 49*, 231-279.

Shoemaker, D. J., South, D. R., & Lowe, J. (1973). Facial stereotypes of deviants and judgments of guilt or innocence. *Social Forces, 51,* 427-433.

Smith, E. R., Miller, D. A., Maitner, A. T., & Crump, S. A., Garcia-Marques, T., & Mackie, D. M. (2006). Familiarity can increase stereotyping. *Journal of Experimental Social Psychology, 42*, 471-478.

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

Stangor, C. & McMillan, D. (1992). Memory for expectancy-congruent and expectancy-incongruent information: a review of the social and social developmental literatures. *Psychological Bulletin, 11*(1), 42-61.

Steblay, N. K., Dysart, J., Fulero, S., & Lindsay. R. C. (2003). Eyewitness accuracy rates in police showup and lineup presentations; a meta-analytic comparison. *Law and Human Behavior, 27*(5), 523-540.

Stirrat, M. & Perrett, D. I. (2010). Valid facial cues to cooperation and trust: Male facial width and trustworthiness. *Psychological Science, 21*, 349-354.

Stoker, J. I., Gerretsen, H., & Spreeuwers, L. J. (2016). The facial appearance of CEOs: Faces signal selection but not performance. *PLoSONE, 11*(7), e0159950.

Sussman, A. B., Petkova, K., & Todorov, A. (2013). Competence ratings in US predict presidential election outcomes in Bulgaria. *Journal of Experimental Social Psychology, 49*, 771-775.

Todorov, A. (2017). *Face value: the irresistible influence of first impressions*. Princeton, NJ: Princeton University Press.

Todorov, A., Mandisodza, A. N., Goren, A., & Hall, C. C. (2005). Inferences of competence from faces predict election outcomes. *Science, 308*, 1623-1626.

Todorov, A., Olivola, C. Y., Dotsch, R., & Mende-Siedlecki, P. (2015). Social attributions from faces: Determinants, consequences, accuracy, and functional significance. *Annual Review of Psychology, 66*, 519-545.

Todorov, A.,Said, C. P., Engell, A. D., & Oosterhof, N. N. (2008). Understanding evaluation of faces on social dimensions. *Trends in Cognitive Sciences, 12*, 455-460.

United States v. Wade, 388 U.S. 230, 288 (1967).

Valla, J. M., Ceci, S. J., & Williams, W. M. (2011). The accuracy of inferences about criminality based on facial appearance. *Journal of Social, Evolutionary, and Cultural Psychology, 5*, 66 – 91.

Wells, G. L. (1984). The psychology of lineup identifications. *Journal of Applied Social Psychology, 14*, 89-103.

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

Wells, G. L., Ferguson, T. J., & Lindsay, R. C. (1981). The tractability of eyewitness confidence and its implication for triers of fact. *Journal of Applied Psychology, 66*, 688-696.

Wells, G. L., Lindsay, R. C., & Ferguson, T. J. (1979). Accuracy, confidence, and juror perceptions in eyewitness identification. *Journal of Applied Psychology, 64*, 440-448.

Wells, G. L., Memon, A., & Penrod. S. D. (2006). Eyewitness evidence: Improving its probative value. *Psychological Science in the Public Interest, 7*(2), 45-75.

Wells, G. L. & Olson, E. A. (2003). Eyewitness testimony. *Annual Review of Psychology, 54*, 277-295.

Wells, G. L., Rydell, S. M., & Seelau, E. P. (1993). On the selection of distractors for eyewitness lineups. *Journal of Applied Psychology, 78*, 835-844.

Willis, J. & Todorov, A. (2006). First impressions: Making up your mind after 100 ms exposure to a face. *Psychological Science, 17,* 592-598.

Wilson, J. P., Hugenberg, K., Bernstein, M. J. (2013). The cross-race effect and eyewitness identification: how to improve recognition and reduce decision errors in eyewitness situations. *Social Issues and Policy Review, 7*(1), 83-113.

Wilson, J. P. & Rule, N. O. (2016). Hypothetical sentencing decisions are associated with actual capital punishment outcomes: the role of facial trustworthiness. *Social Psychological and Personality Science, 7*, 331-338.

Wilson, J. P. & Rule, N. O. (2015). Facial trustworthiness predicts extreme criminal sentencing outcomes. *Psychological Science, 26,* 1325-1331.

Wixted, J. T. (2004). The psychology and neuroscience of forgetting. *Annual Review of Psychology, 55*, 235-269.

Wixted, J. T., Mickes, L., Dunn, J. C., Clark, S. E., & Wells, W. (2016). Estimating the reliability of eyewitness identifications from police lineups. *Proceedings of the National Academy of Sciences of the United States, 113*(2), 304-309.

Yarmey, D. A. (1993). Stereotypes and recognition memory for faces and voices of good guys and bad guys. *Applied Cognitive Psychology, 7*, 419-431.

Young, S. G. & Hugenberg, K. (2012). Individuation motivation and face experience can operate jointly to produce the Own Race Bias. *Social Psychological and Personality Science, 3*, 80-87.

Young, S. G., Hugenberg, K. Bernstein, M. J., & Sacco, D. F. (2012). Perception and motivation in face recognition: a critical review of theories of the cross-race effect. *Personality and Social Psychology Review, 16*, 116-142.

Zebrowitz, L. A. & McDonald, S. M. (1991). The impact of litigants' babyfacedness and attractiveness on adjudication in small claims courts. *Law and Human Behavior, 15*, 603-623.

Table 1

Abbreviation Key and Summary of Variables

Face Variables	Description	Scale used
Match	Whether the race of participant (ALL African American participants) matched the race of each picture they rated (AA or not)	0 = no match 1 = match
Feature	Number of "criminal" facial features each face has as rated by 2 independent raters. Ratings were highly correlated, so they were averaged.	1 – 17
Perp	Refers to the status of each rated face. Specifically whether or not the rated face was that of an actual perpetrator	0 = not a perpetrator 1 = a perpetrator
Exon	Refers to the status of each rated face. Specifically whether or not the rated face was that of an exonerate	0 = not an exonerate 1 = an exonerate
Rater Variables		
Victim	Refers to whether or not the subject has been a victim of a crime before	0 = no 1 = yes
AnxVic	Refers to how anxious the subject is about becoming a victim of a violent crime in the future	5 point scale (1 = never; 5 = always)
AnxInt	Refers to the subject's level of anxiety about interracial contact	Responses to 31 items were averaged to create an overall interracial anxiety score, with higher scores indicating more interracial anxiety.
Outcome Variables		
Crim	Perceived criminality of each rated face	1 – 7 scale (1 = not at all; 7 = very much so)
Dom	Perceived dominance of each rated face	1 – 7 scale (1 = not at all; 7 = very much so)
Trust	Perceived trustworthiness of each rated face	1 – 7 scale (1 = not at all; 7 = very much so)
Other		
FaceID	Refers to the specific face rated	String variable
RaterID	Subject number	

Table 2

Correlation of Outcome and Predictor Variables

	1	2	3	4	5	6	7	8	9
1Crim	--								
1Trust	-.46***	--							
3Dom	.29***	.01	--						
4Perp	.14***	-.11***	.13***	--					
5Exon	.15***	-.05***	.11***	-.34***	--				
6Feat	.16***	-.09***	.15***	.24***	.29***	--			
7Match	-.16***	.19***	.12***	.14***	.08***	.14***	--		
8Vic	.04**	-.05**	-.00	-.00	-.00	.01	-.00	--	
9AnxVic	-.04**	.07***	.01	-.01	-.01	-.00	-.00	-.05***	--
10AnxInt	.03	-.01***	.04**	.02	-.01	-.01	-.01	-.02	.01

* p < .05, ** p < .01, *** p < .001.

Table 3

Means and Standard Deviations for Outcome and Predictor Variables

Variable	Mean	SD
Criminality	3.65	1.90
Trust	2.62	1.53
Dominance	3.62	1.76
Feature	4.92	1.55
Match	.47	.50
Victim	.82	.39
AnxVic	2.82	1.22
AnxInt	3.48	1.07

Table 4

Mean Trait Ratings and Number of Criminal Features by Conviction Status

Face Type	Criminality	Dominance	Trustworthiness	Features
Perpetrator	4.17 (1.86)	4.06 (1.79)	2.28 (1.39)	5.62 (1.44)
Exonerate	4.09 (1.95)	3.89 (1.75)	2.49 (1.51)	5.61 (1.51)
Filler	3.17 (1.76)	3.27 (1.68)	2.84 (1.57)	4.20 (1.28)

Note: SDs are shown in parentheses.

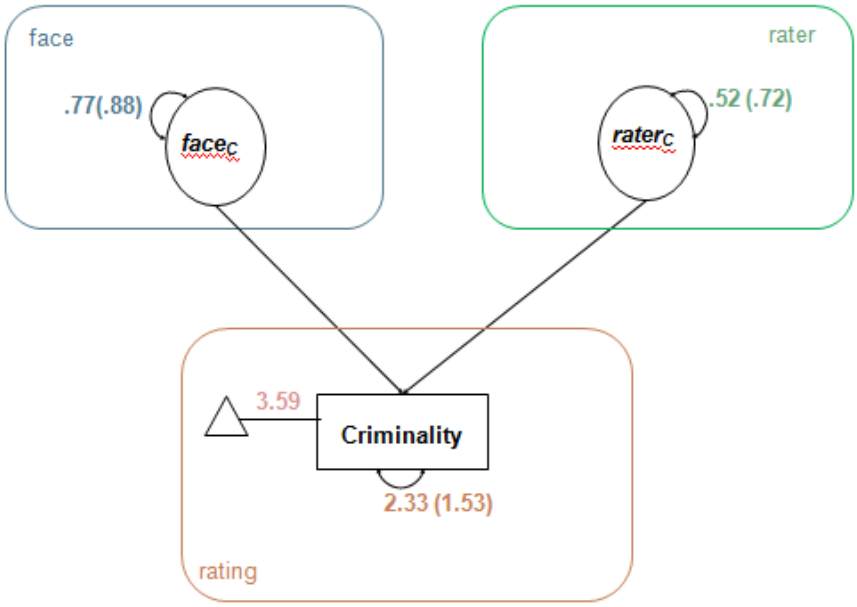

Figure 1

Baseline Model of Criminality Ratings

Note: This is a multilevel SEM diagram (Mehta, 2013) of the cross-classified regression (Baayen, Davidson, Bates, 2008).

Note: These are unstandardized estimates.

Note: Standard deviations appear in parentheses.

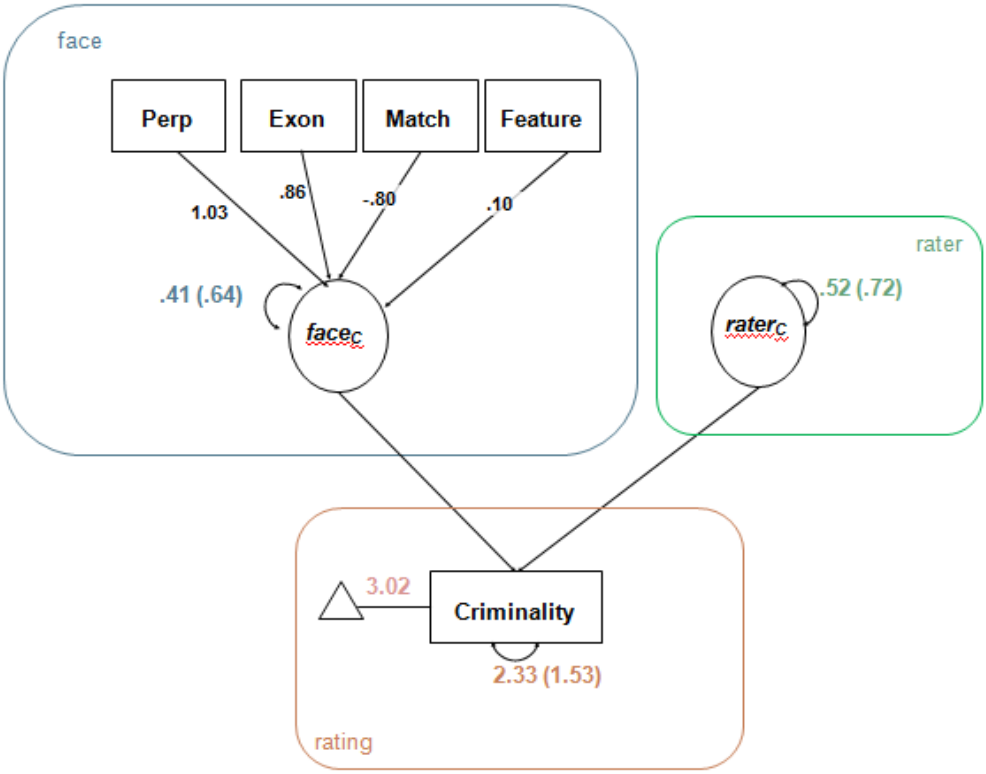

Figure 2

Basic Model of Criminality Ratings

Note: These are unstandardized estimates.

Note: Standard deviations appear in parentheses.

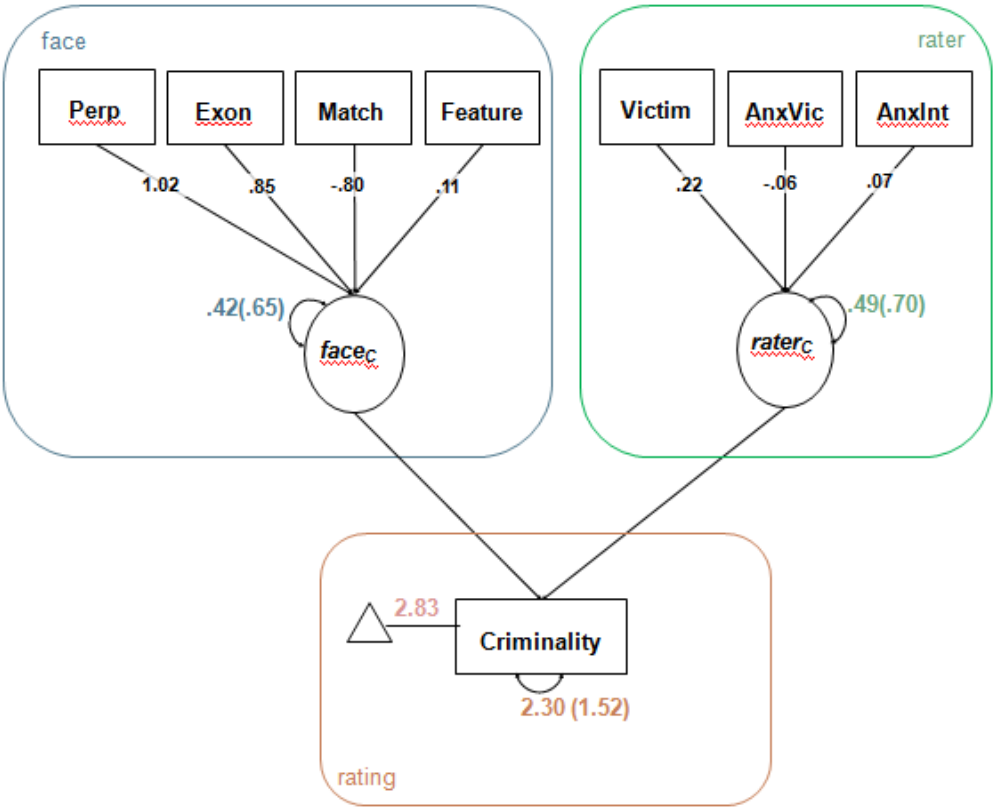

Figure 3

Full Model of Criminality Ratings

Note: These are unstandardized estimates.

Note: Standard deviations appear in parentheses.

Table 5

E*stimates for Baseline, Basic, and Full Models of Criminality*

Criminality		Baseline			Basic Model			Full Model		
Fixed Effects	Variable	Estimate	Std. Error	t value	Estimate	Std. Error	t value	Estimate	Std. Error	t value
	(Intercept)	3.59	0.10	37.69	3.02	0.22	13.79	2.83	0.25	11.29
Face Variables	Match				-0.80	0.13	-6.28	-0.80	0.13	-6.21
	Features				0.10	0.04	2.33	0.11	0.05	2.40
	Exonerate				0.86	0.16	5.28	0.85	0.16	5.20
	Perpetrator				1.03	0.18	5.73	1.02	0.18	5.56
Rater Variables	Victim							0.22	0.14	1.55
	Anx-Victim							-0.06	0.04	-1.40
	Anx-Interracial							0.07	0.05	1.31
Random Effects		Variance	Std.Dev.		Variance	Std.Dev.		Variance	Std.Dev.	
	RaterID	0.52	0.72		0.52	0.72		0.49	0.70	
	FaceID	0.77	0.88		0.41	0.64		0.42	0.65	
	Residual	2.33	1.53		2.33	1.53		2.30	1.52	

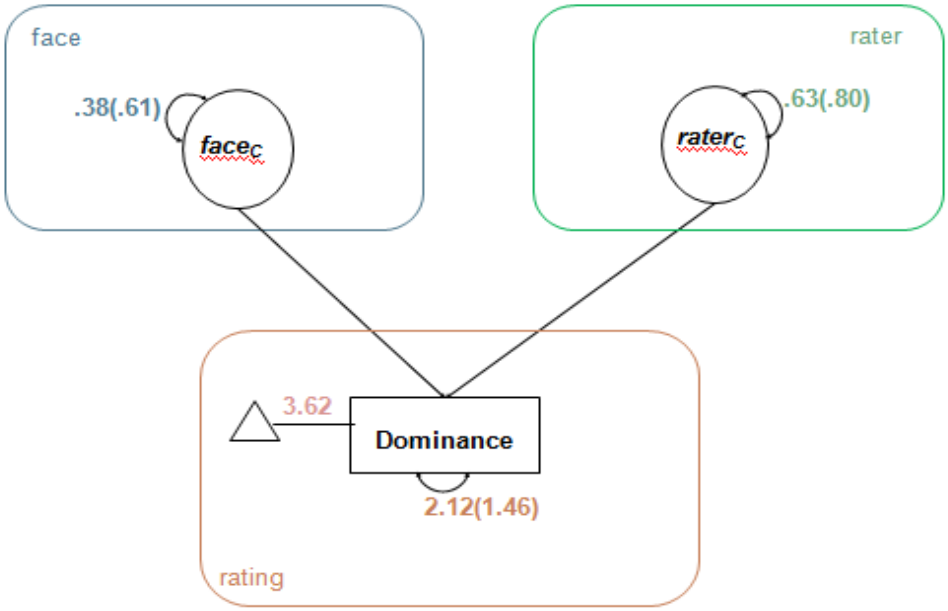

Figure 4

Baseline Model of Dominance Ratings

Note: These are unstandardized estimates.

Note: Standard deviations appear in parentheses.

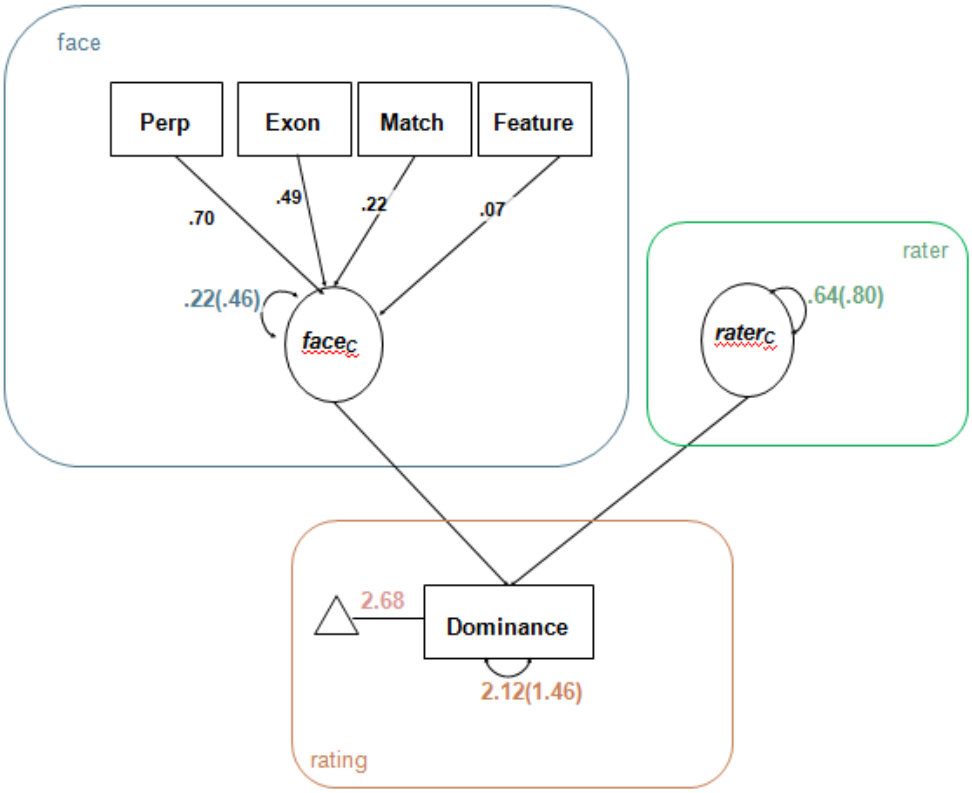

Figure 5

Basic Model of Dominance Ratings

Note: These are unstandardized estimates.

Note: Standard deviations appear in parentheses.

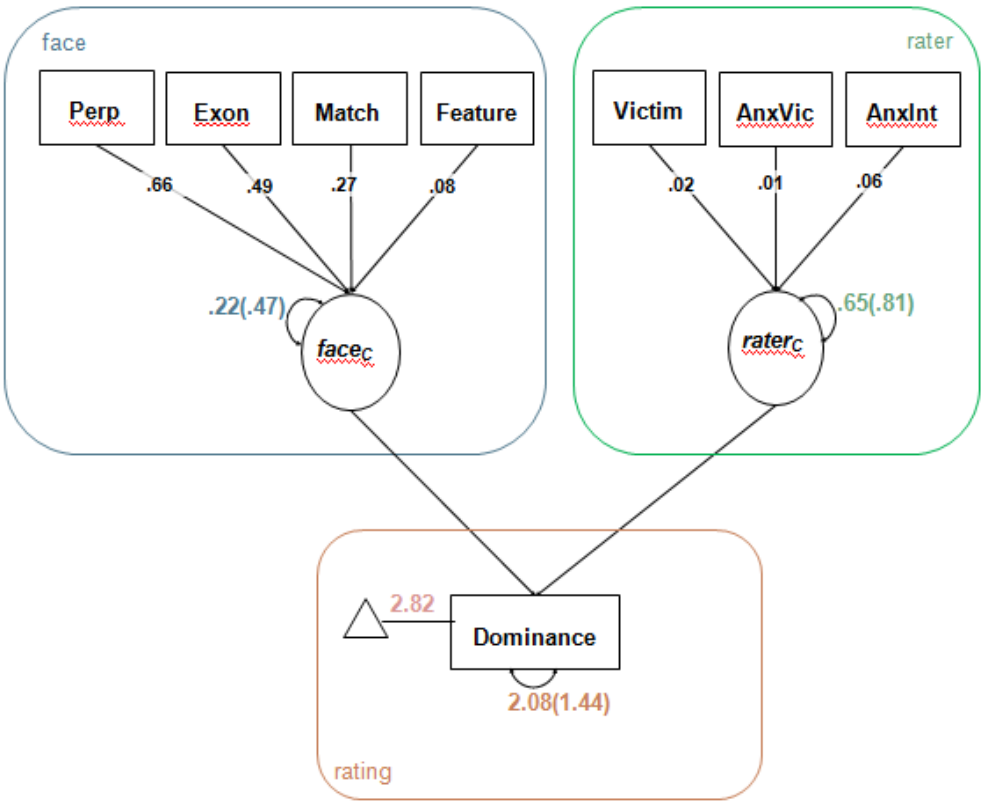

Figure 6

Full Model of Dominance Ratings

Note: These are unstandardized estimates.

Note: Standard deviations appear in parentheses.

Table 6

Estimates for Baseline, Basic, and Full Models of Dominance

Dominance		Baseline			Basic Model			Full Model		
Fixed Effects	Variable	Estimate	Std. Error	t value	Estimate	Std. Error	t value	Estimate	Std. Error	t value
	(Intercept)	3.62	0.08	45.21	2.86	0.17	16.81	2.82	0.22	12.99
Face Variables	Match				0.22	0.10	2.33	0.27	0.10	2.79
	Features				0.07	0.03	2.19	0.08	0.03	2.32
	Exonerate				0.49	0.12	3.99	0.49	0.12	3.97
	Perpetrator				0.70	0.14	5.15	0.66	0.14	4.83
Rater Variables	Victim							0.02	0.16	0.14
	Anx-Victim							0.01	0.05	0.29
	Anx-Interracial							0.06	0.06	1.13
Random Effects		Variance	Std.Dev.		Variance	Std.Dev.		Variance	Std.Dev.	
	RaterID	0.63	0.80		0.64	0.80		0.65	0.81	
	FaceID	0.38	0.61		0.22	0.46		0.22	0.47	
	Residual	2.12	1.46		2.12	1.46		2.08	1.44	

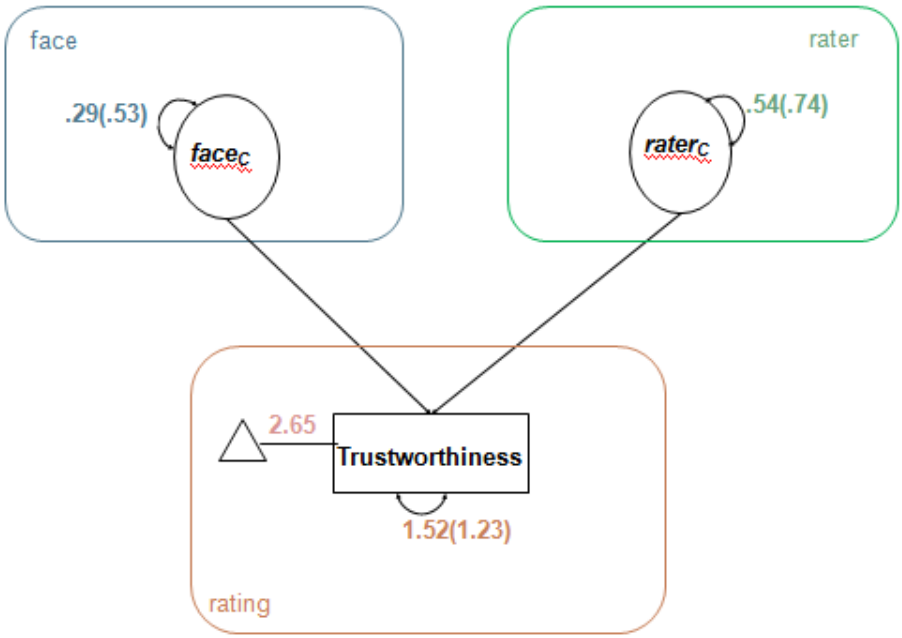

Figure 7

Baseline Model of Trustworthiness Ratings

Note: These are unstandardized estimates.

Note: Standard deviations appear in parentheses.

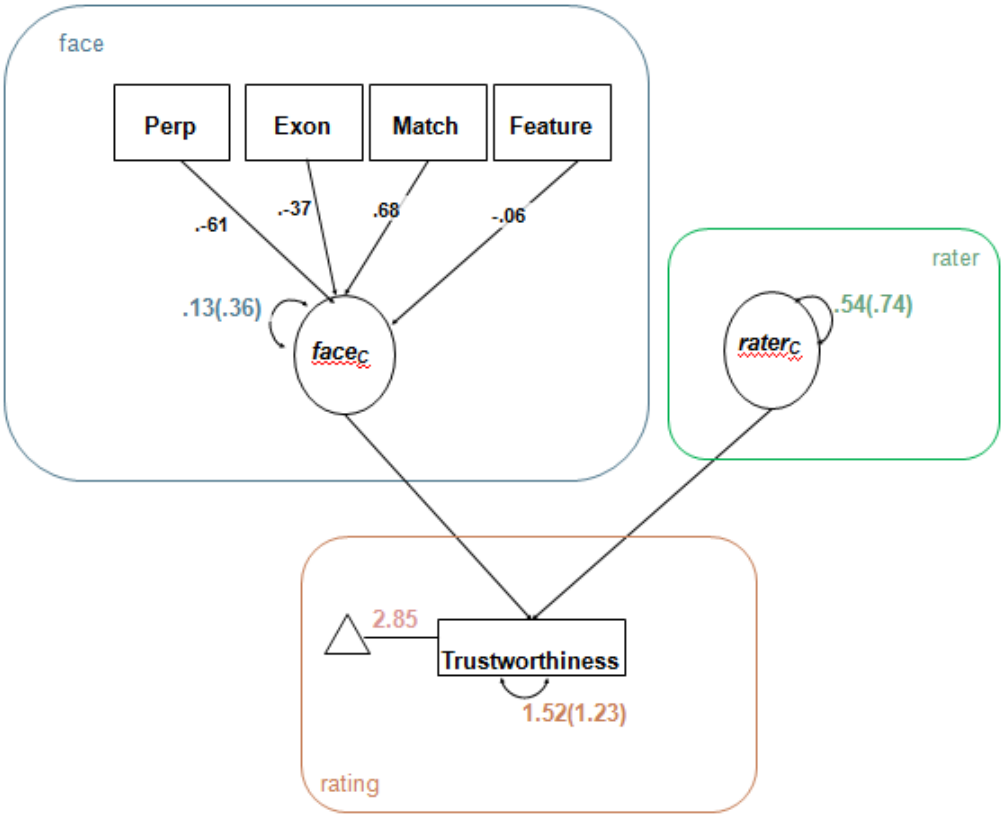

Figure 8

Basic Model of Trustworthiness Ratings

Note: These are unstandardized estimates.

Note: Standard deviations appear in parentheses.

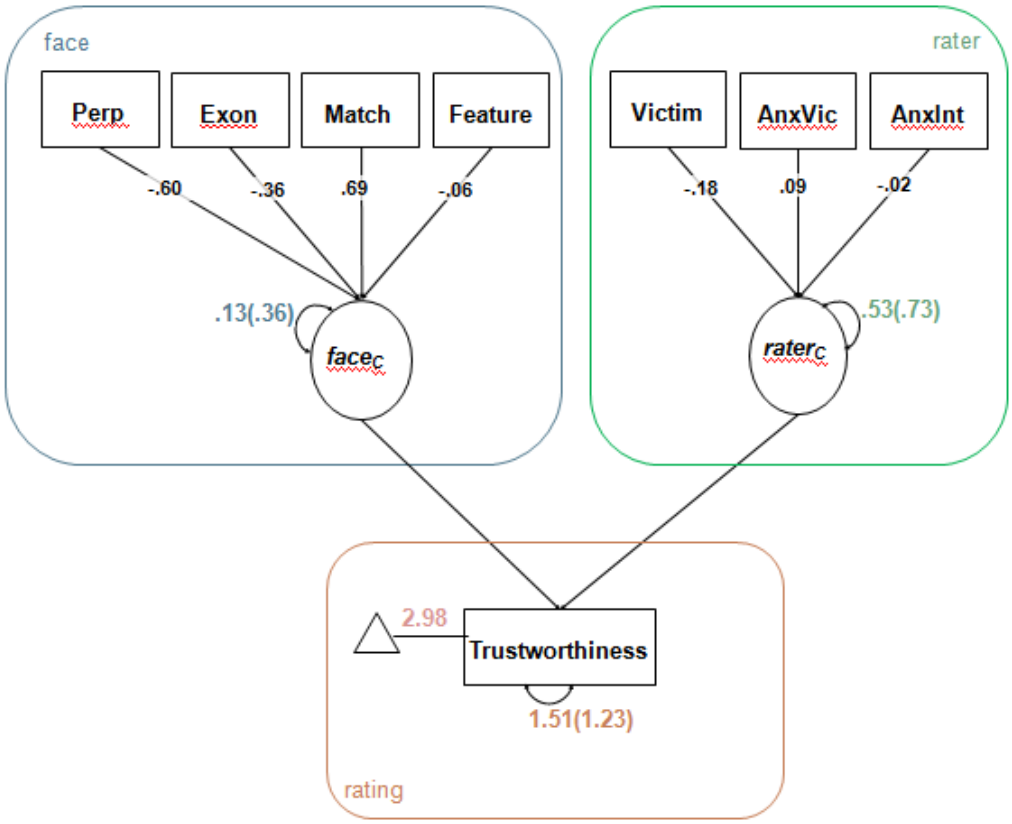

Figure 9

Full Model of Trustworthiness Ratings

Note: These are unstandardized estimates.

Note: Standard deviations appear in parentheses.

Table 7

Estimates for Baseline, Basic, and Full Models of Trustworthiness

Trust		Baseline			Basic Model			Full Model		
Fixed Effects	Variable	Estimate	Std. Error	t value	Estimate	Std. Error	t value	Estimate	Std. Error	t value
	(Intercept)	2.65	0.07	37.05	2.85	0.14	20.66	2.98	0.18	16.30
Face Variables	Match				0.68	0.08	8.86	0.69	0.08	9.01
	Features				-0.06	0.03	-2.28	-0.06	0.03	-2.19
	Exonerate				-0.37	0.10	-3.75	-0.36	0.10	-3.65
	Perpetrator				-0.61	0.11	-5.60	-0.60	0.11	-5.47
Rater Variables	Victim							-0.18	0.14	-1.27
	Anx-Victim							0.09	0.04	1.93
	Anx-Interracial							-0.02	0.05	-0.34
Random Effects		Variance	Std.Dev.		Variance	Std.Dev.		Variance	Std.Dev.	
	RaterID	0.54	0.74		0.54	0.74		0.53	0.73	
	FaceID	0.29	0.53		0.13	0.36		0.13	0.36	
	Residual	1.52	1.23		1.52	1.23		1.51	1.23	

APPENDICES

Appendix A

Appendix A.1 Criminal Facial Feature Inventory

Male	Beady eyes
Shaggy hair	Small eyes
Long hair	Drooping eyes
Dark hair	Prominent chin
Scars	Protruding jaw
Pockmarks	Lowered eyebrows
Tattoos	Large ears
Facial hair	Flat nose
Dark pigmentation	

Appendix A.2 Survey of Community Violence Exposure

1. How many times have you yourself actually been the victim of any type of violence such as being beaten, mugged, raped, or attacked with a weapon? (circle only one)

(a) never (d) 3 or 4 times (g) at least once a month

(b) 1 time (e) 5 or 6 times (h) at least once a week

(c) 2 times (f) 7 or 8 times (i) almost every day

2. Who committed the violent act?

(a) someone of your own gender (b) someone of a different gender

3. Who committed the violent act?

(a) someone of your own race (b) someone of a different race

4. When was the last time this happened? (circle only one)

(a) about a week ago (d) about 6 months ago (g) between 1 and 2 years ago

(b) about a month ago (e) about 9 months ago (h) between 3 and 5 years ago

(c) about 3 months ago (f) about a year ago (i) more than 5 years ago

5. How many times have you seen someone else being victimized by some form of violence such as being beaten, mugged, raped, or attacked with a weapon?
(circle only one)

(a) never (d) 3 or 4 times (g) at least once a month

(b) 1 time (e) 5 or 6 times (h) at least once a week

(c) 2 times (f) 7 or 8 times (i) almost every day

6. Who committed the violent act?

(a) someone of your own gender (b) someone of a different gender

7. Who committed the violent act?

(a) someone of your own race (b) someone of a different race

8. When was the last time you saw this happen? (circle only one)

(a) about a week ago (d) about 6 months ago (g) between 1 and 2 years ago

(b) about a month ago (e) about 9 months ago (h) between 3 and 5 years ago

(c) about 3 months ago (f) about a year ago (i) more than 5 years ago

Appendix A.3 Anxiety about Violence Questionnaire

1. Do you ever carry a gun?

Never Occasionally Sometimes Often Always

2. Do you feel anxious if you carry a gun?

Never Occasionally Sometimes Often Always I never carry a gun

3. How concerned are you that you will be a victim of a violent crime?

Not at all A little bit Somewhat Very Always

4. How concerned are you that you will be a victim of a petty crime?

Not at all Occasionally Somewhat Very Always

5. Are you concerned about your personal safety with regard to terrorism?

Not at all Occasionally Somewhat Very Always

6. Do you ever feel threatened while on line?

Not at all Occasionally Somewhat Very Always

7. How often have you threatened others on line?

Never Rarely Occasionally Often Very Often

8. When you are interacting on line, how self-assured do you feel?

Not at all A little Somewhat Very much Extremely

9. When you are interacting face-to-face, how self-assured do you feel?

Not at all A little Somewhat Very much Extremely

10. How concerned are you that a gun may be used on you?

Not at all A little Somewhat Very much Extremely

11. How threatened do you feel when a police officer comes your direction in a low-crime neighborhood?

Not at all A little Somewhat Very much Extremely

12. How threatened do you feel when a police officer comes your direction in a high-crime neighborhood?

Not at all A little Somewhat Very much Extremely

13. How safe do you feel in neighborhoods where you are in the racial majority?

Not at all A little Somewhat Very much Extremely

14. How safe do you feel in neighborhoods where you are in the racial minority?

Not at all A little Somewhat Very much Extremely

15. How safe do you feel in the presence of police officers?

Not at all A little Somewhat Very much Extremely

16. How safe do you feel in the presence of police officers who are of your own race?

Not at all A little Somewhat Very much Extremely

17. How safe do you feel in the presence of police officers who are of a different race from yours?

Not at all A little Somewhat Very much Extremely

How concerned are you for your own safety if you see a person with a concealed gun . . . ..

18. In a store?

Not at all A little Somewhat Very much Extremely

19. Walking past you on the street?

Not at all A little Somewhat Very much Extremely

20. In an office?

Not at all A little Somewhat Very much Extremely

21. In a retail store where you work?

Not at all A little Somewhat Very much Extremely

Appendix A.4 Interracial Anxiety Scale

All items were rated on the following scale:

1	2	3	4	5	6	7
Strongly Disagree						Strongly Agree

1. In the past, I have interacted with people of other races in many areas of my life (e.g., school, friends, work, clubs).

2. The neighborhood(s) I grew up in had mostly people of my own race.

3. The high school I attended had mostly students of my own race.

4. In the past, I have rarely interacted with people of different races than my own.

5. In the past, my experiences with people from other races has been pleasant.

6. Over the course of my life, I have had many friends that were a different race than mine.

7. I have had many positive experiences with people from different races than my own.

8. I am confident that stereotypes don't affect how I interact with people of different races.

9. Even if we hadn't met before, a person of a different race would expect me to be prejudiced.

10. When interacting with a person of a different race than mine, he or she would see me as prejudiced no matter what I did.

11. When interacting with a person of a different race than mine, I would be unsure how to act in order to show him or her that I am not prejudiced.

12. Sometimes stereotypes come to my mind when interacting with a person of a different race, even when I wish they wouldn't.

13. If I were interacting with a person of another race, regardless of my behavior he or she would interpret my behavior as prejudiced.

14. When interacting with a person of another race, I would imagine that he or she would be watching my behavior closely for prejudice.

15. People of non-White races do not look for prejudice in White people's behavior.

16. I am confident that I can respond without prejudice when interacting with a person of another race.

17. Sometimes people of non-White races view normal behavior of Whites as prejudiced.

18. I would feel awkward when interacting with a person of a different race than mine.

19. I would feel uncomfortable when interacting with a person of a different race than mine.

20. When interacting with a person of another race, I would feel relaxed.

21. When interacting with a person of another race, I would feel nervous.

22. If I had a choice, I would rather not interact with people of different races than my own.

23. If I can avoid interacting with people of different races, I do.

24. I like interacting with people of different races than mine.

25. I would look forward to interacting with people of different races.

26. I would want to avoid interacting with people of different races.

27. I would find interacting with a person of a different race annoying.

28. I would be angry if I had to interact with a person of a different race than mine.

29. I would find interacting with a person of another race frustrating.

30. I would feel hostile when interacting with a person of another race.

31. Interacting with a person of another race would be irritating.