CONVICTABLE FACES: ATTRIBUTIONS OF FUTURE CRIMINALITY FROM FACIAL

APPEARANCE

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Caisa Elizabeth Royer

December 2018

CONVICTABLE FACES: ATTRIBUTIONS OF FUTURE CRIMINALITY FROM FACIAL
APPEARANCE

Caisa Elizabeth Royer, J.D., Ph.D.

Cornell University 2018

Can participants accurately determine whether someone will later become a criminal based only on the person's high school yearbook photo? This project builds on previous research which has found participants are capable of accurately and reliably assessing personality characteristics—like trustworthiness and dominance— based only on a photograph. This paper discusses a series of studies which examine whether participants are also capable of making accurate predictions of criminality by utilizing high school yearbook photographs of men with later criminal records. In Study 1, participants were able to make accurate predictions of future criminality from high school yearbook photographs. In Study 2, the results from the previous study were replicated and confidence in criminality attributions was found to predict accuracy. In Study 3, participants were less accurate when judging photographs of Black students compared to White students, suggesting cross-race bias. Altogether, these studies demonstrated that participants have accurate stereotypes about what a person with a criminal record looks like. These stereotypes may create a self-fulfilling prophecy in which people who look criminal are treated like criminals and thus end up with criminal records. This theory was tested in Study 4 in which participants were asked to judge guilt based on mugshots of exonerated men and true criminals. Overall, this serious of studies demonstrated that participants can make accurate and consistent predictions of future criminality based only on facial appearance.

BIOGRAPHICAL SKETCH

Caisa Elizabeth Royer is a graduate from the dual J.D./Ph.D. program in Human Development, Psychology, and Law at Cornell University. She graduated with her J.D. from Cornell Law School in May 2017 and completed all doctoral requirements from the Department of Human Development in December 2018. She also holds a B.S. from Iowa State University with a double major in psychology and criminal justice. Her research focuses on a variety of issues at the intersection of psychology and law, including perceptions of criminality, jury decision making, capital punishment, and empathy in the courtroom.

Dedicated to Jeff Valla

ACKNOWLEDGMENTS

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

Opportunities for discretion pervade the criminal justice system: witnesses to crime choose whether to call the police, police choose whether to issue a warning or make an arrest, prosecutors choose what charges (if any) to file and what deals to offer, and judges often decide both verdict and sentence. At all of those moments of discretion, characteristics of the alleged perpetrator can influence decision making. Among those influential characteristics is whether the person looks like a "criminal" or not. In fact, previous research has shown that people can make accurate, consistent, and reliable judgments about another person's criminality after seeing the person's face for only a short period of time (Flowe, 2012; Klatt et al., 2016; Valla et al., 2011). These judgments influence the behavior of others, such as when making an identification from a lineup (Flowe & Humphries, 2011) or during sentencing (Wilson & Rule, 2015).

The judgment of personality based on appearance is not limited to criminality. Physiognomy, the science of making attributions of character based on faces, has existed for most of human history (Oosterhof & Todorov, 2008; Suzuki et al., 2017). People also make accurate, consistent, and reliable judgments about a variety of other personality characteristics, including trustworthiness (Chang et al., 2010; Porter et al., 2008; Rule et al., 2013), aggression (Boshyan et al., 2014; Carre & McCormick, 2008), Big Five traits (Hu et al., 2017; Qin et al., 2016), and political orientation (Carpinella & Johnson, 2013; Lenz & Lawson, 2011; Olivola et al., 2012; Olivola & Todorov, 2010). These consistent and reliable judgments can be made after only 100 milliseconds of exposure to someone (Klatt & Phelps, 2016, Todorov & Porter, 2014; Willis & Todorov, 2006). Furthermore, facially-based attributions of character have real world

consequences, as seen in both politics (Olivola et al., 2012) and business (Olivola et al., 2014; Rule & Ambady, 2008).

According to the ecological theory of social perception, initial assumptions about personality characteristics based on physical attributes developed as a way for humans to quickly determine who to trust and who should be avoided (Carre & McCormick, 2008; Flowe, 2012; Haselhuhn & Wong, 2012; McArthur & Baron, 1983). When making judgments about personality from appearance, people respond to two different dimensions of a face: dominance and trustworthiness (Oosterhoff & Todorov, 2008; Solier et al., 2018). These attributions decrease the use of cognitive resources and can help people efficiently determine how to interact with strangers. For example, someone who appears dominant but not trustworthy would be considered a threat and should be instinctively avoided (Oosterhof & Todorov, 2008). The ability to make these inferences develops early. In fact, children as young as 3-4 years old make attributions of "nice," "mean," and "strong" with consistency and accuracy approaching that of adults (Cogsdill et al., 2014), and newborn babies have facial preferences that are strongly correlated with that of adults (Slater et al., 2000).

Researchers have demonstrated that presumptions of "criminality" are positively correlated with "dominance" but negatively associated with "trustworthiness" (Flowe, 2012; Funk et al., 2017). Participants are particularly sensitive to indicators of trustworthiness (Porter et al., 2008; Chang et al., 2010; but see Rule et al., 2012, which suggests that judgments of trustworthiness are subjective impressions with only high interrater agreement and not high accuracy). This sensitivity may explain why humans developed the ability to recognize potential criminality in other people; after all, criminals—especially violent criminals—may pose a threat, and the ability to identify and avoid threatening situations encourages survival.

This paper will explore two questions related to attributions of criminality from facial appearance. First, is it possible to make accurate predictions of criminal behavior based on appearance? Important research has already examined whether people are able to distinguish between criminal and non-criminal faces, but there is little research that examines whether accurate *predictions* of criminal behavior can be made based upon appearance. Second, do stereotypes about criminality create a self-fulfilling prophecy that makes certain criminal-looking individuals more vulnerable to discretion within the criminal justice system? Although previous research has demonstrated that participants can accurately identify photographs of criminals at levels above chance, it is possible that participants are not responding to some inherent signal of criminality but are instead responding to stereotypes about what a criminal looks like. Those same stereotypes may make certain individuals more likely to be convicted of a crime, thus reinforcing the stereotype and supporting the participant's judgment as accurate. The accuracy of predictions of criminal behavior and the influence of stereotypes on these predictions is an important next step in further understanding the validity of appearance-based attributions of criminality.

Attributions of Criminality

A fair amount of research has examined attributions of criminality based on physical appearance. When shown a series of photographs, participants can accurately and reliably distinguish between criminal and non-criminal faces and are even able to correctly match criminal faces with the type of crime committed (Flowe, 2012; MacLin & Herrera, 2006; Valla et al., 2011). In a classic study, G. R. Thornton showed participants photographs of incarcerated men and asked the participants to guess which of four crimes the men in the photographs had committed (1939). Overall, the participants correctly guessed the type of crime at levels above

chance. This study suggests that not only do stereotypes about what a criminal looks like exist, but so do stereotypes about what different *types* of criminals look like. These judgments remain reliable and consistent even when the photograph is displayed for only 100 milliseconds (Klatt et al., 2016).

Furthermore, participants reach a high level of interrater reliability when asked to sort non-criminal faces into criminal subtypes (Bull & Green, 1980; Yarmey, 1993). For example, a research team replicated the Thornton (1939) study but instead used non-criminal faces to assess the prevalence of stereotypes about subcategories of criminality. In the study, participants selected certain faces more frequently for certain crime subtypes and showed high levels of interrater reliability for these judgments (Bull & Green, 1980). Thus, even when responding to non-criminal faces, the stereotypes about what a certain kind of criminal looks like are strong enough that participants reliably agree with each other. This is despite those judgments not being based on any inherent criminality within the faces as all the photographs depict innocent people. These pervasive stereotypes about criminality could be learned from first-hand interactions with people with a particular facial structure or by watching how others interact with people who look a certain way (Blair et al., 2004; Maclin & Herrera, 2006; Zebrowitz, 2017).

Nevertheless, not all researchers have found that stereotypes about criminal subtypes are always consistent. In one experiment, Jeff Valla and colleagues assessed participants' abilities to distinguish between criminals and non-criminals, violent and non-violent criminals, and specific criminal subtypes (2011). Participants rated the convicted criminals as more likely to have committed a crime than the non-criminals even when controlling for other characteristics like attractiveness and affect. But, participants were unsuccessful at distinguishing between violent and non-violent criminals and sorting the photographs based on criminal subtype. In fact, female

participants struggled more than male participants when asked to identify rapists and, in one phase of the study, found rapists to be less likely than non-criminals to have committed any crime.

Regardless of their consistency, stereotypes about criminal appearance have real influence on decision-making in the criminal justice system. For example, defendants who fit these appearance-based stereotypes are more likely to be judged as guilty in ambiguous situations (Dumas & Teste, 2006; Shoemaker et al., 1973). Participants require less evidence before arriving at a guilty verdict and are more confident in that determination when the defendant looks untrustworthy compared to when the defendant appears trustworthy (Porter et al., 2010). Participants are also more likely to select a suspect who is rated high in "criminality" from a lineup than someone who looks less criminal, even when no description of the perpetrator is available (Flowe & Humphries, 2011). Furthermore, faces that are more stereotypically criminal are more memorable for eyewitnesses and are more easily recognized by others than less criminal-looking faces (Maclin & Maclin, 2004). Thus, even if stereotypes about criminality do not perfectly reflect actual behavior, the stereotypes can make certain people more vulnerable than others during situations that may lead to a conviction. This is especially true in situations where little corroborating evidence exists (Flowe & Humphries, 2011; Porter et al., 2010).

Researchers have attempted to uncover what facial features are implicated in stereotypes about criminality, and there are a variety of possible explanations. Utilizing machine learning algorithms, one research team found that criminality is associated with more prominent chins, smaller eyes, lowered eyebrows, and darker complexion (Funk et al., 2017). Another computer algorithm found that a smaller angle between the tip of the nose and the corner of the mouth, a larger lip curvature, and a narrower distance between inner eye corners were all indicators of

criminality (Wu & Zhang, 2016). Using this information, the program was up to 89.4% accurate at identifying whether a new photograph was of a criminal or a non-criminal.

Similarly, facial width-to-height ratio (fWHR) is correlated with criminal behavior and unethical actions, but only for men (Haselhuhn & Wong, 2012). This feature refers to the ratio of the distance between the left and right zygion to the distance between the brow and upper lip (Wong et al., 2011). fWHR is a sexually dimorphic trait that is independent of body size and is linked to aggression and dominance in men but not in women (Carre & McCormick, 2008; Haselhun & Wong, 2012; Wong et al., 2011; but see Burris & Edwards, 2017, Kosinski, 2017 finding no substantial link between fWHR and self-reported measures of behavioral tendencies or violence). Men with higher fWHRs are more likely to be psychopathic, retaliate against perceived slights, act in their own self-interest, and exploit another person's trust during a game (Anderl et al., 2016; Haselhun & Wong, 2012; Ormiston et al., 2017; Stirrat & Perrett, 2010; Wong et al., 2011). Participants assume men with wider faces have less integrity (Ormiston et al., 2017). Researchers have also found real world impact of fWHR. For example, businesses whose male CEOs have a high fWHR perform better financially than businesses with a CEO who has a lower fWHR (Wong et al., 2011).

One explanation for the association of fWHR with dominance or aggression is that higher width-to-height ratios may mimic the appearance of anger: a lowered brow and raised lip (Carre & McCormick, 2008). A higher fWHR therefore makes a man appear to be more physically imposing and aggressive (Stirrat & Perrett, 2010). This theory is referred to as the emotion overgeneralization hypothesis and may explain how other attributions are also made from facial structure: facial features that mimic an emotional expression may be perceived as indicators of

personality consistent with that emotion (Zebrowitz et al., 2010). Thus, people with a rounder face and big eyes may seem happier, more trustworthy, and less criminal.

Other stereotypes about facial appearance also impact sentencing decisions. For example, defendants who have a baby-faced appearance are punished less severely than defendants with a more mature face (Berry & Zebrowitz-McArthur, 1988; Zebrowitz & McDonald, 1991). This may be because roundness is associated with more "childlike" qualities, such as honesty and naïveté, while people with more mature faces are expected to behave more responsibly (Berry & Zebrowitz-McAurthur, 1988; Dumas & Teste, 2006). Furthermore, masculine facial features are associated with attributions of guilt for both male and female suspects (Ward et al., 2012). Similarly, more attractive defendants are likely to receive less severe punishments (Zebrowitz & McDonald, 1991; Stewart, 1980; Leventhal & Krate, 1977, but see Dumas & Teste, 2006; Sigall & Ostrove, 1975). The impact of attractiveness is independent of age (Zebrowitz & McDonald, 1991). The association between attractiveness and guilt may be influenced by people's belief in a just world (Suzuki et al., 2017). For example, someone may assume that an attractive defendant cannot be guilty because criminals do not deserve a beautiful face, meanwhile an unattractive person is more likely to be devious or untrustworthy.

Furthermore, facial features that are associated with certain races are also linked to stereotypes about criminal behavior, likely due to bias associating certain races with crime. In one study, Irene V. Blair and colleagues showed participants photographs of Black and White incarcerated men with similar prison sentences and criminal histories (2004). They found that participants rated photographs of men with stereotypically Black facial features as more criminal than inmates with less stereotypically Black facial features. This was true regardless of whether the incarcerated man was Black or White. Researchers have also found that people with more

7

stereotypically Black facial features are more likely to receive more severe sentences (Blair et al., 2004; Cormier, 2012), including the death penalty (Eberhardt et al., 2006).

Altogether, the research demonstrates clear stereotypes about how a criminal is supposed to look, and these stereotypes influence decisions made by actors in the legal system. This ability to make quick assessments about another person serves an adaptive function, but it is not a perfect system. First impressions based only on stereotypes can be wrong, even if studies show that attributions of certain characteristics have above-chance accuracy (Zebrowitz, 2017). Furthermore, some studies show that judgments of characteristics like criminality change based on the context in which photographs are displayed, such as in an online dating profile or a political campaign (Todorov & Porter, 2014). Photograph selection also determines what attributions are made, with photographs specifically chosen for a particular purpose (like a political campaign) judged differently than randomly selected photographs of the same person (Todorov & Porter, 2014; Jenkins et al., 2011). These limitations suggest that context is an important contributor to stereotypes about appearance.

The Current Research

Previous research has demonstrated that participants are able to make accurate, reliable, and consistent judgments about criminality based simply on a photograph of someone's face. Nevertheless, it remains unclear whether stereotypes about criminality allow for accurate predictions of criminal behavior and whether these stereotypes make some people more vulnerable to the criminal justice system than others. This paper will discuss four studies that attempt to further explore predictions of criminality based on facial appearance and the influence of these judgments. In Study 1 (Chapter 2), participants were asked to make predictions of future violent criminal behavior from yearbook photographs of White boys in their senior year of

high school. These predictions were compared to the criminal records of the boys later in life to see if the participants were able to make accurate predictions of criminality based only on a photograph taken years before any crime had occurred. Study 2 (Chapter 3) is a replication of Study 1 with different non-criminal foils and introduces the role of confidence into the judgment. Study 3 (Chapter 4) used the same basic yearbook photograph paradigm but expanded the stimuli to include both Black and White students as well as violent and non-violent crime. In Study 4 (Chapter 5), the possibility of a self-fulfilling prophecy of conviction created by stereotypes of criminality was addressed. This study tested whether participants could accurately determine guilt when shown mugshots of both exonerees and convicted criminals. This paper will conclude in Chapter 6 with a discussion of the implications of the research and suggestions for future research.

CHAPTER 2

STUDY 1: PILOT

Study 1 was designed to extend the work done by previous researchers, including Valla et

al. (2011), and to examine whether participants would be able to predict criminal behavior from

photographs taken prior to the occurrence of any crime. In most previous studies, researchers

compared edited mugshots of criminals to photographs of volunteers taken in controlled

laboratory settings. This technique may be vulnerable to the influence of environment on facial

appearance, and these differences caused by the environment may be identifiable to participants

beyond any inherent indicators of criminality (Todorov & Porter, 2014). In the current study,

yearbook photographs of high school seniors were used instead of mugshots. Half of the

photographs used in this study depicted students who would later have violent criminal records

and half of the photographs depicted students who would have no criminal record later in life.

This method allowed both exploration of the predictive nature of personality attributions and the

examination of whether something enduring exists in facial appearance that can be used by

others to make determinations of criminality outside the influence of any environmental

contagion.

The research team hypothesized that participants would be able to distinguish between

the yearbook photographs of future criminals and the non-criminal foils. If the hypothesis is not

rejected, the findings would provide evidence against one critique of the previous research: that

either criminal behavior or the arrest process change facial appearance in a way that is noticeable

to others. If participants can accurately distinguish between criminals and non-criminals before a

criminal record exists, it lends credibility to the idea that something within the person's face

itself suggests criminality and the distinction is not a residual of wear and tear in the body caused

by stress related to the arrest. If the hypothesis is supported, more questions will follow. For example, does a "criminal" face indicate a predisposition for crime or a face that simply aligns with criminal stereotypes and thus a vulnerability to a self-fulfilling prophecy?

Methods

Participants and Materials

A total of 152 participants (108 female, 44 male; aged 18 to 24, M = 20.19; 86 White, 44 Asian, 21 other) were recruited from a large New York university and given research credit for participating.
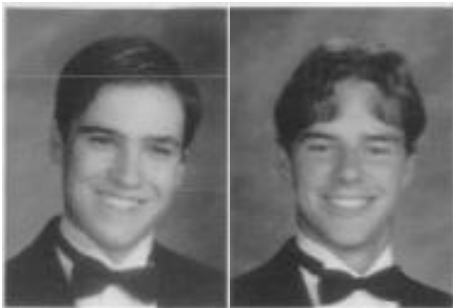

Figure 2.1. Two of the photographs used in this study. The photograph on the left depicts someone with a future violent criminal record and the photograph on the right depicts someone with no future criminal record. For a complete list of photographs for Study 1, see Appendix A.

A research team selected 30 yearbook photographs (15 criminal, 15 non-criminal) from 8 different Utah high school yearbooks (years: 1987-2002). All yearbooks were publicly available on the Internet though classmates.com. To select photographs, the research team systematically searched for the names of all White male students from available yearbooks through the LexisNexis public records database. Criminal records were confirmed by locating the student's future mugshot and comparing its likeness to the yearbook photograph. In the criminal condition, all students had a future violent criminal record (e.g., aggravated assault, battery, forcible sex abuse). This choice was made under the assumption that participants may be more likely to correctly identify someone who is a physical threat than someone who is a non-violent criminal.

11

No student had a conviction until at least two years after their senior year of high school. This choice was made in order to reduce the likelihood that the students were involved in serious crime during high school. For each criminal yearbook photograph chosen, a non-criminal foil was selected by searching the same yearbook for a student with a similar hairstyle who had no criminal record listed on LexisNexis.

All photographs depicted White men in their senior year of high school (approximately 17-19 years of age) with minimal facial hair and no visible scars or tattoos. All students were photographed while wearing tuxedos. By using photographs that controlled for age, gender, clothing style, and race, the materials for this study hopefully address concerns raised by other researchers that participants are more accurate judges of character when the participants are able to rely on knowledge of base-rate frequencies for different demographics as opposed to appearance-based cues (e.g., low-SES men are more likely to commit robbery) (Olivola & Todorov, 2010; Todorov et al., 2015). Judgments in this study were limited to facial cues only.

Procedures

Participants were told they would be shown a series of yearbook photographs and would be asked to rate the likelihood that the person in the photograph would later commit a violent crime. They were not told what proportion of the photographs belonged to each category or how many photographs they would be shown. Participants were then shown all 30 yearbook photographs in random order. Participants were asked to rate the likelihood that each pictured individual would commit a violent crime in the future using a likert scale from 1 (Definitely) to 5 (Definitely Not). Participants then completed demographic questions (including political orientation and whether the participant had ever been a victim of a serious crime) (Gosling et al., 2003).

Results

A mixed effects model was used to test the effect of condition on participants' criminality

ratings, with participant as a random effect and photograph condition (criminal, non-criminal) as

a fixed effect. The main effect of condition was significant, $F(1, 28) = 5.49$, $p = 0.026$. This

suggests that the ratings of photographs of future criminals and non-criminals differed

significantly. Future criminals were more likely to be rated as more likely to commit a crime

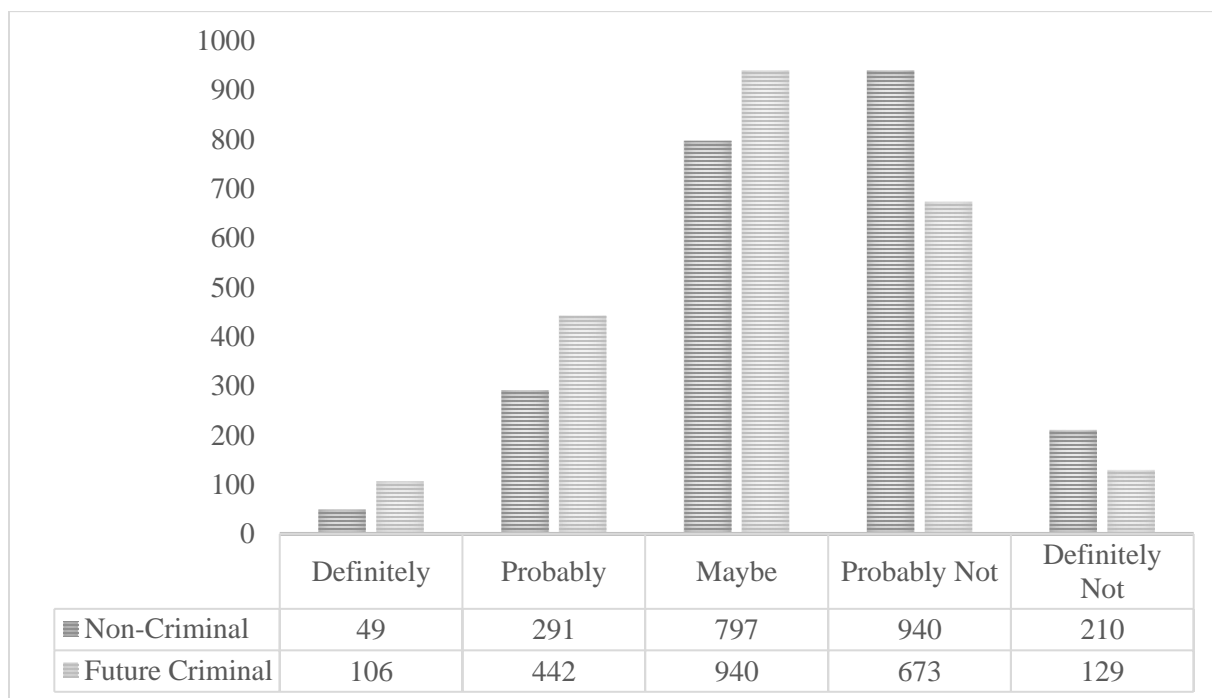than the non-criminal foils. This relationship is displayed in Figure 2.2.

	Definitely	Probably	Maybe	Probably Not	Definitely Not
Non-Criminal	49	291	797	940	210
Future Criminal	106	442	940	673	129

Figure 2.2. Aggregate ratings across all twenty photographs of future criminality for non-criminal and future criminal conditions.

Potential Mediators

In the first phase of the study, participants rated future criminals as more likely to commit

a violent offense than non-criminal foils, but the possibility remained that participants were

responding to some other characteristic of the photographed individuals, such as attractiveness or

perceived socioeconomic status. To test this, a new set of participants rated the same 30 faces on

a series of characteristics stereotypically associated with judgments of criminality: attractiveness,

13

stereotypical whiteness, future salary (a proxy for socioeconomic status), helpfulness, and aggression.

These characteristics were selected by conducting a review of the literature for characteristics that have been shown to influence criminal sentencing. For example, attractiveness is linked to less severe punishments (Mazzella & Feingold, 1994; Stewart, 1980; Zebrowitz & McDonald, 1991), stereotypically black facial features are associated with more severe punishments (Blair et al., 2004; Cormier, 2012; Eberhardt, et al., 2006), defendants with low socioeconomic status are more likely to be found guilty and receive more severe punishments (Mazzella & Feingold, 1994), and defendants with less trustworthy/helpful faces are more likely to be found guilty of a crime (Porter et al., 2010). Furthermore, the ability to detect aggression and helpfulness based only on someone's appearance is considered an evolutionary advantage that can help detect who is worthy of trust and who is a threat (Haselhuhn & Wong, 2012; Flowe, 2012; Carre & McCormick, 2008). Theoretically, the ability to detect threats should lend itself to detecting predispositions to violent crime.

The research team hypothesized that the main effect of criminality found in the prior stage of Study 1 would be mediated by one of these other characteristics. If this hypothesis is supported, then participants are likely partially relying on stereotypes about what types of people are criminals (e.g., unattractive men from poor backgrounds) when making predictions of criminality rather than sensing some inherent predisposition to criminality within the photographed individual. This finding could lend credibility to a theory that appearance-based stereotypes of criminality may make certain people more vulnerable to conviction than other people. In other words, men who fit stereotypes about the type of person who is likely to be a

criminal are more likely to viewed as a criminal and thus are judged by others as guilty, which results in a criminal conviction.

Methods

A total of 257 unique participants were recruited from the same large New York university and given credit for participation. The demographics of the participants were comparable to the previous stage of the study. Participants were asked to rate each of the 30 faces from the previous study on one of five different characteristics: aggression, attractiveness, helpfulness, future salary (a proxy for socioeconomic status), and stereotypical whiteness. Participants were split almost equally among the five characteristics, with between 50 and 54 participants providing ratings of all 30 faces for each characteristic (e.g., 54 participants rated all the faces on expected future salary). None of these participants had participated in the first phase of this study, and each participant only rated the faces for one characteristic.

Each characteristic was measured on a likert scale from 1-5, with 1 being the least like that characteristic (e.g., not at all attractive) and 5 being the most like that characteristic (e.g., extremely attractive), except for future salary. For future salary, participants were asked to indicate what annual salary they expected the student in the photograph to receive in ten years out of a list of possible salary ranges: less than $10,000; $10,000 to $19,999; $20,000 to $39,000; $40,000 to $49,000; $50,000 to $59,000; $60,000 to $69,000; $70,000 to $79,000; $80,000 to $89,000; $90,000 to $99,000; $100,000 to $149,000; $150,000 or more. The ratings were then manipulated to be on the same 1-5 scale as the other characteristics. The ratings from all participants were combined to create the mean ratings for each face for each characteristic. These scores are located in Table 2.1. None of the differences between the criminal and non-criminal conditions reached statistical significance.

Table 2.1
Ratings of other characteristics

Characteristic	Condition	Range	Mean	SD
Aggression	Non-criminal	1.67 – 2.96	2.26	.41
	Criminal	1.59 – 3.73	2.60	.65
Attractiveness	Non-criminal	1.32 – 3.56	2.45	.68
	Criminal	1.34 – 3.80	2.13	.76
Helpfulness	Non-criminal	2.30 – 3.74	2.96	.42
	Criminal	2.04 – 3.48	2.63	.45
Future salary	Non-criminal	2.53 – 3.86	3.34	.39
	Criminal	2.41 – 3.93	3.12	.42
Whiteness	Non-criminal	2.72 – 4.35	3.75	.50
	Criminal	2.47 – 4.33	3.38	.59

Results

A mixed effects model was used to test the effect of photograph condition on criminality ratings, with participant as a random effect and photograph condition (criminal, non-criminal) and mean ratings for each photograph on the five characteristics as fixed effects. This model would explain if any of the characteristics mediate the main effect of criminality seen in the first phase of the study. When the additional characteristics were added into the model, the main effect of criminality disappeared, $F(1, 23) = 1.244$, $p = 0.276$, but the main effects of future salary, $F(1, 23) = 4.767$, $p = .039$, and aggression, $F(1, 23) = 5.229$, $p=.032$, were significant. No other characteristics showed significant main effects. This result suggests that the ratings of criminality were mediated by both perceived future salary and perceived aggressiveness.

Discussion

Overall, the results of Study 1 suggest that participants can accurately predict the likelihood of someone having a criminal record later in life based only on the person's high school yearbook photograph. This distinction lends credibility to the theory that something about a person's face itself indicates criminality and not that crime or the environment of an arrest alters facial appearance in a distinguishable way. If that was the case, photographs of

criminals should not be distinguishable from non-criminals prior to any serious crime occurring, such as in high school yearbook photographs when affect and the environment is constant between criminals and non-criminals.

Furthermore, the second phase of Study 1 found that ratings of perceived future salary and aggressiveness mediated the main effect of criminality. Given this result, participants may be relying on stereotypical assumptions about what criminals look like when rating faces for future criminality: an aggressive-looking teenager from a poor neighborhood is the most likely to become a violent criminal. Interestingly, attractiveness, stereotypical White appearance, and helpfulness did not explain the main effect of criminality despite previous research demonstrating an association between these characteristics and criminality (Blair et al., 2004; Flowe, 2012; Mazzella & Feingold, 1994)

One limitation to this study was the method of photograph selection. All photographs were selected after searching for the criminal records of students in high school yearbooks on the LexisNexis public records database. The research team took great care to verify that each criminal record matched the person in the yearbook photo and that no criminal record existed online for the photographs used for the non-criminal foils. Nevertheless, this process depended upon the assumption that criminal records accurately reflect the criminal history of a person, and there are numerous flaws in that logic. First, it is possible that some of the non-criminal faces have criminal records that have been either sealed or dismissed and are no longer available online. Second, it is possible that some of the criminal faces were falsely convicted of a crime and those men are in fact innocent. Finally, it is also possible that some of the men in the non-criminal photographs committed crimes but were not convicted of those crimes. In effect, the characteristic of "criminality" used in this study was dependent upon whether the man in the

17

photograph had been convicted of a crime, and this measure may not be entirely accurate. This further supports the hypothesis that participants were responding to how "convictable" the students in the photographs looked.

An additional limitation is that the research team do not know if the men in the photographs committed criminal activity prior to turning eighteen. One assumption of this study is that high school seniors are blank slates who have never been involved in serious crime and therefore do not show the influence of past criminal behavior on their faces. Of course, this assumption is flawed. Many high school seniors already have extensive criminal experience, but their age protects them from either formal punishment or longstanding criminal records. Arrests and convictions prior to age 18 would not show up on a LexisNexis public records search, and thus could not be controlled for in this study. In addition, many high school seniors may already be involved in non-criminal activity that is nonetheless predictive of more serious criminal behavior in the future. These limitations are difficult to control for without self-report measures of criminal activity but may still influence the judgments made by participants.

Nevertheless, this study demonstrated a phenomenon that is largely unexplored in the facial attribution literature: participants can accurately predict someone's future behavior based only on the person's photograph. Specifically, participants were able to accurately distinguish between photographs of men who would later have criminal records and those who would not. These decisions were partially influenced by other characteristics of the men, including perceived socioeconomic status and aggressiveness. The next chapter will discuss a replication of this study in order to test if this finding holds with a new set of non-criminal foils. The second study will also explore the influence of confidence on judgments of criminality, which may influence how likely people are to act on their perceptions of a stranger's criminality.

CHAPTER 3

STUDY 2: CONFIDENCE

In Study 1, participants were able to accurately predict whether a high school senior would later have a criminal record or not. A second study was designed to replicate and extend the findings of Study 1. The primary difference between Study 1 and Study 2 was the introduction of new non-criminal foils. The non-criminal foils in Study 1 were chosen based on their similarity to the students in the criminal condition (e.g., matched hair style). In Study 2, new non-criminal foils were randomly selected to see if the findings from Study 1 could be replicated with nonmatching faces. In addition, participants in Study 1 made judgments on a 1-5 likert scale about the likelihood that a student would later engage in a violent crime. Participants' responses trended toward the middle of the scale, indicating that the student "may" later have a criminal record (Figure 2.1). Although this scale allowed for nuanced measurement of predictions, the research team was interested to see how participants would behave if forced to make a binary choice: criminal versus not criminal. Thus, in Study 2, participants were asked to make yes/no predictions of criminal behavior.

Furthermore, Study 2 also explored the influence of confidence on predictions of criminality. There are two primary reasons why confidence ratings were included in this study. First, confidence may make someone more likely to act upon a character judgment. For example, if someone confidently believes that a high school student has a criminal future, the person may act more punitively toward the student and give the student fewer prosocial or redemptive opportunities. This could result in a self-fulfilling prophecy and a criminal conviction. Second, confidence in a judgment can be more convincing to others, even if the judgment is inaccurate. For example, jurors find high-confidence eyewitnesses to be very compelling and indicative of

guilt (Cutler et al., 1988). In fact, over 70% of DNA exonerations involve high-confidence

eyewitness identifications (Wixted et al., 2018). Not only does the appearance of criminality

make it more likely that a suspect will be selected from a lineup (Flowe & Humphries, 2011), but

participants are generally more confident in guilty verdicts when the defendant looks less

trustworthy (Porter et al., 2010). This confidence may then influence other people who interact

with the person, creating a greater likelihood of a criminal conviction.

Altogether, participants' confidence in their judgment can influence how they behave,

how they treat other people, and how much other people believe them. In Study 2, participants

were asked to both rate their confidence in each judgment and guess their overall accuracy for all

30 photographs. These two separate measures allowed for analyses of overall confidence in the

skill of making attributions of criminality from photographs and also confidence for each

specific judgment. In particular, the research team was interested in discovering whether

confidence was related to accuracy and whether specific types of people were more likely to be

accurate or confident in their judgments of criminality. The research team hypothesized that

participants would be able to make predictions of criminality at accuracy levels surpassing

chance and that participants' photograph-level confidence would be predictive of their accuracy.

<div align="center">Methods</div>

Participants and Materials

A total of 146 participants (113 female, 33 male; aged 18 to 35, M = 19.89; 75 White, 36

Asian, 14 Black, 21 other) were recruited from a large New York university and given research

credit for participating.

This study was a replication of Study 1 and included the same 15 criminal photographs.

A random generator was used to randomly select new non-criminal foils from the same

yearbooks as the criminal photographs. The name of each foil was searched in the LexisNexis public records database to verify that the pictured student had no publicly available criminal record. As with the first study, photograph selection controlled for age, gender, clothing, and race. All photographs depicted White men in their senior year of high school who were wearing tuxedos. Judgments in this study were limited to facial cues only. All photographs for Study 2 are listed in Appendix B.

Procedures

Participants were told they would be shown a series of yearbook photographs and would be asked whether the person in the photograph would later commit a violent crime or not. Participants were not told what proportion of the photographs belonged to each condition or how many photographs they would be shown. Participants were then shown all 30 yearbook photographs in random order. In contrast to Study 1, participants were asked to make a binary judgment of future criminality (Yes or No) rather than rating the photographs on a 1-5 likert scale. After making a judgment of future criminality, participants were asked to rate their confidence in their decision using a scale from 1 (Extremely confident) to 5 (Extremely uncertain). After making decisions about all photographs, participants were asked to make a guess about their overall accuracy for the task. Participants then completed the same demographic questions from Study 1.

<div align="center">Results</div>

Participants' judgments were examined to assess what influenced their decisions about criminality. A generalized estimating equation (GEE) was carried out with dichotomous criminality assessments for each photograph as the dependent variable, photograph condition (criminal or non-criminal) as a within-subject factor, and controlling for photograph ID and

<div align="center">21</div>

participant ID. The odds of a participant predicting criminality increased when shown a

photograph in the criminal condition in comparison to the non-criminal condition, B = 0.817, SE

= 0.073, χ^2 (1, n = 145) = 125.007, p < 0.001. Overall, ratings of criminality were 21% in the

non-criminal condition (SD = 0.408) and 38% in the criminal condition (SD = 0.485). Thus,

participants were less likely to find that the non-criminal foils would be criminals in the future

than the men in the criminal photographs (see Figure 3.1).

	Non-criminal	Future Criminal
No	1725	1362
Yes	462	826
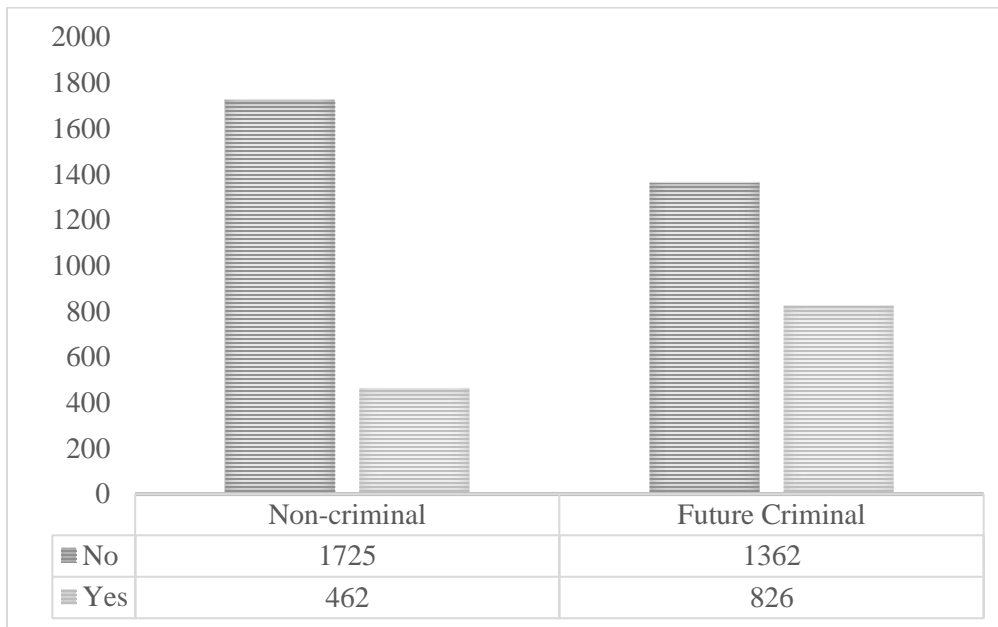
Figure 3.1. Aggregate decisions across all photographs for dichotomous predictions of
criminality. "Yes" indicates a prediction of criminality.

Another GEE was executed to examine the effect of participant characteristics on

criminality assessments. Participants' political orientation was a significant predictor of their

judgment. Libertarians were less likely to say that a student looked like a future criminal than

democrats, B = 0.541, SE = 0.139, χ^2 (1, n = 145) = 16.468, p = 0.002. Participants' other

demographic characteristics (e.g., gender, age, race, victim status) were not found to be

significant predictors of their judgments.

Participants' accuracy was also examined, both overall and for each photograph. The

average overall accuracy was 58.3% correct (SD = 2.64). Mean accuracy for the non-criminal

photographs was 79.2% correct (SD = 2.55) and 37.87% correct (SD = 3.09) for the criminal photographs. Given that each judgment was a 0–1 decision, t-tests were conducted to determine whether participants performed at above chance for each condition by comparing the accuracy scores from each condition and 0.5 as the at-chance level. Participants performed above chance overall, $t(143) = 11.288$, $p < 0.01$. Participants performed above chance for non-criminal photographs, $t(143) = 20.545$, $p < 0.01$. Participants performed below chance for criminal photographs, $t(143) = -7.06$, $p < 0.01$. Overall, the accuracy scores demonstrate that participants had a non-criminal bias, with participants most often giving incorrect responses due to judging the photographs as non-criminal rather than as criminal. This reduced accuracy in the criminal condition and inflated accuracy in the non-criminal condition.

GEE models were used to test whether photograph-level accuracy differed as a measure of category, confidence, and participant demographic characteristics. Participants' odds of giving a correct response increased when shown non-criminal photographs, $B = 1.823$, $SE = 0.144$, χ^2 $(1, n=141) = 158.94$, $p < 0.001$, confident in their decision, $B = -0.115$, $SE = 0.0413$, χ^2 $(1, n=141) = 7.809$, $p = 0.005$, when they had been a victim of a serious crime, $B = 0.261$, $SE = 0.105$, χ^2 $(1, n=141) = 6.159$, $p = 0.013$, and when the participant was male, $B = 0.270$, $SE = 0.080$, χ^2 $(1, n=141) = 11.361$, $p = 0.001$. There was no significant effect of age, race, or political orientation on accuracy. Participants who had previously been the victim of a serious crime had a mean accuracy of 65% (SD = 1.38), and participants who had no history had a mean accuracy of 58% (SD = 2.65). Men had a mean accuracy of 63% (SD = 2.446) and women had a mean accuracy of 57% (SD = 2.564) (Figure 3.2).

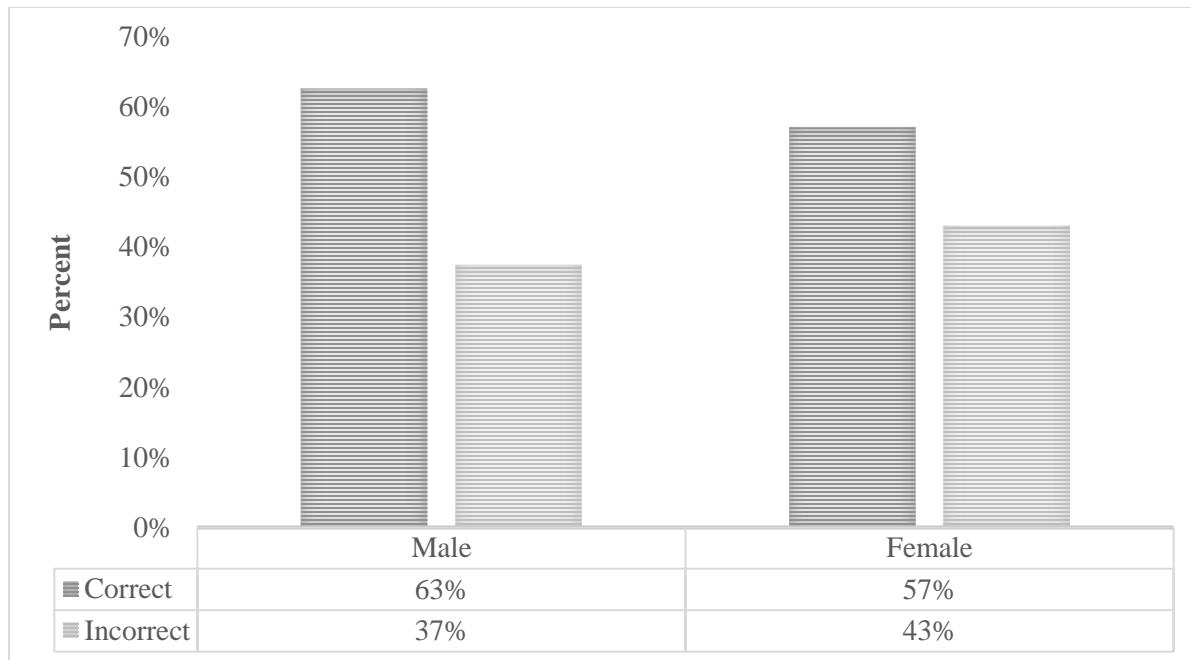|	Male	Female
▤ Correct	63%	57%
▤ Incorrect	37%	43%

Figure 3.2. Gender differences in accuracy.

Finally, the confidence of participants was examined. Confidence was measured for each criminality judgment, but participants also estimated their overall performance across all 30 photographs. There were no significant differences across race, gender, or victim status for overall confidence. A mixed effects model was used to test the effect of participants' judgment on photograph-level confidence. Participants were less confident when they judged a photograph as criminal (M = 3.072, SE = 0.062) than non-criminal (M = 3.146, SE = 0.059), $F (1, 146) = 6.573$, $p = 0.010$. This finding further supports the non-criminal bias found in this study.
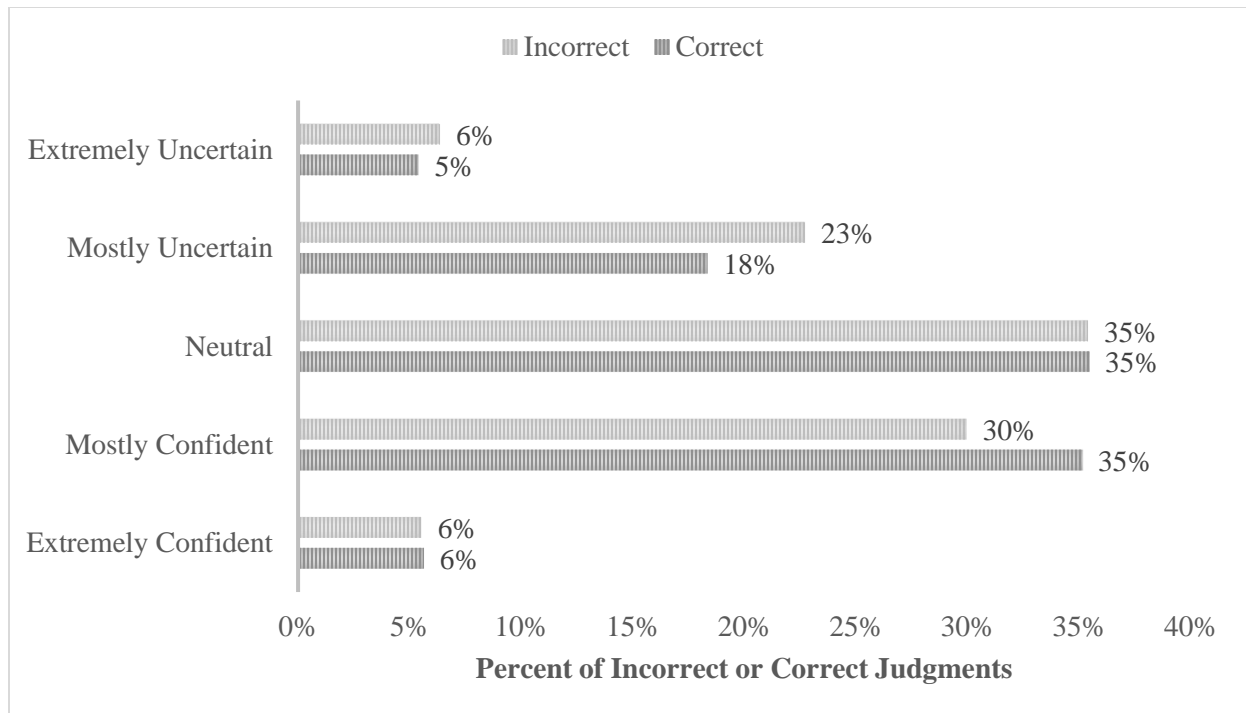
Figure 3.3. Confidence in both accurate and inaccurate judgments.

Another mixed effects model was used to assess the effect of photograph condition and accuracy on confidence. A main effect was found for accuracy, $F(1, 146) = 18.628$, $p < 0.01$. Participants were more confident when they made a correct decision ($M = 3.155$, $SE = 0.060$) than when they made an incorrect judgment ($M = 3.037$, $SE = 0.062$) (Figure 3.3). There was also a significant interaction between accuracy and condition, $F(1, 146) = 9.554$, $p = 0.002$. In the non-criminal condition, participants were significantly more confident when they were correct ($M = 3.190$, $SE = 0.066$) rather than incorrect ($M = 2.981$, $SE = 0.066$). In comparison, there was no significant difference between correct and incorrect judgments in the criminal condition. This may be because accuracy itself was lower in the criminal condition.

<center>Discussion</center>

This study was designed to replicate Study 1 with a few alterations. In Study 2, new randomly selected foils were used, predictions of criminality were measured with dichotomous yes/no judgments, and confidence ratings were collected. The results demonstrated that

participants were able to successfully make predictions of criminality at above-chance levels. This supports the findings from Study 1 and suggests that people have some capacity to predict criminal behavior based on facial appearance. It remains unclear whether this predictive capacity is due to stereotypes about criminality, a self-fulfilling prophecy, or some recognizable criminal predisposition. This study does not dispute the possibility that participants are responding to how convictable a face looks as opposed to the inherent criminality of the photographed person.

Participants in this study demonstrated a non-criminal bias, meaning that participants were less likely to judge that photographs in either condition as criminal as opposed to non-criminal. In addition, participants were less confident in their decisions when they judged a photograph as criminal compared to non-criminal. This suggests that participants were at least somewhat reluctant to make an accusation of criminality, which may have shaped their responses. Overall, participants performed at above-chance levels of accuracy, but participants performed significantly below chance when shown photographs in the criminal condition. Participants' non-criminal bias reduced accuracy in the criminal condition, where the correct response was to say that the man in the photograph would have a criminal future. Perhaps this bias should be expected, as an accusation of criminality can have serious consequences. In this study, without anything but appearance to base the judgment on, participants may have been reluctant to make a definitive claim of criminal behavior. This non-criminal bias may have been exacerbated by the context of the photographs, which depicted smiling high school students wearing tuxedos. This same non-criminal bias may not exist in the real world where other context is available, such as proximity to a crime, prison garb, or presence in a courtroom.

Furthermore, certain characteristics of the participants increased the odds of an accurate judgment of criminality. Male participants were more accurate predictors of criminality, as were

26

participants who had previously been a victim of a crime. This may suggest that familiarity—either with crime or with men— increases accuracy. In other words, those who have been victims of crime may recognize criminality in other people due to their past experiences with criminals. Presumably, some victims may more fleshed out stereotypes about what a criminal looks like because they have at least one real example of a criminal to base their stereotypes on. In addition, men may be more accurate at this task because all the photographic stimuli depicted men. In-group familiarity may allow male participants to recognize nuance in the faces of other men better than women can. More research should be done to examine if familiarity increases accuracy. For example, cross-race effects suggest that out-group members are less able to recognize differences in the faces of people of other races compared to their own race (Bothwell et al., 1989; Shriver et al., 2008; Tanaka et al., 2004). In Chapter 4, photographs of Black men will be added to the experiment to test if cross-race effects alter the performance of participants. The results may lend some support to the theory that men performed better in this study due to their in-group status with the men in the photographs, and not due to some additional sensitivity to criminality in general.

Participants' confidence in their judgments was also examined in Study 2. Participants were more likely to be confident when they were correct, but only in the non-criminal condition. There was no significant difference between correct and incorrect judgments in the criminal condition. Accuracy itself was lower in the criminal condition, suggesting that perhaps more participants were guessing in the criminal condition whereas participants were more confident in their judgments in the non-criminal condition. This may suggest that people do not hold stereotypes about what *criminals* look like, but rather hold stereotypes about what a *non-criminal* looks like. This is supported by previous research which suggests that people are particularly

sensitive to indicators of trustworthiness in others but not dominance or aggression (Porter et al., 2008; Chang et al., 2010). Thus, participants may have instinctively known who among the pictured students would not get wrapped up in the criminal justice system but were less confident about who would. This sensitivity to non-criminal behavior will be further explored in Chapter 5, in which participants were asked to distinguish between photographs of exonerated men and men convicted of violent crimes.

In general, Study 2 further supported the theory that people can make accurate attributions of criminal behavior based only on a person's facial appearance. Nevertheless, this study suffered from many of the same limitations as Study 1. Primarily, participants in this study may still have been responding to how convictable a face looks as opposed to some inherent indicator of true criminal behavior. In fact, the findings suggest that participants may actually be more sensitive to how non-convictable a face looks rather than how criminal the photographed person is. Thus, it is possible that uncertainty about criminality as opposed to confidence in a threat drives behavior in ambiguous situations. If people feel confidence in their ability to detect non-criminals (as this study suggests) but don't feel the same confidence about their ability to detect criminals, everyone who they feel uncertain about poses a potential threat. The next chapter will further explore the implications of this possibility by introducing race into the paradigm and determining whether participants perform differently when shown photographs of Black and White men. This will help explain if familiarity can increase accuracy and will also assess how entrenched stereotypes about the association of race and crime influence performance on the task.

CHAPTER 4

STUDY 3: RACE

In the previous two studies, participants were able to accurately predict the characteristic of criminality based only on yearbook photographs of White men. Study 3 was designed to replicate the first two studies but with all new yearbook photographs. In addition, this study expands previous research by comparing accuracy for photographs of Black and White men. This study also compared accuracy for both non-violent and violent crimes, whereas the previous studies used only photographs of men who were later convicted of violent crimes. This was done in order to test if participants are still accurate when the crime in question does not provide an immediate physical threat.

There are several reasons why it was important to expand the research to look at predictions for Black male faces. First, implicit biases against men of color have been shown to impact decisions at every stage of the criminal justice system, including police shootings (Correll et al., 2006), the setting of bail amounts (Ayers & Waldfogel, 1994), the instigation of investigatory stops (Epp et al., 2014), and sentencing (Mustard, 2001). This influence has partially led to an over-representation of Black men among those with criminal convictions (Mitchell et al., 2005; Pettit & Western, 2004). For this reason, it is necessary to further explore how stereotypes may influence perceptions of people of color in the criminal justice system. It is hypothesized that participants in this study will judge the Black students as more likely to be criminals than the White students. This trend would further support theories about the contribution of bias to racial disparities in convictions.

Second, stereotypes about criminality differ based on the race of the target. Because most research on attributions of criminality use White faces (but see Wu & Zhang, 2016 which

compared a large number of Chinese faces), little is known about accuracy of attributions of criminality for non-White faces. This is especially important because research shows that stereotypes about criminality are intertwined with racial bias (Cormier, 2012; Eberhardt et al., 2006). For example, people with more stereotypic "Black" features tend to be viewed by others as more criminal and are given longer sentences (Blair et al., 2004). Furthermore, stereotypes about race can shift how neutral characteristics (like dominance) are evaluated when interpreting ambiguous situations (Solier et al., 2018). For example, a Black man who appears dominant is more negatively evaluated by participants than an equally-dominant appearing White man. This can result in a misinterpretation of ambiguous behavior as threatening, which could result in a referral to the police and a possible conviction that would not exist for a White man in the same situation.

Third, the introduction of Black faces allows for a more nuanced examination of the performance of participants. In the previous two studies, participants were able to correctly predict the behavior of White men, but it is possible that participants will not show the same level of accuracy for Black men. Research shows that people are better able to recognize faces of others from within their own race as opposed to cross-racially (Bothwell et al., 1989). This is referred to as the cross-race effect and is generally attributed to greater familiarity and experience with people from one's own-race (Shriver et al., 2008; Tanaka et al., 2004). This familiarity also contributes to increased ability recognize subtleties within own-race faces, including emotion recognition. The failure to recognize emotions in members of other races can combine with stereotypes to further complicate attributions of character. For example, White participants are more likely to rate Black faces as looking "angry" or "hostile" than White faces (Halberstadt et al., 2018). These stereotypes may result in different performance when viewing

photographs of different races. It is hypothesized that implicit bias will lead to decreased accuracy for the Black male photographs, especially for White participants who may struggle to separate "criminality" from their stereotypes about "aggression" or "dominance" within Black faces.

Furthermore, in-group membership plays an important role in *how* faces are evaluated. In-group faces are judged as more trustworthy than faces of out-group members, but—even beyond that— the trustworthiness of out-group faces is evaluated differently than in-group faces (Sofer et al., 2017). For in-group members, ratings of trustworthiness have been found to depend on how "typical" the face looks (i.e., whether the face is representative of in-group characteristics) (Dotsch, et al., 2016; Stolier et al., 2018). People who have "typical" or "average" faces are considered more trustworthy. In contrast, out-group faces are evaluated for trustworthiness based on how attractive the face is, rather than how representative the face is of the out-group (Sofer et al., 2017; Stolier, et al., 2018). A participant will judge an attractive out-group member as more trustworthy than a "typical"-looking out-group member. Thus, it is hypothesized that criminality ratings for Black faces will be explained by other characteristics of the photographs, like attractiveness, which have not previously been shown to predict ratings of criminality in White faces.

Due to these reasons, the research team hypothesized that White participants would show decreased accuracy when shown photographs of Black men and asked to make predictions of future criminality. It was also hypothesized that participants would show a criminality bias for the Black faces, even though participants demonstrated a non-criminal bias for the White faces in the previous two studies. This is expected due to the association between Black men and crime, which may reverse the non-criminal bias (Cormier, 2012; Eberhardt et al., 2006). The research

team also predicted that participants would show reduced accuracy when shown photographs of non-violent criminals, because non-violent criminals would be less of an immediate physical threat and therefore less likely to appear aggressive or dominant. Participants were expected to perform at levels above chance overall.

Methods

Participants and Materials

A total of 156 participants (122 female, 32 male, 2 unknown; ages 18 to 22, M=19.76; 77 White, 9 Black, 70 other) were recruited from a large New York university and given research credit for participating.

The research team selected 80 yearbook photographs for this study (10 Black violent criminals, 10 Black non-violent criminals, 20 Black non-criminal foils, 10 White violent criminals, 10 White non-violent criminals, and 20 White non-criminal foils) from 17 different North Carolina high school yearbooks (years: 1990-1999). All yearbooks were publicly available on the Internet though classmates.com. To select photographs, the research team systematically searched for the names of all male students from available yearbooks through the LexisNexis public records database. Photographs were excluded if they did not depict White or Black men. Photographs were also excluded if the man in the photograph had piercings or extreme facial hair.

Within the 17 yearbooks used, 1064 White men and 467 Black men were pictured. Many of these photographs were unusable because no public records were located matching the name, age, and location of the student. Given these constraints, a total of 229 White non-criminal photographs (21.5%) and 35 Black non-criminal photographs (7.5%) were usable. There were also 12 White violent photographs (1.1%), 25 Black violent photographs (5.4%), 11 White non-

violent photographs (1.0%), and 18 Black non-violent photographs (3.9%). Of note, these

proportions demonstrate a different experience between the White and Black men, with far more

of the Black students having later contact with the criminal justice system. This is not

unexpected given the that the likelihood of even minor interactions with the criminal justice

system differ based on race (e.g. Epp et al., 2014; Pettit & Western, 2004). Because more usable

photographs were collected than necessary for the purpose of the study, 10 photographs were

randomly selected from each group of criminal photographs and 20 photographs were randomly

selected from the non-criminal photographs for a total of 80 photographs.
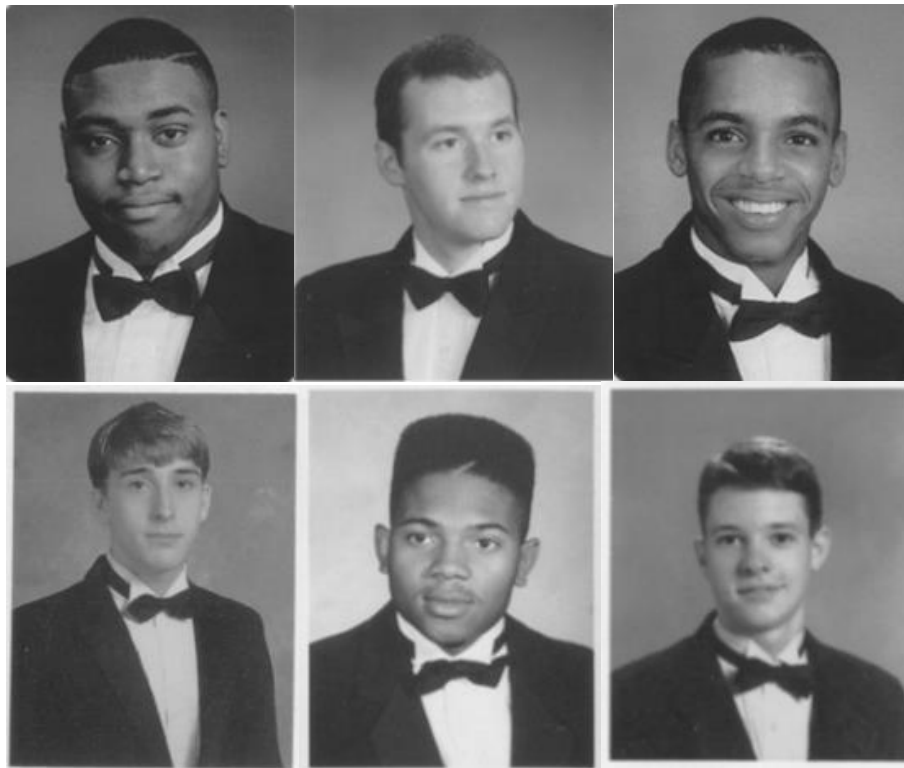
Figure 4.1. Six of the photographs used in this study. In order from first to last: Black non-criminal, White non-criminal, Black non-violent criminal, White non-violent criminal, Black violent criminal, and White violent criminal. For a complete list of photographs from Study 3, see Appendix C.

All photographs depicted men in their senior year of high school (approximately 17-19

years of age) with minimal facial hair and no visible scars or tattoos. All students were

photographed wearing tuxedos. Judgments in this study were limited to facial cues only.

In order to test potential mediators, 70 unique participants were recruited via social media and asked to rate the faces for the following characteristics: aggression, trustworthiness, attractiveness, wealth, and emotional affect. Each participant was randomly shown 20 of the 80 total faces to rate on all characteristics. Each characteristic was measured on a likert scale from 1-5, with 1 being the least like that characteristic (e.g., extremely unattractive) and 5 being the most like that characteristic (extremely attractive). See Table 4.1 for the means for both conditions and Table 4.2 for the means across the interaction of race and criminal conditions.

Table 4.2
Mean ratings for all measured characteristics for both conditions: race and criminality. Standard deviations are included in parentheses.

		Aggressive	Trustworthy	Attractive	Wealth	Affect	fWHR
Race	Black	1.85 (.44)	2.27 (.40)	2.05 (.49)	1.94 (.25)	2.48 (.67)	2.04 (.20)
	White	2.14 (.54)	1.82 (.40)	1.70 (.57)	2.46 (.42)	2.37 (.65)	2.05 (.17)
Criminality	Criminal	2.12 (.57)	1.95 (.52)	1.74 (.58)	2.11 (.43)	2.41 (.63)	2.07 (.17)
	Non-Criminal	1.87 (.43)	2.14 (.38)	2.02 (.52)	2.28 (.42)	2.44 (.69)	2.03 (.20)

An ANOVA was conducted to examine the differences between the criminal and race conditions on the other characteristics. The overall mean rating for aggression was 1.99 (SD = 0.513). Black faces were rated as significantly less aggressive than White faces, $F(1, 79) = 7.681$, $p = 0.007$. Criminal faces were rated as significantly more aggressive than non-criminal faces, $F(1, 79) = 5.777$, $p = 0.019$. There was a significant interaction of criminal/non-criminal and race, $F (1, 79) = 4.228$, $p < 0.043$. Simple main effects analysis showed that White faces were rated as significantly more aggressive when criminal than non-criminal, but there was no difference in aggressiveness between the criminal and non-criminal Black faces.

The overall mean rating for trustworthiness was 2.059 (SD = 0.460). Black faces were rated as significantly more trustworthy than White faces, $F(1, 79) = 26.367$, $p < 0.001$. Criminal

faces were rated as significantly less trustworthy than non-criminal faces, $F(1, 79) = 4.821$, $p <$

0.001. The overall mean rating for attractiveness was 1.883 (SD = 0.56). Black faces were rated

as significantly more attractive than white faces, $F(1, 79) = 2.596$, $p = 0.003$. Criminal faces

were rated as significantly less attractive than non-criminal faces, $F(1, 79) = 5.698$, $p = 0.019$.

The overall mean rating for wealth was 2.200 (SD = 0.431). Black faces were rated as

significantly less wealthy than White faces, $F(1, 79) = 47.985$, $p < 0.001$. Criminal faces were

rated as significantly less wealthy than non-criminal faces, $F(1, 79) = 5.442$, $p = 0.022$. The

overall mean rating for affect was 2.428 (SD = 0.653). There were no significant differences for

affect.

Table 4.2
Mean ratings for all measured characteristics across race and criminality photograph conditions.

		Aggressive	Trustworthy	Attractive	Wealth	Affect	fWHR
Black	Future Criminal	1.865	2.249	1.997	1.903	2.591	2.084
	Non-Criminal	1.828	2.300	2.129	1.975	2.378	2.010
White	Future Criminal	2.377	1.657	1.491	2.321	2.234	2.056
	Non-Criminal	1.904	1.991	1.915	2.600	2.510	2.048

In addition, research assistants used the software ImageJ to measure the fWHR of all 80

photographs following the same methods as used in Carre & McCormick (2008). Inter-rater

reliability was moderately high for all measurements (distance between left and right zygion: $r = 0.921$, $p < 0.001$; distance between upper lip and brow: $r = 0.968$, $p < 0.001$; width-to-height

ratio: $r = 0.769$, $p < 0.001$). The mean fWHR for the photographs was 2.049 (SD = 0.185). There

were no significant differences across conditions for fWHR.

Procedures

The survey design replicated the procedures used in Study 2. Participants were told they

would be shown a series of yearbook photographs and would be asked whether or not the person

in the photograph would later commit a crime. They were not told what proportion of the

photographs belong to each category or how many photographs they would be shown. Due to the

number of photographs required for this study and to reduce fatigue, each participant was

randomly shown 40 of the 80 total photographs. Each photograph was shown to a minimum of

76 participants and a maximum of 79. Finally, participants completed demographic questions,

including age, race, gender, political orientation, and whether they had previously been a victim

of a serious crime.

Results

Participant judgments varied based on the photograph. The photograph with the highest

percentage of ratings of criminality had a mean criminality rating of 66% (SD = 0.497). The

photograph with the lowest percentage of ratings of criminality had a mean criminality rating of

5% (SD = 0.255). For a full list of ratings, see Appendix C.

Participants' judgments were examined to assess what influenced their decisions about

criminality. A generalized estimating equation (GEE) was carried out with dichotomous

criminality assessments for each photograph as the dependent variable, photograph conditions

(non-criminal, violent, or non-violent; Black or White) as a within-subject factors, and

controlling for photograph ID and participant ID. Participants were more likely to predict

criminality in both the violent (χ^2 (1, n = 156) = 4.052, p = 0.044) and non-violent conditions (χ^2

(1, n = 156) = 28.093, p < 0.001) in comparison to the non-criminal condition. Participants were

also more likely to predict criminality when shown a photograph in the White condition

compared to the Black condition, χ^2 (1, n = 156) = 87.294, p < 0.001. Findings for all GEE

models for judgment are located in Table 4.3.

Table 4.3
Generalized estimating equations for judgments of criminality.

Predictors	Comparison Group	Model 1		Model 2 Photograph characteristics	
		B (SE)	p	Estimate (SE)	p
Intercept		1.514 (.082)	.000	-2.54 (.771)	.001
Criminal (Non-criminal)	Non-violent	-.133 (.066)	.044	.179 (.076)	.019
	Violent	-.356 (.067)	.000	-.018 (.079)	.822
Race (White)	Black	-.709 (.076)	.000	1.55 (.936)	.098
Aggressive				-.276 (.162)	.091
Trustworthy				.855 (.200)	.000
Attractive				-.112 (.118)	.339
Wealthy				1.002 (.264)	.000
Affect				.350 (.102)	.001
fWHR				.013 (.197)	.946
Race*Aggressive				-.046 (.194)	.815
Race*Trustworthy				-.812 (.261)	.002
Race*Attractive				.714 (.156)	.000
Race*Wealthy				-.913 (.285)	.001
Race*Affect				.080 (.124)	.520

Overall, ratings of criminality were 24% in the non-criminal condition (SD = 0.430) and

29% in the criminal condition (SD = 0.455). Thus, participants were less likely to find that the

non-criminal foils would be criminals in the future than the men in the criminal photographs (see

Figure 4.1). Furthermore, ratings of criminality were 27% in the non-violent criminal condition

(SD = 0.444) and 31% in the violent criminal condition (SD = 0.464). Participants were less

likely to find that the non-violent criminals were criminal than the violent criminals. Finally,

ratings of criminality were 34% in the White condition (SD = 0.473) and 20% in the Black

condition (SD = 0.400). Participants were more likely to find that White men were criminal than

Black men.

	Non-violent	Violent	Non-criminal	Non-violent	Violent	Non-criminal
		White			Black	
No	63.82%	58.01%	71.79%	82.28%	79.15%	79.28%
Yes	36.18%	41.99%	28.21%	17.72%	20.85%	20.72%
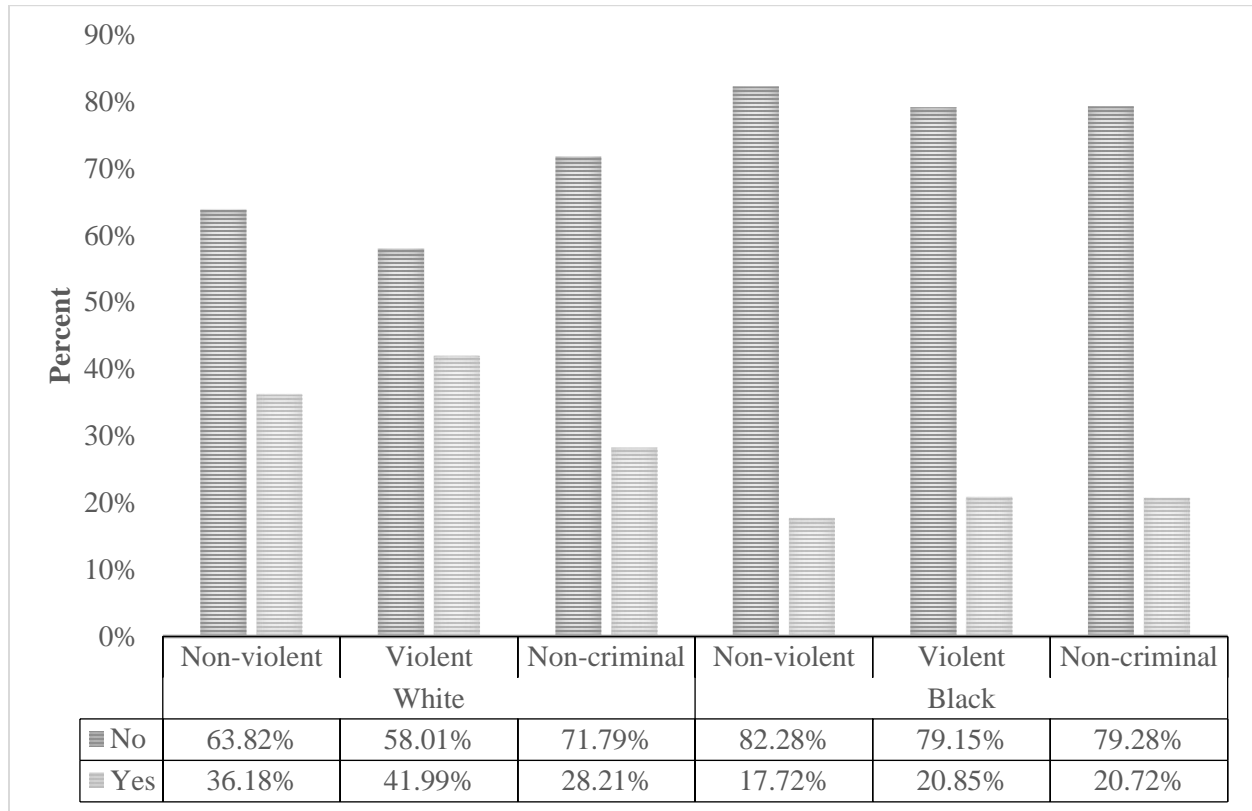
Figure 4.1. Percent of judgments in each condition.

A GEE model was executed to examine the effect of photograph characteristics

(including the interaction between race and other characteristics) on criminality judgment. The

odds of a participant predicting criminality increased when the photograph was rated as less

trustworthy, $\chi^2 (1, n = 156) = 18.260$, $p < 0.001$, less wealthy, $\chi^2 (1, n = 156) = 18.260$, $p < 0.001$,

and less happy, $\chi^2 (1, n = 156) = 11.721$, $p = 0.001$. There was a significant interaction of race

and attractiveness, $\chi^2 (1, n = 156) = 20.827$, $p < 0.001$. Simple main effects analysis showed that

White faces were rated as non-criminal were more attractive (M = 1.78, SD = 0.56) than the

faces rated criminal (M = 1.55, SD = 0.56), whereas there was no difference in attractiveness

between the non-criminal and criminal Black faces (M = 2.08, SD = 0.58; M = 1.97, SD = 0.48).

There was also a significant interaction of race and trustworthiness, χ^2 (1, n = 156) = 9.687, p = 0.002. Simple main effects analysis showed that White faces rated as non-criminal were viewed as more trustworthy (M = 1.88, SD = 0.40) than faces rated criminal (M = 1.72, SD = 0.39), but the difference was greater for Black faces rated criminal versus non-criminal (M = 2.32, SD = 0.39; M = 2.08, SD = 0.39). There was also a significant interaction of race and wealth, χ^2 (1, n = 156) = 10.265, p = 0.001. Simple main effects analysis showed that Black faces rated as non-criminal were viewed as more wealthy (M = 1.96, SD = 0.24) than the faces rated criminal (M = 1.85, SD = 0.22), whereas there was no difference in wealth between the non-criminal and criminal White faces (M=2.48, SD = 0.39; M = 2.41, SD = 0.46).

Another GEE model was executed to examine the effect of participant demographic characteristics on criminality assessment. Participant characteristics were not significant predictors of their judgments of criminality.

Participants' accuracy was also examined, both overall and per photograph. The average overall accuracy was 52.85% correct (SD = 3.06). Mean accuracy for the non-criminal photographs was 75.29% correct (SD = 3.58) and 28.94% correct (SD = 3.27) for the criminal photographs. Given that each judgment was a 0–1 decision, t-tests were conducted to determine whether participants performed at above chance for each condition by comparing the accuracy scores from each condition and 0.5 as the at-chance level. Participants performed above chance overall, t(155) = 3.45, p = 0.001. Performance was above chance for the non-criminal photographs, t(155) = 17.160, p < 0.01. Performance was below chance for the non-criminal photograph, t(155) = -16.093, p < 0.01. This again demonstrates a non-criminal bias, with participants tending to judge faces as non-criminal rather than criminal thus inflating performance in the non-criminal condition. Participants also performed above chance for the

White photographs, t(155) = 4.779, p < 0.01. Mean accuracy for the White photographs was 55.19% (SD = 2.16). However, participants performed at chance for the Black photographs, t(155) = -1.111, p = 0.268. Mean accuracy for the Black photographs was 49.04% (SD = 0.17).

GEE models were used to test whether photograph-level accuracy differed as a measure of condition. All findings for accuracy are located in Table 4.4. Participants were more likely to give a correct response when shown criminal photographs, either non-violent, χ^2 (1, n = 156) = 274.448, p < 0.001, or violent, χ^2 (1, n = 156) = 246.234, p < 0.001. Participants were also more likely to be correct when shown White photographs, χ^2 (1, n = 156) = 31.129, p < 0.001.

Another GEE model was used to test whether photograph-level accuracy differed as a measure of condition and photograph characteristics. Participants' odds of giving a correct response continued to decrease when shown criminal photographs, either non-violent, χ^2 (1, n = 156) = 308.779, p < 0.001, or violent, χ^2 (1, n = 156) = 262.536, p < 0.001. Participants were also more likely to be correct when shown Black photographs, χ^2 (1, n = 156) = 9.749, p = 0.002. In addition, participants' likelihood of giving a correct response decreased when shown a more aggressive-looking photograph, χ^2 (1, n = 156) = 4.522, p = 0.033, increased when shown a more attractive-looking photograph, χ^2 (1, n = 156) = 15.710, p < 0.001, increased when shown a more wealthy-looking photograph, χ^2 (1, n = 156) = 4.348, p = 0.037, decreased when shown a more happy-looking photograph, χ^2 (1, n = 156) = 13.422, p < 0.001, and increased when shown a face with a higher fWHR, χ^2 (1, n = 156) = 3.861, p = 0.049.

Another GEE model was used to test whether photograph-level accuracy differed as a measure of condition, confidence, and participant demographic characteristics. Participants' odds of giving a correct response decreased when the participant was shown criminal photographs, either non-violent, χ^2 (1, n = 156) = 270.247, p < 0.001, or violent, χ^2 (1, n = 156) = 247.878, p <

0.001. Participants odds of a correct response decreased when the participant was shown Black

photographs, χ^2 (1, n = 156) = 30.665, p < 0.001, and increased when the participant was non-

White, χ^2 (1, n = 156) = 6.451, p = 0.011.

Table 4.4
Generalized estimating equations for accuracy

Predictors	Comparison Group	Model 1		Model 2 Photograph Characteristics		Model 3 Participant Characteristics	
		B (SE)	p	B (SE)	p	B (SE)	p
Intercept		-.976 (.068)	.000	-1.940 (.512)	.000	-.719 (.491)	.142
Criminal(Non-criminal)	Non-violent	2.136 (.129)	.000	2.352 (.134)	.000	2.14 (.131)	.000
	Violent	1.918 (.122)	.000	2.061 (.127)	.000	1.945 (.124)	.000
Race (White)	Black	-.314 (.056)	.000	-.256 (.082)	.002	-.315 (.057)	.000
Aggressive				-.196 (.092)	.033		
Trustworthy				.069 (.110)	.530		
Attractive				.295 (.074)	.000		
Wealthy				.203 (.097)	.037		
Affect				-.256 (.070)	.000		
fWHR				.350 (.178)	.049		
Sex (Male)	Female					.031 (.067)	.642
Race (Non-white)	White					.136 (.054)	.011
Age						-.009 (.023)	.687
Political Orientation						-.037 (.025)	.139
Victim (Yes)	No					.234 (.138)	.090
Confidence						.001 (.003)	.783

Discussion

In this study, participants were asked to predict whether high school students would have

criminal records later in life. In contrast to the previous two studies, participants were shown

photographs of both Black and White students, as well as students who later had violent or non-violent criminal records. Overall, participants once again demonstrated the ability to correctly distinguish between criminal and non-criminal photographs. Participants treated both non-violent and violent criminal photographs differently than the non-criminal photographs, suggesting that people have a sensitivity towards all types of criminality, not just physically threatening behaviors. Furthermore, participants demonstrated the same non-criminal bias that was shown in the previous study; participants more frequently rated the photographs as non-criminal as opposed to criminal, despite half of the photographs depicting students who would later have criminal records

Participants also responded differently to the photographs in the Black and White conditions. Participants were more likely to judge the White students as criminal than the Black students. This finding is somewhat surprising given previous research that shows an association between stereotypical Black facial features and criminality (Blair et al., 2004; Kleider-Offutt et al., 2017). One possible explanation for this departure is participants' non-criminality bias. If participants were uncertain about the character of the Black students, they may have been more likely to judge the face as non-criminal by default. Furthermore, the participants from this study came from a university with mandatory implicit bias training for all students. In the context of this experiment, the participants may have been aware of their own implicit racial biases and wished to perform the task in a socially desirable way. This may have resulted in cognitive effort to reduce associations between stereotypical Black features and criminality. More research needs to be done to explore the role of bias in both judgments of criminality and accuracy. For example, future studies should include measures of racial bias in order to see if racial bias would predict more severe judgments in the Black conditions.

The results do not simply show that participants rated the White faces as more criminal than the Black faces. Accuracy also differed between conditions. Participants correctly judged the White criminal faces as more criminal than the White non-criminal faces but incorrectly judged all Black faces as equally non-criminal. One explanation for the disparity in performance between the Black and White conditions is the cross-race effect, which states that people struggle to recognize emotion and subtle differences in other-race faces (Bothwell et al., 1989; Shriver et al., 2008; Tanaka et al., 2004). Participants gave consistent judgments of criminality in the Black condition, despite whether the photograph depicted a criminal or not. Their performance in this condition was at chance, suggesting that the participants were simply guessing when making their judgments. In comparison, in the White condition, participants were accurately able to distinguish between the criminal and non-criminal faces at levels above chance. Thus, participants were able to accurately recognize features of criminality in White faces but failed to recognize the same subtleties in Black faces. The finding of the cross-race effect is further supported by the fact that non-White participants performed better overall at the task than White participants. Non-White participants demonstrated a better grasp of the subtleties of Black faces which led to more accurate performance (Bothwell et al., 1989).

In comparison to the previous studies, characteristics of the photographs had a larger influence on participants' judgments. Participants were more likely to judge a photograph as criminal if the man in the photograph was rated as less trustworthy, less wealthy, and less happy. Furthermore, the race of the photographed men interacted significantly with measures of attractiveness, trustworthiness, and wealth. Altogether, this suggests that participants were relying on other features of the photographed students to make their judgments, including those that make up the stereotypes about what a criminal "looks" like. These stereotypes differed based

43

on the race of the target, with attractiveness influencing judgments of criminality for White faces while perceived wealth influenced judgments of criminality for Black faces. These results differ somewhat from previous research, which found that attractiveness predicted trustworthiness only in out-group members (Sofer et al., 2017; Stolier, et al., 2018). Beyond race, accuracy decreased when the students in the photographs looked more aggressive, more attractive, or happier. This suggests that certain shared stereotypes about what a criminal looks like (aggressive, unattractive, and unhappy) did not predict actual criminality in this study. Thus, although participants performed above chance overall at the task, some of their instincts about criminality were misleading and led to inaccurate judgments.

Performance in this study has implications for the criminal justice system. The results suggest that participants—at least participants from this sample, including a majority of White participants—are able to correctly distinguish between criminal White men and non-criminal White men, but are unable to make the same distinction between Black men. If the ability to make accurate attributions of character is partially owed to the importance of recognizing threats when interacting with new people (Carre & McCormick, 2008; Flowe, 2012; Haselhuhn & Wong, 2012; McArthur & Baron, 1983), then it follows that only certain White men would be considered threats—the ones who look like a criminal. In contrast, because the same ability does not exist when shown Black men, *any* Black man would be perceived as a potential threat. This may lead to increased paranoia and concern about the behavior of Black men, as demonstrated by the recent attention given to White people who have reported Black people to the police for relatively innocuous behaviors, like having a barbeque or sleeping in a library (Fearnow, 2018; Wootson, 2018). The inability to distinguish between someone who is trustworthy and someone who poses a real threat creates a world in which any unknown person becomes threat. If this

inability only exists when judging out-group members, it could lead to arbitrary treatment and disproportionate convictions for members of a minority group (e.g., Mitchell et al., 2005; Pettit & Western, 2004).

This conclusion is supported by the preparation of this study, in which only 7.5% of the usable Black yearbook photographs were found to have no criminal record compared to 21.5% of the White yearbook photographs. The potentiality for any Black man to be perceived as a threat could result in more minor contacts with the police, which in turn would result in a greater likelihood of having any kind of criminal record compared to someone who appears more trustworthy. More research should be done to further explore this conclusion and see if implicit bias training, which is often proposed to counteract the influence of racial bias (Smith, 2015), would reduce this threat-sensitivity at earlier stages of the criminal justice process—like making referrals to the police or making arrests.

The results of this and the previous two studies have shown that participants are able to predict criminality at levels above chance overall. The question remains about whether participants are truly responding to criminality or if the task is capturing a different characteristic like convictableness. The next step must be to explore whether judgments of criminality result in a self-fulfilling prophecy in which criminal-looking men are treated like criminals and thus receive criminal records. To do this, the study discussed in the following chapter will examine how participants respond to the photographs of men who are completely innocent of a crime but nevertheless have a criminal record: exonerees.

CHAPTER 5

STUDY 4: EXONERATED FACES

The previous three studies supported previous research which showed people have reliable and shared stereotypes about criminality, including who looks "criminal" and who doesn't (Funk et al., 2017). These stereotypes are strong enough that participants can predict—at levels above chance—whether a high school senior will later have a criminal record or not. The previous findings suggest that accuracy may be due to a stereotype about what an innocent person looks like rather than what a criminal looks like and that accuracy may decrease when judging members of an out-group. It is possible that these psychological biases or stereotypes may lead someone to treat untrustworthy-looking people as if they are guilty of some sort of illicit act. These stereotypes may contribute to a self-fulfilling prophecy that results in a criminal conviction, thus altering the lives of people who are unlucky enough to resemble the stereotype of a "criminal" or perhaps who do not match the stereotype of an "innocent" person. This may lead to fewer opportunities, more referrals to the criminal justice system, and amplified punitiveness.

One example of how this process may work is found in a study by Ewing et al. (2013). In the study, children were asked to play a game that involved investing money in other players, and the researchers found that children invested less money in players who had untrustworthy faces. Due to their untrustworthy appearance, those players had fewer positive opportunities in the game and were given fewer chances to collaborate with others. The same may be true within the criminal justice system. If someone appears more criminal (or less innocent), others may be more likely to call the police, the police may be more likely to make an arrest, and judges may be

46

more likely to convict. Thus, based on appearance alone, one person may be more likely to have a criminal record than someone else who appears more innocent.

In fact, researchers have found that a criminal appearance does influence judgments in the criminal justice context. For example, the appearance of criminality, as indicated by facial tattoos, led to more severe sentencing decisions in one study (Funk & Todorov, 2013). Furthermore, participants remember stereotypically "criminal" faces better and are more likely to identify "criminal" faces from a lineup when given no description of the perpetrator (Flowe & Humphries, 2011). This bias works in two directions. First, people are more likely to identify someone in a lineup who looks like they may be a criminal. Second, people are less likely to identify someone from a lineup if the person does not look criminal due to concern that the identification may be incorrect. Thus, not only is criminal appearance a risk factor for a conviction, but a more innocent appearance may be protective against the same fate.

In one study of sentencing decisions in real criminal cases, researchers found that convicted criminals who appeared less trustworthy were more likely to be sentenced to death for first-degree murder than convicted criminals who appeared more trustworthy (Wilson & Rule, 2015). This association was true even when the authors compared the sentences of exonerated men. The researchers found that exonerated men with less trustworthy faces received more severe sentences than exonerated men with more trustworthy-appearing faces who were convicted of the same charges (Wilson & Rule, 2015). Because all the exonerated men were innocent of the charged crimes and thus their character did not contribute to the crime itself, this finding suggests that it is not entirely the facts of the case that influence sentencing. Instead, psychological bias against those who look untrustworthy contributes at least somewhat to sentencing decisions. One limitation to this study is that it used photographs of exonerated men

taken after their exoneration and therefore may not be a good indicator of how the men looked during their trials, which often occurred several decades earlier.

The following study was designed to further investigate how a criminal appearance may influence involvement in the criminal justice system. To test this, participants were asked to make judgments of guilt for a series of mugshots depicting either exonerated men or criminal foils. In the previous studies, participants were able to successfully distinguish between criminals and non-criminals. Participants should be able to perform with the same level of accuracy in this task because it is essentially the same as the previous experiments: comparing innocent men to criminal men. If participants are less accurate in this study, then the findings will support the theory that criminal appearance creates a self-fulfilling prophecy that makes some men more vulnerable to arbitrariness in the criminal justice system, including the risk of false convictions. For this study, two different participant samples were used: one group of workers from Amazon Mechanical Turk (MTurk) and one group of students from a large university.

Methods

Materials

In order to study participants' responses to innocent men who may nevertheless look criminal, the research team selected 10 mugshots of men who were convicted of violent crimes (e.g., murder, rape) and who were later exonerated due to DNA evidence. Seven photographs depicted White men, two depicted Black men, and one depicted a Latino man. With the exception of the exonerated crime, none of the men had a prior criminal record. Thus, although the men in the photographs were convicted like criminals, they had no documented history of crime. The research team also selected 10 photographic foils which depicted men who were convicted of either murder or rape due to DNA evidence: the criminal foils. As with the

photographs of exonerated men, seven of the foils were White men, two were Black men, and one was a Latino man. The photographs from both groups were taken between 1978-2000 and were of comparable quality. Due to the limited availability of photographs of exonerated men, several photographs in each group depicted men with moustaches. The photographs were cropped to show mostly the face of the man so that clothing style would not influence participants' judgments. The photographs were pilot tested with a team of undergraduate research assistants to ensure that none of the men were recognizable.

Figure 5.1. Two of the photographs used in this study. At the left is Colin Pitchfork, a foil who was convicted in 1988, and at the right is Craig Coley, who was exonerated in 2017. All photographs used in this study are located in Appendix D.

This study returns to the class criminal attribution paradigm of using booking photographs or mugshots. This was done to reduce any impact of the sentencing decision or time in prison on the appearance of the exonerated men. Nevertheless, there is a possibility of a perceptible ephemeral difference in emotion between the exonerees' mugshots and the criminal foils' mugshots. After all, exonerated men likely experience different emotions during their arrests than men who know that they are guilty of the accused crime. This may result in more negative emotional expressions, different posture, or disparate levels of eye contact.

In order to address this, a separate group of participants rated all photographs for a variety of characteristics other than criminality. A total of 25 participants were recruited from MTurk and were paid for their services. The participants rated all 20 photographs for the following characteristics: aggression, trustworthiness, attractiveness, wealth, and affect. Each characteristic was measured on a likert scale from 1-5, with 1 being the least like that characteristic (e.g., extremely unattractive) and 5 being the most like that characteristic (extremely attractive). Mean ratings for all characteristics are located in Table 5.1

Table 5.1.
Mean ratings of characteristics

	Exonerees		Foils	
	Mean	SD	Mean	SD
Aggression	3.341	.534	3.521	.305
Trustworthiness	2.527	.370	2.41	.353
Attractiveness	2.890	.326	2.769	.378
Wealth	2.604	.389	2.434	.332
Affect	2.629	.454	2.575	.620
fWHR	1.82	.131	1.86	.127

In addition, research assistants used the software ImageJ to measure the fWHR of all 20 photographs following the same methods as in Study 3. Inter-rater reliability was moderately high for all measurements (distance between left and right zygion: $r = 0.997$, $p < 0.001$; distance between upper lip and brow: $r = 0.996$, $p < 0.001$; width-to-height ratio: $r = 0.843$, $p < 0.001$). fWHR did not significantly differ between the exoneree and criminal photographs.

Participants and Procedures

A total of 132 participants were recruited from MTurk and paid for their participation. To qualify, workers had to reside in the United States. Participants were excluded if they missed an attention check or provided answers that suggested they were not human. This left 88 total participants (57 Male, 31 Female; 72 White, 16 non-White). The participants' ages ranged from 22 to 73 (M = 34.48, SD = 9.22).

Participants were told they would be shown a series of mugshots from real criminal cases and that they would be asked to judge whether the men in the photographs were guilty of serious violent offenses or not. Participants were then shown all 20 photographs in random order. They were not told what proportion of the photographs belonged to each category or how many photographs they would be shown. Participants were asked to make a dichotomous judgment of guilt (Yes or No). After each judgment, participants were asked to rate their confidence in their answer on a scale from 1 (Extremely confident) to 5 (Extremely uncertain). At the end of the survey, participants were asked to estimate their overall accuracy. Then, participants completed the same demographic information as in the previous studies.

<div align="center">Results</div>

Participants' responses were examined to see what influenced their judgments of criminality. A generalized estimating equation (GEE) was carried out with dichotomous criminality assessments for each photograph as the dependent variable, photograph condition (exoneree or criminal) as a within-subject factor, and controlling for photograph ID and participant ID. All GEE findings for judgments are located in Table 5.2. The likelihood of a participant saying that the man in the photograph was guilty of a violent crime increased when the participant was shown a criminal in comparison to an exoneree, $\chi^2 (1, n = 880) = 28.866$, $p < 0.001$. Overall, ratings of criminality were 53% in the exoneree condition (SE = 0.023) and 64% in the criminal condition (SD = 0.022). Thus, participants found the criminals to be more likely to be guilty than the exonerated men (see Figure 5.2).

	Exoneree	Criminal
▤ No	414	314
▤ Yes	466	566
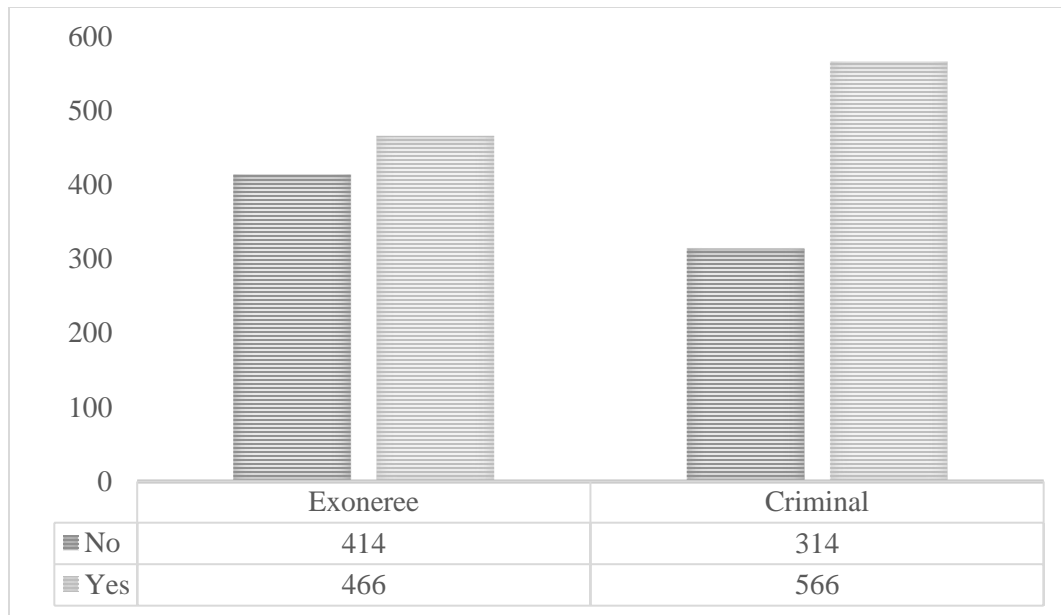
Figure 5.2. Aggregate decisions across all photographs for dichotomous ratings of criminality. "Yes" indicates a judgment of criminality.

A GEE was carried out with dichotomous criminality assessments for each photograph as the dependent variable, photograph condition (exoneree or criminal) as a within-subject factor, and controlling for the possible mediating characteristics of the men in the photographs (aggression, trustworthiness, attractiveness, wealth, affect, and fWHR). The results revealed that photograph condition was a significant predictor of guilt judgments, χ^2 (1, n = 1760) = 4.651, p = 0.031. Participants were more likely to say that the criminal foils had committed serious criminal acts than the exonerees. There was also a significant effect of aggression, χ^2 (1, n = 1760) = 9.789, p = 0.002, and attractiveness, χ^2 (1, n = 1760) = 5.807, p = 0.016. The likelihood of a participant rating a photograph as guilty increased with the more aggressive the man in the photograph appeared and decreased with the more attractive the man in the photograph appeared.

Additional GEEs were carried out to measure the effect of participant characteristics on criminality assessments. There were no effects of participant characteristics on judgments.

Table 5.2
Generalized estimating equations for judgments made by MTurk participants

Predictors	Comparison Group	Model 1		Model 2 Photograph Characteristics	
		B (SE)	p	B (SE)	p
Intercept		.589(.097)	.000	-1.034 (1.661)	.534
Criminal (Criminal)	Exoneree	-.471 (.087)	.000	-.244 (.113)	.031
Aggressive				.782 (.250)	.002
Trustworthy				-.372 (.323)	.249
Attractive				-.397 (.165)	.016
Wealthy				.032 (.217)	.882
Affect				.122 (.250)	.626
fWHR				.270 (.525)	.606

Participants' accuracy was also examined, both overall and for each photograph. The average overall accuracy was 55.68% correct (SD = 1.98). Mean accuracy for the photographs of exonerees was 47.04% correct (SD = 2.21) and 64.31% correct (SD = 2.12) for the criminal photographs. Given that each judgment was a 0–1 decision, t-tests were conducted to determine whether participants performed at above chance for each condition by comparing the accuracy scores from each condition and 0.5 as the at chance level. Participants performed above chance overall, $t(87) = 5.389$, $p < 0.001$. Participants performed above chance for criminal photographs, $t(87) = 6.347$, $p < 0.01$. Participants did not perform significantly different from chance when shown the exoneree photographs.

GEE models were used to test whether photograph-level accuracy differed as a measure of condition. Participants were more likely to be correct when shown a photograph of a criminal foil, χ^2 (1, n = 1760) = 17.166, $p < 0.001$. Results from all GEE models for accuracy are located in Table 5.3.

Table 5.3
Generalized estimating equations for accuracy of MTurk participants

Predictors	Comparison Group	Model 1		Model 2 Photograph Characteristics		Model 3 Participant Characteristics	
		B (SE)	p	B (SE)	p	B (SE)	p
Intercept		-.589 (.097)	.000	-7.674 (1.421)	.000	-.838 (.281)	.003
Criminal (Criminal)	Exoneree	.708 (.171)	.000	1.095 (.184)	.000	.652 (.171)	.000
Aggressive				1.798 (.255)	.000		
Trustworthy				-.082 (.300)	.785		
Attractive				.008 (.164)	.963		
Wealthy				-.790 (.216)	.000		
Affect				1.525 (.256)	.000		
fWHR				-.585 (.551)	.289		
Sex (Male)	Female					.087 (.093)	.351
Race (Non-white)	White					.119 (.108)	.272
Age						-.005 (.006)	.363
Political Orientation						-.012 (.017)	.468
Victim (Yes)	No					.273 (.109)	.013
Confidence (photograph)						.083 (.041)	.044
Confidence (overall)						.014 (.010)	.164

A GEE model was also used to test whether photograph-level accuracy differed as a measure of condition and characteristics of the photograph (aggression, trustworthiness, attractiveness, wealth, affect, and fWHR). The odds of a participant making an accurate judgment of guilt increased when shown a criminal photograph, χ^2 (1, n = 1760) = 35.59, p <

0.001, and when the photograph was rated as more aggressive, χ^2 (1, n=88) = 49.696, p < 0.001, less wealthy, χ^2 (1, n = 1760) = 13.361, p < 0.001, and more positive in mood, χ^2 (1, n = 1760) = 35.533, p < 0.001.

Another GEE model was used to test whether photograph-level accuracy differed based upon participants' characteristics and confidence. The odds of a participant making an accurate judgment of guilt increased when shown a criminal photograph, χ^2 (1, n = 1760) = 14.525, p < 0.001, when the participant had been a victim of a serious crime, χ^2 (1, n = 1760) = 6.222, p = 0.013, and when their confidence in their decision increased, χ^2 (1, n = 1760) = 4.068, p = 0.044.

Replication

This study was replicated using the exact same materials and procedures but utilizing participants from a large New York university instead of MTurk workers. This was done to compare performance across two different populations. A total of 74 participants were recruited from the university and given research credit for their participation (11 male, 63 female; 34 White, 19 Asian, 21 other). The participants' ages ranged from 18 to 24 (M = 19.50, SD = 1.45). They were given the exact same materials and instructions as in the first iteration of this study.

Results

Participants' responses were examined to see if condition influenced judgments of criminality. Overall, ratings of criminality were 56% in the exoneree condition (SD = 0.496) and 57% in the criminal condition (SD = 0.496). There was no significant difference between judgments in the exoneree condition compared to the criminal foil condition. Participants performed almost identically in both conditions and demonstrated a slight preference of judging the men in the photographs as guilty of a crime.

	Exoneree	Criminal
No	322	317
Yes	418	423
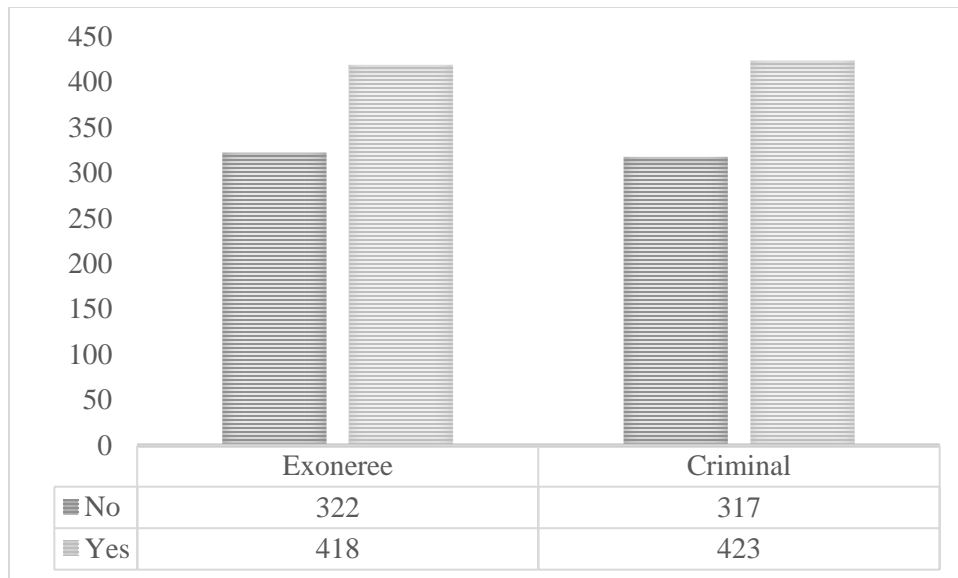
Figure 5.3. Aggregate decisions across all photographs for dichotomous ratings of criminality. "Yes" indicates a judgment of criminality.

To assess whether factors other than criminality influenced participants judgments of guilt, a GEE was carried out with dichotomous criminality assessments for each photograph as the dependent variable, photograph condition (exoneree or criminal) as a within-subject factor, and controlling for the possible mediating characteristics of the men in the photographs (aggression, trustworthiness, attractiveness, wealth, affect, and fWHR). Results for GEE all models for judgments are located in Table 5.4. The results revealed that there was a significant effect of trustworthiness, χ^2 (1, n = 1480) = 5.078, p = 0.024, attractiveness, χ^2 (1, n = 1480) = 27.701, p < 0.001, perceived wealth, χ^2 (1, n = 1480) = 20.149, p < 0.001, and affect, χ^2 (1, n = 1480) = 5.090, p = 0.024. The odds of a participant rating a photograph as guilty increased with the less trustworthy, less attractive, and less wealthy the man in the photograph appeared as well as when the man in the photograph's affect was judged as more positive. This suggests that participants may have relied on certain characteristics of the photographs to make a decision, but those characteristics were not correlated with actual guilt. Additional GEEs were carried out to

measure the effect of participant characteristics on criminality assessments. There were no effects of participant characteristics on judgments.

Table 5.4
Generalized estimating equations for judgments made by university participants

Predictors	Comparison Group	Model 1		Model 2 Photograph Characteristics	
		B (SE)	p	B (SE)	p
Intercept		-.288 (.076)	.000	-3.046 (1.734)	.079
Criminal (Criminal)	Exoneree	.028 (.082)	.738	-.069 (.115)	.548
Aggressive				-.173 (.328)	.598
Trustworthy				.716 (.318)	.024
Attractive				1.056 (.201)	.000
Wealthy				-1.082 (.241)	.000
Affect				.606 (.269)	.024
fWHR				-.128 (.631)	.840

Participants' accuracy was also examined, both overall and for each photograph. The average overall accuracy was 50.35% correct (SD = 1.74). Mean accuracy for the photographs of exonerees was 43.50% correct (SD = 1.73) and 57.20% correct (SD = 1.62) for the criminal photographs. Given that each judgment was a 0–1 decision, t-tests were conducted to determine whether participants performed at above chance for each condition by comparing the accuracy scores from each condition and 0.5 as the at chance level. Overall, participants performed at chance levels. Participants performed below chance for the exoneree photographs, $t(73) = -3.223$, $p = 0.002$, and above chance for the criminal photographs, $t(73) = 3.809$, $p < 0.001$. This performance seems to be driven by a criminality bias that led to participants being more likely to rate all photographs as criminal, thus increasing performance in the criminal condition.

GEE models were used to test whether photograph-level accuracy differed as a measure of the characteristics of the photograph (aggression, trustworthiness, attractiveness, wealth, affect, and fWHR). The odds of a participant making an accurate judgment of guilt increased when shown a criminal photograph, $B = -0.925$, $SE = 0.163$, χ^2 $(1, n = 1760) = 32.462$, $p < 0.001$, and when the photograph was rated as more aggressive, $B = 2.495$, $SE = 0.316$, χ^2 $(1, n=1760) = 62.247$, $p < 0.001$, was rated as more positive in mood, $B = 2.499$, $SE = 0.315$, χ^2 $(1, n = 1760) = 62.79$, $p < 0.001$, and had a higher fWHR, $B = -1.526$, $SE = 0.589$, χ^2 $(1, n = 1760) = 6.723$, $p = 0.10$. The same judgments of aggression and affect predicted photograph-level accuracy for the MTurk participants as well. Another GEE model was used to test whether photograph-level accuracy differed based upon participants' characteristics and confidence. No significant predictors were found.

Discussion

In this study, participants were asked to judge the guilt of men depicted in a series of mugshot photographs. Half of the photographed men were later exonerated for the crimes for which they had been arrested, and the other half were determined to be true criminals. The exonerated men have no inherent criminality but were nevertheless convicted of a crime. The purpose of the study was to see if participants were able to distinguish between the exonerated and criminal men with the same level of accuracy they had shown in previous experiments. If so, that would help reject the hypothesis that the results shown in the previous chapters were due to participants' understanding of what a "convictable" face looks like and those stereotypes contribute to a self-fulfilling prophecy that results in criminal-looking men receiving a criminal conviction. If participants perform at chance in this experiment, then it would suggest certain

men—like exonerated men—are especially vulnerable to criminal accusations due to their facial appearance.

This study used two different participant samples: a sample from MTurk and a sample from a large university. The two samples performed quite differently on the task. The MTurk workers were able to accurately distinguish between the exonerated men and the criminal foils, even when accounting for other characteristics like attractiveness and affect. In contrast, the university sample performed at chance and were essentially guessing when making their determinations of guilt. This is despite the fact that participants from the same population performed above chance in Studies 1-3 when asked to judge non-criminal and criminal faces. The results of the university sample suggest that it is possible that a self-fulfilling prophecy created by criminality stereotypes exists and that criminal sentencing is at least somewhat arbitrarily influenced by facial appearance. This arbitrariness may put certain people at unfair risk of a conviction if, like the exonerated men in this study, they are undistinguishable from criminal foils. This bias may also unfairly benefit others whose appearance do not meet these stereotypes.

There are several possible explanations for why the MTurk workers were more accurate than the university students. First, the composition of the two groups are different, with MTurk workers tending to be older, more White, and more male. The results of the previous studies suggest that in-group status may improve performance on this task. For example, men may be better able to distinguish between facial features of other men than women can. Thus, although participant characteristics did not predict accuracy for either sample in this study, it is possible that the MTurk population was better suited for the task due to being more demographically similar to the men in the photographs. Second, although the photographs were pilot tested with

undergraduate research assistants to ensure that none of the men were recognizable, several MTurk workers mentioned that they recognized some of the criminal foils. It is possible that the men in the mugshots, whose arrests happened largely in the 1980s or 1990s, may be recognizable to an older population of participants but not to college-aged participants. This recognition may have led to better performance in the foil condition and thus allowed for better accuracy overall. This would be consistent with the data, which shows that the improved accuracy for the mTurk workers was largely was attributable to their performance in criminal condition and not the exoneree condition. Finally, it is important to note that the validity of data from MTurk has been a subject of some debate among researchers (Kees et al., 2017; Vakharia & Lease, 2015). Although high performance on this task does not suggest illegitimacy and the research team attempted to remove all workers who failed to meet attention checks, this replication may contribute to the discussion about the merits of different recruitment tools in psychological research.

It appears that all participants were only able to guess when shown the photographs of the exonerated men whereas the MTurk sample made accurate judgments for the criminal foils at levels above chance. In contrast, in the previous experiments, participants were both the most confident and the most accurate when shown innocent faces. Thus, even for the MTurk population, it is possible that exonerated men may be judged differently than other innocent faces. A future direction for this study would be to compare innocent men who have been photographed in similar situations but whom were never convicted of a crime to determine if participants perform differently between exonerees (who were determined to be guilty) and innocent foils (who were not determined to be guilty). If performance between these two groups

differ, it would provide additional evidence that the fates of exonerated men were at least somewhat influenced by their appearance.

For both groups, the odds of a participant accurately rating a photograph as guilty increased with the more aggressive the man in the photograph appeared and increased with the happier the man in the photograph appeared. This suggests that aggression and mood may help distinguish between true criminals and men who may be falsely convicted. After all, the emotions experienced during an arrest must differ greatly between a man who knows he is guilty and a man who knows he is innocent. It is possible that these differences can be detected and may help when making attributions of guilt based on appearance.

In contrast to the previous studies, participants demonstrated a criminal bias instead of a non-criminal bias. Participants more frequently guessed guilty rather than innocent in both conditions. This is likely due to the photograph context. Unlike the yearbook photos which depict tuxedo-wearing students, the photographs in this study were all mugshots. Participants were likely more comfortable making an accusation of guilt because they knew that someone else had already made this accusation as well. In addition, the base rates of criminality between a yearbook population and a mugshot population are different. It is more likely that a man who has been arrested is guilty of a crime than a man pulled randomly from a high school yearbook. This shift in bias—from non-criminal to criminal— may further demonstrate the danger of a convictable face. Once one person has made a judgment about a face, it can encourage other people to rely on these stereotypes as well.

Due to the different performance between the two participant samples, the results of this study were somewhat inconclusive. Nevertheless, the findings suggested that at least some populations may have difficulty determining true guilt when shown photographs of exonerated

61

men and criminals. This lends support to the idea that facial appearance may contribute to the arbitrariness of the criminal justice system, with more "criminal"-looking men more likely to receive convictions. In some cases, this may be true even if the man has not committed any type of crime.

CHAPTER 6

GENERAL DISCUSSION

The studies in this paper explored two questions related to attributions of criminality from facial appearance. First, is it possible to make accurate predictions of criminal behavior based on appearance? Previous research has found that people are able to accurately distinguish between criminal and non-criminal faces with high-levels of interrater reliability. The studies in this paper extended these findings and showed that accurate predictions of criminal behavior can be made absent from potential environmental contagion. The second question addressed in this paper was whether the accuracy of attributions of criminality was due to a self-fulfilling prophecy that makes certain criminal-looking individuals more vulnerable to discretion within the criminal justice system. The findings suggest that looking "convictable" may contribute to at least some arbitrariness in the criminal justice system.

In Studies 1-2, participants were asked to make judgments of future criminality based only on high school yearbook photographs, and participants performed at levels above chance overall. Participants were therefore able to accurately predict future criminal records with only knowledge of a person's facial appearance. Performance was mediated by a variety of characteristics of the men in the photographs, including perceived wealth, aggressiveness, and trustworthiness. In Study 3, the task was complicated with the introduction of a race condition, and participants' accuracy fell to at-chance levels for the photographs of Black men while remaining above chance for the photographs of White men. Finally, in Study 4, participants were asked to make judgments of guilt based on mugshots of exonerated men and criminal foils. On this task, Mechanical Turk (MTurk) participants performed above chance, while a student participant pool performed at chance. The accurate performance by MTurk workers may have

partially explained by recognition of some of the faces. The finding for the student sample supports the theory that, for at least some observers, facial appearance may create arbitrariness that leads to a higher risk of a conviction for people who fit stereotypes of criminality.

Throughout the first three studies discussed in this paper, the factor used to distinguish the "criminal" and "non-criminal" faces was whether the person in the photograph had a criminal record or not. Nevertheless, some people without public criminal records may have a criminal past that is either sealed, dismissed, or did not result in a criminal conviction. It is also possible that some people with criminal records were falsely accused or faced inaccurate charges. Therefore, the measurement of criminality used in this study may not be entirely accurate. In fact, what participants may have been responding to was how convictable the man in the photograph looked, a trait that is more closely linked to whether the man in the photograph had a public criminal record. In other words, participants may understand what the typical person who gets convicted of a crime looks like and these stereotypes may be correlated with who receives a conviction. This association may have inflated participants' accuracy at this task even if these stereotypes are not accurate portrayals of actual criminal behavior. Further research should be done that use self-reported measures of criminal or threatening behavior in order to further explore the possibility that participants are not sensitive to either criminality or innocence, but instead are sensitive to "convictableness" (i.e., how likely someone is to have a criminal conviction).

One interesting possibility suggested by the findings of the current research is that people may be more sensitive to the appearance of innocence than criminality. In Studies 1-3, participants were more accurate when shown non-criminal photographs. Furthermore, in Study 2, participants' confidence predicted accuracy but only for photographs in the non-criminal

condition. This sensitivity to innocence rather than criminality is supported by previous research that has found people are particularly attuned to trustworthiness in a face, as opposed to aggression or dominance (Chang et al., 2010; Porter et al., 2008). The heightened ability recognize trustworthiness rather than dominance is somewhat intuitive. While trustworthiness is a relatively stable characteristic, dominance can hold either positive or negative valence depending on the situation. An in-group leader's dominance may be a beneficial thing, while a stranger's dominance may be threatening. Thus, perception of dominance may be more nuanced than perception of trustworthiness. In terms of this paper, the heightened sensitivity to trustworthiness allowed participants to successfully recognize someone who is unlikely to be convicted of a crime, but participants may still struggle with knowing whether someone who looks like an aggressor is convictable or not.

Additionally, participants in Studies 1-3 demonstrated a strong non-criminal bias. Participants were more likely to say the yearbook photographs depicted someone who was not going to have a criminal record in the future than someone who was. This bias shifted in Study 4 when participants were asked to judge mugshot photographs rather than yearbook photographs. There, participants were more likely to say that the photograph depicted someone guilty rather than innocent of a crime. This difference in default judgment may be explained by the context of the photographs. Unlike yearbook photos, which depict smiling students, the photographs in Study 4 were clearly taken in the context of an arrest and some featured injured or disheveled-looking men. The base rates of criminal records between a yearbook population and a mugshot population are different, and this difference may shape participants' default guesses of criminality (Todorov & Peter, 2014). A participant would be more likely to be correct when guessing that a man in a randomly-selected mugshot was guilty of a crime than a man in a

randomly-selected yearbook photograph. Participants may also have been more comfortable making an accusation of criminality for someone who they know has already been arrested, like the men in the mugshot photographs. This shift in bias demonstrates how photograph context matters, at least when people are uncertain or the situation is ambiguous. If one person has already deemed someone guilty by making an accusation or an arrest, it may be harder for other people to look past their own biases and give the person a second chance. This may be especially true if the person does not have the appearance of an innocent person.

Furthermore, participants in Study 3 were less accurate at predicting criminality when judging the faces of Black men as opposed to White men. In the Black condition, participants performed no better than chance. As supported by the cross-race effect (Bothwell et al., 1989; Shriver et al., 2008; Tanaka et al., 2004), this finding suggests that participants may not have fleshed out stereotypes for what a non-criminal Black man looks like even if they can detect the same type of nuance in White faces. This has serious implications for the interaction of race within the criminal justice system. If people struggle to identify non-criminals amongst out-group members, then all members of an out-group become potential threats. Although ambiguous behaviors may not seem dangerous if conducted by in-group members, because the perceiver does not know what a "safe" out-group member looks like, the behavior becomes potentially threatening. In contrast, the perceiver would be able to accurately and quickly determine if an in-group member doing the same thing was a threat or just participating in an activity with a reasonable, safe explanation.

Some of the results in this paper suggest that familiarity with the target of the judgment may increase performance on this task. For example, in study 2, men performed better than women. Non-white people performed better than White participants on Study 3. This may be

66

because men are better able to recognize nuance other men and non-White people are better able to recognize nuance in Black faces. This finding supports the need for more representation within the criminal justice system at every level, from police officers to prosecutors to juries. If in-group status partially improves ability to recognize trustworthiness in others, then a more representative system would help reduce bias caused by different instinctive standards being used to evaluate different groups, like racial minorities. More research should be done to see if familiarity or in-group status does in fact improve ability to recognize personality characteristics like criminality from appearance.

In each study, other characteristics of the faces also influenced both accuracy and judgments. Some features stereotypically associated with crime, like aggression in Study 3, decreased performance while other features, like trustworthiness, increased performance. Looking at performance across all four studies, the relationships between these other characteristics and criminality remains unclear. Nevertheless, the results show that participants responded to other characteristics when making their judgments. This suggests that stereotypes about criminality are complicated and are influenced by other factors, like perceived wealth and attractiveness. A prediction of criminality is not simple or independent of other features.

More research should also be done to look at how stereotypes about criminality interact with gender and other demographics. Almost all research on criminal attributions have used male stimuli. There is little known about stereotypes of female criminals, and some of the literature suggests that those stereotypes would be different from male criminals. For example, fWHR is correlated with criminality attributions, but it is a sexually dimorphic trait and is not correlated with aggressive behavior for women (Haselhun & Wong, 2012). Although there was very little evidence that fWHR was an accurate predictor of criminality or influenced the judgments of

67

participants in the studies in this paper, this still suggests that women and men may have different facial indicators of deception or aggression. It is an important next step to see how the accuracy of criminality attributions hold up for other populations, including women, other races, and different age groups.

An additional future direction for this research is to see how different populations perform at this task. A finding from Study 2 suggests that victims may be better at recognizing criminality than other groups of people, but the influence of familiarity with potential criminals on accuracy should be further explored. Other populations who interact with potential criminals and are part of the discretionary process that may result in a conviction should be given this task as well. This could include prosecutors, police officers, and judges. Their performance on this task will help explain whether experience improves accuracy.

Finally, it is important to note that while participants performed above chance overall in all four studies discussed in this paper, there is no indication that perception of criminality is a reliable or perfect system of judgment. Although performance was above chance in all studies but the university replication of Study 4, overall accuracy never eclipsed 60% correct. Performance for some faces was much higher than other faces, but even on the photograph level participants never approached perfect judgment for any face tested. Furthermore, previous research suggests that context of the photograph and photograph selection alters judgments (Todorov & Peter, 2014). First impressions based on stereotypes can also be entirely wrong (Zebrowitz, 2017). The results of these studies are helpful for examining the potential arbitrary nature of discretionary judgments relevant to criminal records but should not be used to support the notion that criminality should be—or could be—determined based on someone's appearance.

Overall, the studies in this paper suggest that participants have some ability to predict who will end up with a criminal conviction and who will not. This ability may result in a more arbitrary criminal justice system given that any suspect of a crime must go through several stage of judgment before they receive a conviction. At each of these stages, there is some level of discretion that could be influenced by the suspect's appearance. The appearance of criminality may encourage a bystander to call the police, a police officer to make an arrest rather than give a warning, a prosecutor to pursue certain charges, and a judge to make a conviction. These appearance-based biases can be difficult to control, even if the decision maker is aware of and educated about the bias (Kleider-Offutt et al., 2017). Nevertheless, this field of research sheds important light on how appearance-based stereotypes can influence real world outcomes. As research in this field continues, it will be important to think about what influence appearance-based bias may have in different realms of the criminal justice system and what, if anything, can be done to protect those who have convictable faces.

REFERENCES

Anderl, C., Hahn, T., Schmidt, A., Moldenhauer, H., Notebaert, K., Clement, C. C., &
Windmann, S. (2016). Facial width-to-height ratio predicts psychopathic traits in males.
Personality and Individual Differences, *88*, 99-101.

Ayers, I., & Waldfogel, J. (1993). A market test for race discrimination in bail setting. *Standford
Law Review*, *46*, 987.

Berry, D. S., & Zebrowitz-McArthur, L. (1988). What's in a face? Facial maturity and the
attribution of legal responsibility. *Personality and Social Psychological Bulletin*, *14*, 23-
33.

Blair, I. V., Judd, C. M., & Chapman, K. M. (2004). The influence of afrocentric facial features
in criminal sentencing. *Psychological Science*, *15*(10), 674-679.

Boshyan, J., Zerowitz, L. A., Franklin, R. G. J., McCormick, C. M., & Carre, J. M. (2014). Age
similarities in recognizing threat from faces and diagnostic cues. *Journal of Gerontology:
Psychological Sciences and Social Sciences*, *69*(5), 710-718.

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

Bull, R. H. C., & Green, J. (1980). The relationship between physical appearance and
criminality. *Medical Science & Law, 20*(2), 79-83.

Burris, C., & Edwards, S. (2017). Does facial width-to-height ratio predict male offender
aggression?. *Journal of Criminal Psychology*, *7*(4), 280-286.

Carpinella, C. M., & Johnson, K. L. (2013). Appearance-based politics: Sex-typed facial cues
communicate political party affiliation. *Journal of Experimental and Social Psychology*,
49, 156-160.

Carre, J. M., & McCormick, C. M. (2008). In your face: facial metrics predict aggressive

    behaviour in the laboratory and in varsity and professional hockey players. *Proceedings*

    of the Royal Society B, *275*, 2651-2656.

Chang, L. J., Doll, B. B., van't Wout, M., Frank, M. J., & Sanfey, A. G. (2010). Seeing is

    believing: Trustworthiness as a dynamic belief. *Cognitive Psychology*, *61*(2), 87-105.

Cogsdill, F. J., Todorov, A.T., Spelke, F.S., & Banaji, M.R. (2014). Inferring character from

    faces: a developmental study. *Psychological Science*, *25*, 1132-1139.

Cormier, J. (2012). The influence of phenotypic variation on criminal judgment (PhD

    dissertation).

Correll, J., Urland, G. R., & Ito, T. A. (2006). Event-related potentials and the decision to shoot:

    The role of threat perception and cognitive control. *Journal of Experimental Social*

    Psychology, *42*(1), 120-128).

Cutler, B. L., Penrod, S. D., & Stuve, T. E. (1988). Juror decision making in eyewitness

    identification cases. *Law and Human Behavior*, *12*(1), 41-55.

Dotsch, R., Hassin, R. R., & Todorov, A. (2016). Statistical learning shapes face evaluation.

    Nature: Human Behavior, *1*, 1-6.

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. L. (2006). Looking

    deathworthy: Perceived stereotypicality of black defendants predicts capital-sentencing

    outcomes. *Psychological Science*, *17*, 383-386.

Ewing, L., Caulfield, F., Read, A., & Rhodes, G. (2015). Perceived trustworthiness of faces

    drives trust behaviour in children. *Developmental Science*, *18*(2), 327-334.

Flowe, H.D. (2012). Do characteristics of faces that convey trustworthiness and dominance

    underlie perceptions of criminality. *PLoS ONE*, *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*, 265-73.

Epp, C. R., Maynard-Moody, S. & Haider-Markel, D. P. (2014). *Pulled over*: *How police stops

    define race and citizenship*. University of Chicago Press.

Fearnow, B. (2018, May 10). Video: White woman calls police on Black family's BBQ for

    'trespassing' in Oakland park. Retrieved from https://www.newsweek.com/lake-merritt-

    bbq-barbecue-video-oakland-racist-charcoal-east-bay-black-family-919355.

Funk, F., & Todorov, A. (2013). Criminal stereotypes in the courtroom: Facial tattoos affect guilt

    and punishment differently. *Psychology, Public Policy, and Law*, *19*(4), 466-478.

Funk, F., Walker, M., & Todorov, A. (2017). Modelling perceptions of criminality and remorse

    from faces using a data-driven computational approach. *Cognitive and Emotion*, 1-13.

Gosling, S. D., Rentfrow, P. J., & Swann, W. B. (2003). A very brief measure of the Big-Five

    personality domains. *Journal of Research in Personality*, *37*, 504-528.

Halberstadt, A. G., Castro, V. L., Chu, Q., Lozada, F. T., & Sims, C. M. (2018). Preservice

    teachers' racialized emotion recognition, anger bias, and hostility attributions.

    Contemporary Educational Psychology, *54*, 125-138.

Haselhuhn, M. P., & Wong, E. M. (2012). Bad to the bone: facial structure predicts unethical

    behavior. *Proceedings of the Royal Society B*, *279*, 571-676.

Hu, S., Xiong, J., Fu, P., Qiao, L., Tan, J., & Tang, K. (2017). Signatures of personality on dense

    3D facial images. *Nature*, *7*(1), 73-83.

Jenkins, R., White, D., Van Montfort, X., & Burton, A. M. (2011). Variability in photos of the same face. *Cogntition*, *121*, 313-323.

Kees, J., Berry, C., Burton, S., & Sheehan, K. (2017). An analysis of data quality: professional panels, student subject pools, and Amazon's Mechanical Turk. *Journal of Advertising*, *46*(1), 141-155.

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

Kleider-Offutt, H., Bond, A. D., & Hegerty, S. E. A. (2017). Black stereotypical features: When a face type can get you in trouble. *Cognitive Directions in Psychological Science*, *26*(1), 28-33.

Kosinski, M. (2017). Facial width-to-height ratio does not predict self-reported behavioral tendencies. *Psychological Science*, 1-8.

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*(3), 574-589.

Leventhal, G., & Krate, R. (1977). Physical attractiveness and severity of sentencing. *Psychological Reports*, *40*, 315-318.

MacLin, M. K., & Herrera, V. (2006). The criminal stereotype. *North American Journal of Psychology*, *8*(2), 197-207.

Maclin, O. H., & MacLin, M. K. (2004). The effect of criminality on face attractiveness, typicality, memorability, and recognition. *North American Journal of Psychology*, *6*(1), 145-154.

73

Mazzella, R., & Feingold, A. (1994). The effects of physical attractiveness, race, socioeconomic status, and gender of defendants and victims on judgments of mock jurors: A meta-analysis. *Journal of Applied Social Psychology*, *24*(15), 1315-1344.

McArthur, L. Z., & Baron, R. M. (1983). Toward an ecological theory of social perception. *Psychological Review*, *90*(3), 215-238.

Mitchell, T. L., Haw, R. M., Pfeifer, J. E., & Meissner, C. A. (2005). Racial bias in mock juror-decision making: A meta-analytic review of defendant treatment. *Law and Human Behavior*, *29*(6), 621-637.

Mustard, D. B. (2001). Racial, ethnic, and gender disparities in sentencing: Evidence from the U.S. federal courts. *The Journal of Law and Economics*, *44*(1), 285-314.

Olivola, C. Y, Eubanks, D. L, & Lovelace, J. B. (2014). The many (distinctive) faces of leadership: inferring leadership domain from facial appearance. *The Leadership Quarterly*, *25*(5), 817-834.

Olivola, C. Y., Sussman, A. B., Tsetsos, K., Kang, O. E., & Todorov, A. (2012). Republicans prefer republican-looking leaders: Political facial stereotypes predict candidate electoral success among right-leaning voters. *Social Psychology and Personality Science, 3*(5), 605-613.

Olivola, C. Y., & Todorov, A. (2010). Fooled by first impressions? Reexamining the diagnostic value of appearance-based inferences. *Journal of Experimental Social Psychology*, *46*, 315-324.

Ormiston, M. E., Wong, E. M., & Haselhuhn, M. P. (2017). Facial-width-to-height ratio predicts perceptions of integrity in males. *Personality and Individual Differences*, *105*, 40-42.

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *PNAS*, *105*(32), 11087-11092.

Pettit, B., & Western, B. (2004). Mass imprisonment and the life course: race and class inequality in U.S. incarceration. *American Sociological Review*, *69*, 151-169.

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

Porter, S., ten Brinke, L., & Gustaw, C. (2010). Dangerous decisions: the impact of first impressions of trustworthiness on the evaluation of legal evidence and defendant culpability. *Psychology, Crime, & Law*, *16*, 477-491.

Qin, R., Gao, W., Xu, H., & Hu, Z. (2016). Modern physiognomy: an investigation on predicting personality traits and intelligence from the human face. *Science China Information Science*, *61*(5), 1-27.

Rule, N. O., & Ambady, N. (2008). The face of success: inferences from chief executive officers' appearance predict company profits. *Psychological Science*, 19(2), 109-111.

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*, *45*, 1245-1251.

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

Shriver, E. R., Young, S. G., Hugenberg, K., Bernstein, M. J., & Lanter, J. R. (2008). Class, race, and the face: social context modulates the cross-race effect in face recognition. *Personality and Social Psychology Bulletin*, *34*(2), 260-274.

Sigall, H., & Ostrove, N. (1975). Beautiful but dangerous: Effects of offender attractiveness and

    nature of the crime on juridic judgment. *Journal of Personality and Social Psychology*,

    31, 410-414.

Slater, A., Bremner, G., Johnson, S. P., Sherwood, P., Hayes, R., & Brown, E. (2000). Newborn

    infants' preference for attractive faces: The role of internal and external facial features.

    Infancy, *1*(2), 265-274.

Smith, R. J. (2015). Reducing racially disparate policing outcomes: Is implicit bias training the

    answer? *University of Hawaii Law Review*, *37*, 295.

Sofer, C., Dotsche, R., Oikawa, M., Oikawa, H., Wigboldus, D. H. J., & Todorov, A. (2017). For

    your local eyes only: Culture-specific face typicality influences perceptions of

    trustworthiness. *Perception*, *46*(8), 914-928.

Stewart, J. E. (1980). Defendant's attractiveness as a function in the outcome of criminal trials:

    An observational study. *Journal of Applied Psychology*, *10*, 348-361.

Stirrat, M., & Perrett, D. I. (2010). Valid facial cues to cooperation and trust: male facial width

    and trustworthiness. *Psychological Science*, *21*(3), 349-354.

Stolier, R. M., Hehman, E., & Freeman, J. B. (2018). A dynamic structure of social trait space.

    Trends in Cognitive Science, *22*(3), 197-200.

Suzuki, A., Tsukamoto, S., & Takahashi, Y. (2017). Faces tell everything in a just and

    biologically determined world: Lay theories behind face reading. *Social Psychological*

    and Personality Science, 1-11.

Tanaka, J. W., Kiefer, M., Bukach, C. M. (2004). A holistic account of the own-race effect in

    face recognition: evidence from a cross-cultural study. *Cognition*, *93*, B1-B9.

Thornton, G. (1939). The ability to judge crimes from photographs of criminals. *Journal of Abnormal and Social Psychology, 34,* 378-383.

Todorov, A., & Porter, J. M. (2014). Misleading first impressions: different for different facial images of the same person. *Psychological Science*, *25*(7), 1404-1417.

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.

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

Vakharia, D., & Lease, M. (2015). Beyond Mechanical Turk: An analysis of paid crowd work platforms. *Proceedings of the iConference*.

Ward, C., Flowe, H., & Humphries, J. (2012). The effects of masculinity and suspect gender on perceptions of guilt. *Applied Cognitive Psychology*, *26*, 482-488.

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

Wilson, J. P., & Rule, N. O. (2015). Facial trustworthiness predicts extreme criminal-sentencing outcomes. *Psychological Science*, 1-7.

Wixted, J. T., Mickes, L., & Fisher, R. P. (2018). In the DNA exoneration cases, eyewitness memory was not the problem: A reply to Berkowitz and Frenda (2018) and Wade, Nash, and Lindsay (2018). *Perspectives on Psychological Science*, *13*(3), 343-345.

Wong, E. M., Ormiston, M. E., & Haselhuhn, M. P. (2011). A face only an investor could love: CEOs' facial structure predicts their firms' financial performance. *Psychological Science*, *22*(12), 1478-1483.

Wootson, C. R. (2018, May 11). A black Yale student fell asleep in her dorm's common room. A white student called the police. Retrieved from https://www.washingtonpost.com/news/grade-point/wp/2018/05/10/a-black-yale-student-fell-asleep-in-her-dorms-common-room-a-white-student-called-police/?utm_term=.e5ae800314ff.

Wu, X., & Zhang, X. (2016). Automated Inference on Criminality using Face Images.

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

Zebrowitz, L. A. (2017). First impressions from faces. *Current Directions in Psychological Science*, *26*(3), 237-242.

Zebrowitz, L. A., & McDonald, S. M. (1991). The impact of litigants' baby-facedness and attractiveness on adjudications in small claims courts. *Law and Human Behavior*, *15*(6), 603-623.

Zebrowitz, L.A., Kikuchi, M., & Fellous, J-M. (2010). Facial resemblance to emotions: group differences, impression effects, and race stereotypes. *Journal of Personality and Social Psychology*, *98*, 175-189.

All photographs from Study 1. Criminality scores were made on a scale from 1 (definitely) to 5 (definitely not).
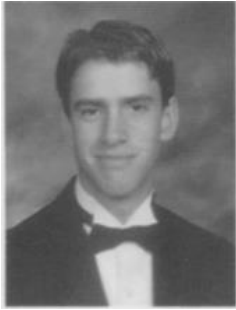
Non-criminal
Criminality: 3.52
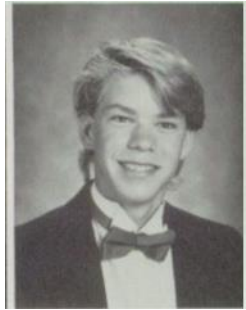
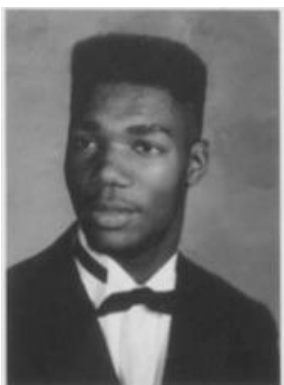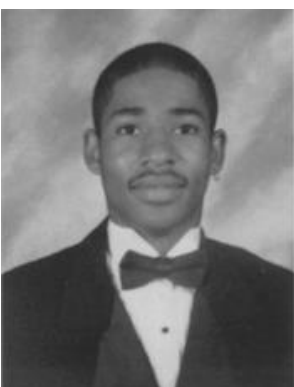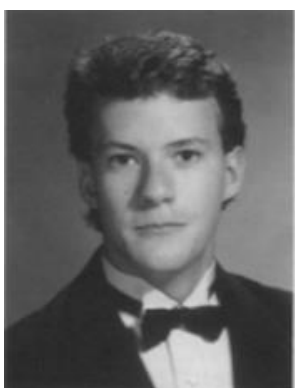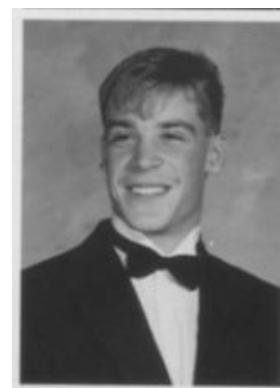**Non-criminal**
Criminality: 3.82

Non-criminal
Criminality: 3.72
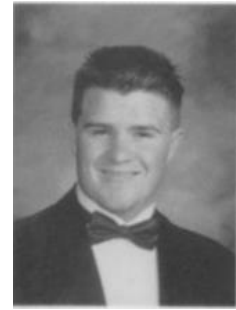
Non-criminal
Criminality: 3.42

Non-criminal
Criminality: 2.79
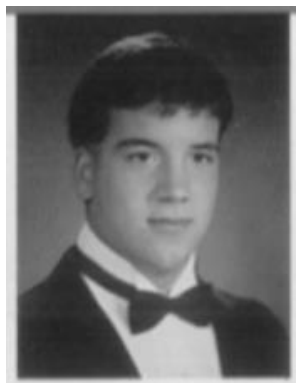
Non-criminal
Criminality: 3.50
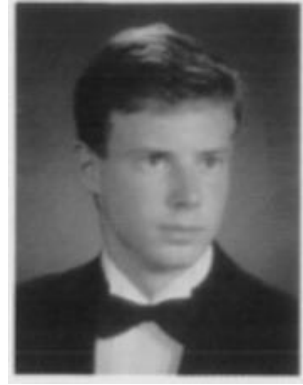
Non-criminal
Criminality: 3.69

Non-criminal
Criminality: 3.84

Non-criminal
Criminality: 3.69
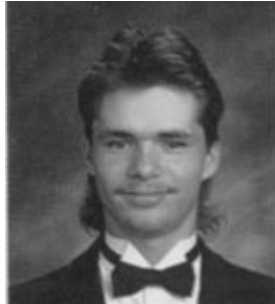
Non-criminal
Criminality: 2.96
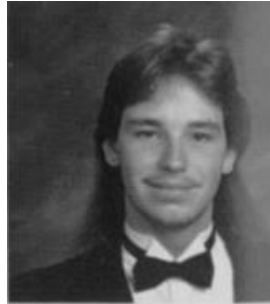
Non-criminal
Criminality: 3.05
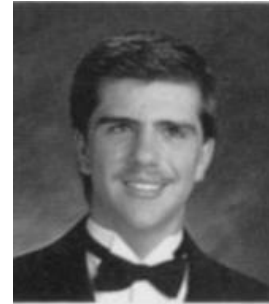
Non-criminal
Criminality: 3.29
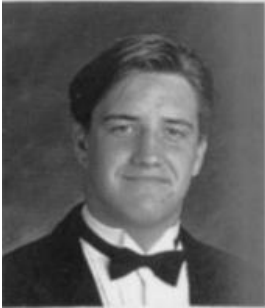
Non-criminal
Criminality: 3.41
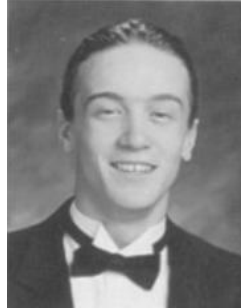
Non-criminal
Criminality: 2.95

Non-criminal
Criminality: 3.71

Criminal
Criminality: 3.05

Criminal
Criminality: 3.82

Criminal
Criminality: 3.02

Criminal
Criminality: 2.61

Criminal
Criminality: 2.65

Criminal
Criminality: 2.97

Criminal
Criminality: 3.64

Criminal
Criminality: 3.65

Criminal
Criminality: 2.96

Criminal
Criminality: 3.14
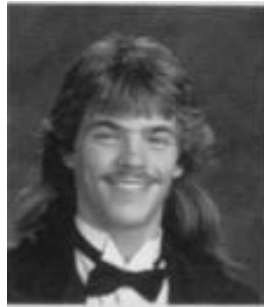
Criminal
Criminality: 3.18
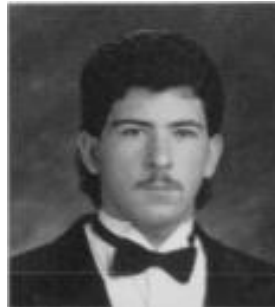
Criminal
Criminality: 2.92

Criminal
Criminality: 2.92

Criminal
Criminality: 2.73

Criminal
Criminality: 3.18

APPENDIX B

The new foils used in Study 2. Criminality ratings are the percent of "yes" judgments for each photograph.
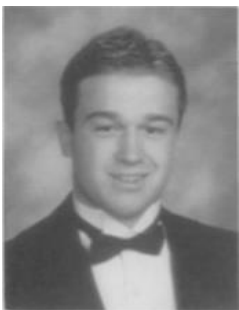
Non-criminal
Criminality: 17%

Non-criminal
Criminality: 10%

Non-criminal
Criminality: 11%
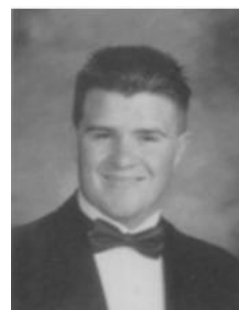
Non-criminal
Criminality: 18%

Non-criminal
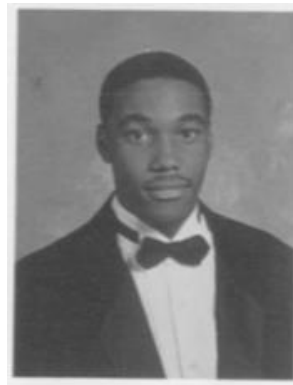Criminality: 34%

Non-criminal
Criminality: 38%

Non-criminal
Criminality: 10%

Non-criminal
Criminality: 20%

Non-criminal
Criminality: 38%

Non-criminal
Criminality: 20%
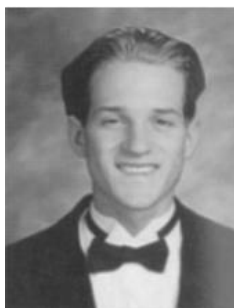
Non-criminal
Criminality: 30%
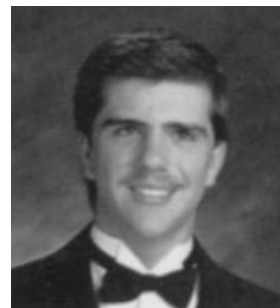
Non-criminal
Criminality: 34%

Non-criminal
Criminality: 10%
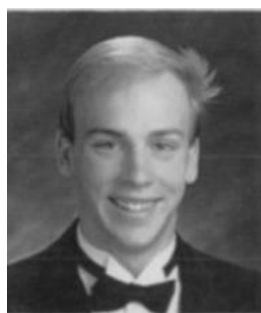
Non-criminal
Criminality: 18%

Non-criminal
Criminality: 10%

Criminal
Criminality: 12%

Criminal
Criminality: 12%

Criminal
Criminality: 38%

Criminal
Criminality: 62%

Criminal
Criminality: 60%

Criminal
Criminality: 49%

Criminal
Criminality: 19%

Criminal
Criminality: 12%

Criminal
Criminality: 29%

Criminal
Criminality: 36%
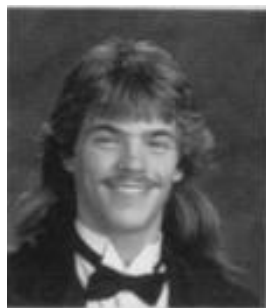
Criminal
Criminality: 36%
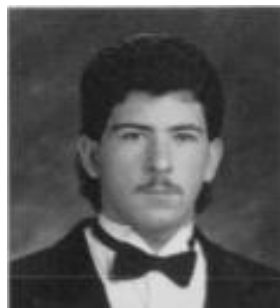
Criminal
Criminality: 47%

Criminal
Criminality: 51%

Criminal
Criminality: 60%

Criminal
Criminality: 43%

All photographs from Study 3. Criminality ratings are the percent of "yes" judgments.

**Non-criminal
Black**
Criminality: 8%

**Non-criminal
Black**
Criminality: 19%

**Non-criminal
Black**
Criminality: 6%
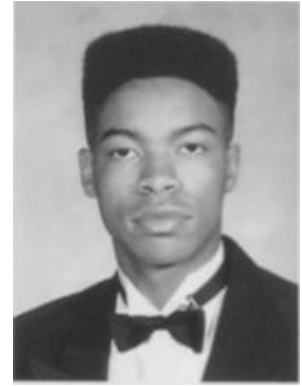
**Non-criminal
Black**
Criminality: 51%
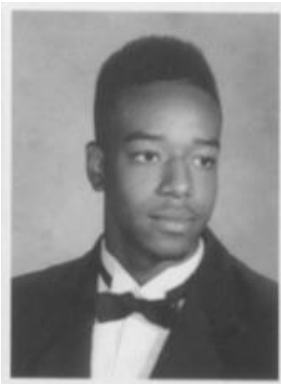
**Non-criminal
Black**
Criminality: 14%
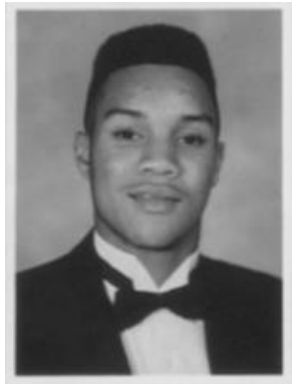
**Non-criminal
Black**
Criminality: 46%

**Non-criminal
Black**
Criminality: 14%

**Non-criminal
Black**
Criminality: 9%

**Non-criminal
Black**
Criminality: 25%
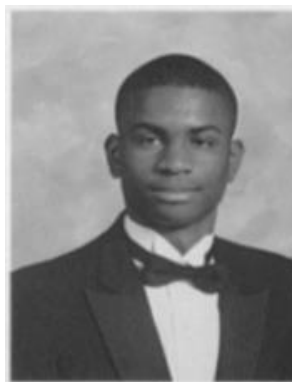
**Non-criminal
Black**
Criminality: 22%

**Non-criminal
Black**
Criminality: 5%
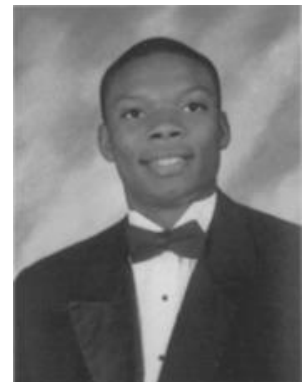
**Non-criminal
Black**
Criminality: 12%

**Non-criminal
Black**
Criminality: 9%

**Non-criminal
Black**
Criminality: 9%

**Non-criminal
Black**
Criminality: 6%

**Non-criminal
Black**
Criminality: 37%

**Non-criminal
Black**
Criminality: 13%

**Non-criminal
Black**
Criminality: 58%

**Non-criminal
Black**
Criminality: 22%

**Non-criminal
Black**
Criminality: 28%

**Non-criminal
White**
Criminality: 14%

**Non-criminal
White**
Criminality: 43%

**Non-criminal
White**
Criminality: 36%

**Non-criminal
White**
Criminality: 15%

Non-criminal
White
Criminality: 16%

Non-criminal
White
Criminality: 40%

Non-criminal
White
Criminality: 47%

Non-criminal
White
Criminality: 21%

Non-criminal
White
Criminality: 8%

Non-criminal
White
Criminality: 50%

Non-criminal
White
Criminality: 29%

Non-criminal
White
Criminality: 18%

Non-criminal
White
Criminality: 32%

Non-criminal
White
Criminality: 10%

Non-criminal
White
Criminality: 22%

Non-criminal
White
Criminality: 24%

**Non-criminal
White**
Criminality: 36%

**Non-criminal
White**
Criminality: 23%

**Non-criminal
White**
Criminality: 31%

**Non-criminal
White**
Criminality: 49%

**Non-violent
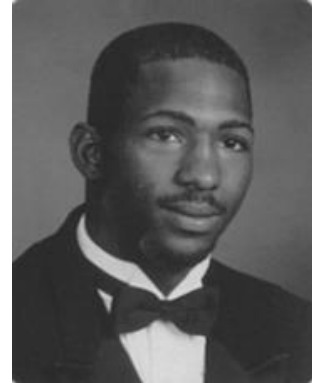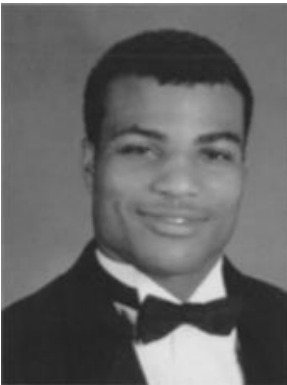Black**
Criminality: 18%

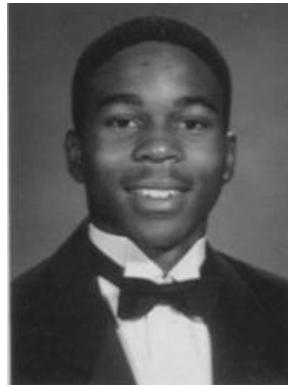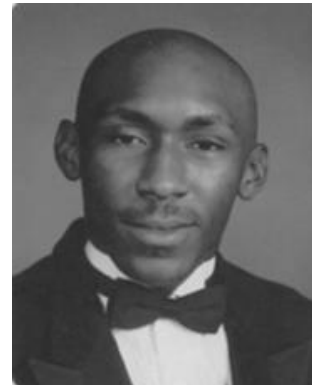**Non-violent
Black**
Criminality: 26%

**Non-violent
Black**
Criminality: 10%

**Non-violent
Black**
Criminality: 11%

**Non-violent
Black**
Criminality: 12%
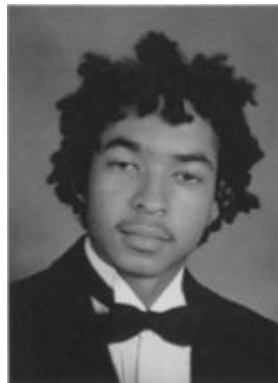
**Non-violent
Black**
Criminality: 55%

**Non-violent
Black**
Criminality: 14%
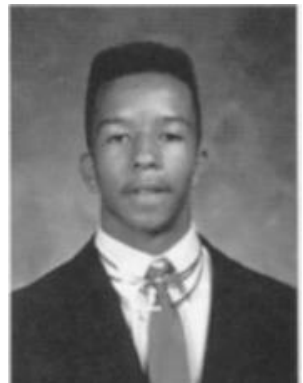
**Non-violent
Black**
Criminality: 6%

**Non-violent
Black**
Criminality: 16%

**Non-violent
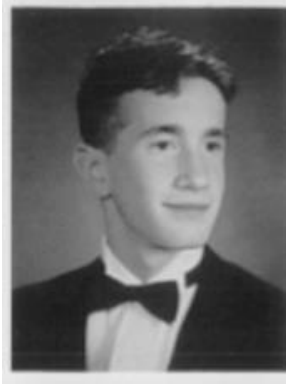Black**
Criminality: 9%

**Non-violent
White**
Criminality: 56%
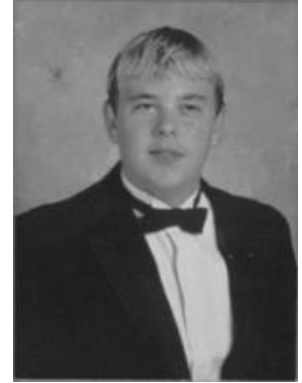
**Non-violent
White**
Criminality: 44%

**Non-violent
White**
Criminality: 19%
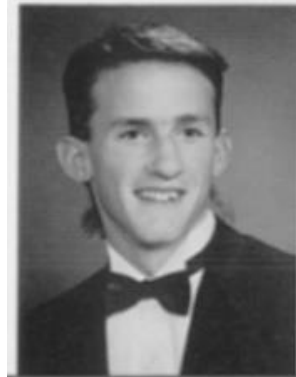
**Non-violent
White**
Criminality: 51%
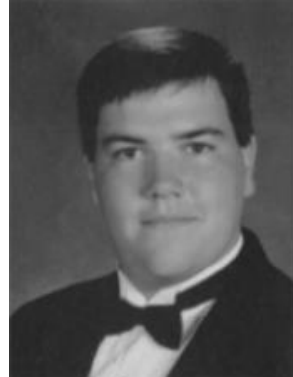
**Non-violent
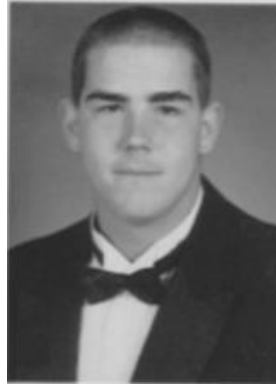White**
Criminality: 26%
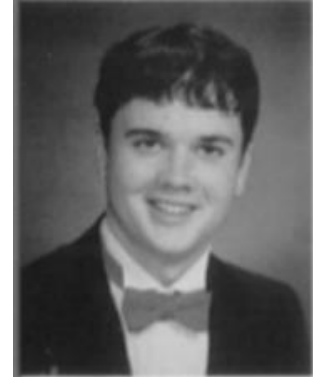
**Non-violent
White**
Criminality: 39%

**Non-violent
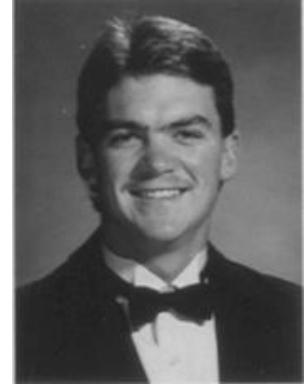White**
Criminality: 39%

**Non-violent
White**
Criminality: 28%
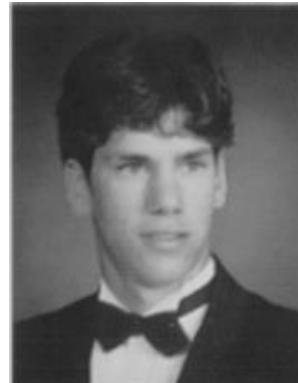
**Non-violent
White**
Criminality: 35%
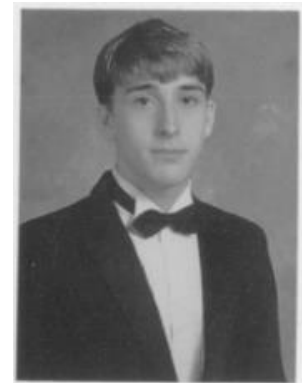

**Non-violent
White**
Criminality: 24%
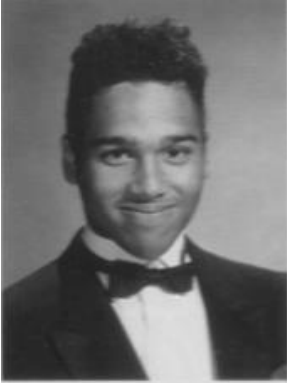
Violent
Black
Criminality: 24%

Violent
Black
Criminality: 37%
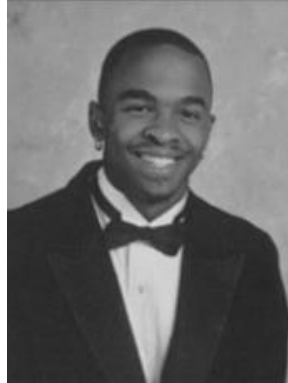
Violent
Black
Criminality: 19%

Violent
Black
Criminality: 17%

Violent
Black
Criminality: 18%

Violent
Black
Criminality: 12%

Violent
Black
Criminality: 9%

Violent
Black
Criminality: 51%

Violent
Black
Criminality: 8%

Violent
Black
Criminality: 14%

Violent
White
Criminality: 58%

Violent
White
Criminality: 28%

Violent
White
Criminality: 41%

Violent
White
Criminality: 13%

Violent
White
Criminality: 58%

Violent
White
Criminality: 39%
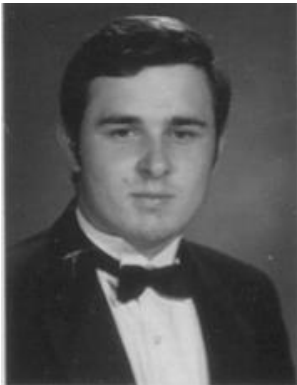
Violent
White
Criminality: 59%

Violent
White
Criminality: 66%

Violent
White
Criminality: 40%

Violent
White
Criminality: 16%

All photographs from Study 4. Criminality ratings are the percent of "yes" judgments for each photograph. The first percentage is for the mTurk workers and the second percentage is for the university sample.

Angel Gonzalez
Exoneree
Criminality:
38% - mTurk
32% - university

Craig Coley
Exoneree
Criminality:
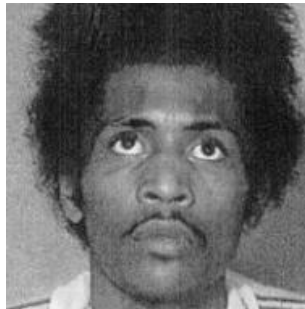52% - mTurk
62% - university

Damon Thibodeaux
Exoneree
Criminality:
86% - mTurk
86% - university

David Camm
Exoneree
Criminality:
49% - mTurk
81% - university

Dennis Maher
Exoneree
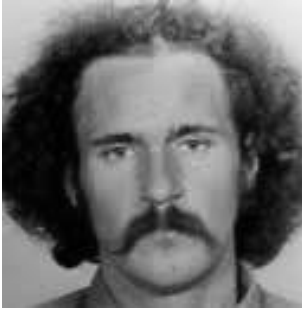Criminality:
68% - mTurk
76% - university

John Thompson
Exoneree
Criminality:
49% - mTurk
41% - university

Leslie Vass
Exoneree
Criminality:
43% - mTurk
35% - university

Michael Morton
Exoneree
Criminality:
42% - mTurk
41% - university

Randall Adams
Exoneree
Criminality:
68% - mTurk
70% - university

Wilton Dedge
Exoneree
Criminality:
34% - mTurk
41% - university

Andrew Cunahan
Criminal
Criminality:
59% - mTurk
38% - university

Colin Pitchfork
Criminal
Criminality:
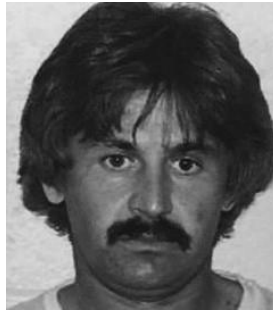77% - mTurk
88% - university

Gary Ridgway
Criminal
Criminality:
66% - mTurk
49% - university

Kenneth Troyer
Criminal
Criminality:
61% - mTurk
72% - university

Paul Carr
Criminal
Criminality:
68% - mTurk
61% - university

Richard Ramirez
Criminal
Criminality:
66% - mTurk
62% - university

Ricky McGinn
Criminal
Criminality:
76% - mTurk
66% - university

Steven Pennell
Criminal
Criminality:
70% - mTurk
68% - university

Timothy Spencer
Criminal
Criminality:
39% - mTurk
16% - university

Tommie Andrews
Criminal
Criminality:
60% - mTurk
53% - university