5-24-2000

Criminal appearance and legal decision-making

Marisa Evelyn Collett
Florida International University

DOI: 10.25148/etd.FI14060876

FLORIDA INTERNATIONAL UNIVERSITY

Miami, Florida

CRIMINAL APPEARANCE AND LEGAL DECISION-MAKING

A thesis submitted in partial fulfillment of the

requirements for the degree of

MASTER OF SCIENCE

in

PSYCHOLOGY

by

Marisa Evelyn Collett

2000

To: Dean Arthur W. Herriott
College of Arts and Sciences

This thesis, written by Marisa Evelyn Collett, and entitled Criminal Appearance and Legal Decision-Making, having been approved in respect to style and intellectual content, is referred to you for judgement.

We have read this thesis and recommend that it be approved.

Margaret Kovera

Brian Cutler

Ronald Fisher, Major Professor

Date of Defense: May 24, 2000

The thesis of Marisa Evelyn Collett is approved.

Dean Arthur W. Herriott
College of Arts and Sciences

Dean Richard L. Campbell
Division of Graduate Studies

Florida International University, 2000

DEDICATION

I dedicate this thesis to my parents, John and Leanne Hill and Timothy and Lourdes Collett, to my grandmother, Lee Bowling Hines, and my voice of sanity, Vincent Burke. Without their endless patience, love, and most of all, their belief in me, completion of this thesis would never have been possible.

ACKNOWLEDGMENTS

I wish to thank the members of my committee, Dr. Margaret Kovera and Dr. Brian Cutler, for their insight and guidance. I also wish to thank Mark Phillips for his assistance and encouragement through every stage of this project. I would also like to thank Zion Mass, who patiently assisted me in developing the stimulus materials and with data collection for this project.

I also want to extend a very special thank you to Dr. Ronald Fisher, who from the beginning was wonderful with his enthusiasm, support, guidance, patience, and most of all his steadfast encouragement.

ABSTRACT OF THE THESIS

CRIMINAL APPEARANCE AND LEGAL DECISION-MAKING

by

Marisa Evelyn Collett

Florida International University, 2000

Miami, Florida

Professor Ronald P. Fisher, Major Professor

Previous research has found that people are able and willing to assess whether an individual is a criminal or a non-criminal based on facial observations. What has not been looked at is whether an attribution of criminality could influence decisions as verdict choice, culpability, or punishment severity. The present study examined the effects of target photos that depicted pre-determined "bad guys" and "good guys" on legal decision-making. Participants viewed a case file of an armed robbery and attempted murder. Half the participants viewed a photo of a defendant who was previously deemed a "bad guy" and the other half a "good guy." No differences were found in verdict preference; however, target photos of "bad guys" elicited higher estimates of the future likelihood that the defendant would commit this type of crime than target photos of good guys. Results indicate that target photos are perceived congruent to their pre-determined categories, but those perceptions were disregarded and participants based their decisions on other factors when making crucial legal decisions.

TABLE OF CONTENTS

LIST OF TABLES

"He did it. I know it. Just look at him," is something that many of us have either said ourselves or have heard others say, either in response to a mystery movie depicting a crime and many potential assailants or news clips that show the alleged criminal being taken away by the police after capture. Judgments like these are made in the absence of other relevant information about the person being perceived.

The illustration presented is an example of physiognomy, which is the practice of judging character and other psychological qualities by merely observing faces. Physiognomy is an ancient practice with the earliest known work being linked to Aristotle (Alley, 1988). Darwin (1872) made mention of physiognomy and how the contraction of different facial muscles could indicate the disposition of the person. After continuous use of the particular facial muscles, the face would contain conspicuous lines or furrows that would display the specific disposition. Sociobiologists have ascribed a purpose to numerous facial characteristics (i.e. certain characteristics depict health or fitness; Alley, 1988). However, scientific literature, with few exceptions, has failed to find relationships between normal variations in facial features and psychological characteristics. There is good evidence that, even knowing this, the general public continues to believe and in some instances, practices physiognomy (Brandt, 1980).

Studies have shown that people tend to make enduring judgments about an individual's personality, occupation, potential behavior, intelligence, and much more, based solely on their perceptions of others' facial characteristics (e.g. Fiske & Taylor, 1984; Hamilton, 1981; Schneider, Hastorf, & Ellsworth, 1979). A few studies have shown that in some situations people can make accurate judgments of personality and

social dimensions of another based simply on photographs (Thornton, 1939; Terry & Snider, 1972; Mazur, Mazur, & Keating, 1984). The fact that in some situations people may be accurate about an occupational category or a personality characteristic is not as interesting as the fact that people are quite willing to make judgments of others based on just a photograph. It may be that an individual's behavior, in this regard, is reinforced by the fact that many times he or she is correct and this reinforces the practice of physiognomy.

Terry and Krantz (1993) empirically evaluated the salient features involved in physiognomy by varying different features of the face. They specifically examined the effects of eyeglasses, men's facial hair, and women's hair length on the traits attributed to facial appearance. Subjects rated Photo-Ident composites of target persons using 20 different pairs of adjectives (i.e. aloof-caring, bad-good). Results of the experiment were that eyeglasses on both men and women were associated with attributions of decreased forcefulness and heightened mental competence. Beards on men were associated with lessened mental competence and women's long hair was associated with diminished forcefulness.

To give another example of physiognomy and the influence of different facial components, Zebrowitz & Berry (1985) looked at the components of babyfacedness and the trait attributions of those components. The researchers found that large, round eyes, high eyebrows, and a small chin yielded the perception of a babyish appearance in photographs of individuals bearing these features. These facial features also elicited impressions from participants that the stimulus person was naïve, kind, warm, and honest. Zebrowitz & Berry (1985) also tested to see if the previous results could be reproduced

cross-culturally with Korean subjects and found the same results as with the previous experiment with American participants. These results demonstrate that impressions of baby-faced individuals are similar across cultures.

Just as in the examples above, other studies on stereotyping and facial appearance have shown that judgments made by subjects are highly consensual (Berry & Zebrowitz, 1988; Brownlow & Zebrowitz; Goldstein, Chance, & Gilbert, 1984; Kahlick, Zebrowitz, Langlois, & Johnson, 1988; Perrett & Yoshikawa, 1994; Shoemaker, South, & Lowe, 1973; Terry & Davis, 1976; Terry & Krantz, 1993; Zebrowitz, Voinescu, & Collins, 1996; Zebrowitz & McDonald, 1991). Subjects make these judgments in the absence of any other type of descriptive or historical information about the target except a photograph and they seem to agree as to which photograph reveals which personality dimension or belongs in which occupational category. One such study by Goldstein, Chance, & Gilbert (1984) tested the hypothesis that faces can serve as stimuli that trigger a consensual stereotypical response. They displayed facial photographs of white, middle aged men in five separate arrays. They instructed participants that each array contained a portrait of one of the following "occupations": mass murderer, armed robber, rapist, doctor, clergyman, and engineer. Participants were then asked to choose one picture for each of the six occupations. The results of the study were that a small number of portraits were often selected as criminals and a small number of other portraits were often selected as non-criminals. Participants' choices were more highly consensual among the criminal occupations (rapist or mass murderer) than the non-criminal occupations (doctor, minister, and lawyer).

In two different experiments, Shoemaker, South, & Lowe (1973) found a similar effect to Goldstein et al. The first experiment was similar to Goldstein et al. in that participants were asked to select from among a set of twelve photographs the picture that looked the most and next to the most likely and the least and next to least likely to have committed one of four types of crime. The four types of crimes were pedophilia, murder, robbery, and treason. In the second experiment they used contrived vignettes that presented ambiguous evidence of the four types of crime and had participants evaluate the extent of guilt or innocence of the fictitious person presented in the vignette (the target pictures were not presented at this point). These same participants were then asked to view four of the twelve pictures (either individually or consecutively) and evaluate the extent of guilt or innocence of the person in one of the previously read vignettes. Shoemaker et al. demonstrated that certain persons in pictures appear guilty of specific crimes. They also found that negative and positive stereotypes are correlated with judgments of guilt or innocence for the four crimes considered for the four crimes evaluated. The results also indicated that negative stereotypes were more important for appraising guilt than were positive stereotypes for appraising innocence.

A vast majority of studies in facial stereotyping demonstrate that people are able and willing to make judgments of others based solely on a photograph, but researchers have not looked at whether these stereotypical responses influence decision-making. We may make judgments of other people but when deciding critical issues that involve these individuals, do these judgments influence our decision-making processes?

Stereotyping has been noted as one of many different processes used to clarify our environment and economize on mental effort (Reisberg, 1997; Fiske & Taylor, 1984).

Stereotypes function by categorizing a target as part of a particular group, within which there is a basis of knowledge that has been built. The activation of a stereotype elicits a selective processing strategy in which stereotype-relevant information is attended to and processed (e.g. Rothbart, Fulero, Jensen, Howard, & Birrell, 1978; Hastie, 1980; Locksley, Borgida, Brekke, & Hepburn, 1980; Bodenhausen & Wyer, 1985). Thus, the ability of stereotypes to influence information processing comes from selectively attending to only stereotype-congruent information and disregarding all other information about the target. Research has demonstrated that stereotypes can function as judgment heuristics, i.e. simplifying rules of thumb that are used to interpret the behavior of others (Bodenhausen & Wyer, 1985). More often than not, our judgement heuristics are correct and this reinforces their use. It is this proclivity to use judgment heuristics that leads to utilizing stereotypes and stereotyped information when making decisions.

Bodenhausen and Wyer (1985) evaluated the effects of stereotypes on reactions to behavioral transgressions and the recall of information about the transgression and transgressor. Participants read a case file that described the transgression committed by the transgressor. The researchers manipulated the name of the transgressor in order to reveal which ethnic group the transgressor belonged to and they manipulated the transgression so that it was stereotypically congruent with the transgressor's ethnic group. Participants were asked to judge the probability that the transgression would recur and they were asked to prescribe a punishment for the infraction. Participants were also asked to recall all the information they could about the case. The results indicated that transgressions that are stereotypic of a transgressor's ethnic group were seen as more likely to recur than non-stereotypic transgressions. Stereotypic transgressions were also

punished more severely. Recall data show a differential recall of presented information, with stereotypic information more likely to be recalled than non-stereotypic information. This suggests that when an applicable stereotype-based explanation was available to explain the transgressor's offense participants were likely to think about this information more or view this information as more diagnostic than other information presented. This in turn led to a reduced recall of the information that was not congruent with the stereotype elicited. The recall data obtained by Bodenhausen and Wyer support the notion of the "confirmation bias" (Nisbett & Ross, 1980), in which once a stereotype-based impression of the incident and its components was formed participants better recalled information that was congruous with the stereotype.

Going beyond ethnic stereotypes, Berry and Zebrowitz (1988) evaluated the impact of facial maturity and the attribution of legal responsibility. The researchers used a simulated trial format and found that a babyfaced individual was found guilty more often than the more mature-faced individual on charges stemming from behavior which resulted from negligence. Babyfaced individuals were also more often acquitted of crimes that involved intentional criminal actions. The results of the experiment also yielded evidence to show that subjects recommended less severe punishment for babyfaced individuals than to mature-faced individuals.

Zebrowitz and McDonald (1991) evaluated the impact of litigants' attractiveness and facial maturity on adjudications in small claims court. Their results demonstrated that facial maturity and attractiveness had a significant impact on adjudications. As defendants decreased in facial maturity (i.e. they increased in baby-facedness) they were more likely to win cases involving intentional actions and less likely to win cases involving negligent

actions (Zebrowitz & McDonald, 1991). This is a similar pattern of results as those produced in Berry and Zebrowitz (1988) described previously. As plaintiffs increased in attractiveness, they were more likely to win the case. Also, as defendants increased in facial maturity they were required to pay larger monetary sums or awards to (specifically) baby-faced individuals.

Berry and Zebrowitz (1988) and Zebrowitz and McDonald (1991) both demonstrated that dimensions of attractiveness and facial maturity play roles in adjudication decisions made by participants. Bodenhausen and Wyer (1985) also showed that stereotype-congruent behavior plays a role in legal decision-making. Other researchers have found that people are able and willing to assess whether an individual belongs to either criminal or non-criminal occupations and judge specific trait attributions based solely on their observations of a face. Research on stereotypes and decision-making show that people are harsher in their judgments of culpability and punishment severity when behavior is stereotypically congruent with the individual's racial or ethnic background. What research has not examined is whether faces deemed most likely to be armed robbers or murderers are more likely to be convicted, viewed more culpable, punished more severely, or deemed more likely to commit a crime that is congruent with their perceived criminal occupation. It is interesting to know that people categorize others into different criminal and non-criminal occupations, but it is more useful to know whether that classification affects decision-making, especially legal decision-making. Do people use the stereotype that "bad people do bad things" when it comes to legal decision-making? This is the purpose of the present study.

This study consists of three experiments. The purpose of Experiment 1 and 2 was to identify a set of target pictures for which there is consensual agreement about good and bad guy targets. The use of two experiments to identify target photographs of good and bad guys allowed for convergence of good and bad guy targets across different experimental methods. Experiment 3 tested the legal decision-making of participants when exposed to either the good or bad guy targets.

In the first experiment, participants were asked to look through 60 pictures of white males, approximately 25-55 years of age. This part of the experiment was undertaken in order to identify which of the 60 pictures participants consensually agree are most indicative of "good" and "bad" guys. The hypothesis of Experiment 1 is that certain faces elicit consensually similar notions that an individual is either a good guy or a bad guy, i.e. individuals would feel safe opening the door for one individual over another simply based on their appearance.

In the second experiment a separate group of participants imagined that they were managers of a convenience store and they made a hiring decision about a target individual depicted in their file. Two of the pictures depicted individuals that participants in Experiment 1 consensually agreed were bad and the other two pictures depicted individuals that participants consensually agreed were good. The purpose of this experiment was to converge on the selection of the target pictures from Experiment 1 by using a different experimental method. It was predicted that participants would be more likely to hire the target good guy than they would the target bad guy.

In the final experiment participants were shown a criminal case file. The case file is ambiguous and when run in mock trials it typically ends up with half of the juries

acquitting and half convicting the defendant (J.A. Tanford, personal communication, February 24, 1998). The case file viewed by participants involved an alleged armed robbery of a convenience store and an attempted murder of the convenience store manager. The material presented to participants was similar in every way except the picture of the defendant. The pictures obtained in the first experiment and then verified in the second experiment were used.

Upon completing the case file presentation, participants were asked to fill out a questionnaire with several different measures. Measures within the questionnaire included: verdict choice; current probability that the defendant committed the crimes charged; punishment sentence ranging from no probation or prison time to 25 years in prison; and future probability that the defendant will commit the same type of crime in the future. Following those measures, participants indicated the top five most diagnostic pieces of information used in making their verdict decisions. Participants also rated the defendant on different personality trait dimensions and recalled as much of the case file as they could remember.

The experimental hypothesis was that participants would be more likely to convict the bad guy than the good guy. This is because the behavior depicted in the case file was more congruent with the stereotype elicited by the picture of the bad guy (i.e. bad guys do bad things) than the good guy. Participants were also expected to be harsher in their sentencing of the bad guy. As for the current and future probabilities that the defendant committed or will commit these types of crimes in the future, I expected that participants would rate each of the probabilities higher for the bad guy than the good guy condition.

Experiment 1

Method

Participants. Thirty participants were drawn from the Miami, FL area. The average age of the participants was 29 years (SD = 15.41; range = 18-82 years old). Slightly over half of the participants (58%) were female and slightly less than half (43%) were male. The participant population was split in terms of race and ethnic background. Nearly half of the participants indicated they were non-Hispanic white (47%) and the other half indicated they were Hispanic white (53%). Nearly two-fifths of the participants had high school diploma or less (37%), a little over a quarter had some college (27%), and one-third had a college degree and more (30%). Half the participants noted that their full time occupation was that of a student, almost two-fifths noted they were in sales/ customer service type jobs (37%), and the remaining (13%) were employed in other fields.

Materials. The target faces were collected from photo lineups conducted by the Hollywood, Florida Police Department in 1991. Photos were limited to white males, ranging in age from 25 to 50 years of age. They were arranged such that participants viewed one photo per page, creating 60 pages in the binder. The photos were in random order and each of the participants viewed the pictures in the same order.

Procedure. Participants were asked to volunteer in a short study that looked at people's feelings and impressions of other people. Those who agreed to participate were asked to look through a binder of 60 pictures and evaluate their feelings and impressions of the people depicted in the pictures. They were asked to think in terms of whether they would open the door for this individual if the individual were ringing the doorbell asking for donations. For each picture, they were asked rate their feelings and impressions of the

person depicted in the picture by indicating whether the person impressed them as a "good" person, a "bad" person, or it was undeterminable. They indicated this by writing "G" for good person, "B" for bad person, or "M" for mixed or undeterminable in the numbered space that corresponded with the picture.

Results

Participants showed a high consensual agreement between which targets appeared to be good guys or bad guys (i.e. whether they would open the door for this person if he were at their front door). Participants showed higher agreement for the exemplars of bad guys than they did of good guys in this set of sixty individuals. Results for the bad guys were as follows: Pictures 10 and 20 had the highest consensual agreement as to whether they looked like bad guys, i.e. someone the participants would not be likely to open their door for (see Figure 1 to view pictures selected and see Table 1 for results).

For the pictures depicting good guys, Pictures 15 and 48 had the highest consensual agreement as to whether they looked like good guys, i.e. someone that participants felt comfortable opening their door for (see Figure 1 to view pictures selected). Consensual agreement as to which pictures depicted good guys was not as high as for bad guys, but more than half the participants agreed that Picture 15 and 48 depicted good guys (see Table 1 for results).

Discussion

These results are similar to those obtained by Goldstein et al. in that a small number of portraits were often selected as bad guys and a small number of other portraits were selected as good guys. Unlike Goldstein et al., participants did not then further categorize these pictures into criminal or non-criminal occupations. These results, like

Goldstein et al., are compatible with the idea that people may be processing perceptual stimuli by means of a non-conscious comparison of the target face with either an exemplar or prototype they have of a good guy or a bad guy (1984).

Overall results for the sixty pictures indicate that participants were inclined to find most of the pictures to be either bad guys or they had mixed/ undeterminable reactions to the pictures. There were very few instances in which participants felt that a picture depicted a good guy. This may reflect the source of these pictures, police photo line-ups. More than likely many of the people depicted in the photos have been arrested or suspected for a crime.

Experiment 2

Method

Participants. Thirty-two participants were drawn from different psychology classes taught at the North Campus of Florida International University. The average age of participants was 24.10 years (SD = 3.88; range = 18-34 years old). Participants were predominantly female (78%). Race and ethnic background was diverse as nearly one-fifth of the sample (19%) was non-Hispanic white, nearly one-fifth (19%) was non-Hispanic black, nearly two-fifths (38%) was Hispanic white, approximately one-tenth of the sample was Hispanic black (9%), and the remaining (16%) indicated they were of Asian descent. The majority of the sample (81%) had at least some college. The remaining one-fifth (19%) had received a college degree. All participants indicated that they were full-time students

Participants were given extra credit in their psychology classes in return for their involvement in this part of the study. Participants were told that they would be

participating in a personnel management study that evaluates how managers make decisions. Participants were tested in groups of one to eight people.

Materials. The materials within this portion of the experiment consisted of: instructions to participants; a photocopy of a mock application for employment; a mock California driver's license; and response sheets which included a demographics section (Appendix 1 contains a sample of the materials used in this portion of the experiment). One of four different pictures (two pictures depicting "good guys and two pictures depicting "bad" guys) was inserted into the photo area of the driver's license. The selection of the four pictures was pre-determined from the outcome of Experiment 1. The application is a standard application form that can be found at office supply stores. The information filled out on the application form was the same for all four pictures. The application listed the current address of the target (Wayne Walker), previous employers, previous education, a list of references, and the target's signature attesting to the accuracy of the information provided in the application.

The response section of the application contained an initial question that asked participants to indicate on a scale of 1 to 10 whether they would hire Wayne Walker (the target), 1 indicating that they would be least likely to hire Wayne Walker and 10 indicating they would be most likely. After the initial hiring question, participants were asked to provide five pieces of information that were most diagnostic in making their decision. In the next portion of the experiment, participants were asked about particular personality traits Wayne Walker might possess. On a scale of 1 to 7 (1 indicating that in no way did Wayne Walker possess the trait asked about and 7 indicating that Wayne Walker very much possessed the trait), participants were asked to indicate whether Wayne Walker

possessed the particular trait. The traits were responsible, honest, sincere, reliable, moral, conscientious, and likable. After completing this section, participants indicated what they thought this experiment was about. This was an awareness question used to determine whether there was a difference between participants who were aware and those who were unaware of the experimental paradigm. The final portion of the response sheet was the demographics section. The demographics asked participants to indicate their age, gender, race/ethnic origin, education level, and occupation.

Procedure. Participants were given packets and asked to wait for further instructions from the researcher. The instructions were read out loud to participants. The instructions indicated that participants should pretend to be a manager of a convenience store and they were reviewing this particular application in order to make a decision as to whether they would hire this particular individual for the cashier position. Upon completing their review of the application and the driver's license they filled out the response sheets that followed.

After participants finished filling out the response sheets they were asked to wait silently until others in the room had completed their packets. When everyone in the room was finished, the packets were collected and the participants were debriefed and thanked for their time.

Results

Participants were asked to indicate on a scale of 1 to 10 (1 being least likely and 10 being most likely) whether they would hire the target person. Participants indicated they were more likely to hire targets who were initially selected as good guys ($M = 7.25$, $SD = 2.02$) than they would targets selected as bad guys ($M = 4.25$, $SD = 1.77$), t (30) = -

4.47, SE = .67 ($p < .01$).

A one-way multiple analysis of variance of the personality traits (i.e. the Likert scale in which one indicates not possessing the quality asked about and seven indicates very much possessing quality) indicated that there were no significant differences between the means of each of the seven traits for the good or bad guy conditions, all F's (1,28) ≤ 1.59.

Participants were asked to write down the most important pieces of information they used in determining their likelihood of hiring the target. Recall information was scored using the criterion that the "general meaning" of a particular behavior had been accurately conveyed by the participant (Rothbart, Evans, & Fulero, 1979). For example, if participants wrote, "He never finished college," this statement was given one point as it conveyed a basic piece of information on the target. There were no differences in amount recalled between the bad guy condition (M = 3.69, SD = 1.02) and the good guy condition (M =3.88, SD = 0.89), t (30) = -.56, SE = .34. Recall information was also scored for positivity or negativity of the information. Examples of positive information are, "He has had some college," and "He has been an attendant at a gas station before." Examples of negative information are, "He has never been a cashier," and "He was out of work for a long time." There was no difference in the amount of negative information recalled between the bad guy condition (M = 1.81, SD = 1.47) and the good guy condition (M = 1.69, SD = 1.20), t (30) = .26, SE = .47. There was also no difference in the amount of positive information recalled between the bad guy condition (M = .56, SD = 1.09) and the good guy condition (M = .44, SD = .96), t (30) = .34, SE = .36.

Participants who indicated they were aware of the experimental paradigm were not

excluded from the analysis of the results, rather separate analyses were run to indicate whether there was a difference between the aware and unaware participants. There was a marginal difference in likelihood of hiring the target between subjects who were aware of the experimental paradigm (M = 5.10, SD = 2.57) and those who were not (M = 6.80, SD = 2.15), t (30) = -1.80, SE = .95 (p = .08). Participants who were aware of the experimental paradigm were slightly less likely to hire the target than those who were not aware. One anomaly in this awareness data is that participants in the bad guy condition were overwhelmingly more likely to be aware (93%) of the experimental paradigm than those in the good guy condition (40%).

Discussion

The premise of this experiment was to verify the pictures selected in the first experiment using alternative methods. Converging with Experiment 1, Experiment 2 demonstrated that participants reacted differently to the targets in the bad guy condition than the targets in the good guy condition. The results indicate that a separate sample of participants, being given a hiring task, were less likely to hire the bad guy target than the good guy. Participants demonstrated in this experiment that when given the choice to hire the target for the position of cashier, a position that involves some level of trust and responsibility, participants were less likely to hire the targets in the bad guy condition than in the good guy condition. This affirms the selection of the pictures in the first experiment and allows for the implementation of the third and crucial experiment.

Method

Participants. Ninety-three participants were drawn from different psychology classes taught at the North Campus of Florida International University. Participants averaged 21.80 years (SD = 5.10; range = 17 – 48 years). As in the previous experiment a majority of the sample was female (86%). The racial and ethnic background of the sample was diverse: non-Hispanic white (17%), Hispanic white (17%), Asian (17%), non-Hispanic black (45%), and Hispanic black (3%). Most of the sample (88%) had had some college experience but had not yet received their degree; the remaining had received their college degree and/or had post-graduate experience or degrees (12%). Most of the sample were full-time students (62%), a little over one-fifth (23%) of the sample were in sales or customer service, and the remaining portion of the sample (15%) were in other occupations such as law, social service, medical, or education.

Participants were given extra credit in their psychology classes in return for their involvement in this part of the study. Participants were told that they would be participating in a jury decision-making study that evaluates how jurors make decisions. Participants were tested in groups of one to five people.

Materials. In the third experiment participants were shown a criminal case file. The case file involved an alleged armed robbery of a convenience store and an attempted murder of the convenience store manager. The case file was selected because when run in mock trials it typically ended up with half the mock juries acquitting and half convicting the defendant and as such would maximize the likelihood of finding effects of the manipulated variable, facial appearance (J.A.Tanford, February 24, 1998 personal

communication). The case file included several statements including one from the defendant, two separate statements from alibi witnesses for the defense, a statement from the detective investigating the case, and a hearing that involved questioning of the manager of the convenience store by both the prosecution and the defense. Within the case file there were also pictures of all the major players in the case, maps, fingerprints, and various other pieces of evidence accumulated in the investigation. The material presented to participants was similar in every way except the picture of the defendant. The pictures of the defendant used in the separate conditions were the ones obtained in the first experiment and then verified in the second experiment.

The case file was viewed on Microsoft PowerPoint 7.0. Participants saw slides on a 15 in., Dell Computer color monitor and listened to narration that accompanied each slide on Altec Lansing ©, multi-media speakers. There were 43 slides in total. For slides that contained just one picture, the average picture size was 6.5 in X 6.5 in. and each picture was centered on the screen. Slides containing scanned copies of the pseudo legal paperwork involved in a case (i.e. indictments, maps, fingerprint cards, firearm registration, etc.) were 7 in. X 7.5 in. and were centered on the screen. The total running time of the presentation was 54 minutes and 45 seconds. The defendant's photo was shown three times for a total viewing time of 7 minutes and 57 seconds, once at the beginning when photos of all the key players in the case file were introduced (the viewing time for these pictures was 7 seconds), during the defendant's statement (5 minutes and 40 seconds), and at the end of the case file when participants were asked to fill out the questionnaire (2 minutes and 10 seconds). All other photos, except the judge's and the witness for the prosecution, were shown twice. The judge's photo was shown four times,

once at the beginning of the case file to introduce the case, one time during the presentation of the key players, during the hearing, and at the end of the case file presentation. The photo of the defense witness, also the alleged victim, was viewed a total of three times. His photo was viewed during the presentation of key players in the case, during his statement, and in the hearing

At the end of the presentation participants filled out response sheets (see Appendix 2). The response sheets began with having participants indicate their verdict choice. Following this, participants indicated on a scale of zero to one hundred percent what their personal estimation was of the probability that the defendant committed the crimes of armed robbery and attempted murder. Then they assigned a sentence preference for the defendant. There were eight different sentences that participants could choose to assign the defendant. The sentence selections ranged from no sentence, to probation and counseling, to one year in prison and four years probation and counseling, and then to prison time ranging from five years up to 25 years (see Appendix 2 for the specific measure). After assigning sentence severity participants wrote down their personal estimation, on a scale of zero to one hundred percent, of whether the defendant would commit this type of crime in the future. On a lined sheet following these measures, participants recalled all the facts of the case they could remember. Following the recall portion, participants were asked to provide the five most important pieces of information they used in making their verdict decision. They were asked to rank order the information with one indicating that the piece of information was most important in making their decision and five indicating that it was least important. In the next portion of the experiment, participants were asked about particular personality traits Wayne Walker may

or may not possess. On a scale of one to seven (one indicating that in no way did Wayne Walker possess the trait asked about and seven indicating that Wayne Walker very much possessed the trait), participants were asked to indicate whether Wayne Walker possessed the particular trait. The traits were responsible, honest, sincere, reliable, moral, conscientious, and likable. After completing this section, participants filled out what they thought this experiment was about. This was an awareness question used to determine whether there was a difference between participants who were aware and those who were unaware of the experimental paradigm. The final portion of the response sheet was the demographics section.

Design. This experiment was a single factor design. The independent variable was the type of picture viewed for the defendant (i.e., good guy versus bas guy). Within each condition there were two pictures (i.e., two pictures of bad guys and two pictures of good guys). The dependent variables were verdict choice, the probability that the defendant had committed the crime, sentence assignment, the probability that the defendant will commit the same crime sometime in the future, and recall.

Procedure. Participants, in groups of one to five people, were assigned randomly to view one of the four versions of the case file presented on Microsoft PowerPoint. Participants were seated in front of the computer. One volunteer member in the group manually advanced each of the 43 slides. An oral narrative that briefly described the slide accompanied each slide. On slides in which witnesses gave statements there was a brief description of what the slide was and then the witness gave the statement (in a separate voice from the narrator). To give an example of these descriptions that follow each slide, the following is taken from the slide showing the defendant Wayne Walker while he reads

his deposition. The description says, "This is the defendant Wayne Walker. You will now hear a statement from the defendant. Once he is finished please advance to the next slide."

There were slides in which the judge gave instructions in a separate voice from the narrator. Also, there were two slides in which participants viewed a probable cause hearing. The probable cause hearing exposed participants to the victim/ witness' testimony of the armed robbery and attempted murder and their subsequent identification of the defendant. On these two slides were pictures of the prosecuting attorney, defense attorney, the witness to the crime, and the judge. Participants waited until the completion of the narration to advance to the next slide. Participants were only allowed to advance the slides; they were not allowed to go back to earlier slides. The final slide instructed participants to begin filling out response sheets.

Upon completion of the presentation, participants filled out the response sheets, which contained the dependent measures described previously. Participants were not timed when filling out the response sheets. Questionnaires were filled out individually without deliberation.

Results

There were similar patterns of results for the two pictures within each condition, so the following results were collapsed across photographs within each condition.

Verdict. Regardless of condition participants overwhelmingly opted to acquit the target (see Table 3). There was no difference, however between the conditions, χ^2 (1, N = 93) = .02, $p > .05$.

Current and Future Estimation. A repeated measure analysis of variance was performed on the current and future estimations. There was a significant interaction

between time of estimation and condition, F (1,91) = 14.89, MSE = 406.44 ($p < .01$). For the current estimation the bad guy was equivalent to the good guy, t (91) = -.80, SE = 6.11 ($p = .43$; see Table 4), whereas for the future estimation the bad guy was rated as being more likely to commit the crimes of armed robbery and attempted murder than the good guy, t (91) = 3.08, SE = 5.83 ($p < .01$; see Table 4).

Sentence Assignment. Sentence assignment was not related to condition (good versus bad guy), χ^2 (7, N = 93) = 11.48, $p > .05$. The modal sentence assignment in both conditions was "No sentence, I would allow him to go free." Nearly half (48%) of the participants in the bad guy condition and nearly two-fifths (38%) of the participants in the good guy condition selected this sentence option. The remaining participants in each condition were approximately equally distributed across the other seven sentence options (see Table 5).

Personality Traits. A one-way multiple analysis of variance was conducted for condition and the seven personality traits of responsible, honest, sincere, reliable, moral, conscientious, and likable (see Table 6). The personality traits of conscientious and reliable were the only traits that showed a significant difference between conditions with the bad guy targets receiving considerably lower ratings for these traits, F (1, 91) = 4.61, MSE = 1.69 and F (1, 91) = 3.92, MSE = 1.57 ($p \leq .05$), respectively; all other F's < 2.02.

Recall. Similar to Experiment 2, recall information and diagnostic information were scored using the criterion that the "general meaning" of a particular behavior had been accurately conveyed by the participant (Rothbart, Evans, & Fulero, 1979) and hence, a point was given for that recall information. For example, if the participant wrote, "The defendant wore a red bandana over his mouth," or "The Nite Owl convenience store was

robbed," a point was given. Diagnostic information was coded for whether it was information that would acquit or convict the defendant. For example, if participants wrote that the night manager stated the doors had not been cleaned in a few days and the fingerprint (one of the key pieces of evidence presented by the prosecution) could have been from a previous visit by the defendant, the diagnostic information was scored as acquitting the defendant. If participants wrote that the defendant was identified by the night manager (i.e. the alleged victim), the diagnostic information was scored as convicting the defendant. The total amount of information recalled was equivalent between the bad guy condition (M = 9.10, SD = 4.36) and good guy condition (M = 9.67, SD = 5.43); t (91) = 0.55, SE = 1.02. There was no difference between the bad guy condition (M = 2.96, SD = 1.79) and good guy condition (M = 2.69, SD = 1.89) in information that would acquit the target, t (91) = .71, SE = .38. There was also no difference between the bad guy condition (M = .88, SD = 1.72) and good guy condition (M = 0.76, SD = 1.42) in diagnostic information that would convict the defendant, t (91) = .36, SE = .33.

Awareness. A large proportion of participants (86%) were not aware of the experimental paradigm, i.e. the influence of the defendants' picture. Participants who indicated that they were aware of the experimental paradigm (i.e., they were able to indicate that experiment had something to do with the appearance of the defendant) were removed and the data were reanalyzed. With aware participants removed, the same pattern of results as prior analysis was obtained for verdict, current and future estimation, and recall. Results not containing the aware participants differed from overall results on sentence assignment. Without the aware participants, the sentence assigned the target was

significantly related to condition, χ^2 (7, N = 80) = 14.38, $p < .05$. Nearly half (48%) the unaware participants in the bad guy condition chose, "No sentence, I would allow him to go free." The remaining participants in the bad guy condition distributed their sentence assignments almost equally across the five other selections and only one person selected the harshest sentence of 20- 25 years in prison. Unlike the bad guy condition where sentence assignments, besides the "no sentence" selection, were evenly distributed, participants in the good guy condition were more selective in their sentence assignments. Nearly two-fifths (38%) of the unaware participants in the good guy condition chose, "No sentence, I would allow him to go free." Nearly one-fifth (18%) of the remaining participants in this group chose the harshest sentence of 20- 25 years and prison. A small minority of participants (8%) in this group chose the option of 10- 15 years in prison and the remaining participants (36%) evenly distributed their selections over the sentences that included a period of probation.

Reanalysis of the data, after removing aware participants, indicated that only the personality trait of conscientious showed any significant difference between conditions, F (1, 78) = 4.70, MSE = 1.80 ($p < .05$). All other personality traits were not significant between conditions for these reanalyzed data, all other F's (1, 78) $\leq$ 1.82. This was unlike previous results in which both the personality traits of responsible and conscientious showed significant differences between groups

In analyzing the aware versus the unaware groups, the only difference between participants who indicated that they were aware of the paradigm and those who did not was the estimation that the defendant would commit this type of crime in the future. There was a marginally significant difference between these two groups, t (91) = 1.83, SE

= 8.67 (p = .07). The aware group (M = 50.77, SD = 31.68) deemed the target as more likely to commit this crime in the future than the unaware group (M = 30.93, SD = 28.55). In further analysis of the aware group (n = 13) on the dependent variable of future estimation, there was a significant difference between the bad guy condition (M = 65.63, SD = 30.29) and the good guy condition (M = 27.00, SD = 16.05). Participants in the bad guy condition were more likely to rate the target as having a higher probability of committing this type of crime in the future than participants in the good guy condition, t (11) = 2.60, SE = 14.84 ($p < .05$). This pattern of results was the same for the unaware group. As such, awareness of the experimental paradigm did not influence the pattern of results other than that the estimation that the defendant would commit this type of crime in the future was higher for the aware group than for the unaware group.

General Discussion

The hypothesis that faces deemed bad or criminal will elicit different responses than faces deemed good or non-criminal has been generally confirmed in several early studies (Goldstein, et al., 1984; Saladin, Saper, & Breen, 1988; Shoemaker, et al., 1973; Yarmey, 1993). The current experiments substantiate this hypothesis and take it one step further to demonstrate that faces deemed bad or criminal elicit prejudice behaviors and educe stereotypes which participants can reject or implement depending on the specific decision. The current experiments show that individuals have consensual stereotypic responses to different faces, but those stereotypes do not influence all types of decision-making.

Experiment 3 demonstrated that participants possessed a negative stereotype of the bad guy target, as evidenced by the fact that they estimated a higher probability that

the bad guy target would commit similar crimes in the future. Nevertheless, they were unlikely to convict the bad guy target. Participants in the current experiment were equally as likely to convict the bad guy target as the good guy target.

It is reassuring to know that even with a picture of a person previously determined as bad by over 80% of the participants in the first experiment, a majority of the participants in the third experiment would still not convict the bad guy target given the evidence presented. Participants' overwhelmingly chose the not guilty verdict. This is contrary to the admonition given by Goldstein, et al. who noted that their findings suggested, "the definite possibility that the degree to which a particular individual's face invites facial stereotyping may influence the outcomes of any legal process in which they may become involved" (p. 552). The current results indicate that a "particular individual's face may invite stereotyping," but that a negative stereotype does not influence the outcome of the legal decision-making in which it is involved.

The results of this experiment, along with other experiments on facial stereotyping (Goldstein, et al., 1984; Saladin, Saper, & Breen, 1988; Shoemaker, et al., 1973; Yarmey, 1993), indicate that participants have an automatic stereotypic response to the target photograph deemed bad or criminal. This experiment, unlike others, went further and showed that participants can disregard this "automatic response". Participants deemed the bad guy target as more likely to commit this type of crime in the future and as less responsible and conscientious than the good guy target, yet they overwhelmingly acquitted the defendant regardless of condition. These results are similar to Devine (1989) who examined automatic stereotypic responses in high and low prejudice people. She found that all participants elicited these automatic stereotypic responses, but that under

circumstances that allowed controlled processing (i.e., more normal day to day functions), low-prejudice people appeared to actively reject the automatic stereotypic responses and replaced them with more equality-oriented thoughts. Participants in the present experiment exhibited prejudice reactions to the bad guy target, i.e., they were more likely to think he would commit this type of crime in the future, but were unwilling to convict the defendant based on the evidence presented. This does not necessarily indicate that participants were low prejudice, but rather that they can control the processing of information.

Given the results of Experiment 3, it could be that the experimental conditions might have allowed participants to control their information processing. This may have not been the case in previous studies on facial stereotyping. One obvious difference between the current study and previous studies on facial stereotyping of criminal versus non-criminal faces was the use of a full criminal case file. These previous studies either did not include any sort of description of a crime (Goldstein, et al. 1984 & Yarmey, 1993) or they included a contrived vignette of a crime and asked participants to select a picture that depicted a person who was most likely to have committed the crime described (Saladin, Saper, & Breen, 1988; Shoemaker, et al., 1973). Contrived vignettes provide minimal amounts of information. As such they diminish an individual's ability to systematically evaluate information and in turn they increase the likelihood that individuals will look to other cues for processing, e.g., "the guy looks bad so he must be bad" (Chaiken, Liberman, & Eagly, 1989; Petty & Caccioppo, 1989). So, providing participants with more information about the case may have enhanced their ability to scrutinize the information and weigh both sides of the case and decreased their likelihood of utilizing

peripheral cues to make a decision. It could be that the use of a more elaborate case file in Experiment 3 of the current study resulted in the finding that people were not so easily influenced by the elicitation of a stereotype, which is contrary to previous findings on facial stereotyping.

Another difference between the current experiment and previous experiments on facial stereotyping is that the tasks that participants engaged in during previous experiments were different than the tasks that participants in the current experiment engaged in. Participants in previous experiments simply identified the picture that best depicted an individual guilty of a specific crime or they read a contrived vignette and then selected the picture that best depicted an individual guilty of a specific crime. Participants in Experiment 3 of the present study viewed a full case file, which included arguments and evidence from both the prosecution and defense and within this case file the picture of the defendant was embedded amongst other pictures of other major players in the case. Embedding the picture should have lowered the demand characteristics of the current experiment compared to previous experiments and as such, participants should have been less likely to figure out the experimental paradigm and act according to what they believed were the desired results of the experimenter, i.e. the "good-subject tendency" (McBurney, 1994). Because awareness questions were either never posed to participants or never reported, it is difficult to say definitively whether the experimental methods had an influence or not.

The current experiment also differed from previous experiments on facial stereotyping of criminals in that the current experiment was more similar than other experiments to a real trial. It could be that this led participants into taking their role as

decision-makers more seriously than previous experiments. Because of this perception of their roles, participants may have engaged in more scrutiny of the information and weighed the results of their actions more seriously when making decisions about the defendant.

It could be argued that the failure to find effects of the stereotype on the verdict measure indicates ceiling levels of performance. This argument can be rejected because the pattern of results (no difference between the bad guy and good guy conditions) was the same for the two other measures, current probability of guilt and sentence assignment, yet there were no ceiling effects on either of these measures. The overall mean for the current probability of guilt was .37 (SD = .29). For sentence assignment nearly half (46%) of all the participants selected sentences that were in the middle of the scale and not at either extreme of the scale (see Table 5 for results). Furthermore, in earlier pilot studies a similar case file was used which yielded conviction rates that were intermediate (overall conviction rates for the bad guy and good guy conditions in the pilot studies were .36 and .42, respectively). However, the principal results were the same as the present experiment: there was no effect of appearance on verdict and the present estimation of guilt, but there was a strong effect of appearance on the future probability of committing a similar crime.

The results of Experiment 3 of the current study were not congruent with the initial hypothesis that participants would be more likely to convict the bad guy than the good guy. The results supported the null hypothesis that participants would be equally as likely to convict the bad guy as the good guy. Several important factors support the acceptance of the null hypothesis for the results of Experiment 3. The first factor that

indicated the results supported the null hypothesis and were not due to remnants of the design or stimulus materials was the power of the test. Given the large number of participants, Experiment 3 was powerful enough to detect a difference should there have been one, $\phi = .96$. The second factor that supports the acceptance of the null hypothesis was the control exerted over the experimental materials and conditions. Experiment 2 established that there was a significant difference between the way participants perceived the bad and good guy targets. Also, measures within Experiment 3 indicated that participants had different perceptions of the two targets. So, the independent variables, the target faces, were distinct enough from one another that participants differentially perceived the faces and perceived them as initially hypothesized. In terms of the stimulus materials, the case file had been used previously in mock trials that resulted in half of the juries acquitting and half convicting the defendant (J.A. Tanford, personal communication, February 24, 1998). So, the materials were ambiguous enough to maximize the likelihood of finding effects of the manipulated variable, facial appearance. Participants went through the same procedures and were exposed to the same materials across all conditions, except for the target picture they viewed. The third factor that supports the acceptance of the null hypothesis is the sensitivity of the dependent measures. With the dependent measure verdict, there is no way to determine whether other researchers had obtained comparable results to those in Experiment 3 because this measure had not been used in previous experiments. Rather, researchers used ordinal scales to evaluate participants' likelihood of convicting the defendant. The verdict measure in and of itself is not a very sensitive measure and that was why there was a comparable dependent measure, current estimation of guilt. One reason the second, more sensitive measure of

likelihood to convict was not similar to results in previous research is that it consistently came after participants chose their individual verdict (this was done in order to maintain ecological validity). Participants simply used their verdict preference as an anchor in which to base their current estimation of guilt. Other dependent measures, such as future estimation of guilt and the personality traits of reliable and conscientious, indicated differences between the two conditions. So, differences in perceptions between the good and bad guy targets existed, but how participants acted on those perceptions were counter to the initial experimental hypothesis.

The results of this experiment must be looked at in their entirety. There were differential responses depending on condition and differential responses depending on the dependent measure. Obtaining different responses to the targets depending on condition was expected and even hypothesized, but obtaining different responses to each of the dependent measures was not. One possible way to understand the variation between dependent measures is to examine the gravity of the question. It could be argued that the gravity of the question posed to participants elicited varying decision-making operations. The measures used in Experiment 3 differed in the direct effect they could have on the mock defendants' life. Verdict was the measure that had the greatest direct effect on the defendant, whereas future probability that the defendant would commit the crime had the smallest potential impact. Convicting the defendant meant that participants felt the defendant was a danger to society and should have his freedom revoked in some way. This is a very serious consideration knowing the value placed on personal freedom. The measure future probability was simply a conjecture that had no impact on the defendant's personal freedom and as such was a less serious decision for participants. Participants

were not necessarily asked to be jurors, but they were asked to make judgments about the defendant, as would jurors. In this situation they may have weighed the direct effects of their responses to different measures on the defendant, i.e., choosing guilty would revoke the freedom of the defendant, whereas indicating the defendant had a high probability of committing the crime in the future would not revoke his freedom. So, the more directly a measure affected the defendant, i.e., verdict and current estimation, the less impact there was of the stereotype. The less directly a response impacted the defendants' life, the greater impact there was of the stereotype.

One limiting factor in this study was the use of students, who may be low prejudiced individuals. University curriculum and the diversity of the student body at Florida International University can potentially attenuate the effects of stereotypes and prejudice by merely shedding light on their existence and as Devine (1989) demonstrated low prejudice individuals actively reject the automatic elicitation of a stereotype. It could be that merely shedding light on the existence of stereotypes attenuates the existence of stereotypes, which in turn would classify students as low prejudice individuals. The majority of the participants in these experiments were currently enrolled university students. The results of Experiment 3 may simply be a reflection of a participant pool of low prejudiced individuals who rejected the stereotype on more serious questions and reflected the influence of the stereotype on less serious questions. People who are not currently enrolled students may be higher prejudiced people because they are not exposed to university curricula. It could be that using individuals who are not university students might generate different results than those found in this study, i.e., they would be more

likely to convict the bad guy target and deem it more probable that he committed the current crime. Again, this is a separate issue to be looked at in future research.

This study takes previous research one step further and looks at the impact of a full case file on legal decision-making and criminal appearance. As stated previously, it is uplifting to know that people are not so easily influenced by negative stereotypes when making decisions that have ramifications on another's life. All three experiments indicated that criminal facial stereotypes exist, but Experiment 3 demonstrated that other factors could outweigh the influence of the negative facial stereotype. This study as other studies on jury decision-making indicated that individuals do take these types of legal decisions seriously and are not so easily influenced by the many extra-legal factors that legal scholars have for so long cautioned about.

References

Alley, T. R. (ed.) (1988). Social and Applied Aspects of Perceiving Faces. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Berry, D. S. & Zebrowitz-McArthur, L. (1985). Some components and consequences of a babyface. Journal of Personality and Social Psychology, 48, 312-323.

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

Bodenhausen, G. V. (1990). Second- guessing the jury: Stereotypic and hindsight biases in perceptions of court cases. Journal of Applied Social Psychology, 20, 1112-1121.

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

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

Chaiken, S., Liberman, A., & Eagly, A. H. (1989). Heuristic and Systematic information processing within and beyond the persuasion context. In J.S. Uleman & J.A. Bargh (Eds.), Unintended Thought (pp. 212-252). New York: Guilford Press.

Devine, P.G. (1989). Stereotypes and prejudice: Their automatic and controlled components. Journal of Personality and Social Psychology, 56, 5-16.

Dion, K., Berscheid, E., & Walster, E. (1972). What is beautiful is good. Journal of Personality and Social Psychology, 24, 285-290.

Dolan, M. (1994, September 25) Trials: Jurors' flaws reflect the frailties of human nature. The Los Angeles Times, pp. A1, A22-A23.

Fiske, S. T., & Neuberg, S. L. (1990). A continuum model of impression formation from category-based to individuating processes: Influences of information and motivation on attention and interpretation. In M. P. Zanna (Ed.), Advances in Experimental Social Psychology (Vol. 3, pp. 1-74). San Diego, CA: Academic Press.

Fiske, S. T., & Taylor, S. E. (1984). Social Cognition. Reading, MA: Addison-Wesley.

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

Hamilton, D. L. (1981). Cognitive processes in stereotyping and intergroup behavior. Hillsdale, NJ: Earlbaum.

Hastie, R. (1980). Memory for behavioral information that confirms or contradicts a personality impression. In R. Hastie, T. Ostrom, E. Ebbeson, R. Wyer, D. Hamilton, & D. Carlston (Eds.), Person memory: The cognitive basis of social perception (pp. 155-177). Hillsdale, NJ: Earlbaum.

Levine, J. P. (1992). Juries and politics. Pacific Grove, CA: Brooks/ Cole.

Light, L. L., Hollander, S. & Kayra-Stuart, F. (1981). Why attractive people are harder to remember. Personality and Social Psychology Bulletin, 7, 269-276.

Locher, P., Unger, R., Sociedade, P., & Wahl, J. (1993). At first glance: Accessibility of the physical attractiveness stereotype. Sex Roles, 28, 729-743.

Locksley, A., Borgida, E., Brekke, N., & Hepburn, C. (1980). Social stereotypes and judgments of individuals: An instance of the base-rate fallacy. Journal of Experimental Social Psychology, 18, 23-42.

Macrae, C. N.; Bodenhausen, G. V.; & Milne, A. B. (1995). The dissection of selection in person perception: Inhibitory processes in social stereotyping. Journal of Personality and Social Psychology, 69, 397-407.

Macrae, C. N.; Milne, A. B.; & Bodenhausen, G. V. (1994). Stereotypes as energy-saving devices: A peek inside the cognitive toolbox. Journal of Personality and Social Psychology, 66, 37-47.

McBurney, D. A. (1994). Research Methods (3rd Ed.). Pacific Grove, CA: Brooks/Cole Publishing Co.

Nisbett, R. & Ross, L. (1980). Human inference: Strategies and shortcomings. Englewood Cliffs, NJ: Prentice- Hall.

Petty, R. E., & Cacioppo, J. T. (1989). The elaboration likelihood model of persuasion. In Petty, R. E., & Cacioppo, J. T. (Eds.), Communication and persuasion: Central and peripheral routes to attitude change (pp. 123-203). New York: Springer-Verlag.

Perret, D. I., May, K. A., & Yoshikawa, S. (1994). Facial shape and judgments of female attractiveness. Nature, 368, 239-242.

Rothbart, M., Fulero, S., Jensen, C., Howard, J. & Birrell, P. (1978). From individual to group impressions: Availability heuristics in stereotype formation. Journal of Experimental Social Psychology, 14, 237-255.

Reed, S. K. (1972). Pattern recognition and categorization. Cognitive Psychology, 3, 382-407.

Saladin, M., Saper, Z., & Breen, L. (1988). Perceived attractiveness and attributions of criminality: What is beautiful is not criminal. Canadian Journal of Criminology, 251-259.

Schneider, D. J., Hastorf, A. H., & Ellsworth, P. C. (1979). Person perception (2nd ed.). Reading, MA: Addison- Wesley.

Smith, E. R., & Zarate, M. A. (1990). Exemplar and prototype use in social categorization. Social Cognition, 8, 243-262.

Terry, R. L., & Krantz, J. H. (1993). Dimensions of trait attributions associated with eyeglasses, men's facial hair, and women's hair length. Journal of Applied Social Psychology, 23, 1757-1769.

Terry, R. L. & Davis, J. S. (1976). Components of Facial Attractiveness. Perceptual and Motor Skills, 42, 918.

Zarate, M. A., & Smith, E. R. (1990). Person categorization and stereotyping. Social Cognition, 8, 161-185.

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, 603-623.

Zebrowitz, L. A., Voinescu, L. & Collins, M. A. (1996). "Wide-eyed" and "crooked-faced": Determinants of perceived and real honesty across the life span. Personality and Social Psychology Bulletin, 22, 1258-1269.

Personnel Management Study

Please wait until you have heard instructions from the researcher.

Response Sheets

On a scale of 1 to 10 (with **1** indicating that it would be **LEAST LIKELY** to hire Wayne Walker and **10** indicating that you would be **MOST LIKELY** to hire Wayne Walker), please indicate how likely it would be for you to hire Wayne Walker as a cashier in your convenience store. Please circle your choice.

1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Least Likely									**Most Likely**

In the following spaces please write down the various pieces of information you used in making your decision. Please write **one** piece of information in each of the spaces provided.

1. _____

_____

2. _____

_____

3. _____

_____

4. _____

_____

5. _____

_____

When responding to the questions contained in this section, you will be asked to use a scale with values ranging from 1 to 7 to rate your opinions. A **1** indicates that you feel Wayne Walker in no way possesses the quality asked about. A 7 indicates that Wayne Walker very much possesses the quality asked about.

In the following section, please circle the number that best describes your opinion of Wayne Walker.

1. How responsible is the employee, Wayne Walker?

1	**2**	**3**	**4**	**5**	**6**	**7**
Not responsible at all					Very responsible	

2. How honest is the employee, Wayne Walker?

1	**2**	**3**	**4**	**5**	**6**	**7**
Not honest at all						Very honest

3. How sincere is the employee, Wayne Walker?

1	**2**	**3**	**4**	**5**	**6**	**7**
Not sincere at all						Very sincere

4. How reliable is the employee, Wayne Walker?

1	**2**	**3**	**4**	**5**	**6**	**7**
Not reliable at all						Very reliable

5. How moral is the employee, Wayne Walker?

1	**2**	**3**	**4**	**5**	**6**	**7**
Not moral at all						Very moral

6. How conscientious is the employee, Wayne Walker?

1	**2**	**3**	**4**	**5**	**6**	**7**
Not conscientious at all					Very conscientious	

7. How likable is the employee, Wayne Walker?

1	**2**	**3**	**4**	**5**	**6**	**7**
Not likable at all						Very likable

In your own words, please tell us what you think this experiment is about.

DEMOGRAPHIC INFORMATION

1. Date of birth: _____

2. Gender:

   ☐ Female
   ☐ Male

3. Race/Ethnic background:

   ☐ Non-Hispanic White
   ☐ Hispanic White
   ☐ Non-Hispanic Black
   ☐ Hispanic Black
   ☐ Other: _____

4. Please indicate how much formal education you have received:

   ☐ Less than High School
   ☐ Some High School
   ☐ Received High School Diploma
   ☐ Some College
   ☐ Received College Degree
   ☐ Some Post-Graduate
   ☐ Received Post-Graduate Degree

5. Please list your current occupation.

_____

Thank you for your participation in this study.

Questionnaire
Jury Decision-Making Study

Instructions

On the following pages are a number of questions about the case. Please complete this questionnaire independently. We are interested in your opinions and impressions. This is a research project not a test. There are no right or wrong answers. Please consider each item carefully, and be as honest as possible in your answers. Please answer each question in the order that they appear.

If you have any questions about these instructions or any other aspect of the questionnaire, feel free to ask the researcher. If you do not have any questions at this moment please turn the page and begin filling out the questionnaire.

What is your verdict in the case of the State of Columbia vs. Wayne B. Walker? *Please indicate your choice by circling either verdict.*

GUILTY **NOT GUILTY**

(2) Please provide your personal estimation of the probability that the defendant did in fact commit the crimes of armed robbery and attempted murder given the set of evidence you just viewed.

Please indicate your personal estimation of the probability that the defendant is guilty on a scale of 0% to 100%. 0% indicates that there is no likelihood that the defendant did commit the crimes of armed robbery and attempted murder and 100% indicates that it is certain that the defendant did commit these crimes.

_____%

(3) The maximum sentence for the crime in this case (armed robbery and attempted murder) is 25 years in prison. If you could assign the sentence, what sentence would you give Wayne Walker, the defendant? *Please mark your sentence preference in the space provided.*

_____ 20 to 25 years in prison
_____ 15 to 20 years in prison
_____ 10 to 15 years in prison
_____ 5 to 10 years in prison
_____ 1 year in prison and four years probation and counseling
_____ 5 years probation and counseling
_____ One year probation only
_____ No sentence, I would allow him to go free

(4) Please provide your personal estimation of the probability that the defendant will commit this type of crime in the future.

Please indicate your personal estimation on a scale of 0% to 100%. 0% indicates that there is no likelihood that the defendant will commit these types of crimes in the future and 100% indicates that it is certain that the defendant will most definitely commit these types of crimes in the future.

_____%

(5) Please list all the facts in the case that you can remember. *Please list the facts on the lines provided below and on the back if needed.*

In the following spaces please write down the top five most important pieces of information you used in making your verdict decision. Please write one piece of information in each of the spaces provided. Then rank the pieces of information, on a scale of 1 through 5, on their importance in helping you make a verdict decision. A **"1"** indicates that this piece of information was **most important** and **"5"** indicates that this information was the **least important** of you top five pieces of information.

1. _____

_____

2. _____

_____

3. _____

_____

4. _____

_____

5. _____

_____

When responding to the questions contained in this section you will be asked to use a scale with values ranging from 1 to 7 to rate your opinions. A **"1"** indicates that you feel Wayne Walker **in no way** possesses the quality asked about in the question and a **"7"** indicates that Wayne Walker **very much** possesses the quality asked about.

In the following section, please circle the number that best describes your opinion Wayne Walker.

8. How responsible is Wayne Walker?

1 **2** **3** **4** **5** **6** **7**
Not responsible at all Very responsible

9. How honest is Wayne Walker?

1 **2** **3** **4** **5** **6** **7**
Not honest at all Very honest

10. How sincere is Wayne Walker?

1 **2** **3** **4** **5** **6** **7**
Not sincere at all Very sincere

11. How reliable is Wayne Walker?

1 **2** **3** **4** **5** **6** **7**
Not reliable at all Very reliable

12. How moral is Wayne Walker?

1 **2** **3** **4** **5** **6** **7**
Not moral at all Very moral

13. How conscientious is Wayne Walker?

1 **2** **3** **4** **5** **6** **7**
Not conscientious at all Very conscientious

14. How likable is Wayne Walker?

1 **2** **3** **4** **5** **6** **7**
Not likable at all Very likable

DEMOGRAPHIC INFORMATION

5. Date of birth: _____

6. Gender:

- ☐ Female
- ☐ Male

7. Race/Ethnic background:

- ☐ Non-Hispanic White
- ☐ Hispanic White
- ☐ Non-Hispanic Black
- ☐ Hispanic Black
- ☐ Other: _____

8. Please indicate how much formal education you have received:

- ☐ Less than High School
- ☐ Some High School
- ☐ Received High School Diploma
- ☐ Some College
- ☐ Received College Degree
- ☐ Some Post-Graduate
- ☐ Received Post-Graduate Degree

6. Please list your current occupation.

_____

Thank you for your participation in this study.

DEBRIEFING

This page is to be used during the debriefing section of the experiment. Ask these questions out loud in the order presented and record the answers in the space provided. *Please do not hand this page to participants.*

1. In your own words, describe what you feel this experiment is about.

2. Do you think the way the defendant looks in any way influenced your verdict decision?

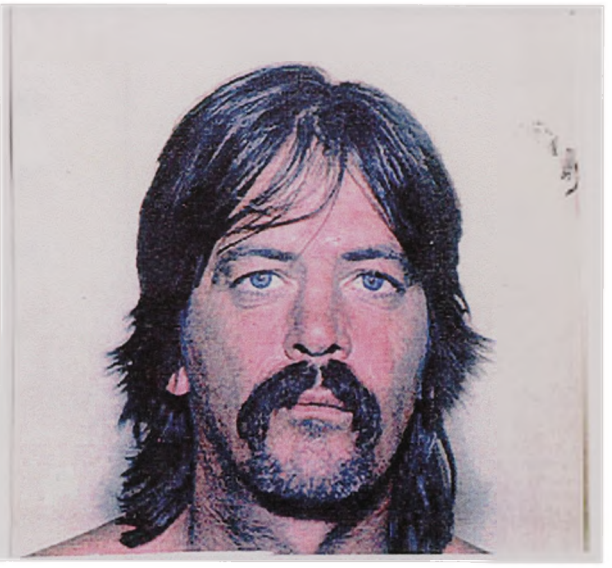
YES **NO**

3. Explain experimental paradigm to participants. Ask if they have any questions and thank them for their time when you are finished.

Figure 1. Target Pictures Obtained in Experiment 1

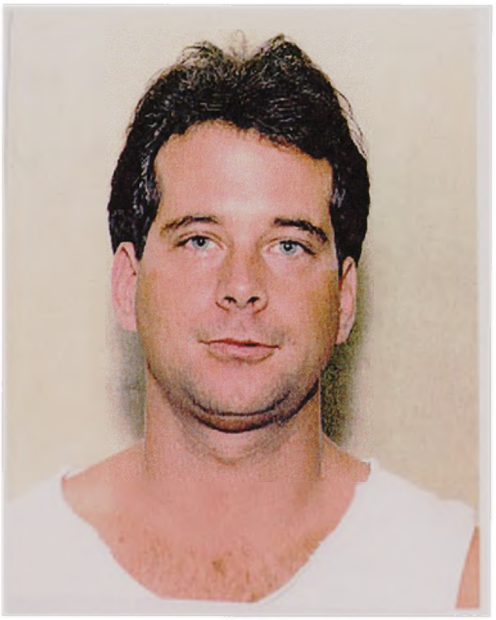
Bad Guys

Picture 10

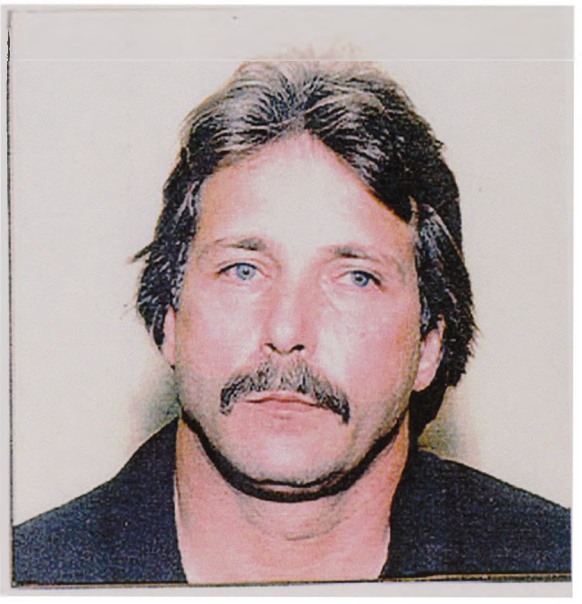


Picture 20

Good Guys



Picture 48



Picture 15

Table 1. Proportion of Good, Bad and Mixed Responses to Pictures Selected in Experiment 1

Picture	Good	Bad	Mixed/ undetermined
10	.07	.83	.10
20	.03	.80	.17
15	.66	.03	.37
48	.66	.17	.23
Overall	.19 (SD = .20)	.40 (SD = .13)	.41 (SD = .20)

Table 2. Personality Trait Means by Condition in Experiment 2

Trait	Bad Guy [a]		Good Guy [b]		
	M	SD	M	SD	F (1, 28), p > .05
Responsible	3.60	0.99	4.06	1.18	1.08
Honest	4.07	1.22	3.63	1.20	0.70
Sincere	4.20	1.08	3.75	1.24	1.59
Reliable	3.40	1.12	3.94	1.53	1.03
Moral	4.00	0.88	3.75	1.00	0.69
Conscientious	4.00	1.13	3.88	1.26	1.16
Likable	3.80	1.52	3.88	1.03	0.04

[a] n = 14

[b] n = 16

Table 3. Proportion of Verdict Choices by Condition in Experiment 3

		Verdict	
Condition	n	Guilty	Not guilty
Bad Guy	48	.19	.81
Good Guy	45	.18	.82

Table 4. Current and Future Estimates by Condition in Experiment 3

	Estimation			
Condition	Current	SD	Future	SD
Bad Guy	34.58	32.54	45.83	30.04
Good Guy	39.44	25.74	27.87	25.82

Table 5. Proportion of Sentence Assignments by Condition in Experiment 3

Sentence Assigned	Bad Guy	Good Guy
20 to 25 years in prison	.04	.18
15 to 20 years in prison	.04	.02
10 to 15 years in prison	.08	.07
5 to 10 years in prison	.10	.00
1 year in prison and 4 years probation and counseling	.08	.18
5 years probation and counseling	.06	.09
One year probation only	.10	.09
No sentence, I would allow him to go free	.48	.38

Table 6. Personality Trait Means by Condition in Experiment 3

Trait	Bad Guy [a]		Good Guy [b]		
	M	SD	M	SD	F (1, 91)
Responsible	3.40	1.28	3.91	1.22	3.92*
Honest	3.90	1.51	4.31	1.29	2.02
Sincere	3.83	1.60	4.27	1.44	1.88
Reliable	3.67	1.40	3.76	1.38	0.09
Moral	3.29	1.15	3.56	1.37	1.02
Conscientious	3.35	1.16	3.93	1.44	4.61*
Likable	3.50	1.37	3.87	1.41	1.62

Note. Judgements were made on 7-point scales (1 = target in no way possesses trait, 7 = target very much possesses the trait

[a] n = 48

[b] n = 45

* $p \leq .05$