9-12-2024

A mock juror investigation of the influence of extralegal factors of juvenile defendants: Gender, attractiveness, psychopathology, and race

Jennifer A. Ranjit
University of Windsor

A MOCK JUROR INVESTIGATION OF THE INFLUENCE OF EXTRALEGAL FACTORS OF JUVENILE DEFENDANTS: GENDER, ATTRACTIVENESS, PSYCHOPATHOLOGY, AND RACE

By

Jennifer A. Ranjit

A Thesis
Submitted to the Faculty of Graduate Studies
through the Department of Psychology
in Partial Fulfillment of the Requirements for
the Degree of Master of Arts at the
University of Windsor

Windsor, Ontario, Canada

2024

A MOCK JUROR INVESTIGATION OF THE INFLUENCE OF EXTRALEGAL FACTORS OF JUVENILE DEFENDANTS: GENDER, ATTRACTIVENESS, PSYCHOPATHOLOGY, AND RACE

by

Jennifer A. Ranjit

APPROVED BY:

D. Levin

School of Social Work

K. Lafreniere

Department of Psychology

C. M. Langton, Advisor

Department of Psychology

August 20, 2024

DECLARATION OF ORIGINALITY

I hereby certify that I am the sole author of this thesis and that no part of this thesis has been published or submitted for publication.

I certify that, to the best of my knowledge, my thesis does not infringe upon anyone's copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my thesis, published or otherwise, are fully acknowledged in accordance with the standard referencing practices. Furthermore, to the extent that I have included copyrighted material that surpasses the bounds of fair dealing within the meaning of the Canada Copyright Act, I certify that I have obtained a written permission from the copyright owner(s) to include such material(s) in my thesis and have included copies of such copyright clearances to my appendix.

I declare that this is a true copy of my thesis, including any final revisions, as approved by my thesis committee and the Graduate Studies office, and that this thesis has not been submitted for a higher degree to any other University or Institution.

ABSTRACT

Extralegal factors are those that do not pertain to the facts of a case in a court of law. In Canada, the youth criminal justice system incarcerates racialized youth, males, and those with psychiatric diagnoses at disproportionate rates. The mock juror paradigm is one way to investigate implicit biases arising from extralegal factors that might affect decisions such as guilt or innocence. Although jury trials are not used in the youth criminal justice system, a range of professions exercise judgment about whether youth enter the system and how far they progress in it; therefore, it is worthwhile to investigate extralegal factors that may give rise to implicit biases. This project included three studies, each investigating the effects of attractiveness and gender. Additionally, in Study One, a possible biasing effect for crime type (assault or fraud) was investigated; in Study Two the possibly biasing effect of psychiatric diagnosis (conduct disorder, psychopathic traits, schizophrenia, or no diagnosis) was also investigated; and in Study Three the biasing influence of race (Black or White) was investigated. Participants were female undergraduate students at the University of Windsor. Across the three studies, attractiveness and gender did not affect decision-making in isolation. Attractiveness and crime type were associated with higher guilt ratings for attractive defendants accused of assault. Attractiveness was associated with higher guilt ratings for attractive defendants with psychopathic traits and lower guilt ratings for attractive defendants with schizophrenia. Race did not interact with attractiveness or gender to produce biased guiltiness ratings. These results contribute to the body of mock juror research, particularly as it pertains to youth involved in the criminal justice system. With further methodological refinement, replication and extension of these findings are needed with representative samples of the jury eligible as well as those whose employment brings them into contact with youth at risk of involvement in the justice system.

DEDICATION AND ACKNOWLEDGEMENTS

This thesis is dedicated to my parents, for their faith in me that I can do more and for taking it for granted that I will succeed. Dad, I'm halfway there.

I would like to extend my deepest thanks to my research supervisor, Dr. Calvin Langton, for your guidance through this project, the ones that came before, and the ones ahead. Thank you for your unwavering support and patience, encouragement, and believing in me. I have learned so much from you and I appreciate how you have helped me grow more than I can say. I would also like to thank my committee members, Dr. Kathryn Lafreniere and Dr. Dana Levin, for your feedback and thoughtful comments.

Thank you to my sister, Kathryn, for helping me with search terms, teaching me all the cool stuff Word can do, and supporting me through everything – school and life. To my child clinical cohort crew, Alana, Carly, Emily, and Taffy – I'm so grateful we all started here at the same time. I couldn't have asked for a better group of women to journey with through graduate school. To Janet, thank you for your phrasing tips, validation and friendship. And Jen, thank you for always having my back.

Finally, I would like to thank my husband, Chris, and daughter Arya. Chris, thank you for taking care of us (e.g., making sure we always have clean laundry) and for all your patience and comfort. Arya, thank you for reminding me that story time is very important. I love you both the most and would not be where I am without you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

CHAPTER 1 : INTRODUCTION

Overview

Section 11(f) of the Canadian Charter of Rights and Freedoms (1982) includes, "The guarantee of the benefit of a trial by jury implies that the jury will be impartial and representative." Although this is a guaranteed right, extralegal factors have repeatedly been found to influence verdicts made by mock jurors. The practical implications of findings such as these can be found in the overrepresentation of Black and Indigenous youth incarcerated in Canada. For example, Black youth comprised 4% of the youth population and 18% of youth custody admissions in 2020/2021 (Statistics Canada, 2022). Indigenous youth comprised 8% of the youth population but accounted for 50% of custody admissions in 2020-2021 (Statistics Canada, 2022). Society pays a hefty toll for the youth criminal justice system. In 2010, the estimated costs were $1.4 billion (Zhang & Hoddenbagh, 2013), with approximately $52k spent annually per youth in secure custody (Gabor, 2015). Given the disproportionate treatment of certain youth and the societal costs, it seems worthwhile to investigate which extralegal factors are associated with higher rates of youth incarceration.

Mock-jury decision making has been a subject of research since the 1960s. In this type of research, participants are asked to make decisions based on descriptions of court proceedings as though they were real jury members. Researchers can evaluate how mock juror decision-making varies based on manipulated variables within the descriptions of court proceedings (Bornstein et al., 2017). Some researchers have found mock juror decisions to be associated with the attractiveness of the defendant (Landy & Aronson, 1969; Rice et al., 2020). Some have shown an interaction between defendant race and juror race (Abwender & Hough, 2001; Stevenson & Bottoms, 2009; Pica et al., 2017). Others have shown that different mental health diagnostic

labels are associated with varying degrees of risk, culpability, and capacity for behaviour change (Boccaccini et al., 2008; Taylor et al., 2019). Still others have demonstrated that it is the type of crime a defendant has been accused of that is associated with the willingness to convict (Esses & Webster, 1988). The interaction of mock juror and defendant gender (Agthe et al., 2016) have also been associated with mock juror decision making.

The present project was undertaken to investigate the effects of these extralegal factors across three studies using the mock juror paradigm and focusing on the attractiveness and gender of the defendant as common variables throughout. The first study was focused on the effects of attractiveness, gender, and type of crime of the defendant on mock juror decision making. Foci in the second study were the effects of attractiveness, gender, and psychiatric diagnoses of the defendant on mock juror decision making. For the third study, the effects of attractiveness, gender, and race of the defendant on mock juror decision making was investigated.

Youth Criminal Justice Act

The *Youth Criminal Justice Act* (YCJA, 2002) was enacted to address concerns about the previous legislation governing Canada's youth justice system, the *Young Offenders Act* (YOA, 1985). These concerns were regarding the overuse of the courts, incarceration for minor offences, transfers to adult court, and discriminatory sentencing practices. Under the YOA, 80% of youth were serving custodial sentences for minor offenses leading to Canada having the highest rate of youth incarceration in the Western world (Bala et al., 2009). The changes brought about by the YCJA increased the use of extrajudicial measures by police officers in the youth justice system. These measures include taking no further action, informal police warnings, police cautions (e.g., a letter from police to a youth and their parents, or a meeting with the police), crown cautions (prosecutors provide a letter to the youth and their parents), referrals (e.g.,

community programs, counseling), or extrajudicial sanctions (e.g., compensating the victim, volunteer work, or specialized programs). As a result of the YCJA, the rate of youths in provincial correctional services decreased from 15.39 per 10 000 youth in 2002-2003 to 2.37 per 10 000 youth in 2021-2022 (Statista Research Department, 2023).

The YCJA is used for youth aged 12 to 17 years. The maximum length of a sentence can range from two to ten years depending on the crime, and these sentences can be served in the community or in custody. Custodial sentences are typically reserved for those who have committed violent or serious repeat offences. With the YOA, a youth charged with first- or second-degree murder could choose trial by superior court judge and jury. The YCJA diverges from the previous YOA in that regardless of offence type, a jury is never used for determining guilt. In some cases, when a violent offence (i.e., murder) has been committed, a judge may deem it necessary to impose an adult sentence, but this always happens after guilt has been determined by the judge in Youth Court.

Although youth are not subjected to judgment by a jury in Canada, other countries do rely on jury judgments for youth (e.g., United States). Additionally, they are subject to the court of public opinion. As with adult court settings, youth court hearings are open to the public and media. In a review of media depictions of youth crime in Canada, Silcox (2022) found that news reports tend to follow youth crime trends in Canada. However, some spikes in coverage were associated with perpetrator class, gender, race, and age. For example, one spike in coverage was regarding societal concerns about impoverished neighbourhoods, visible minorities and gang affiliations following the murder of Jordan Manners in 2007. It was reported as a common problem with poor minority neighbourhoods, rather than a serious but isolated incident (Silcox, 2022). This type of reporting can create or bolster negative implicit attitudes, which can affect explicit attitudes (Arendt & Northup, 2015). The effect of biased reporting of crime trends

suggests it is worthwhile to investigate the implicit biases that some may hold, with consideration of the influence certain members of society may have regarding public policy and legislation. The mock-juror design is one such way of investigating these biases.

Discrimination in the Age of the YCJA

Despite provisions put in place to limit discriminatory sentencing practices, Indigenous and Black youth are still disproportionately incarcerated in Canada. Indigenous youth comprise 8% of Canada's youth population and accounted for 50% of youth custody admissions and Black youth comprise 4% of the youth population and 18% of youth custody admissions in 2020/2021 (Statistics Canada, 2022). The YCJA includes specific provisions for Indigenous youth, stating that sanctions other than custody should always be used when appropriate, especially when considering the circumstances of young Indigenous people. The systemic factors that may predispose these youth to adverse circumstances may limit their culpability.

Adolescence is also a time of significant change. Some estimates state that approximately one third of youth will engage in delinquency, but most will desist as they mature (Lambie & Randell, 2013). Because adolescents are particularly susceptible to peer influences, it makes sense that incarceration may inhibit rehabilitation. Incarceration, especially for minor offences or prolonged incarceration, prevents youth from engaging with more prosocial peers, a significant protective factor for desistence from crime, or naturally maturing out of delinquent behaviour (Lambie & Randell, 2013). It also exposes them to more risk factors that predict recidivism, such as victimization (verbal, physical, sexual, and emotional), isolation (Ashkar & Kenny, 2008), and stressful conditions of confinement, which may lead to suicidal ideation and exacerbate mental health problems (Bonner, 2011). The youth most at risk of incarceration are those who are systemically disadvantaged, as evidenced by the abovementioned statistics. Thus, these youth

may be the least likely to be rehabilitated, which is, or at least should be, a primary objective of sentencing (Verbrugge, 2003). Indeed, section 718 of the Criminal Code (1985) states the foremost purpose of sentencing is to protect the public, and to contribute to the maintenance of a just and safe society, which is achieved through one or more of six objectives: (a) to denounce unlawful conduct and the harm done to victims or to the community that is caused by unlawful conduct; (b) to deter the offender and other persons from committing offences; (c) to separate offenders from society where necessary; (d) to assist in rehabilitating offenders; (e) to provide reparations for harm done to victims or to the community; and (f) to promote a sense of responsibility in offenders, and acknowledgement of the harm done to victims or to the community.

An important point is that it is police who give youth the option of an alternative (i.e., a warning or diversion) to receiving a formal charge and going through youth court. Issues can arise when police discretion is biased in any way. Samuels-Wortley (2019) investigated the association of police discretion on the number of charges, diversions, and warnings given to youth with 6479 cases from a police service in Ontario, Canada. The results showed a small but significant association between race, gender, and charging decisions, where Black males were charged more often, and given warnings less often than any other race. Moreover, when the offense was cannabis possession, Black males were much more likely to be formally charged when compared to "Other" minority males and White males (Samuels-Wortley, 2019). Given these findings, the disparate incarceration rates of Canada's youth, and that of the influence of media on implicit attitudes and explicit actions, it seems useful to study what factors are associated with ratings of guiltiness and how they interact.

Mock Juror Design

One way in which the extralegal factors that influence jury decision making has been studied has been with the mock-juror (experimental simulation) design (Bornstein et al., 2017). In this design, participants (i.e., the mock jurors) are presented with court proceedings and then asked to make decisions as though they are real jurors. Researchers can manipulate characteristics of the case such as the type of crime, and race or gender of the defendant and victim to discern their effects on juror decision making.

Given the high experimental control in these simulations, mock-juror designs can have high internal validity (Bornstein et al., 2017). Undergraduate student samples are often used in mock-juror studies. Using this type of sample has drawn criticism citing a lack of generalizability and verisimilitude. Despite these criticisms, this design is considered acceptable for several reasons. Researchers can collect information on the 'how' and 'why' in addition to the 'what' of juror behaviour. Feasibility is another benefit, as researchers often have greater access to undergraduate student samples compared to community samples or jury pools. This type of research also contributes to a vast literature on behaviour and psychological processes in simulations in lieu of real-world situations. Perhaps most importantly, mock-juror designs using undergraduate student samples appear to be representative of community samples regarding sentencing severity since 1993 and guilt ratings (Bornstein et al., 2017), suggesting that resolving issues in mock-juror research should focus on factors beyond the sample type.

A limitation of research into the extralegal factors associated with mock-juror decision making is that researchers are often limited to a small number of variables that can be investigated within a single sample pool without being insufficiently powered. In many studies up to three independent variables are manipulated (e.g., defendant and juror race, and defendant attractiveness, Cothran et al., 2017; defendant attractiveness and offense history, Esses &

Webster, 1988). With statistical power in mind and keeping in mind the difficulty of recruiting a sufficiently large sample size for adequate statistical power, the present project involved three studies using the same sample pool to collect information about the effect of several extralegal factors on decision making within one population.

Extralegal Characteristics

Attractiveness

Physical attractiveness has long been associated with benefits for those with high attractiveness. The tendency to attribute positive psychological traits to attractive individuals is known as the attractiveness bias. These attributes include social and intellectual competence, dominance, and a predisposition to cooperate in social dilemma situations. Additionally, those with less attractive faces may be found to be less sociable, altruistic, and intelligent when compared with more attractive faces. For example, Dion et al. (1972) examined the perceptions of undergraduate students to determine the existence of an attractiveness stereotype. They found that attractive individuals were considered more socially desirable, more likely to have important careers, be better spouses with happier marriages, and to have generally happier lives than unattractive individuals. Seligman et al. (1974) found that attractive females were considered to be more responsible for good outcomes and unattractive females to be more responsible for bad outcomes. Landy and Harold (1974) found that high attractiveness increased perceptions of talent, specifically when the talent was perceived as poor in absence of attractiveness information.

Attractiveness has also been found to elicit emotions from observers. In one study, 66 children and 73 adults observed attractive and unattractive faces while their facial electromyography was measured for affective response (Principe & Langlois, 2011). Both child

and adult observers experienced greater disgust and negative affect when observing unattractive faces when compared to attractive faces. Only the adult observers experienced less negative affect or more positive affect when observing the attractive faces. Trustworthiness is another attribute associated with attractiveness. Zhao et al. (2015) demonstrated implicit trust responses to attractive faces with a sample of 59 undergraduate students. Participants were shown images of attractive and unattractive faces paired with words related to trust and distrust. Participants were timed on how quickly they chose the correct word. The findings showed a large effect of participants taking a significantly longer time to select trust words when paired with unattractive faces compared to attractive ones.

The benefits of physical attractiveness have been found to extend to defendants in some mock jury research, although the findings have been inconsistent. For example, Mazzella and Feingold's (1994) meta-analysis showed that attractive defendants have been found significantly less guilty and have received more lenient sentences for crimes of robbery, rape and cheating than unattractive defendants, and significantly more guilty for crimes of negligent homicide than unattractive defendants. Ahola et al. (2010) found that legal practitioners gave harsher judgments to attractive males and unattractive females than attractive females.

These benefits also extend to real people involved with the criminal justice system. For example, Downs and Lyons (1991) found that individuals with average or above-average attractiveness received lower bail and fine amounts than individuals with below-average attractiveness. In a longitudinal study, Beaver et al. (2019) found that high attractiveness was associated with less criminal behaviour, fewer arrests, and fewer convictions than low attractiveness.

Taken together, these findings indicate that facial attractiveness is associated with biases that can influence behaviour. However, these studies often fail to define attractiveness. Some

previous research on attractiveness in the mock jury paradigm has relied on subjective participant ratings of images to determine which images are considered the most and least attractive (e.g., Jacobson, 1981; McKelvie & Coley, 1993; Wareham et al., 2019). Other research has described the images as either attractive or unattractive without explaining what (un)attractive means (e.g., Sigall & Ostrove, 1975). This suggests that there is an objective definition of attractiveness that researchers have used to select images for evaluation. And indeed, some features seem to be considered universally beautiful, suggesting there are specific criteria by which faces are measured (Langlois et al., 2000). Evolutionary psychology indicates facial averageness, symmetry and masculinity/femininity may be biologically based preferences which may communicate mate quality to the observer (Little et al., 2011). Information processing may contribute to these preferences. For example, preferences for average faces may be a preference for familiarity (Rhodes, 2006). Because beauty-based stereotyping ascribes positive attributes to more attractive people, familiarity heuristics may contribute to these preferences (Monin & Oppenheimer, 2005).

However, globalization has influenced societal values. Millennials and Gen Z comprise more than half of the working population and are more diverse and educated than previous generations (Statistics Canada, 2022). Because of their culturally diverse experiences and exposures, millennials and Gen Zers have been found to have less tolerance for injustice (Pichler et al., 2021). Furthermore, most also value increasing diversity (Pichler et al., 2021). To this end, 40% of Gen Zers consider diversity and inclusion to be paramount for beauty brands (Petruzzi, 2022). As societal values shift, there is greater acceptance of and appreciation for people with bodies and faces that do not fall into the abovementioned preferences. Moreover, as advertisers and social media influencers use more people with diverse bodies and features, they increase in familiarity through multiple exposures, and attractiveness takes on a new meaning.

Given these broad meanings of attractiveness, it is necessary to operationalize how it is used in this project. The Chicago Face Database (Ma et al., 2015) is a repository of norm-referenced photos (see the Measures section below for more information). The photos have been subjectively rated on 14 facets, including attractiveness, trustworthiness, baby-facedness, masculinity, femininity and more. Aside from these subjective ratings, the photos have also been objectively measured based on measurements gleaned from evolutionary psychology, and social perception literature associated with attractiveness.

Attractiveness and Punishment. Putz and colleagues (2016) investigated the significance of attractiveness as a factor that may influence how a mock juror metes out punishment. With 197 undergraduate participants, the researchers investigated the effects of attractiveness on norm-enforcement behaviour and emotions using an experimental computer game they developed for the study. The Third-Party Punishment and Reward game required participants to remain neutral towards, punish, or reward two images of "players". One was labelled a cooperator, and the other was labelled a free rider. Over four stereotype consistent rounds (i.e., four attractive cooperators and four unattractive free riders) and four stereotype inconsistent rounds (i.e., four unattractive cooperators and four attractive free riders), Putz and colleagues found that attractive and cooperative players earned higher rewards and free riders who were unattractive received more punishment. Attractive free riders received significantly less punishment than unattractive free riders. Finally, attractive cooperators elicited more contentment than unattractive cooperators and unattractive free riders evoked more intense anger than attractive free riders. An important part of this study design is that participants were compensated with $4 to $20, which was the amount leftover after administering punishments and rewards to players. These rules were made clear at the outset, suggesting participants were willing to reward attractiveness, and punish unattractiveness even at their own expense.

Attractiveness and the Beauty Penalty. Although it appears that people generally believe that what is beautiful is good, there may be a cost to beauty, such that when an attractive person betrays the ideal to which they are held the price they must pay for their transgression is higher than if they had been average looking. Wilson and Eckel (2006) examined this with a group of 206 students from three southeastern universities. Participants were virtually partnered with each other, and one was required to send the other an amount of money up to ten dollars. After viewing a picture of the sender, the receiver was required to guess the amount of money they were sent and if the guess was correct, the receiver would be paid one dollar. Receivers would then refund an amount of money up to three times what they received. The researchers found that whether unattractive senders met or failed to meet receivers' expectations regarding dollar amounts, they were refunded approximately 35% of what they gave. Attractive senders were penalized when they did not send higher dollar amounts to receivers who expected more money from attractive senders, with refunds decreasing from approximately 35% to approximately 26% of what senders gave. These findings demonstrate the beauty penalty in action when implicit expectations go unmet in a victimless situation.

The beauty penalty is also found when attractive people use their looks to manipulate or take advantage of others. When the positive traits and associated expectations attributed to more attractive people are intentionally used to victimize others, people are more willing to severely punish the transgressor. Indeed, Yang et al. (2019) found this effect with female participants who were significantly more likely to severely punish attractive male defendants compared to unattractive male defendants who swindled a new romantic partner. Moreover, female participants did not render disparate punishments to attractive or unattractive male defendants who swindled in an online shopping scam. These findings indicate a beauty penalty only when appearance is perceived as a manipulation tool.

Gender

Another factor associated with disparate sentencing practices is gender. In Canada's youth justice system, boys made up 82% of custody admissions in 2021/2022 (Statistics Canada, 2023). An examination of data from the United States Sentencing Commission from 2001 to 2003 revealed that when all other defendant demographics and legal factors (i.e., offence history, number of charges, trial or guilty plea, guidelines minimum sentence, offence type, and receipt of departure) were comparable, the odds of a female defendant's incarceration was approximately 39% lower than and 23% shorter than a male defendant's (Doerner & Demuth, 2014). Furthermore, extralegal factors were associated with differential sentencing outcomes where the odds of Hispanic men and women versus White men and women were higher by 44% and 13%, respectively. Black women had the lowest odds of being incarcerated compared to all other races. Having an offence history increased the odds of being incarcerated for women more than men, and having less than a high school education increased the odds of being incarcerated for men more than women.

Given the way the abovementioned legal and extralegal factors interact with gender to affect real-world outcomes, it seems worthwhile to investigate how the interactions of gender, attractiveness and other factors contribute to the behaviours of jury-eligible individuals.

Gender and Attractiveness. Defendant and mock-juror gender may also interact with attractiveness in such a way that attractiveness may either increase or decrease ratings of culpability, depending on the genders of the raters and defendants. Indeed, Abwender and Hough (2001) found that participating women were more likely to give longer sentences to unattractive versus attractive female defendants, while participating men were more likely to give longer sentences to attractive versus unattractive female defendants. These findings may represent both the attractiveness bias and the beauty penalty, such that female participants attribute more

negative traits and culpability to unattractive defendants, and male participants consider attractive defendants to have broken their end of the implicit social contract in which they have more positive attributes and behaviours.

Crime Type

The type of crime that has been committed has been found to be associated with more severe penalties when the crime is against a person compared to property (Sanderson et al., 2000). Sentencing disparities based on other extralegal factors have also been found to become insignificant as the severity of the crime increases. For example, Hester and Hartman (2017) examined 17, 671 cases from a southern U.S. Circuit Court, and found that as violent offence histories increased, racial and gender disparities decreased. Thus, racial and gender disparity was more likely to be found for drug, property, and other offences, and not violent (including drug trafficking) offences, especially when the offender did not have an extensive history of criminal behaviour.

Crime Type and Gender. The differential treatment of male and female defendants has been observed in several studies (e.g., Farnworth & Teske, 1995; Franklin & Fearn, 2008; Steffensmeier & Demuth, 2006) and has been attributed to different factors. One of these is the *chivalry thesis,* which posits that women are weaker, passive, and maternal and therefore need the protection of men (Visher, 1983). This may lead to fewer women being formally charged, having their charges reduced to less serious offences, or receiving guilty verdicts less often and receiving more lenient sentences than their male counterparts. This effect has been challenged when female defendants have been accused of a crime that is inconsistent with stereotypical gender roles. For example, with a sample of 394 undergraduate students, Meaux and colleagues (2018) found female mock jurors gave more lenient sentences to female versus male defendants

compared to male mock jurors, who gave harsher sentences to female versus male defendants in an ambiguous assault case. All mock jurors gave female defendants harsher sentences when the victim was male. Additionally, Rodriguez et al. (2006) found that female offenders benefitted from more lenient sentencing in property and drug offences, but not violent offences compared to male offenders. However, other studies have found no such gender-based sentencing leniency effects (Maeder et al., 2018). With 200 jury-eligible U.S. citizens, Maeder et al. (2018) compared mock-juror impressions of defendants in a case of shoplifting or grand-theft auto. All defendants were more likely to be perceived as guilty in the case of grand theft auto, regardless of gender and there were no gender differences in guilt ratings of shoplifting defendants.

These inconsistent findings suggest it is worthwhile to continue investigating the effects of crime type and gender on mock-juror perceptions of guilt.

Crime Type and Attractiveness. Another interaction that has been found in the literature is that between attractiveness and crime type. Esses and Webster (1988) examined the role of attractiveness in participant perceptions of convicted offenders' dangerousness. Participants were given a criminal offence history that included either sexual or nonsexual offences and a photograph of an attractive, average, or unattractive person, and they were asked to answer questions about the dangerousness of the person. Despite all the offence histories being perceived as equally severe, participants perceived the unattractive photograph with sexual offences as being the most dangerous and the most unlikely to restrain future violent behaviour. For nonsexual offences, this effect was not seen. Taken with the characteristics attributed to less attractive people, the implicit bias that an unattractive person will not or cannot restrain future violent behaviour makes sense.

The beauty penalty has also been found in studies of attractiveness and crime type. Yang et al. (2019) investigated the effect of attractiveness and trustworthiness of male defendants in

two different swindles with a sample of university students. In the first swindle, participants rated guiltiness of a male defendant in a blind-date swindle. Female participants convicted the attractive, untrustworthy-looking defendants at a rate significantly higher than the male participants did. Moreover, they imposed significantly harsher punishments to the attractive defendants. In the second swindle, participants rated the guiltiness of a male defendant in an online scam swindle. In this case, there were no significant differences in conviction rates between more attractive and less attractive defendants or more trustworthy and less trustworthy-looking defendants for male or female participants. These findings support the notion that when good looks are used to manipulate another, the transgression is seen as more severe than if the offence was not associated with attractiveness and the associated attributes.

Indeed, in their meta-analysis, Mazella and Feingold (1994) found that participants rated attractive defendants as less guilty than unattractive defendants. Further, attractive defendants were given significantly less punishment for crimes of robbery, rape, and cheating, and significantly more punishment for the crime of negligent homicide when compared to unattractive defendants. These disparate ratings of guilt and punishment in consideration of crime type suggest it is worthwhile to continue examining these factors.

Crime Type, Gender, and Attractiveness. As described above, attractiveness-based biases have been found to influence mock jurors across crime type or gender. Some research has found that participant gender significantly interacts with characteristics of the defendant to affect ratings of guilt and punishment severity. For example, with a sample of 320 undergraduate students, Wuensch et al. (1991) found that female participants were likely to give significantly longer sentences to unattractive male defendants who committed a swindle compared to attractive male defendants who committed a swindle (Wuensch et al., 1991). For the crime of burglary, female participants gave consistent ratings and punishments regardless of

attractiveness. These findings support the attractiveness-leniency effect for one type of crime, but not the other. Given the inconsistent findings, it seems worthwhile to continue investigating how these factors may interact to affect mock-juror decision-making.

Present Project

Previous research indicates mock juror decisions are influenced by characteristics of the defendants, victims, and themselves. Despite efforts of the youth justice system in Canada to address discriminatory practices leading to youth incarceration, more research is needed to determine which characteristics of defendants and jurors are associated with attitudes and beliefs that may contribute to higher ratings of guilt and blameworthiness and how these characteristics interact.

CHAPTER 2 : STUDY ONE

The purpose of the first study was to replicate and extend previous research showing that attractiveness is associated with less guilt and blameworthiness, female defendants are perceived as less guilty than male defendants, unattractive defendants of a violent crime are perceived as more dangerous than attractive defendants, and attractive defendants who have committed fraud are perceived as more guilty than unattractive defendants who have committed fraud. A 2 (attractiveness) x 2 (gender) x 2 (crime type) factorial design was employed for this purpose. This study extended the current literature by considering the interaction effects of attractiveness, gender, and crime type. We included measures of socially desirable responding, legal authoritarianism, and participants' past experiences with the criminal justice system as covariates to determine how much variability in the model was due to social desirability response bias, authoritarian attitudes, and domain-specific response bias, respectively.

Hypotheses

Based on the reviewed literature, five hypotheses were made. The first three considered the main effects of attractiveness (high or low), gender (male or female), and type of crime (assault or fraud). The fourth considered the interaction between attractiveness and type of crime. The fifth considered the interaction between attractiveness, gender, and type of crime. See Table 2.1 for a summary of the first study's hypotheses, the studies they are replicating and extending, and the data-analytic strategies used to test them.

1. It is hypothesized that attractive defendants will be perceived as not guilty more often and less guilty than unattractive defendants.
2. It is hypothesized that female defendants will be perceived as not guilty more often and less guilty than male defendants.
3. It is hypothesized that defendants who have committed fraud will be perceived as not guilty more often and less guilty than defendants who have committed an assault.
4. It is hypothesized that attractive defendants who have committed fraud will be perceived as guilty more often and more guilty than unattractive defendants who have committed fraud and attractive defendants who have committed an assault.
5. It is hypothesized that attractive female defendants will be perceived as not guilty more often and less guilty of either crime than male unattractive defendants.

Table 2.1

Hypotheses, Justification, and Corresponding Data-Analytic Plan

Hypothesis	Empirical Basis	Data-Analytic Strategy
1. Attractive defendants will be perceived as not guilty more often and less guilty than unattractive defendants.	Several studies have shown an attractiveness-based leniency effect, where more attractive defendants are perceived more positively, less guilty, and less deserving of punishment (Abwender & Hough,	Chi-square analysis and ANOVA will be used to assess the dichotomous and Likert-type rating for guilt for the attractive defendants and unattractive defendants. An ANCOVA will be used to adjust for the covariates.

	2001; Esses & Webster, 1988; Putz et al., 2016; Rice et al., 2020)	
2. Female defendants will be perceived as not guilty more often and less guilty than male defendants.	Several studies have shown a gender-based leniency effect, where female defendants are rated as guilty less often and given lighter sentences compared to their male counterparts (Meaux et al., 2018; Rodriguez et al., 2006	Chi-square analysis and ANOVA will be used to assess the dichotomous and Likert-type rating for guilt for the male and female defendants. An ANCOVA will be used to adjust for the covariates.
3. Defendants who have committed fraud will be perceived as not guilty more often and less guilty than defendants who have committed an assault.	Some research shows that crimes against persons is associated with more severe penalties compared to crimes against property (Sanderson et al., 2000). An examination of court records in the U.S. showed that more severe crimes are associated with more consistent incarceration.	Chi-square analysis and ANOVA will be used to assess the dichotomous and Likert-type rating for guilt across crime type. An ANCOVA will be used to adjust for the covariates.
4. Attractive defendants who have committed fraud will be perceived as innocent less often and more guilty than unattractive defendants who have committed fraud and attractive defendants who have committed an assault.	Research has shown evidence of a beauty penalty, wherein more attractive people are punished more severely than unattractive people when attractive people fail to meet the expectations of participants (i.e., in a case of fraud, and in a case of sending lower than expected dollar amounts to their partners) (Wilson & Eckel, 2006; Yang et al., 2019)	A two-way ANOVA will be conducted to assess how attractiveness affects the perceptions of guilt in consideration of type of crime committed. An ANCOVA will be used to adjust for the covariates.
5. Attractive female defendants will be perceived not guilty more often and as less guilty of either crime than unattractive male defendants.	An examination of the National Longitudinal Study of Adolescent to Adult Health demonstrated that attractive females are less likely to be arrested and convicted compared to unattractive male and females and attractive males (Beaver et al., 2019)	A three-way ANOVA will be conducted to assess how attractiveness and age affect the perceptions of guilt in consideration of type of crime committed. An ANCOVA will be used to adjust for the covariates.

METHODS

Participants

An a priori power analysis was conducted using G*Power version 3.1.9.6 (Faul et al., 2007) to determine the minimum sample size required to test the hypotheses. In order to detect a minimum effect size of a partial $\eta = 0.05$ for a 2x2x2 factorial ANOVA, a sample size of 152

participants was needed for the current study. Two hundred and five undergraduate students from the University of Windsor participated in this study. Twenty-four cases provided responses to less than 80% of the items in the set of measures comprising the online survey; these were removed from the analysis. Also removed were seven duplicate cases. The remaining participants responded correctly to the validity check items embedded in the online set of measures, including the transcript-related questions. This left 173 self-identified female participants and one participant who identified as non-binary. The mean age of the sample was 21.47 years (*SD* = 5.46). The majority of participants self-identified as heterosexual (81%). The majority of participants self-identified as Caucasian (64.9%). Other ethnicities representing ≥ 5% were South Asian (6.3%), Middle Eastern (19.5%), and Latin American (5.2%).

Measures

Demographics

A demographics questionnaire was given to participants requesting information about their age, gender, ethnicity, and previous jury experience.

Short Transcript of a Court Case

A short transcript of a court case (amended, with permission, from Blais and Forth, 2014) was used. The summary described a case of assault or fraud perpetrated by an acquaintance, in which the defendant's attractiveness, gender, and crime type were manipulated. In the assault condition, the victim claimed they were pushed and then stabbed by the defendant during a card game. The defendant claimed they acted in self-defence after the victim choked them. In the fraud condition, the victim claimed the defendant cheated during a card game, defrauding them of $1500. The defendant denied cheating and claimed to run away with the money after the

victim threatened them. In both versions, an expert witness gave testimony identifying the defendant as being of sound mind. The vignettes ended with instructions from the judge on providing a verdict considering reasonable doubt. Participants were given one of eight vignettes describing the assault or fraud, and that varied in attractiveness (attractive or unattractive) and gender (male or female). Following the reading of the court summary, participants were given a short questionnaire to evaluate their understanding of the case (see Appendices A1 and B).

Verdict and Related Questionnaires

Questions regarding verdict, participants' perceptions of defendant guiltiness, credibility of the defendant and victim, treatment amenability of the defendant, and risk for future crime or violence were taken, with permission, from those used by Belton et al. (2022), which form a set adapted, with permission, from Blais and Forth (2014). Aside from the guilty verdict (i.e., yes or no) these questions are on a 7-point Likert-type scale (see Appendix C).

Attractive and Unattractive Images

The Chicago Face Database (Ma et al., 2015) is a repository of high-resolution images that have been subjectively norm-referenced for attractiveness and ethnicity and measured for objective physical measurement. The images have been rated for various qualities on a 7-point Likert-type scale, with the qualities of interest for this study being attractiveness and baby-facedness. Lower ratings coincide with less attractiveness and baby-facedness, and higher ratings coincide with more attractiveness and baby-facedness. The interrater reliabilities of the subjective norm-referencing have been found to be high ranging from .89 to .99 (Ma et al., 2015). Additionally, the inter-rater reliability of the objective physical measurements has been found to be high, $r \geq .74$. Faces norm-referenced to be attractive or unattractive defendants have been selected from this database and were included in the vignettes for participants to view (see

Appendix A1). The selected images were chosen based on baby-facedness and the lowest and highest attractiveness ratings for White and Black (for study 3) males and females aged from 16 to 22 years old. The baby-faced ratings ranged from 3 to 4.37, and attractiveness ratings ranged from 2.4 to 5.1.

Social Desirability Response Bias

The *Balanced Inventory of Desirable Responding Short Form* (BIDR-16; Hart et al., 2015) measures socially desirable responding with 16 questions measured on a 7-point Likert-type scale ranging from 1 (Strongly disagree) to 7 (Strongly agree). It contains two subscales measuring Self-Deceptive Enhancement (SDE) and Impression Management (IM), which can be combined for a total score as a general index of socially desirable responding. SDE measures the tendency to provide true but positively biased self-reports (e.g., "I never regret my decisions.") and IM measures the tendency to provide inflated self-descriptions to acquire a socially desirable image, feeding into a need for approval (e.g., "I don't gossip about other people's business.") (Hart et al., 2015). The BIDR-16 subscales have been shown to have adequate test-rest reliability (SDE, $r = .79$, $p < .001$; IM, $r = .74$, $p < .001$; Hart et al., 2015). Additionally, it's been found to have adequate construct validity on par with other measures of socially desirable responding (e.g., BIDR-40 subscales, $r = .84$ and $r = .87$; Marlowe-Crowne Social Desirability Short, $r = .53$; Hart et al., 2015). For the present study, internal consistency was found to be good, ($\alpha = .82$).

Authoritarian Attitudes

The *Revised Legal Attitudes Questionnaire – 23* (RLAQ23; Kravitz et al., 1993) measures juror bias towards legal authoritarianism with 23 questions measured on a six-point Likert-type scale ranging from 1 (Strongly Disagree) to 6 (Strongly Agree). Eight of these items

comprise the Authoritarian scale (e.g., “The law coddles criminals to the detriment of society.”), six comprise the Antiauthoritarian scale (e.g., “Unfair treatment of underprivileged groups and classes is the chief cause of crime.”), and nine items comprise the Equalitarian scale (e.g., “Citizens need to be protected against excess police power as well as against criminals.”). Higher scores are positively correlated with more authoritarian attitudes. It has been found to have adequate internal reliability of .83 (Kravitz et al., 1993) and .70 (Lundrigan et al., 2015), and adequate concurrent validity when compared with other measures of authoritarian attitudes (e.g., Balanced F-Scale, $r = .57$; Attitudes toward the death penalty, $r = .51$; conviction proneness and reasonable doubt subscales of the Pre-trial Jury Attitudes Questionnaire (PJAQ), $r = .53$ and $r = .37$, respectively; Kravitz et al., 1993; Lundrigan et al., 2015). For the present study, internal consistency was found to be very low, ($\alpha = .42$).

Experiences with the Criminal Justice System

The *Experience with Criminal Justice System Measure* (Henderson et al., 1997) contains four dichotomously coded items, all questions to which the respondent answers Yes or No. Two of these items comprise the rationally grouped Contact with the Criminal Justice System scale (e.g., “In the past two years, do you know anyone personally who was arrested for committing a crime?”). The other two items comprise the rationally grouped Hassled by the Police scale (e.g., In the past two years, have you ever been stopped or watched closely by a police officer, even when you had done nothing wrong?”). Internal consistency for these brief scales was reported by Henderson et al. and found to be adequate: $\alpha = .67$ for the Contact scale, and $\alpha = .78$ for the Hassled scale. The internal consistency was found to be low for the Hassled scale ($\alpha = .58$). Item two of the Contact scale had no variance; therefore, a Cronbach’s alpha was not produced. The internal consistency for the three-item scale was low ($\alpha = .58$).

Procedure

Ethical clearance was obtained from the University of Windsor's Research Ethics Board. Undergraduate students were given the option to sign up for any of these studies after viewing an advertisement on the participant pool website. It was expected that the studies should take approximately 30 minutes. Once they registered for the study, they were directed to the Qualtrics website where they provided their consent to participate in the research. Once consent was obtained, they were presented with the demographics questionnaire followed by one of the vignettes which were distributed randomly. Once they read through the vignette, they answered questions pertaining to their understanding of the court case to ensure validity of their responses. Then they were given the verdict questions and Likert scales pertaining to guiltiness, credibility, treatment amenability, and risk. After they completed those questions, they were given a measure of socially desirability responding, legal authoritarianism attitudes, and participants' experiences with the criminal justice system. Once the questionnaires were completed, participants were invited to enter their email address to receive their compensation in the form of course credit and then they were given a list of community resources in the event of any adverse reactions.

RESULTS

Statistical analyses were run with SPSS Version 28. Data were checked for missing data then assumptions for ANOVA, ANCOVA and Chi-Square analysis were run.

Missing Value Analysis

Little's MCAR test indicated that missing data were missing completely at random; $\chi^2(1310) = 1250.77$, $p = .877$.

Attractiveness Manipulation

Attractiveness was manipulated using images from the Chicago Face Database. To verify the effectiveness of the attractiveness manipulation, participants were asked to rate the attractiveness of the photo attached to the transcript they read. Although the unattractive images were given lower ratings than the attractive images, attractiveness ratings in the Chicago Face Database set of images had larger differences between unattractive and attractive pairings. Descriptive statistics can be seen in Table 2.2.

Table 2.2

Descriptive Statistics of Attractiveness Ratings of Defendant Photographs

	M	*SD*	Range	CFD *M*
Unattractive male	3.10	1.07	1-4	2.65
Attractive male	4.00	.79	2-5	4.12
Unattractive female	3.57	.88	2-5	2.68
Attractive female	4.14	.42	3-5	5.09

Note. CFD = Chicago Face Database

Verdict

Participants were randomly assigned one of eight vignettes in which attractiveness, gender, and crime type were varied. Descriptive statistics are given in Table 2.3.

Table 2.3

Vignette Classification and Sample Size

	n	% of sample per group
UMF	21	12.1
AMF	23	13.2
UFF	26	14.9
AFF	19	10.9
UMA	22	12.6
AMA	20	11.5
UFA	25	14.4
AFA	18	10.3

Note. UMF = unattractive male fraud, AMF = attractive male fraud, UFF = unattractive female fraud, AFF = attractive female fraud, UMA = unattractive male assault, AMA = attractive male assault, UFA = unattractive female assault, AFA = attractive female assault

Overall, 56.9% of participants gave a guilty verdict. A series of chi-square tests of independence were conducted to ascertain if attractiveness, gender, and crime type were associated with differences in guilty or not guilty ratings (see Table 2.4). Key assumptions were met for the individual chi-square tests comparing attractiveness, gender, and crime type. There was not a statistically significant association between attractiveness and guilty/not guilty verdicts, $\chi^2(1) = .21$, $p = .649$. This is inconsistent with hypothesis one, that attractive defendants would be perceived as not guilty more often than unattractive defendants. There was not a statistically significant association between gender and guilty/not guilty verdicts, $\chi^2(1) = .88$, $p = .347$, which is inconsistent with hypothesis two, that female defendants would be perceived as not guilty more often than male defendants. There was not a statistically significant association between crime type and guilty/not guilty verdicts $\chi^2(1) = 1.06$, $p = .303$. This was inconsistent with hypothesis three, that defendants accused of fraud would be considered not guilty more often than defendants accused of assault.

Table 2.4

Frequencies and Chi-Square Results for Guilty and Not Guilty Verdict Associations with Attractiveness, Gender, and Crime Type (N = 174)

Transcript Condition	Yes		No		$\chi^2(1)$
	n	%	n	%	
Attractiveness	99	56.9	75	43.1	.21
Attractive	47	58.8	33	41.3	
Unattractive	52	55.3	42	44.7	
Gender	99	56.9	75	43.1	.88
Female	47	53.4	41	46.6	
Male	52	60.5	34	37.1	

Crime Type	99	56.9	75	43.1	1.06
Fraud	54	60.7	35	39.3	
Assault	45	52.9	40	47.1	

Participant Perceptions of The Defendant

A series of ANOVAs and ANCOVAs were conducted to determine if levels of guiltiness varied as a function of attractiveness, type of crime, and gender. Exploratory ANCOVA analyses of participant confidence in ratings, and defendant and victim credibility can be seen in Appendix D. Given the absence of hypotheses and in the interests of parsimony, exploratory analyses were restricted to ANCOVAs and pertinent post hoc testing. All of the ANOVAs and ANCOVAs evaluated mean differences of scores based on: attractiveness, gender, crime type, attractiveness and crime type, and attractiveness and gender.

Outliers were found for all of the variables, as assessed by inspection of boxplots. After winsorizing outliers by assigning to them the next highest value plus one, some outliers remained which were kept in the analyses. Shapiro-Wilk tests indicated that the data were not normally distributed for several groups. However, visual inspection of histograms revealed approximately normal distributions, and the skewness and kurtosis showed most z-scores within ± 1.96, with a few within ± 3.29, suggesting the data were acceptable (Kim, 2013). Additionally, ANOVA is considered robust to violations of normality. There was homogeneity of variances for most of the analyses; Welch ANOVA was used when homogeneity of variances was violated (as indicated in the text).

Level of Guiltiness

Hypothesis 1: Attractive Defendants Will Be Perceived As Less Guilty Than Unattractive Defendants. To evaluate mean differences in participant ratings of guiltiness based

on attractiveness, a one-way ANOVA was conducted. The assumption of normality was not met, as indicated by a significant Shapiro-Wilk test of normality ($p < .001$). However, visual inspection of histograms and normal Q-Q plots showed data that were approximately normally distributed. Homogeneity of variances was met. Inconsistent with hypothesis one, there was not a statistically significant mean difference in ratings of guiltiness by attractiveness (see Table 2.5).

Table 2.5

Means, Standard Deviations, and One-Way Analyses of Variance in Ratings of Guiltiness based on Attractiveness

	Attractive (n=80)		Unattractive (n = 94)		$F(1, 172)$	η^2
	M	SD	M	SD		
Level of Guiltiness	4.18	1.58	4.21	1.36	.03	.000

An ANCOVA was attempted to determine the effect of attractiveness on ratings of guiltiness after controlling for socially desirable responding (BIDR-16) and authoritarian legal attitudes (RLAQ). There was a nonlinear relationship between the BIDR-16 and RLAQ and levels of guiltiness, as assessed by visual inspection of scatterplots. Attempts to transform the covariates did not improve linearity, therefore, these results should be interpreted with caution. There was homogeneity of regression slopes. Standardized residuals for the ratings of guiltiness and the overall model were violated, as assessed by Shapiro-Wilk's tests; however, visual inspection of histograms showed approximately normal distributions. There was homoscedasticity. The assumption of homogeneity of variances was violated, as assessed by a significant Levene's test. A Log10 transformation was applied to attempt to resolve the homogeneity of variance violation. Although, Levene's test improved, it was still significant (p = .040). Therefore, Huber-White's robust standard errors were used. There were no outliers, as assessed by no cases with standardized residuals ± 3 standard deviations. After controlling for socially desirable responding and authoritarian legal attitudes, there was not a statistically

significant difference in ratings of guiltiness based on attractiveness, $t(158) = .45$, $p = .655$, $\eta^2 = .00$ (see Table 2.6).

Table 2.6

Analysis of Covariance, Means, Adjusted Means, Variances, and Robust Standard Errors of Ratings of Guiltiness Based on Attractiveness After Controlling for The BIDR-16 and RLAQ

	M	SD	M_{adj}	SE	$F(1, 158)$	Partial η^2	Robust *SE*
Attractiveness					.22	.00	.03
Unattractive	.59	.17	.59	.02			
Attractive	.58	.19	.58	.02			

Hypothesis 2: Female Defendants Will Be Perceived As Less Guilty than Male Defendants. To evaluate mean differences in participant ratings of guiltiness based on gender, a one-way ANOVA was conducted. Key assumptions were met except for the assumption of normality; however, visual inspection of histograms and normal Q-Q plots showed data that were approximately normally distributed. Inconsistent with hypothesis two, there was not a statistically significant difference in ratings of guiltiness by gender (see Table 2.7).

Table 2.7

Means, Standard Deviations, and One-Way Analysis of Variance in Ratings of Guiltiness based on Gender

	Male (n=86)		Female (n= 88)		$F(1, 172)$	η^2
	M	SD	M	SD		
Level of Guiltiness	4.28	1.48	4.11	10.45	.56	.003

An ANCOVA was performed to determine the effect of gender on ratings of guiltiness after controlling for socially desirable responding (BIDR-16) and authoritarian legal attitudes (RLAQ). The assumption of normality was violated for the male and female conditions, as well as the overall model; visual inspection of histograms showed approximately normal distributions. Linearity was violated for the BIDR-16 or the RLAQ and ratings of guiltiness and a Log10 transformation was applied, after which linearity was established for the RLAQ. The BIDR-16

was left out of the analysis. Other key assumptions were met. After controlling for authoritarian legal attitudes, there was not a significant difference in ratings of guiltiness based on gender, $F(1, 161) = 1.04$, $p = .310$, $\eta^2 = .01$ (see Table 2.8).

Table 2.8

Means, Standard Deviations, Adjusted Means, and Standard Errors of Ratings of Guiltiness Based on Gender After Controlling for Authoritarian Legal Attitudes

	Male	Female
M	.60	.57
SD	.18	.18
M_{adj}	.60	.57
SE	.02	.02

Hypothesis 3: Defendants Who Have Committed Fraud Will Be Perceived As Less Guilty Than Defendants Who Have Committed An Assault. To evaluate mean differences in participant ratings of guiltiness based on crime type, a one-way ANOVA was conducted. The assumption of normality was not met, as indicated by a significant Shapiro-Wilk test of normality ($p < .001$). However, visual inspection of histograms and normal Q-Q plots showed data that were approximately normally distributed. Inconsistent with hypothesis three there was not a statistically significant mean difference in ratings of guiltiness by gender (see Table 2.9).

Table 2.9

Means, Standard Deviations, and One-Way Analyses of Variance in Ratings of Guiltiness based on Crime Type

	Fraud (n=89)		Assault (n= 85)		$F(1, 172)$	η^2
	M	*SD*	*M*	*SD*		
Level of Guiltiness	4.38	1.33	4.00	1.57	3.01	.02

An ANCOVA was performed to determine the effect of crime type on ratings of guiltiness after controlling for socially desirable responding (BIDR-16) and authoritarian legal attitudes (RLAQ). The assumption of normality was violated; however, visual inspection of

histograms showed approximately normal distributions. The assumption of linearity was violated for the BIDR-16; therefore, it was dropped from the analysis. Homogeneity of variances was violated, as assessed by a significant Levene's test (p = .044). After applying a Log10 transformation, the homogeneity of variances did not improve. Therefore, Huber-White's robust standard errors were used. There were no outliers. After controlling for authoritarian legal attitudes, ratings of guiltiness did not differ significantly between crime types, $t(163) = 1.75$, p = .082, η^2 = .02 (see Table 2.10).

Table 2.10

Analysis of Covariance, Means, Adjusted Means, Variances, and Robust Standard Errors of Ratings of Guiltiness Based on Crime Type After Controlling for the RLAQ

	M	*SD*	M_{adj}	*SE*	*F*(1, 158)	Partial η^2	Robust *SE*
Crime Type					3.18	.02	.23
Fraud	4.35	1.30	4.35		.15		
Assault	3.95	1.55	3.95		.16		

Hypothesis 4: Attractive Defendants Who Have Committed Fraud Will Be Perceived As More Guilty Than Unattractive Defendants Who Have Committed Fraud and Attractive Defendants Who Have Committed An Assault. To evaluate mean differences in participant ratings of guiltiness based on crime type and attractiveness, a two-way ANOVA was conducted. Shapiro-Wilk tests of normality indicated the data were not normally distributed; however, z-scores of skewness and kurtosis were within ± 1.96, suggesting the data were approximately normally distributed. The homogeneity of variances was violated, as assessed by Levene's test for equality of variances. To address this violation, a robust ANOVA using Huber-White standard errors was performed. There was a significant interaction between attractiveness and crime type, $t(173) = 2.05$, p = .042. An analysis of simple main effects for attractiveness was performed with a Bonferroni adjustment. Pairwise comparisons showed there was a

statistically significant difference in level of guiltiness ratings between assault and fraud for unattractive defendants (Mean difference = .81, 95% *CI* =.22, .14, $p = .007$), who were considered significantly more guilty when accused of fraud compared to assault, $F(1, 170) = 7.38$, $p = .007$, $\eta^2 = .04$ (see Table 2.11 and Figure 2.1). However, hypothesis four was not supported, as attractive defendants accused of fraud were not considered significantly more guilty than unattractive defendants accused of fraud or attractive defendants accused of assault.

Table 2.11

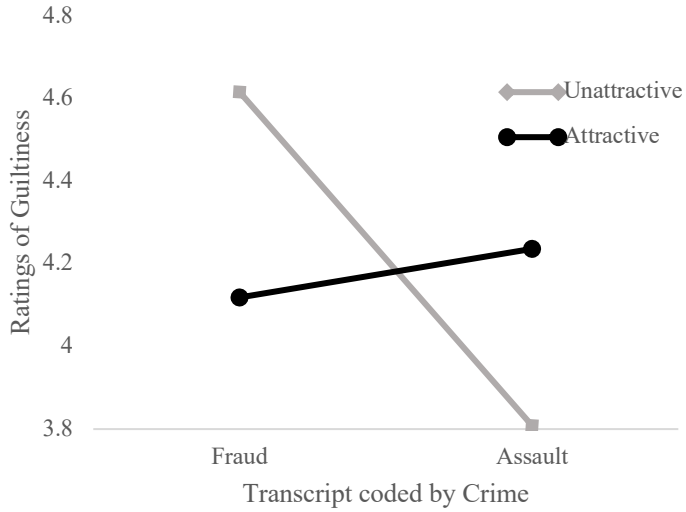
Sample Size, Means, Standard, Deviations, and Two-Way Analysis of Variance for Ratings of Guiltiness based on Attractiveness and Crime Type

Variable	Fraud			Assault			ANOVA		
	N	M	SD	N	M	SD	*F* ratio	*df*	η^2
Attractiveness									
Unattractive	47	4.62	1.03	47	3.81	1.53	7.38*	1, 170	.04
Attractive	42	4.12	1.57	38	4.24	1.62	.13	1, 170	.00

* $p < .05$

Figure 2.1

Changes in Ratings of Guiltiness as a Function of Attractiveness and Crime Type



A two-way ANCOVA was conducted to ascertain the effects of crime type and attractiveness on ratings of guiltiness, after controlling for the BIDR-16 and RLAQ. Key assumptions were met except for homogeneity of variances was violated, as determined by a significant Levene's test (p = .003), therefore Huber-White robust standard errors were used. Studentized residuals were not normally distributed for the attractive and unattractive fraud groups (p = .045, p = .038, respectively). However, visual inspection of histograms, and normal Q-Q plots revealed approximately normal distributions. After controlling for socially desirable responding and authoritarian legal attitudes, there was a statistically significant two-way interaction between attractiveness and crime type on ratings of guiltiness, $t(161) = 2.33$, p = .021, η^2 = .03 (see Table 2.12). An analysis of simple main effects for attractiveness and crime type was performed with a Bonferroni adjustment. The effect of crime type in the unattractive condition was statistically significant, $F(1, 156) = 8.874$, p = .003, η^2 = .05. Pairwise comparisons showed significantly higher mean scores for those accused of fraud compared to those accused of assault (Mean difference = .90, 95% CI = .30, 1.50, p = .003) (see Table 2.13 and Figure 2.2).

Table 2.12

Analysis of Covariance Results of Guiltiness Ratings Based on Attractiveness and Crime Type when Controlling for Socially Desirable Responding and Authoritarian Legal Attitudes

	$F(1, 156)$	Partial η^2	Robust *SE*
Attractiveness	.01	.00	.37
Crime Type	2.46	.02	.39
Attractiveness * Crime Type	5.98*	.04	.47*

*$p < .05$

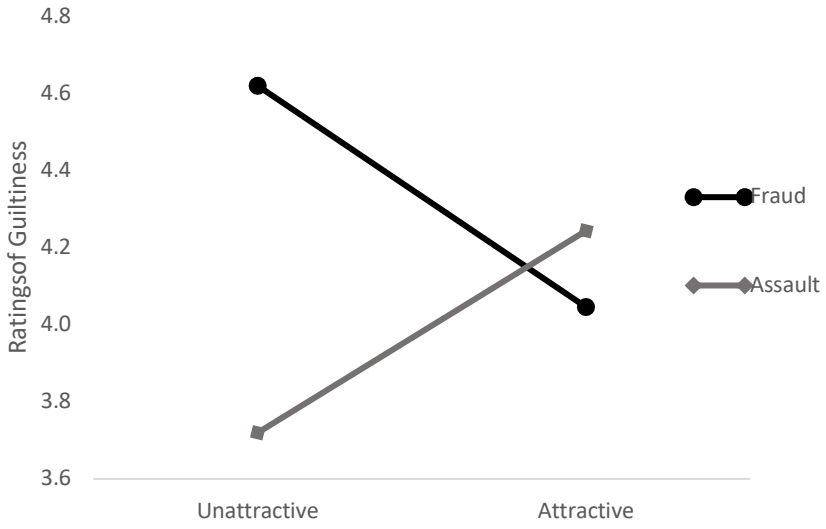
Table 2.13

Means, Adjusted Means, Standard Deviations, and Standard Errors for Ratings of Guiltiness for the Attractiveness and Crime Type Groups

	Unattractive		Attractive	
Crime Type	Fraud	Assault	Fraud	Assault
M	4.62	3.72	4.05	4.24
SD	1.01	1.45	4.54	1.67
M_{adj}	4.62	3.72	4.05	4.24
SE	.21	.22	.22	.24

Figure 2.2

Changes in Ratings of Guiltiness as a Function of Attractiveness and Crime Type After Controlling for Socially Desirable Responding and Authoritarian Legal Attitudes



Hypothesis 5: Attractive Female Defendants Will Be Perceived As Less Guilty Of Either Crime Than Male Unattractive Defendants. To evaluate mean differences in ratings of guiltiness based on attractiveness, gender, and crime type, a three-way ANOVA was conducted. Key assumptions were met, except for normality for three groups (unattractive fraud male, p = .015; unattractive fraud female, p = .006; attractive fraud male, p = .018), as assessed by Shapiro-Wilk's test of normality. However, visual inspection of normal Q-Q plots revealed approximately normal distributions. There was homogeneity of variances. Inconsistent with hypothesis five, there was no statistically significant three-way interaction between attractiveness, gender, and crime type, $F(1, 166) = .72$, p = .398 (see Tables 2.14 and 2.15).

Table 2.14

Sample Size, and Standard Deviations for Ratings of Guiltiness by Attractiveness, Gender, and Crime Type

Variable	Male			Female		
	N	M	SD	N	M	SD
Unattractive						
Fraud	21	4.62	1.07	26	4.62	1.02
Assault	22	3.91	1.44	25	3.72	1.62
Attractive						
Fraud	23	4.04	1.67	19	4.21	1.48
Assault	20	4.60	1.60	18	3.83	1.58

Table 2.15

Three-Way Analysis of Variance for Ratings of Guiltiness

	$F(1, 166)$	p	η^2
Attractiveness	.04	.842	.00
Crime Type	3.60	.109	.02
Gender	.80	.371	.01
Attractiveness *Crime Type	4.08	.045	.02
Attractiveness*Gender	.21	.646	.00
Crime Type*Gender	1.60	.207	.01
Attractiveness*Gender*Crime Type	.72	.398	.00

A three-way ANCOVA was conducted to examine the effects of attractiveness, gender, and crime type on ratings of guiltiness, after controlling for the BIDR-16 and RLAQ. Linearity was not established for the covariates at each level of the independent variables; therefore, a reflect and log10 transformation was applied. All other key assumptions were met except for normality. Visual inspection of histograms and normal Q-Q plots showed approximately normal distributions. Means, adjusted means, standard deviations and standard errors are presented in Table 2.16. After controlling for the BIDR-16 and RLAQ, there was not a statistically significant three-way interaction between attractiveness, gender, and crime type on ratings of guiltiness, $F(1, 152) = .75$, $p = .386$, $\eta^2 = .01$ (see Table 2.17). There were no significant main effects of after controlling for the BIDR-16 and RLAQ; however, there was a significant two-way interaction between attractiveness and crime type for unattractive defendants, $F(1, 152) = 5.72$ $p = .018$, $\eta^2 = .04$. Pairwise comparisons with a Bonferroni adjustment for multiple comparisons showed that adjusted mean scores were significantly higher for unattractive defendants accused of fraud than unattractive defendants accused of assault (Mean difference = .91, 95% *CI* =.30, 1.51, $p = .003$). This was true for both male defendants ($F(1, 152) = 4.10$, $p = .045$, $\eta^2 = .03$;

Mean difference = .91, 95% *CI* =.02, 1.80, p = .045) and female defendants (F(1,152) = 4.77, p = .030, η^2 = .03; Mean difference = .90, 95% *CI* =.09, 1.71, p = .030).

Table 2.16

Means, Adjusted Means, Standard Deviations, and Standard Errors for Ratings of Guiltiness by Attractiveness, Gender, and Crime Type while Controlling for Socially Desirable Responding and Authoritarian Legal Attitudes

		M	SD	M_{adj}	SE
Unattractive					
Fraud	Male	4.70	1.03	4.69	.32
	Female	4.56	1.00	4.56	.28
Assault	Male	3.80	1.44	3.78	.32
	Female	3.65	1.50	3.66	.30
Attractive					
Fraud	Male	4.05	1.70	4.10	.31
	Female	4.06	1.35	3.99	.34
Assault	Male	4.67	1.68	4.66	.33
	Female	3.75	1.57	3.77	.36

Table 2.17

Analysis of Covariance Results of Attractiveness, Crime Type, and Gender when Controlling for Socially Desirable Responding and Authoritarian Legal Attitudes

	F(1, 152)	Partial η^2
Attractiveness	.04	.00
Crime Type	2.60	.02
Gender	1.91	.01
Attractiveness * Crime Type	5.72*	.04
Attractiveness * Gender	.69	.00
Crime Type * Gender	.72	.01
Attractiveness * Crime Type * Gender	.75	.01

*p < .05

DISCUSSION

This study examined whether attractiveness and gender had an effect on participant perceptions of guilt across two crime types: fraud and assault. The findings were broadly inconsistent with the hypotheses. However, given the mixed findings in the literature, they

contribute to the body of mock jury research. The results of this study did not change when the potential effects of measures of socially desirable responding (SDR) and authoritarian legal attitudes were statistically controlled. The role of SDR in mock jury research is unclear. Because mock juror research such as this is aiming to identify the implicit biases of its participants, it is valuable to account for how much of their behaviour can be attributed to SDR; however, it appears that the variables of interest in this study were not affected to a statistically significant degree by SDR.

Previous research has shown that authoritarian legal attitudes influence verdicts. For example, McGowen and King (1982) found that people with authoritarian attitudes are more punitive to defendants with whom they are similar, compared to dissimilar. In a 1993 meta-analysis, Narby et al. (1993) showed that high legal authoritarianism has been associated with more guilty verdicts than low legal authoritarianism, specifically for felonies, murder, and rape. Jones et al. (2015) also found higher ratings of guilt by people with authoritarian attitudes when sentencing severity was high. However, high authoritarian legal attitudes were associated with significantly less conviction-proneness for the least severe sentences and not associated with moderately severe sentences. Therefore, although legal attitudes may affect decision-making, specific moderating variables are relevant.

Guilt and Attractiveness

The analyses of categorical guilty/not guilty verdicts and dimensional ratings of guiltiness based on attractiveness did not show any statistically significant differences. Although attractiveness has been associated with leniency towards more attractive people when they behave in ways inconsistent with expectations, there can be an adverse effect that may result in more harsh punishment than if they had been unattractive (Wilson & Eckel, 2006) or average

looking (Voit et al., 2023). The attractiveness manipulation in this study indicated that participants believed all of the defendant images were broadly average in appearance. In light of this, although the attractive defendants were given guilty verdicts less often, it makes sense that these differences were not statistically significant. Our first hypothesis was based on previous work that identified an attractiveness-based leniency effect (Abwender & Hough, 2001; Esses & Webster, 1988; Putz et al., 2016; Rice et al., 2020). Perhaps this effect would have reached significance if the attractiveness manipulation had been more successful. Additionally, some research suggests that sentencing severity is associated with attractiveness, whereas, guilty/not guilty verdicts are not (Mazzella & Feingold, 1990; DeSantis & Wesley, 1997).

Attractiveness and Gender

Gender also did not have a significant effect on ratings of guiltiness. Although female defendants were considered less guilty than male defendants, the difference was not statistically significant. Our findings must be taken with the understanding that we did not have self-identified male participants in this study (this was not the result of a deliberate recruitment strategy but rather an unexpected outcome of the use of the Psychology Participant Pool, which usually facilitates recruitment of self-identified males albeit at lower numbers than self-identified females. Our hypotheses were based on previous research which showed differential treatment of male and female defendants (Meaux et al., 2018; Rodriguez et al., 2006). These differences have also been found between male and female participants (Pozzulo et al., 2017). For example, in a case of assault, female mock jurors were shown to consider the defendant significantly more guilty than did the male mock jurors (Pozzulo et al., 2017). In a case of intimate partner violence, male mock jurors were more likely to consider a female defendant to be guilty than were female mock jurors, who were shown to consider male defendants to be significantly more guilty than

female defendants (Pica et al., 2019). Because of this major limitation, the findings must be taken specifically to represent female undergraduate psychology students at the University of Windsor.

Attractiveness and Crime Type

Contrary to our expectations, based on previous studies and statistics (Sanderson et al., 2000), there was not a statistically significant difference in ratings of guiltiness between fraud and assault; however, there was a small effect, which was higher for those accused of fraud than those accused of assault. This may be because participants may be less likely to render guilty verdicts for more serious crimes for fear of making an error in situations when sentencing may be more severe (Freedman et al., 1994). Additionally, Maedar et al. (2018) found that although crime type did not influence perceptions of guilt, it did influence perceptions of the defendant, which was consistent with our findings – crime type did affect defendant and victim credibility ratings (see Appendix D, Tables D.6 and D.10). Because participants considered defendants accused of fraud to be less credible and those accused of assault to be more credible, the nonsignificant differences in guilt ratings make sense – if someone is perceived to be credible, their testimony may be more believable and they will be perceived as less guilty and the inverse may also be true (Pozzulo et al., 2015).

Attractiveness, Gender, and Crime Type

Attractiveness, gender, and crime type did not have effects on ratings of guiltiness in isolation; however, attractiveness and crime type did have a significant effect on ratings of guiltiness. Unattractive defendants accused of assault were considered significantly less guilty than unattractive defendants accused of fraud. Although the difference was not statistically significant, attractive defendants accused of fraud were considered less guilty than those accused

of assault. This was contrary to our expectations based on Yang et al. (2019), who found that attractiveness was associated with lenience for those accused of an online scam compared to a blind-date swindle. This may be because ratings of guiltiness in the fraud condition may not have been influenced by the defendant's appearance due to the average attractiveness ratings given to defendants across high and low attractiveness. It may also be because the defendant would not have had to use their appearance to commit the type of fraud in the vignette.

When comparing unattractive and attractive defendants accused of fraud, those who were unattractive were considered guiltier. The opposite was true for the crime of assault – those who were attractive were considered guiltier. These findings make sense in consideration of Mazella and Feingold (1994). Even if participants believed the attractive defendant accused of assault behaved in self-defence, the higher standard to which they are held may have led to harsher treatment when they failed to meet it through poorly controlled behaviour. These findings demonstrate the beauty bias in the fraud condition and the beauty penalty in the assault condition. Attractive defendants accused of violence may have been considered in breach of the implicit social contract that their fraud-accused counterparts were not. Consistent with Wuensch et al. (1991), female participants in their study judged unattractive defendants accused of a swindle statistically significantly more harshly than attractive defendants accused of a swindle. Additionally, rather than a swindle, participants in this study may have considered the fraud condition to be a case of robbery and/or cheating, which Mazzella and Feingold (1994) demonstrated were associated with lenience for attractive defendants.

There are several possible reasons why the results contradicted our expectations. Regarding attractiveness, although the images selected from the Chicago Face Database were the most attractive and unattractive within the age range of interest for this study, participants did not consider them to be "extremely attractive" or "extremely unattractive," therefore, it is possible

they did not evoke the intended response. Despite the approximately average ratings, the unattractive defendants had lower mean scores of attractiveness, which suggests that was enough to increase ratings of guilt when comparing crime types. Another possible explanation is that behaviour may affect the perception of attractiveness. Some research suggests that aggressive behaviour decreases ratings of facial attractiveness of male targets (Niimi & Goto, 2023). Perhaps if participants rated the defendant images for attractiveness prior to reading the court transcript, their ratings would have been more reflective of those in the CFD.

Additionally, perhaps the societal standards of beauty are shifting. Body positivity influencers are gaining popularity on social media. Advertisements for popular brands use people with different body types, skin conditions, and ages to sell their products. These advertisements spark conversations about the definition of beauty (Feng et al., 2019), which brings beauty as a concept to awareness, which may change beliefs about beauty over time – just as the hour-glass figure has become a desirable body shape, in contrast to the 'heroin chic' of the early 1990s (Hunter et al., 2021).

Regarding gender, some research has shown a leniency effect for women (Mazzella & Feingold, 1994; Meaux et al., 2018; Rodriguez et al., 2016) and others have shown gender to be inconsequential (Devine & Caughlin, 2014). Our null gender-based findings may have resulted from no self-identified men in our sample. Previous research has shown that men and women have been more or less punitive with female defendants based on other attributes, such as attractiveness (Abwender & Hough, 2001), crime type (Meaux et al., 2018;) and a combination of both (Yang et al., 2019).

Therefore, these findings contribute to this greater body of research.

Exploratory Analyses

Defendant credibility was significantly affected by attractiveness and crime type (see Appendix D, Table D.7). Attractive defendants were considered significantly less credible than unattractive defendants, and this held true after controlling for socially desirable responding and authoritarian legal attitudes (see Table D.5). Consistent with previous studies, the beauty penalty may be found when the attributes ascribed to those with higher attractiveness are found to be violated (Wilson & Eckel, 2006). Those accused of fraud were considered significantly less credible than those accused of assault (see Table D.6). Because fraud is necessarily deceptive, it makes sense that those accused of fraud would be considered less credible than those accused of assault.

Crime type and attractiveness together also influenced ratings of defendant credibility. Attractive defendants accused of fraud were considered less credible than unattractive defendants accused of fraud and attractive defendants accused of assault. Attractive defendants accused of assault were considered less credible than their unattractive counterparts (see Table D.7). These findings exemplify the beauty penalty – although the hypothesis that attractive defendants accused of fraud would be considered guiltier was not supported, it appears that when attractive people commit crimes, the standards they have been held to means that a more severe infraction has been committed than if they had been unattractive.

Victim credibility was not affected by defendant attractiveness or gender in isolation; however, victims of fraud were considered significantly more credible than victims of assault (see Table D.10). When attractiveness was included with crime type, only victims of unattractive defendants accused of fraud were considered more credible than unattractive defendants accused of assault; this held true when controlling for socially desirable responding and authoritarian legal attitudes (see Table D.11 and Figures D.2 and D.3). One reason for these findings may be

the Belief in a Just World, which refers to a belief that good things happen to good people, and bad things happen to people who deserve it (Furnham, 2003). Reich et al. (2021) examined victim blaming following violent crimes and sexual assault. They found that when crimes increase in severity, although victim blaming may not increase, other negative social reactions may, such as treating the victim differently or discouraging them from talking (Reich et al., 2021). This may explain why the victims of aggravated assault were considered less credible than victims of fraud.

Implications

Research such as this contributes to the growing body of research on extralegal factors that affect decision-making. Importantly, investigating these factors with youth provides evidence for the types of biases that may affect youth who are involved in the criminal justice system.

Studies such as these have real-world implications for policymakers regarding factors that police, judges, and decision-makers should consider when in contact with youth who are involved with the criminal justice system. Although real juries are not involved in youth court in Canada, this research identifies the types of biases that all who come into contact with youth may be susceptible to. Training and education for the first points of contact up to and including judges may reduce disparate arrest and sentencing practices. However, more research is needed.

Limitations and Future Directions

This study had several limitations. First, the sample was all undergraduate women and one nonbinary person. Undergraduate men, nonbinary, and trans folks may have different implicit biases that were not considered in this study due to the absence of male and other gender participants. Additionally, the defendants in the vignettes were male or female and did not

include nonbinary or trans people, which is an important demographic for research, as these communities are underserved and face unique challenges in the criminal justice system. For example, Hebert et al. (2022) discussed problems such as discriminatory, stigmatizing, and victimizing encounters at each stage of the legal system experienced by trans, Two-Spirit, and non-binary people. Additionally, much of their sample indicated first contact with the legal system was in childhood. Future research should include the perspectives of nonbinary and trans participants, as well as perceptions of nonbinary and trans defendants and victims.

Another limitation of the study was the gender congruent defendant and victim pairings. Although this provides important information, gender incongruent pairings have been shown to affect perceptions of defendants (e.g., Pozzulo et al., 2009). Future research should include gender victim incongruent pairings.

CHAPTER 3 : STUDY TWO

Study two continues with the investigation into extralegal factors that have been found to influence mock juror decision making. Along with attractiveness and gender, the effect of psychiatric diagnostic labels was evaluated.

Psychiatric Diagnosis

Over 90% of justice-involved youth have been found to meet criteria for one psychiatric disorder (Drerup et al., 2008). Sixty percent of incarcerated girls and 32% of incarcerated boys have been found to meet criteria for three or more disorders (Drerup et al., 2008). Problems arise when decisions about blameworthiness, dangerousness, deserved punishment and treatment amenability are influenced by diagnostic labels without understanding their meaning.

Psychiatric Diagnosis and Labelling Effects. Labelling effects have been found to influence mock juror decision making across a variety of psychiatric diagnoses. Boccaccini and

colleagues (2008) considered these effects with a sample of 891 jury-pool members evaluating juvenile defendants with a vignette of expert testimony. Three variables were manipulated: history of antisocial behaviour (substantial vs minimal), psychopathic personality traits (present or absent), and a diagnostic label (is a psychopath, meets criteria for psychopathy, meets criteria for conduct disorder, or does not meet criteria for any diagnosis). Participants rated juvenile defendants who had a substantial history of antisocial behaviour as having the highest risk, especially when they were labelled as a "psychopath" or had no diagnosis. The rating was also significantly higher when the youth was labelled as a "psychopath" versus meets criteria for psychopathy. When there was a minimal history of antisocial behaviour, participants gave a higher risk rating for youth who were labelled as a "psychopath" or given a diagnosis of meets criteria for psychopathy versus no diagnosis or conduct disorder. However, when there was a minimal history of antisocial behaviour, the risk rating was significantly lower than when the history of antisocial behaviour was substantial. Participants gave harsher punishments to youth who had psychopathic traits whether they had a history of antisocial behaviour or not. They also gave harsher punishments to youth who were labelled a psychopath versus meets criteria for psychopathy. These findings indicate that laypersons who are not familiar with the criteria of psychiatric disorders may rely on assumptions of what they mean to make decisions. The diagnostic label of "psychopath" has the same meaning of "meets criteria for psychopathy". However, it is noteworthy that psychopathy is not a psychological diagnosis in the Diagnostic and Statistical Manual of Mental Disorders (DSM-5-TR; American Psychiatric Association, 2022), or in the International Classification of Diseases (ICD-11; World Health Organization, 2019). Therefore, the variance in risk between these labels suggests a lack of understanding by participants which can lead to harmful real-world outcomes for adjudicated individuals.

Taylor and colleagues (2019) investigated the effect of several diagnostic labels (i.e., conduct disorder (CD), major depressive disorder (MDD), attention deficit/hyperactivity disorder (ADHD), schizophrenia, and antisocial traits/behaviours) of juvenile defendants on judgments of blameworthiness, punishment severity, capacity for behaviour regulation and dangerousness by a sample of 252 jury-eligible US residents. They found youth with a schizophrenia diagnosis were considered the least blameworthy, given the least severe punishment, and considered least capable of behaviour regulation. Youth with antisocial traits/behaviours were considered the most blameworthy, given the harshest punishment, were considered most dangerous, and were considered less capable of behaviour regulation than those with CD, MDD, and ADHD.

Because many of the antisocial behavioural issues that occur in adolescence are outgrown or change over time, it can be especially harmful for life-altering decisions to be made based on temporary conditions. Indeed, Dyck and colleagues (2013) observed a decrease in offending over time by adjudicated youth with low, moderate, or high psychopathic traits. Another issue is the effect of detention on the likelihood of recidivism. The use of pretrial detention to ensure youth attend court hearings was used for 61% of custody admissions in Canada in 2021/2022 (Statistics Canada, 2023), and the choice to use pretrial detention is determined by several risk factors (i.e., a serious offence has been committed or the youth has outstanding criminal charges or convictions, to ensure the youth attends court, or releasing the youth would not sufficiently address the court's concerns; Department of Justice, n.d.). However, the use of pretrial detention has been associated with a 33% increase in felony recidivism and an 11% increase in misdemeanor recidivism when youth have committed no or few prior offences (Walker & Herting, 2020). These findings are especially troubling given that it is at the discretion of police officers to decide who should be released with a court date, or who should be arrested and detained for a bail hearing (John Howard Society of Ontario, 2021). As discussed above, police

discretion can be biased, and harmful decisions have been made based on these biases as evidenced by the inconsistent use of diversionary measures (Samuels-Wortley, 2019).

Psychiatric Diagnosis and Gender. Given the rates of incarcerated youth who meet criteria for one or more psychiatric diagnoses, and how they differ between males and females (Drerup et al., 2008), the interaction of psychiatric diagnosis and gender is another factor to consider when evaluating mock-juror decision-making. A recent investigation into implicit gender biases using a male versus female case vignette among 180 licensed psychologists revealed a significant tendency for psychologists to give the diagnosis of borderline personality disorder to females and an antisocial personality disorder diagnosis to males (Bruce & Weinraub, 2023). With all else being equal except for gender in these vignettes, it is concerning that an implicit gender bias exists in the people most qualified to give these diagnoses.

Other research has shown that although mock jurors are more likely to render guilty verdicts to female defendants with mental illness, they also believe female defendants have less ability to manage their illness (Breheney et al., 2007). These findings suggest incarcerated females with mental illness carry a double burden of being perceived as more guilty and unable to care for themselves which may explain the higher rates of incarceration for this subset of women. This may be in part due to the chivalry hypothesis, because incarceration provides these women with structure of some kind.

Another study investigating the effects of gender and mental illness on mock-juror decision-making did not reveal disparate guilt ratings based on gender (Mossière & Maeder, 2016). However, defendant attributions varied as a function of gender. Mock jurors perceived the female defendant's actions as being caused by her personality, rather than external factors. The researchers suggest this may be due to the fundamental attribution error (FAE) which is more likely to impact women when considering mental illness (Mossière & Maeder, 2016).

Psychiatric Diagnosis and Attractiveness. The interaction of psychiatric diagnosis and attractiveness has not been evaluated in juvenile defendants. However, across two studies Jones and colleagues (1978) examined how participants evaluated attractive and less attractive target photographs based on descriptions of them either having a serious psychological disturbance requiring inpatient psychiatric care or having a mild psychological crisis and receiving counseling. In the first study, 58 male and female undergraduates perceived the photographs of unattractive faces as having significantly more psychological disturbance compared to the attractive faces. In the second study, 40 undergraduates rated the photographs, but half of them were instructed to attend to cues other than attractiveness. Male participants in both conditions continued to evaluate less attractive photographs as significantly more psychologically disturbed, while female participants in both conditions did not consider unattractive photographs to be more psychologically disturbed. These results suggest that the more attractive a defendant is, the less their psychiatric diagnosis may impact guiltiness ratings.

Present Project

The purpose of the second study was to replicate previous research showing that diagnostic labelling is associated with differing ratings of guilt, treatability, and capacity for behaviour change. We also extended the current literature on attractiveness and gender with consideration of diagnostic labelling effects on mock-juror decision making. A 2 (attractiveness) x 2 (gender) x 4 (diagnostic label) factorial design was employed for this purpose. As in study one, we included measures of socially desirable responding, legal authoritarianism, and participants' past experiences with the criminal justice system as covariates. We also included a measure of prejudice towards those with mental illness as an additional covariate.

Hypotheses

Based on the reviewed literature, five hypotheses were made. The first two were consistent across all three studies and considered the main effects of attractiveness (high or low), and gender (male or female). The third considered the main effects of diagnostic label (schizophrenia, psychopathic traits, conduct disorder, no diagnosis). The fourth considered the interaction between attractiveness and diagnostic label. The fifth considered the interaction between attractiveness, gender, and diagnostic labelling effects. See Table 3.1 for a summary of the second study's hypotheses, the studies they replicated, and the data-analytic strategy used to test them.

1. It is hypothesized that attractive defendants will be perceived as less guilty than unattractive defendants.
2. It is hypothesized that female defendants will be perceived as less guilty than male defendants.
3. It is hypothesized that guilty ratings will vary based on psychiatric diagnosis where those diagnosed with schizophrenia will be considered less guilty than the rest, and those with psychopathic traits will be considered more guilty than the rest.
4. It is hypothesized that attractive defendants with a psychiatric diagnosis will be perceived as less guilty than unattractive defendants with a psychiatric disorder.
5. It is hypothesized that unattractive female defendants with a psychiatric diagnosis will be perceived as more guilty than all other defendant types.

Table 3.1

Hypotheses, Justification, and Corresponding Data-Analytic Plan

Hypothesis	Empirical Basis	Data-Analytic Strategy

1. Attractive defendants will be perceived as more innocent and less guilty than unattractive defendants.	Several studies have shown an attractiveness-based leniency effect, where more attractive defendants are perceived more positively, less guilty, and less deserving of punishment (Abwender & Hough, 2001; Esses & Webster, 1988; Putz et al., 2016; Rice et al., 2020)	Chi-square analysis and ANOVA will be used to assess the dichotomous and Likert-type rating for guilt for the attractive defendants and unattractive defendants. An ANCOVA will be used to adjust for the covariates
2. Female defendants will be perceived as more innocent and less guilty than male defendants.	Studies have shown a gender-based leniency effect, where female defendants are rated as guilty less often and given lighter sentences compared to their male counterparts (Meaux et al., 2018; Rodriguez et al., 2006	Chi-square analysis and ANOVA will be used to assess the dichotomous and Likert-type rating for guilt for the male and female defendants. An ANCOVA will be used to adjust for the covariates
3. Guilty ratings will vary based on psychiatric diagnosis where those diagnosed with schizophrenia will be considered not guilty more often and less guilty than the rest, and those with psychopathic traits will be considered innocent least often and more guilty than the rest	Several studies have shown the impact of psychiatric diagnosis labelling effects where those with psychopathic traits are perceived as the most guilty (Blais & Forth, 2014; Boccaccini et al., 2008; Taylor et al., 2019) and those with schizophrenia are perceived as the least guilty (Taylor et al., 2019).	Chi-square analysis and ANOVA will be used to assess the dichotomous and Likert-type rating for guilt across psychiatric diagnoses. An ANCOVA will be used to adjust for the covariates
4. Attractive defendants with a psychiatric diagnosis will be perceived as less guilty than unattractive defendants with a psychiatric disorder.	Research has shown evidence that unattractive individuals described as having a serious psychological disturbance or having a mild psychological crisis are perceived as significantly more psychologically disturbed than attractive individuals in either condition (Jones et al., 1978).	A two-way ANOVA will be conducted to assess how attractiveness affects the perceptions of guilt in consideration of psychiatric diagnosis. An ANCOVA will be used to adjust for the covariates.
5. Unattractive female defendants with a psychiatric diagnosis will be perceived as more guilty than all other defendant types.	As described above, there is an attractiveness-leniency effect, and a gender-based leniency effect that favours attractive women (see Hypotheses 1 and 2). Conversely, unattractive female defendants are likely to receive harsher penalties by female mock jurors (Abwender & Hough, 2001). Additionally, research has shown female defendants with mental illness are perceived as more guilty and less stable (Breheney et al., 2007).	A three-way ANOVA will be conducted to assess how attractiveness and gender affect the perceptions of guilt in consideration of psychiatric diagnosis. An ANCOVA will be used to adjust for the covariates.

METHODS

Participants

An a priori power analysis was conducted using G*Power version 3.1.9.6 (Faul et al., 2007) to determine the minimum sample size required to test the hypotheses. In order to detect a minimum effect size of a partial $\eta = 0.05$ for a 2x2x4 factorial ANOVA, a sample size of 179 participants was needed for the current study. Two hundred and forty-two undergraduates from the University of Windsor participated in this study. Eleven cases with less than 80% of the study completed were removed from the dataset and five duplicate cases were removed. Fourteen additional cases were removed due to failed validity checks. This left 209 female and 3 non-binary participants. The mean age of the sample was 21.48 years (*SD* = 3.97). The majority of participants self-identified as heterosexual (76.4%). The majority of the participants were self-identified Caucasian (63.2%). Other ethnicities representing ≥ 5% were South Asian (5.2%), Middle Eastern (17%), and Latin American (5.2%).

Measures

The measures across all studies were consistent, with two exceptions: the summary of a court case and a measure of mental illness prejudice, which are described below. The manipulated variables in the summary for study two are attractiveness, gender, and psychiatric diagnosis. Please see the descriptions of the measures used across all three studies above, for Study 1, and the adapted/additional measures below. The internal consistency for the BIDR-16 was adequate ($\alpha = .79$). The internal consistency of the RLAQ-23 was low ($\alpha = .63$). The internal consistency of the ECJS was low ($\alpha = .58$). For the Contact scale, internal consistency was unacceptable ($\alpha = .24$), and the Hassled scale was low ($\alpha = .61$).

Short Transcript of a Court Case

Participants were randomly assigned to read (displayed in the Qualtrics survey) one of sixteen short court transcripts developed for this study (amended, with permission, from Blais and Forth, 2014). The summary described an assault, in which the victim claimed they were pushed and then stabbed by the defendant. The defendant claimed they acted in self-defence after being choked by the victim. After testimonies of the crime were given, an expert witness provided their impressions of the defendant. Finally, the judge provided instructions to the jury to provide verdicts with consideration of reasonable doubt. Participants were randomly assigned to one of 16 vignettes that varied the defendant's attractiveness (attractive or unattractive), gender (male or female), and psychiatric diagnosis (conduct disorder, psychopathic traits, schizophrenia, or no diagnosis). Following the reading of the court summary, participants were given a short questionnaire to evaluate their understanding of the case (see Appendices A2 and B).

Mental Illness Prejudice

The *Prejudice Towards People with Mental Illness Scale* (PPMI; Kenny et al., 2018) measures mental illness prejudice with 28-items rated on a 9-point Likert-type scale, ranging from -4 (very strongly disagree) to +4 (very strongly agree). It contains 4 subscales measuring Fear/Avoidance (8 items; e.g., "I would find it hard to talk to someone who has a mental illness."), Malevolence (8 items; e.g., "People who are mentally ill are avoiding the difficulties of everyday life."), Authoritarianism (6 items; "People who are mentally ill need to be controlled by any means necessary."), and Unpredictability (6 items; "People with mental illness often do unexpected things."). The PPMI Scale has been found to have adequate reliability, $\alpha = .91$, as have its subscales, fear/avoidance $\alpha = .87$; malevolence $\alpha = .83$; authoritarianism $\alpha = .82$; $\alpha = .79$

(Kenny et al., 2018) and fear/avoidance α = .83/.89; malevolence α = .68/.74; authoritarianism α = .72/.85; α = .77/.85 (α = Time 1/Time 2; Poulgrain et al., 2022). It also has adequate concurrent validity when compared with the *Community Attitudes toward the Mentally Ill scale* (CAMI; α = .78*), Social Dominance Orientation* scale (SDO; α = .52) and the *Authoritarianism-Conservatism-Traditionalism scale* (ACT; α = .51; Poulgrain et al., 2022). The internal consistency of the PPMI was good (α = .91).

Procedure

The procedure is consistent across all studies and is described above for Study 1.

RESULTS

Statistical analyses were run with SPSS Version 28. Data were checked for missing data then assumptions for Chi-Square, ANOVA, and ANCOVA were run.

Missing Value Analysis

Little's MCAR test indicated that missing data were missing completely at random; $\chi^2(252) = 284.53$, $p = .078$.

Attractiveness Manipulation

Attractiveness was manipulated using images from the Chicago Face Database. To verify the effectiveness of the attractiveness manipulation, participants were asked to rate the attractiveness of the photo attached to the transcript they read. The unattractive images were given lower ratings than the attractive images and was more effective for females than males; however, attractiveness ratings in the Chicago Face Database set of images had larger differences between unattractive and attractive pairings. Descriptive statistics can be seen in Table 3.2.

Table 3.2

Descriptive Statistics of Attractiveness Ratings of Defendant Photographs

	M	SD	Range	CFD *M*
Unattractive male	3.06	1.24	1-6	2.65
Attractive male	3.98	1.09	1-6	4.12
Unattractive female	3.30	1.01	1-5	2.68
Attractive female	4.28	.95	1-6	5.09

Note. CFD = Chicago Face Database

Verdict

Participants were randomly assigned one of 16 vignettes that manipulated attractiveness, gender, and psychiatric diagnosis. Descriptive statistics can be seen in Table 3.3.

Table 3.3

Vignette Classification and Sample Sizes

Vignette Type	N	% of sample per group
AFCD	12	5.7
AFPT	15	7.1
AFS	12	5.7
AFND	16	7.5
UFCD	13	6.1
UFPT	18	8.5
UFS	13	6.1
UFND	12	5.7
AMCD	14	6.6
AMPT	15	7.1
AMS	10	4.7
AMND	11	5.2
UMCD	13	6.1
UMPT	10	4.7
UMS	15	7.1
UMND	13	6.1

Note. AFCD = attractive female conduct disorder, AFPT = attractive female psychopathic traits, AFS = attractive female schizophrenia, AFND = attractive female no diagnosis, UFCD = unattractive female conduct disorder, UFPT = unattractive female psychopathic traits, UFS = unattractive female schizophrenia, UFND = unattractive female no diagnosis, AMCD = attractive male conduct disorder, AMPT = attractive male psychopathic traits, AMS = attractive male schizophrenia, UMCD = unattractive male conduct disorder,

UMPT = unattractive male psychopathic traits, UMS = unattractive male schizophrenia, UMND = unattractive male no diagnosis

Overall, 71.7% of participants gave a guilty verdict. A series of chi-square tests of independence were conducted to ascertain if attractiveness, gender, and psychiatric diagnosis were associated with differences in guilty or not guilty ratings (see Table 3.4). Key assumptions were met for the individual chi-square tests comparing attractiveness, gender, and psychiatric diagnosis. Inconsistent with hypothesis one, there was not a statistically significant difference in guilty/not guilty verdicts based on attractiveness, $\chi^2(1) = .00$, $p = 1.00$. Inconsistent with hypothesis two, there was not a statistically significant difference in guilty/not guilty verdicts based on gender, $\chi^2(1) = .04$, $p = .848$. Consistent with hypothesis three, there was a statistically significant difference in guilty/not guilty verdicts based on psychiatric diagnosis, $\chi^2(3) = 15.08$, $p = .002$. Defendants with schizophrenia were given not guilty verdicts more than any other condition and those with psychopathic traits were given not guilty verdicts least often.

Table 3.4

Frequencies and Chi-Square Results for Guilty and Not Guilty Verdict Associations with Attractiveness, Gender, and Psychiatric Diagnosis (N = 210)

Transcript Condition	Yes		No		χ^2	df
	n	%	n	%		
Attractiveness	152	72.4	58	27.6	.00	1
Attractive	76	72.4	29	27.6		
Unattractive	76	72.4	29	27.6		
Gender	152	72.4	58	27.6	.04	1
Female	79	71.8	31	28.2		
Male	73	73	27	27		
Psychiatric Diagnosis	152	72.4	58	27.6	15.08*	3
Conduct Disorder	42	80.8	10	19.2		
Psychopathic Traits	50	86.2	8	13.8		
Schizophrenia	29	59.2	20	40.8		
No Diagnosis	31	60.8	20	39.2		

*$p < .01$

Participant Perceptions of the Defendant

A series of ANOVAs and ANCOVAs were conducted to determine if levels of guiltiness varied as a function of attractiveness, gender, and psychiatric diagnosis. Exploratory analyses of confidence in ratings, defendant and victim credibility, the likelihood of recommending treatment, whether the defendant would benefit from treatment, risk of future violence by the defendant, and the threat they pose to society can be seen in Appendix E. Given the absence of hypotheses and in the interests of parsimony, exploratory analyses were restricted to ANCOVAs and pertinent post hoc testing. All of the ANOVAs and ANCOVAs evaluated mean differences of scores based on: attractiveness, gender, psychiatric diagnosis, attractiveness and psychiatric diagnosis, and attractiveness, gender and psychiatric diagnosis.

Outliers were found for all the dependent variables as assessed by inspection of boxplots. After winsorizing extreme cases by reducing them to the next highest value plus one, some outliers remained which were kept in the analyses. Shapiro-Wilk tests indicated that the data were not normally distributed for several groups. However, visual inspection of histograms and normal Q-Q plots revealed approximately normal distributions, and the skewness and kurtosis showed most z-scores within ± 1.96, with a few within ± 3.29, suggesting the data were acceptable (Kim, 2013). Additionally, ANOVA is considered robust to violations of normality.

Level of Guiltiness

Hypothesis 1: Attractive Defendants Will Be Perceived As Less Guilty Than Unattractive Defendants. To evaluate mean differences in participant ratings of guiltiness based on attractiveness, a one-way ANOVA was conducted. The assumption of normality was violated, as assessed by a significant Shapiro-Wilk test. Visual inspection of histograms and normal Q-Q plots showed approximately normal distributions. There were no outliers. Homogeneity of

variances was violated; therefore, a Welch's ANOVA was conducted. Inconsistent with hypothesis one, there was no significant difference in mean ratings of guiltiness by attractiveness (see Table 3.5). When BIDR-16, RLAQ, and PPMI scores were included as covariates, the effect remained nonsignificant, $F(1, 182) = .43$, $p = .514$, partial $\eta^2 = .00$ (see Table 3.6 for means, adjusted means, and variances of guiltiness ratings).

Table 3.5

Means, Standard Deviations, and One-Way Analyses of Variance in Ratings of Guiltiness based on Attractiveness

	Attractive (n=105)		Unattractive (n = 107)		$F(1, 201.94)$	η^2
	M	SD	M	SD		
Level of Guiltiness	4.70	1.36	4.96	1.13	2.42	.01

Table 3.6

Means, Adjusted Means, and Variances of Ratings of Guiltiness Based on Attractiveness After Controlling for the BIDR-16, RLAQ, and PPMI

	M	SD	M_{adj}	SE
Attractive	4.81	1.32	4.82	.13
Unattractive	4.94	1.11	4.93	.12

Hypothesis 2: Female Defendants Will Be Perceived As Less Guilty than Male Defendants. A one-way ANOVA was conducted to evaluate mean differences in participant ratings of guiltiness based on gender. The assumption of normality was violated, as assessed by a significant Shapiro-Wilk test. Visual inspection of histograms and normal Q-Q plots showed approximately normal distributions. Eleven outliers were found and were winsorized. Some outliers remained but were not extreme and were genuinely unusual values; therefore, they were kept them in the analyses. Homogeneity of variances was met. Inconsistent with hypothesis two, there was not a statistically significant mean difference in ratings of guiltiness by gender, $F(1, 210) = 1.50$, $p = .222$, $\eta^2 = .01$ (see Table 3.7 for means and standard deviations of ratings of

guiltiness). When BIDR-16, RLAQ, and PPMI scores were included as covariates, the effect remained nonsignificant, $F(1, 182) = 1.15$, $p = .29$, partial $\eta^2 = .01$ (see Table 3.8 for means, adjusted means, and variances of ratings of guiltiness).

Table 3.7

Means and Standard Deviations of Ratings of Guiltiness based on Gender

	Male (n=101)		Female (n= 111)	
	M	SD	M	SD
Level of Guiltiness	4.94	1.33	4.73	1.18

Table 3.8

Means, Adjusted Means, and Variances of Ratings of Guiltiness Based on Gender After Controlling for the BIDR-16, RLAQ, and PPMI

	M	SD	M_{adj}	SE
Female	4.77	1.15	4.78	.13
Male	5.00	1.28	4.98	.13

Hypothesis 3: Guilty Ratings Will Vary Based On Psychiatric Diagnosis; Those Diagnosed With Schizophrenia Will Be Considered Less Guilty Than The Rest, And Those With Psychopathic Traits Will Be Considered More Guilty Than The Rest. A one-way ANOVA was conducted to evaluate mean differences in ratings of guiltiness based on psychiatric diagnosis. The assumption of normality was violated, as assessed by a significant Shapiro-Wilk test. Visual inspection of histograms and normal Q-Q plots showed approximately normal distributions. Outliers were found and were winsorized by adding one to the next highest value. Some outliers remained but were not extreme and were genuinely unusual values; therefore, these were retained in the analyses. Homogeneity of variances was violated; therefore, a Welch's ANOVA was conducted. There was a statistically significant difference in mean guiltiness ratings based on psychiatric diagnosis (see Table 3.9). This effect remained statistically significant when controlling for the BIDR-16, RLAQ, and PPMI, $F(3, 180) = 5.05$, p

= .002, partial η^2 = .08. Post-hoc analyses with a Bonferroni adjustment showed partial support for hypothesis three; those with psychopathic traits were given significantly higher ratings of guiltiness than those with schizophrenia (Mean difference = 8.22, 95% *CI* =2.15, 14.30, *p* = .002) and no diagnosis (Mean difference = 6.33, 95% *CI* =.14, 12.51, *p* = .042) (see Table 3.10).

Table 3.9

Means, Standard Deviations, and One-Way Analysis of Variance for Guiltiness and Psychiatric Diagnosis

				ANOVA		
	N	M	SD	Welch's *F*	*df*	η^2
Psychiatric Diagnosis	212	4.83	1.25	6.46*	3, 112.70	.08
Conduct Disorder	52	4.94	1.06			
Psychopathic Traits	58	5.33	1.00			
Schizophrenia	50	4.58	1.28			
No Diagnosis	52	4.40	1.47			

*$p < .01$

Table 3.10

Means, Adjusted Means, Standard Deviations, and Standard Errors for Ratings of Guiltiness by Psychiatric Diagnosis with the BIDR-16, RLAQ, and PPMI as Covariates

	M	*SD*	M_{adj}	*SE*
Conduct Disorder	26.34	10.71	26.35	1.64
Psychopathic Traits	29.45	10.20	29.42	1.54
Schizophrenia	21.24	11.07	21.20	1.68
No Diagnosis	23.02	12.81	23.10	1.73

Hypothesis 4: Attractive Defendants With A Psychiatric Diagnosis Will Be Perceived As Less Guilty Than Unattractive Defendants With A Psychiatric Disorder. A two-way ANOVA was conducted to determine if attractiveness and psychiatric diagnosis had an effect on ratings of guiltiness. Homogeneity of variances was violated, and a square transformation was applied which satisfied the assumption of homogeneity of variances. The assumption of normality was also violated. Skewness and kurtosis values showed that z-scores were within ± 3 standard deviations, which suggests the data were approximately normally

distributed (Kim, 2013). There were no outliers. The interaction effect between attractiveness and psychiatric diagnosis was not statistically significant, $F(3, 204) = 1.79$, $p = .151$, partial $\eta^2 = .03$ (see Table 3.11). An analysis of main effects for psychiatric diagnosis was significant. This remained true when controlling for the BIDR-16, RLAQ, and PPMI, which showed a significant effect of psychiatric diagnosis, $F(3, 176) = 5.17$, $p = .002$, partial $\eta^2 = .08$. Pairwise comparisons using a Bonferroni correction showed significantly higher ratings of guiltiness for defendants with psychopathic traits compared to those with schizophrenia (Mean difference = 8.34, 95% *CI* =2.24, 14.44, $p = .002$) and no diagnosis (Mean difference = 6.40, 95% *CI* =.21, 12.58, $p = .038$). This was only true for attractive defendants with schizophrenia (Mean difference = 11.51, 95% *CI* =2.54, 20.49, $p = .005$) and no diagnosis (Mean difference = 9.74, 95% *CI* =.82, 18.67, $p = .024$) (see Table 3.12 for means, adjusted means, and variances of guiltiness ratings and Figure 3.1).

Table 3.11

Means, Standard Deviations, and Two-Way ANOVA Statistics for Ratings of Guiltiness by Attractiveness and Psychiatric Diagnosis

	M	*SD*	Effect	*F* ratio	*df*	η^2
Attractive			A	2.28	1, 204	.01
Conduct disorder	4.65	1.06	PD	5.25*	3, 204	.07
Psychopathic traits	5.47	1.11				
Schizophrenia	4.41	1.18	CDxA	3.15	1, 204	.01
No diagnosis	4.11	1.65	PTxA	1.41	1, 204	.01
Unattractive			ScxA	1.00	1, 204	.01
Conduct disorder	5.23	.99	NDxA	1.75	1, 204	.01
Psychopathic traits	5.18	.86	AtxPD	6.00**	3, 204	.08
Schizophrenia	4.71	1.36	UnxPD	1.17	3, 204	.02
No diagnosis	4.72	1.21				

Note. A = attractiveness, PD = psychiatric diagnosis, CD = conduct disorder, PT = psychopathic traits, Sc = schizophrenia, ND = no diagnosis, At = attractive, Un = unattractive
* $p < .05$, ** $p < .001$.

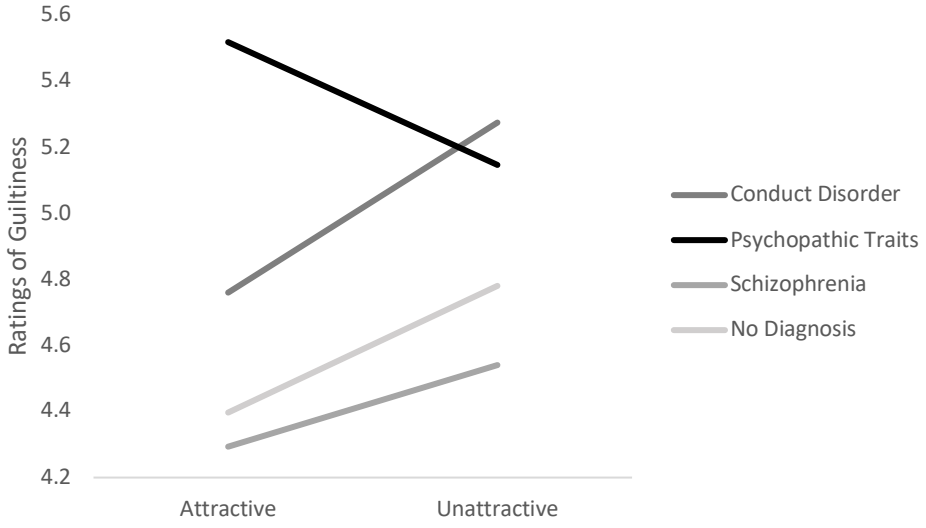
Table 3.12

Means, Adjusted Means, and Variances for Ratings of Guiltiness by Attractiveness and Psychiatric Diagnosis while Controlling for the BIDR-16, RLAQ, and PPMI

	M	*SD*	M_{adj}	*SE*
Attractive				
Conduct disorder	4.74	1.10	4.76	.24
Psychopathic traits	5.52	1.05	5.52	.23
Schizophrenia	4.32	1.25	4.29	.27
No diagnosis	4.37	1.61	4.40	.27
Unattractive				
Conduct disorder	5.29	1.00	5.28	.24
Psychopathic traits	5.15	.88	5.15	.23
Schizophrenia	4.54	1.24	4.54	.23
No diagnosis	4.78	1.20	4.78	.24

Figure 3.1

Estimated Marginal Means of Guiltiness by Attractiveness and Psychiatric Diagnosis After Controlling for the BIDR-16, RLAQ, and PPMI



Hypothesis 5: Unattractive Female Defendants With A Psychiatric Diagnosis Will Be Perceived As More Guilty Than All Other Defendant Types. A three-way ANOVA was attempted to determine if attractiveness, gender, and psychiatric diagnosis had an effect on ratings of guiltiness. There were six non-extreme outliers that were kept in the analysis. The assumptions of normality and homogeneity of variances were violated. A square transformation

was applied which resolved the violation of homogeneity of variances. Visual inspection of histograms and normal Q-Q plots showed approximately normal distributions. There was not a statistically significant interaction between attractiveness, gender, and psychiatric diagnosis, $F(3, 196) = .50$, $p = .685$, partial $\eta^2 = .01$. After adjustment for the BIDR-16 and PPMI, the non-significant interaction remained, $F(3, 177) = .57$, $p = .635$, $\eta^2 = .01$. Inconsistent with hypothesis five, unattractive females with a psychiatric diagnosis did not have significantly higher mean scores than all other conditions (see Table 3.13 for means and standard deviations). An examination of main effects indicated statistically significant differences based on psychiatric diagnosis, $F(3, 196) = 5.15$, $p = .002$, partial $\eta^2 = .07$ (see Table 3.14). Pairwise comparisons with a Bonferroni adjustment showed that attractive defendants with psychopathic traits were considered significantly more guilty than those with schizophrenia (Mean difference = 9.83, 95% *CI* =.72, 18.93, $p = .027$) and no diagnosis (Mean difference = 11.57, 95% *CI* =2.90, 20.23, $p = .003$). Attractive defendants with conduct disorder were considered significantly less guilty than unattractive defendants with conduct disorder (Mean difference = 6.66, 95% *CI* =.01, 13.31, $p = .050$). Male defendants with psychopathic traits were considered significantly more guilty than male defendants with no diagnosis (Mean difference = 12.67, 95% *CI* =3.30, 22.03, $p = .002$). Unattractive males were considered significantly more guilty than attractive males (Mean difference = 5.14, 95% *CI* =.31, 9.96, $p = .037$). Attractive males with psychopathic traits were given significantly higher guiltiness ratings than attractive males with no diagnosis (Mean difference = 14.41, 95% *CI* =1.56, 27.26, $p = .019$). Unattractive females with psychopathic traits were considered significantly less guilty than unattractive males with psychopathic traits (Mean difference = 11.22, 95% *CI* =1.77, 20.67, $p = .020$).

Table 3.13

Means, Standard Deviations, and Sample Size for Guiltiness Ratings by Psychiatric Diagnosis, Attractiveness, and Gender

Psychiatric Disorder	Attractiveness	Gender	*M*	*SD*	*N*
Conduct Disorder	Attractive	Female	21.58	7.14	12
		Male	23.71	12.91	14
	Unattractive	Female	29.08	13.98	13
		Male	29.54	8.53	13
Psychopathic Traits	Attractive	Female	29.60	11.52	15
		Male	31.13	14.29	15
	Unattractive	Female	22.78	9.01	18
		Male	34.00	7.39	10
Schizophrenia	Attractive	Female	20.58	8.67	12
		Male	20.50	12.32	10
	Unattractive	Female	21.15	13.16	13
		Male	26.00	16.13	15
No Diagnosis	Attractive	Female	20.88	15.89	16
		Male	16.73	11.11	11
	Unattractive	Female	23.92	11.90	12
		Male	23.08	12.53	13

Table 3.14

Three-Way Analysis of Variance for Ratings of Guiltiness

	F	*p*	η^2
Attractiveness	3.37	.068	.02
Psychiatric Diagnosis	5.15	.002	.07
Gender	1.25	.265	.01
Attractiveness * Psychiatric Diagnosis	1.25	.293	.02
Attractiveness*Gender	1.45	.230	.01
Psychiatric Diagnosis *Gender	1.20	.311	.02
Attractiveness*Gender* Psychiatric Diagnosis	.50	.685	.01

A three-way ANCOVA was conducted to examine the effects of attractiveness, gender, and psychiatric diagnosis on ratings of guiltiness, after controlling for socially desirable responding (BIDR-16), authoritarian legal attitudes (RLAQ), and mental illness prejudice (PPMI). Linearity was not established for the RLAQ which was omitted from the analysis. Homogeneity of variances was violated, and a square transformation was applied, which resolved the homogeneity violation. Key assumptions were met, except for normality of

standardized residuals for all but two groups; however, visual inspection of normal Q-Q plots showed approximately normal distributions. Pair-wise comparisons with a Bonferroni adjustment showed that attractive males with psychopathic traits were considered significantly more guilty than attractive males with schizophrenia (Mean difference = 13.89, 95% *CI* =.46, 27.32, p = .038) and no diagnosis (Mean difference = 20.65, 95% *CI* =6.36, 34.95, $p < .001$) (see Figure 3.2). Unattractive males with psychopathic traits were considered significantly more guilty than unattractive females with psychopathic traits (Mean difference = 12.47, 95% *CI* =2.85, 22.09, p = .011) (see Figure 3.3). Unattractive males with no diagnosis were considered significantly more guilty than attractive males with no diagnosis (Mean difference = 11.17, 95% *CI* =.41, 21.94, p = .042) (see Table 3.15 for adjusted and unadjusted means and variability and Figure 3.4).

Figure 3.2

Estimated Marginal Means of Guiltiness Ratings for Attractive Defendants

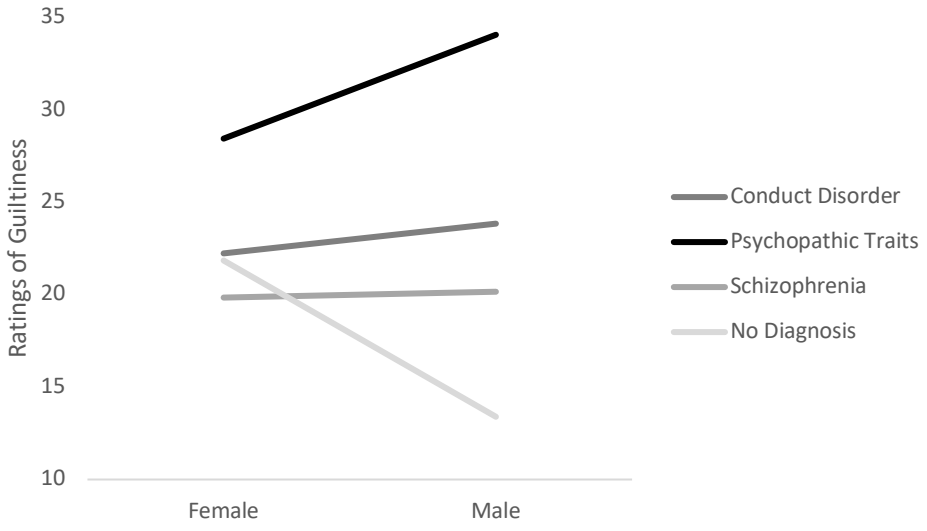


Figure 3.3

Estimated Marginal Means of Guiltiness Ratings for Unattractive Defendants

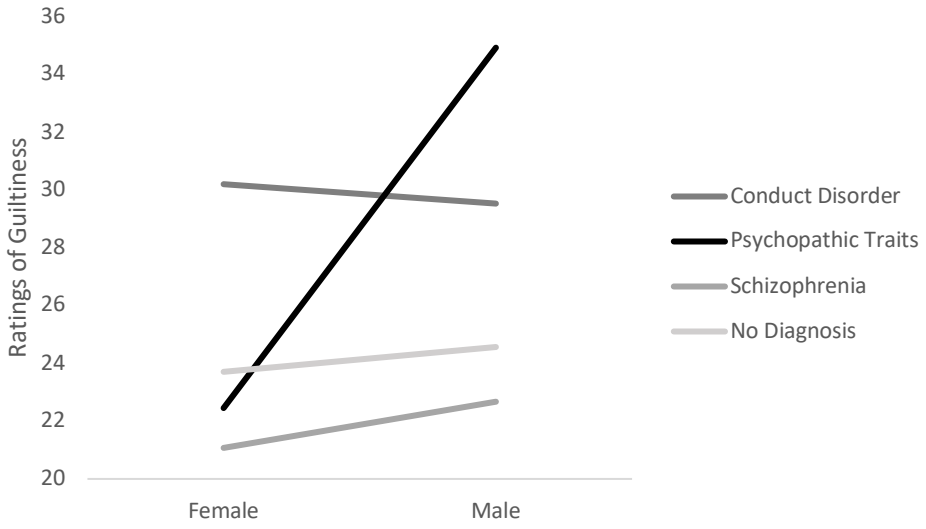


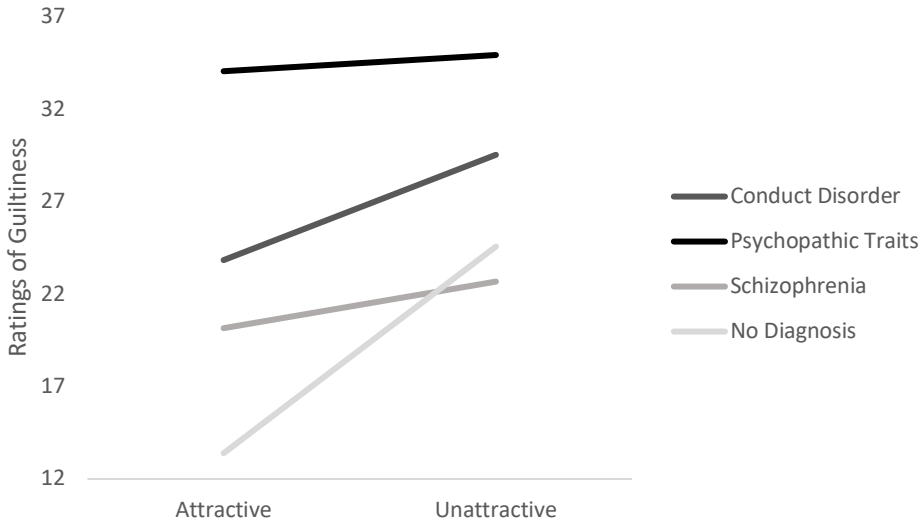
Table 3.15

Adjusted and Unadjusted Means for Guiltiness Ratings Based on Attractiveness, Gender, and Psychiatric Diagnosis with the BIDR-16 and PPMI as Covariates

	N	*M*	*SD*	M_{adj}	*SE*
Attractive					
Female					
Conduct Disorder	11	22.09	7.26	22.23	3.62
Psychopathic Traits	14	28.21	10.58	28.43	3.19
Schizophrenia	10	19.70	9.31	19.84	3.80
No Diagnosis	14	21.57	16.95	21.84	3.20
Male					
Conduct Disorder	14	23.71	12.91	23.83	3.19
Psychopathic Traits	13	34.00	13.02	34.05	3.31
Schizophrenia	10	20.50	12.32	20.15	3.79
No Diagnosis	8	13.38	10.03	13.39	4.22
Unattractive					
Female					
Conduct Disorder	11	30.64	14.60	30.20	3.65
Psychopathic Traits	18	22.78	9.01	22.45	2.83
Schizophrenia	13	21.15	13.16	21.07	3.31
No Diagnosis	12	23.92	11.90	23.71	3.46
Male					
Conduct Disorder	13	29.54	8.53	29.53	3.33
Psychopathic Traits	9	35.00	7.09	34.92	3.97
Schizophrenia	13	22.46	14.20	22.67	3.32
No Diagnosis	12	24.25	12.32	24.56	3.45

Figure 3.4

Estimated Marginal Means of Ratings of Guiltiness for Male Defendants



DISCUSSION

Guilt, Attractiveness and Gender

This study examined whether attractiveness, gender, and psychiatric diagnosis had an effect on participant perceptions of guilt. Measures of SDR, authoritarian legal attitudes and prejudice against people with mental illness were used to control for their effects on decision making. The first two hypotheses were not supported, as attractiveness and gender did not have an effect on the dichotomous guilty/not guilty rating or the Likert-type rating of guiltiness. As in study one, although mean ratings of attractiveness were in alignment with the intended manipulation, the differences were not as large as the mean ratings in the CFD. This suggests that the attractiveness manipulation may not have been as effective as intended. The ratings were approximately average, with the unattractive images being just below average and the attractive images being just above average. These findings may have been different had the sample not been entirely made up of self-identified female participants.

Psychiatric Diagnosis

The study did find partial support for hypothesis three, that those with schizophrenia would be perceived not guilty more often and less guilty than the rest, and those with psychopathic traits would be perceived as guilty more often than the rest. However, those with psychopathic traits were considered significantly more guilty than those with schizophrenia and no diagnosis, not those with conduct disorder, which remained true after controlling for socially desirable responding, authoritarian legal attitudes, and mental illness prejudice. These findings were consistent with previous research regarding diagnostic labelling effects (e.g., Blais & Forth, 2014; Boccaccini et al., 2008; Taylor et al., 2019). Some evidence indicates these labelling effects are also generalizable to real courts. The psychopathic traits label has been used increasingly in the juvenile court system to infer high risk or low treatment amenability (Viljoen et al., 2010). Furthermore, psychopathy evidence has been introduced in cases involving giving adult sentences to youth, sexual offences, and to bolster criminal responsibility arguments (Viljoen et al., 2010).

Those with schizophrenia were not considered significantly less guilty than those with no diagnosis or conduct disorder. These findings suggest participants considered those with psychopathic traits to be more similar to those with conduct disorder than those with schizophrenia and no diagnosis. Because participants in this study were undergraduate students registered in psychology courses, they may have been exposed to diagnostic criteria and lecture material for conduct disorder and psychopathic traits, which may have affected the nonsignificant differences in ratings of guiltiness between psychopathic traits and conduct disorder. Indeed, undergraduate students in majors related to mental health have been found to be significantly more aware of mental disorders and resources than other undergraduate students (Kalkbrenner et al., 2020)

The increased ratings of guiltiness for psychopathic traits may also be due to media portrayals of mental illness, which rarely represent reality and often paint those with some mental illnesses to be more dangerous and likely to victimize others (Ma, 2017). Additionally, previous research has shown there to be a higher rate of psychopathic traits in prison populations compared to the general population, which may contribute to a heuristic that associates psychopathic traits with criminal behaviour. However, it may be that the inclusion of criminal/antisocial behaviour on psychopathy measures artificially inflates those rates (Boduszek et al., 2019). When those behaviours are excluded, one study showed that some dimensions of psychopathy are higher among adolescents (i.e., underdeveloped emotional reactions and cognitive engagement with others), university students (i.e., interpersonal manipulation and egocentrism), and community adults (i.e., interpersonal manipulation) compared to incarcerated people (Boduszek et al., 2019). This suggests that criminal/antisocial behaviour is a possible outcome of psychopathic traits, rather than a dimension of psychopathy. This is especially important when considering adolescents, because they have greater deficits in some facets of psychopathy (i.e., affective and cognitive responsiveness) due to the underdeveloped prefrontal cortex (Boduszek et al., 2019).

Attractiveness and Psychiatric Diagnosis

Attractiveness interacted with psychopathic traits, such that attractive defendants with psychopathic traits were considered guiltier than any other defendant. This was inconsistent with hypothesis four. All other attractive defendants had lower ratings of guiltiness compared to their unattractive counterparts, but these differences were not significantly different. The psychopathic traits finding may be due to the description of psychopathic traits given by the expert witness, which included language consistent with fraud (i.e., … tries to charm, con, and manipulate others

to obtain what he/she wants). The beauty penalty may have been responsible for these increased ratings of guiltiness, because participants may have viewed the attractive defendant as having used their good looks to manipulate others (Yang et al., 2019).

Attractiveness, Psychiatric Diagnosis and Gender

Hypothesis five was not supported. Unattractive females with a psychiatric diagnosis were not considered more guilty than all other defendants. Males with psychopathic traits were given the highest ratings of guiltiness compared to all other conditions. These findings were in contrast to expectations, which have shown women to be considered more guilty when they have a mental illness or when they are unattractive (Breheney et al., 2007; Abwender & Hough, 2001). However, because the participants were all women, it is possible there was an in-group leniency effect, or that the attractiveness manipulation was not effective enough to elicit more extreme attractiveness-based judgments. Interestingly, the attractive male with no diagnosis was considered the least guilty. Participants may have responded in this way because the beauty bias may be especially salient for opposite-gender attractive individuals (Agthe et al., 2016). This is especially interesting in light of exploratory analyses which showed that attractive males with no diagnosis were considered the least likely to benefit from treatment (see below).

Exploratory Analyses

Exploratory analyses were also run to investigate the possible effects of attractiveness, gender, and psychiatric diagnosis on confidence ratings, defendant and victim credibility ratings, the likelihood of recommending treatment, treatment benefit, the effectiveness of a treatment aimed at reducing violence, the defendant's risk for future violence, the defendant's threat to society, and the defendant's trustworthiness (see Appendix E).

Confidence Ratings

Participants' confidence in their ratings was affected by attractiveness, gender, and psychiatric diagnosis. Participants had lower confidence when the defendant was attractive and had conduct disorder compared to their unattractive counterparts, regardless of gender (see Table E.1 and Figure E.1). Confidence was lower for participants in the unattractive female defendants with psychopathic traits group compared to their male counterparts and unattractive females in the conduct disorder group (see Figure E.4.). Confidence was also lower for attractive males with conduct disorder compared to any males with psychopathic traits (see Figure E.5). Confidence ratings may have been lower when participants were attractive with conduct disorder because of the beauty bias – when unattractive people also have a psychiatric diagnosis that is associated with negative attributes, people may feel more confident in their decisions about them. Conversely, when attractive people have the same psychiatric diagnosis, the incongruency of positive attributes associated with attractiveness and negative attributes associated with the diagnosis may cause a cognitive distortion that affects confidence in decision-making (Curley et al., 2022).

Defendant Credibility

Defendant credibility varied as a function of gender, psychiatric diagnosis and attractiveness (see Table E.7). Females who were unattractive with no diagnosis were considered significantly more credible than those with conduct disorder or psychopathic (see Table E.6). Given the description of the defendants with psychopathic traits, it makes sense that the attractive and unattractive defendants with psychopathic traits were considered the least credible. This may also be due to female defendants generally receiving more leniency compared to male defendants, as previous research has shown (Doerner & Demuth, 2014). After controlling for

socially desirable responding, authoritarian legal attitudes, and mental illness prejudice, males with psychopathic traits were considered significantly less credible than the unattractive female with no diagnosis.

Treatment Recommendation

The likelihood that participants would recommend treatment was affected by psychiatric diagnosis, gender, and attractiveness (see Table E.12). Attractive male defendants with no diagnosis were the least likely to receive a treatment recommendation. There are several reasons why this may be. Perhaps it is an example of the beauty penalty – because an attractive person behaved in a way inconsistent with expectations, they are undeserving of help (Wilson and Eckel, 2006). Additionally, because they did not have a diagnosis, but still allegedly engaged in violent behaviour, they did not have an excuse for their behaviour, whereas a person with a psychiatric diagnosis may have a reason for behaving violently (Berryessa et al., 2015). Further, some research shows that participants are more likely to recommend psychiatric commitment to any defendant with a mental illness compared to defendants with no diagnosis (Kortright, 2019). Perhaps participants believed treatment would be ineffective for those with no diagnosis because they did not have a labelled problem to treat.

There was also a significant effect of psychiatric diagnosis, gender and attractiveness, such that attractive females with schizophrenia were less likely to be recommended treatment than attractive males with schizophrenia, and unattractive females with psychopathic traits were significantly less likely to receive a treatment recommendation than unattractive males with psychopathic traits. These findings are consistent with Mossière and Maeder (2016), who found that participants perceived female defendant behaviours to result from their personality rather than their circumstances. Despite a growing body of research that personality traits change over

the lifespan, the assumption that they are static is still common (Bleidorn et al., 2021). Because of this belief, it is possible that participants believed enduring traits led to the female defendants with schizophrenia and psychopathic traits alleged delinquency, such that treatment recommendation would not be worthwhile. And indeed, attractive females with schizophrenia were considered significantly less likely to benefit from treatment than their male counterparts (see Table E.15).

Treatment Benefit

Attractive males with no diagnosis were considered significantly less likely to benefit from treatment than those in the unattractive female schizophrenia group, attractive male schizophrenia group and the unattractive male conduct disorder group (see Tables E.15 and E.16). This was inconsistent with Taylor et al. (2019) who found that defendants with schizophrenia were considered the least likely to benefit from treatment. In each of the groupings by attractiveness and gender, defendants with no diagnosis were considered less likely to benefit from treatment than their same gender and attractiveness counterparts with a psychiatric diagnosis. This trend of considering attractive males with no diagnosis to be least likely to benefit from treatment extended to when a treatment was specifically aimed at reducing violence (see Tables 19 and 20).

Risk for Future Violence

Despite participants' expectations that those with no diagnosis would benefit least from treatment, they also rated those with no diagnosis to be at the lowest risk for future violence. Specifically, they rated attractive defendants with no diagnosis to be at the lowest risk for future violence (see Table E.22). Perhaps because they assumed attractive defendants with no diagnosis are at less risk for violence, they simply did not need treatment, and the benefit they would get

from treatment would be negligible because they were already considered low risk. Similarly, participants considered attractive defendants with no diagnosis to be the lowest threat to society, and attractive defendants with psychopathic traits to be at a significantly higher threat to society than those with no diagnosis (see Table E.22). Indeed, attractiveness has been found to influence review board members when a detainee has psychopathy, such that they are more likely to remain detained than if they had been unattractive (Denomme et al., 2020). These findings imply that labelling effects are not solely responsible for decision-making, as unattractive defendants with psychopathic traits were considered a lower threat to society than those with conduct disorder and schizophrenia. Therefore, it may be the attractiveness component that increases threat. Perhaps because an attractive person with psychopathic traits may be more successful in manipulating others than an unattractive person with psychopathic traits.

Threat to Society

Finally, attractive defendants with psychopathic traits were considered a greater threat to society than those who were unattractive and those who were attractive with schizophrenia or no diagnosis (see Table E.26 and Figures E.7 and E.8). Research has shown this effect, such that attractiveness moderates the effect of psychopathic traits in a study of forensic patients' detainment or release decisions by a review board (Denomme et al., 2020). It may be explained by the beauty penalty regarding punishment towards attractive people who manipulate others. Another possible explanation is that the tendency to be drawn towards attractiveness increases attention to all of the information associated with that attractiveness (Cann et al., 1981); this may have led participants to pay more attention to the description of the psychopathic traits mentioned in the transcript, increasing awareness of the dangers associated with those traits.

Unattractive defendants with psychopathic traits were not significantly less trustworthy than unattractive defendants with no diagnosis.

Implications

This study contributes to the research on extralegal factors that affect decision-making. Research that examines the impact of diagnostic labelling on youth is especially important, given their development through adolescence and emerging adulthood. It also indicates the effect of a label on decision-making when all else is equal. The expert witness in the vignettes stated they spoke with the defendant for several hours which led them to believe the defendant had psychopathic traits. However, there is no indication that this is a valid diagnosis based on assessment. Therefore, life-altering decisions should not be based on potentially inaccurate labels. This suggests that mental health labels should not be factored into the judicial process without confirmation of their validity, as judges may also be influenced by the testimony of expert witnesses (Zinger & Forth, 1998). The findings of this study that showed the effect of attractiveness on diagnostic labels is also important. These findings, if replicated, can contribute to training for decision-makers who are involved with justice-involved youth.

Limitations and Future Directions

This study shares in common with Study 1 the same set of methodological limitations. As well, another limitation pertains to the chosen psychiatric diagnoses. Although the chosen diagnoses were appropriate based on previous research, there are other diagnoses that are especially common in justice-involved youth (e.g., attention-deficit/hyperactivity disorder, major depressive disorder). Future studies should incorporate the diagnoses that are most prevalent in justice-involved youth.

Another limitation pertains to participant characteristics. Some research has shown that experiences of victimization or the presence of psychopathic traits in the participant have an effect on decision-making (Lilley et al., 2023). Future research should incorporate participant experiences of victimization and personality traits, as these are relevant in people who may have contact with and particularly exercise decision-making responsibilities for justice-involved youth.

CHAPTER 4 : STUDY THREE

The third study continued the investigation into possible influences of extralegal factors on mock-juror decision making. Along with attractiveness and gender, race was considered.

Race

Race is a well-established correlate in discriminatory sentencing practices as evidenced by the overrepresentation of racialized groups incarcerated in Canada (Statistics Canada, 2022) and a number of American archival studies indicating longer or death penalty sentences for racial minorities (Austin & Allen, 2000; Mustard, 2001; Williams & Holcomb, 2001). Meta-analytic reviews of studies using the mock-juror design have demonstrated inconsistent results, with one suggesting there is a small but significant effect of racial bias on jury decision-making with White participants giving Black defendants longer sentences compared to White defendants (Sweeney & Haney, 1992). Conversely, Mazzella and Feingold (1994) did not find significantly disparate ratings of guilt between White and Black participants. However, defendants received longer sentences for crimes against White victims compared to Black victims, and White defendants received harsher sentences for crimes of fraud, while Black defendants received harsher sentences for crimes of negligent homicide (Mazzella & Feingold, 1994). These findings

indicate that race alone may not determine disparate sentencing practices, but the interaction of race and several other variables appear to.

Indeed, Mitchell and colleagues (2005) examined the literature pertaining to racial bias on verdict and sentencing practices and found a small but significant effect of racial bias in verdict and sentencing practices favouring racial in-group members, but this effect was moderated by the date of publication (i.e., more verdict disparity associated with race for studies conducted in the 1970s), whether standard jury instructions were given or not (i.e., instructions were associated with less racial disparity), the race of the participant (i.e., Black participants were more lenient with in-group members), and whether the verdict was based on a continuous or a dichotomous scale (i.e., more verdict disparity associated with race with a continuous scale).

These findings suggest that real jury members may exhibit less racial bias in sentencing practices than do participants in mock-juror studies. They also suggest the presence of implicit biases that undergraduates and community members may have towards racial out-group members that may be malleable through instruction. It is also worth noting the greater lenience of Black participants compared to White participants for in-group members. This may be due to the importance or strength of identification with one's racial group, which may be stronger for minority groups (Grier & Deshpande, 2001).

Another consideration is race salience, which is overt in many studies investigating racial bias in mock-juror decision-making. Consistent with the covert nature of modern racism, when race is made salient, participants may overcorrect for their racial bias leading to no significant differences between guilty verdicts of Black and White defendants (Sommers & Ellsworth, 2001). The converse is also true; when race is not salient, Black defendants receive guilty verdicts by mock jurors at a significantly higher rate than White defendants. This may explain many of the nonsignificant findings of racial bias in the research using mock juror design with

mock trial vignettes in which descriptions of defendants are overtly race salient (Sommers & Ellsworth, 2001).

Race and Attractiveness. The effect of race on ratings of guiltiness and sentencing severity has been found to be moderated by several variables. Attractiveness is one of these. Cothran and colleagues (2017) investigated this interaction with a group of 363 undergraduate students across a variety of crimes (i.e., theft, home invasion, aggravated robbery, and aggravated assault). They found that Black participants were less likely to find any defendant guilty compared to White participants, and all participants were more lenient towards Black defendants unless they were unattractive. In this case, White participants were more likely to rate unattractive Black defendants as guilty compared to unattractive White, or attractive Black or White defendants. These findings suggest White participants were likely to have an in-group bias towards racial majority members, and an attractiveness bias towards racial minority members.

Race and Gender. Racial bias has also been found in mock juror studies of juvenile defendants. Stevenson and Bottoms (2009) considered the effect of race of juvenile defendants with male and female participants. They found that men were likelier to give guilty verdicts for murder when the defendant was Black rather than White. The opposite was true for women who were more likely to give guilty verdicts for murder when the defendant was White rather than Black. However, women were more likely to give longer sentences for murder when the defendant was Black rather than White. These findings suggest it is not only the gender of the defendant that influences mock-juror ratings, but also the gender of the participant.

Race, Gender, and Attractiveness. As described above, attractiveness-based biases have been found to influence mock jurors across race and gender. This effect may be more salient for within-race evaluation compared to between-race evaluation. Agthe and colleagues (2016) found this to be true with a sample of 2557 White university students in Germany who

were asked to evaluate the attractiveness-based social evaluation biases of targets from five racial backgrounds (i.e., White, Middle Eastern, Asian, Latin American, and Black). Across races, the participants perceived the attractive targets as better looking compared to the less attractive targets. Female participants provided more positive attributes to attractive male targets than unattractive male targets, and less positive attributes to attractive female targets than unattractive female targets. Conversely, male participants provided more positive attributes to attractive female targets than unattractive female targets, and less positive attributes to attractive male targets than unattractive male targets. However, these findings were only significant when participants were evaluating White targets, indicating attractiveness may not be as salient as race when cross-race evaluations are being made.

Taken with the findings of Abwender and Hough (2001) and Cothran et al. (2017) it seems worthwhile to investigate the effects of attractiveness, race, and gender together in the context of judgments of guiltiness, given their isolated effects inside the legal context and their combined effects outside of the legal context.

Present Study

The purpose of the third study was to replicate previous research showing that minority participants treat minority defendants with more leniency. The purpose was also to extend that research by evaluating attractiveness and gender with consideration of race. A 2 (attractiveness) x 2 (gender) x 2 (race) factorial design will be employed for this purpose. As in studies one and two, measures of socially desirable responding, legal authoritarianism, and participants' past experiences with the criminal justice system were included as covariates. A measure of racial prejudice was also included as a covariate.

Hypotheses

Based on the reviewed literature, five hypotheses were generated. The first three consider the main effects of attractiveness (high or low), race (White or Black), and gender (male or female). The third and fourth consider interactions between the variables. See Table 4.1 for a summary of these hypotheses, the empirical basis for each, and the intended data-analytic strategies to test them.

1. It is hypothesized that attractive defendants will be perceived as not guilty more often and less guilty than unattractive defendants.
2. It is hypothesized that female defendants will be perceived as not guilty more often and less guilty than male defendants.
3. It is hypothesized that minority participants will be more lenient to Black defendants compared to White defendants.
4. It is hypothesized that attractive opposite-gender defendants will be perceived as not guilty more often and less guilty compared to all other defendants.
5. It is hypothesized that attractive opposite gender defendants of other races will be perceived as not guilty more often and less guilty compared to attractive same gender defendants of other races.

Table 4.1

Hypotheses, Justification, and Corresponding Data-Analytic Plan

Hypothesis	Empirical Basis	Data-Analytic Strategy
1. Attractive defendants will be perceived as not guilty more often and less guilty than unattractive defendants.	Several studies have shown an attractiveness-based leniency effect, where more attractive defendants are perceived more positively, less guilty, and less deserving of punishment (Abwender & Hough, 2001; Esses & Webster, 1988; Putz et al., 2016; Rice et al., 2020)	Chi-square analysis and ANOVA will be used to assess the dichotomous and Likert-type rating for guilt for the attractive defendants and unattractive defendants. An ANCOVA will be used to adjust for the covariates.

2. Female defendants will be perceived as not guilty more often and less guilty than male defendants.	Several studies have shown a gender-based leniency effect, where female defendants are rated as guilty less often and given lighter sentences compared to their male counterparts (Meaux et al., 2018; Rodriguez et al., 2006	Chi-square analysis and ANOVA will be used to assess the dichotomous and Likert-type rating for guilt for the male and female defendants. An ANCOVA will be used to adjust for the covariates.
3. Minority participants will be more lenient to Black defendants compared to White defendants.	Studies have shown the impact of race on mock-juror decision-making, such that there is a racial in-group bias (Mitchell et al., 2005).	Chi-square analysis and ANOVA will be used to assess the dichotomous and Likert-type rating for guilt across psychiatric diagnoses. An ANCOVA will be used to adjust for the covariates.
4. Attractive opposite-gender defendants will be perceived as not guilty more often and less guilty compared to all other defendants.	Research has shown that male and female participants attribute more positive attributes to opposite-gender attractive targets compared to unattractive or same gender targets (Agthe et al., 2016).	An ANOVA will be used to assess the effect of mock juror's gender on guiltiness ratings with consideration of the defendant's gender and attractiveness. An ANCOVA will be used to adjust for the covariates.
5. Attractive opposite gender defendants of other races will be perceived as not guilty more often and less guilty compared to attractive same gender defendants of other races.	Research has shown an in-group bias when White participants evaluate unattractive defendants of same and other races and an attractiveness bias when evaluating attractive defendants of other races (Cothran et al., 2017).	An ANOVA will be used to assess the effect of mock juror's gender on guiltiness ratings with consideration of the defendant's gender, attractiveness, and race. An ANCOVA will be used to adjust for the covariates.

METHODS

Participants

An a priori power analysis was conducted using G*Power version 3.1.9.6 (Faul et al., 2007) to determine the minimum sample size required to test the hypotheses. In order to detect a minimum effect size of a partial $\eta = 0.05$ for a 2x2x2 factorial ANOVA, a sample size of 152 participants was needed for the current study. Two hundred and five undergraduates from the University of Windsor participated in this study. Eleven cases that failed 50% or more of the two validity checks and transcript-related questions were removed from the dataset. Also removed were five duplicate cases. This left 194 women, and one nonbinary participant. The mean age of the sample was 21.11 years ($SD = 3.53$). The majority of participants self-identified as

heterosexual (82.1%). The majority of participants self-identified as Caucasian (63.1%). Other ethnicities representing ≥ 5% were Black (5.1%), South Asian (5.6%), and Middle Eastern (20%).

Measures

The measures across all studies were consistent, with two exceptions: the short summary of a court case and a measure of racial bias (described below). The manipulated variables in the summary for study three were attractiveness, gender, and race. Please see the descriptions of the measures used across all the studies above, and the short summary of a court case and racial bias measure below. The internal consistency of the BIDR-16 was good (α = .83). The internal consistency of the RLAQ23 was poor (α = .59). The internal consistency of the ECJS was low (α = .58). The internal consistency was very low for the Contact scale (α = .47) and the Hassled scale (α = .49).

Short Transcript of a Court Case

The short transcript for this study was amended, with permission, from Blais and Forth (2014). The summary described an assault perpetrated by an acquaintance, in which the victim claimed the defendant stabbed them after an altercation during a card game. The defendant claimed they acted in self-defence after being choked by the victim. An expert witness provided testimony that their impressions were that the defendant was of sound mind. Finally, the judge provided instructions for reaching a verdict with consideration of the meaning of reasonable doubt. Participants were randomly assigned to read one of eight short court transcripts in which defendants attractiveness (attractive or unattractive), gender (male or female), and race (Black or White) were manipulated. Following the reading of the court summary, participants were given a short questionnaire to evaluate their understanding of the case (see Appendices A3 and B).

Racial Bias

The *Bayesian Racism Scale* (Uhhmann et al., 2010) is a measure of discriminatory attitudes against others based on racial stereotyping. It is a 6-item measure rated on a 7-point Likert-type scale from 1 (Strongly Disagree) to 7 (Strongly Agree). Items include, "If your personal safety is at stake, it's sensible to avoid members of ethnic groups known to behave more aggressively." It has been found to have adequate reliability, $\alpha = .74$, and convergent validity when compared with measures of prejudice and rational thinking (Social Dominance Orientation, $r = .55$, belief that stereotyping is rational, $r = .53$, and finding humour in racist jokes, $r = .41$; Uhhman et al., 2010). For the present study, internal consistency was found to be low ($\alpha = .68$).

Procedure

The procedure was consistent across all three studies (see description given for Study 1, above).

RESULTS

Statistical analyses were run with SPSS Version 28. Data were checked for missing data then assumptions for chi-square analysis, ANOVA, and ANCOVA were run.

Missing Values Analysis

Little's MCAR test indicated that missing data were missing completely at random; $\chi^2(159) = 156.72$, $p = .536$.

Attractiveness Manipulation

To evaluate the effectiveness of the attractiveness manipulation, participants were asked to rate the attractiveness of the defendant on a 7-point Likert-type scale from 1 (very

unattractive) to 7 (very attractive). For females and White males, the unattractive images scores were given lower scores than the attractive images, although not as disparate as mean scores in the Chicago Face Database. For Black males, the scores were approximately equal for the unattractive and attractive images. Descriptive statistics can be seen in Table 4.2.

Table 4.2

Descriptive Statistics of Attractiveness Ratings of Defendant Photographs

	M	*SD*	Range	CFD *M*
White				
Unattractive male	2.96	1.33	1-5	2.65
Attractive male	3.69	1.04	1-5	4.12
Unattractive female	3.26	1.26	1-5	2.68
Attractive female	4.27	1.04	1-6	5.09
Black				
Unattractive male	3.54	1.23	1-6	2.91
Attractive male	3.69	.95	1-5	4.56
Unattractive female	3.45	1.10	1-5	2.66
Attractive female	4.67	.86	3-6	4.78

Verdict

Participants were randomly assigned to one of eight vignettes that manipulated attractiveness, gender, and crime type. Descriptive statistics can be seen in Table 4.3.

Table 4.3

Vignette Classifications and Sample Sizes

	N	%	White participants	Non-White participants
Attractive Black female	21	10.8	14	7
Attractive White female	27	13.8	19	8
Unattractive Black female	22	11.3	12	10
Unattractive White female	27	13.8	14	13
Attractive Black male	17	8.7	9	8
Attractive White male	29	14.9	16	13
Unattractive Black male	28	14.4	20	8
Unattractive White male	24	12.3	14	10

Overall, 49.2% of participants gave guilty verdicts. A series of chi-square tests of independence were conducted to ascertain if attractiveness, gender, and race were associated with differences in guilty or not guilty verdicts (see Table 4.4). Key assumptions were met for the individual chi-square tests comparing attractiveness, gender, and race. There was not a statistically significant association between attractiveness and guilty/not guilty verdicts, $\chi^2(1) = .61$, $p = .435$. There was not a statistically significant association between gender and guilty/not guilty verdicts, $\chi^2(1) = .87$, $p = .352$. There was not a statistically significant association between race and guilty/not guilty verdicts, $\chi^2(1)$ 3.31, $p = .069$.

Table 4.4

Frequencies and Chi-Square Results for Guilty and Not Guilty Verdict Associations with Attractiveness, Gender, and Race (N = 195)

Transcript Condition	Yes		No		$\chi^2(1)$
	n	%	*n*	%	
Attractiveness	96	49.2	99	50.8	.61
Attractive	49	52.1	45	47.7	
Unattractive	47	46.5	54	53.5	
Gender	96	49.2	99	50.8	.87
Female	51	52.6	46	47.4	
Male	45	45.9	53	54.1	
Race	96	49.2	99	50.8	3.31
White	59	55.1	48	44.9	
Black	37	42.0	51	58	

Participant Perceptions of the Defendant

A series of ANOVAs and ANCOVAs were conducted to determine if levels of guiltiness varied as a function of attractiveness, gender, and race. Exploratory analyses of participants confidence in ratings, and defendant and victim credibility can be found in Appendix F. Given the absence of hypotheses and in the interests of parsimony, exploratory analyses were restricted to ANCOVAs and pertinent post hoc testing. All of the ANOVAs and ANCOVAs evaluated

mean differences of scores based on attractiveness, gender, race, attractiveness and gender, attractiveness and race, gender and race, and attractiveness, race, and gender.

Level of Guiltiness

Hypothesis 1: Attractive Defendants Will Be Perceived As Not Guilty More Often And Less Guilty Than Unattractive Defendants. To evaluate mean differences in ratings of guiltiness based on attractiveness a one-way ANOVA was conducted. Key assumptions were met, except for normality. Visual inspection of histograms and normal Q-Q plots showed approximately normal distributions. Inconsistent with hypothesis one, there was not a statistically significant difference in ratings of guiltiness by attractiveness (see Table 4.5). This remained true when controlling for the RLAQ23 and BRS, $F(1, 172) = .39$, $p = .536$, $\eta^2 = .00$ (see Table 4.6 for means, adjusted means, and variances of guiltiness ratings).

Table 4.5

Means, Standard Deviations, and One-Way Analysis of Variance in Ratings of Guiltiness based on Attractiveness

	N	*M*	*SD*	*F*(1, 192)	η^2
Attractiveness				.30	.00
Attractive	93	3.94	1.60		
Unattractive	101	4.06	1.55		

Table 4.6

Means, Adjusted Means, and Variances of Ratings of Guiltiness Based on Attractiveness After Controlling for the RLAQ, and BRS

	M	*SD*	M_{adj}	*SE*
Unattractive	4.03	1.53	4.01	.16
Attractive	3.85	1.56	3.87	.17

Hypothesis 2: Female Defendants Will Be Perceived As Not Guilty More Often And Less Guilty Than Male Defendants. To evaluate mean differences in ratings of guiltiness based on gender a one-way ANOVA was conducted. Key assumptions were met, except for normality.

Visual inspection of histograms and normal Q-Q plots showed approximately normal distributions. Hypothesis 2 was not supported; there was not a statistically significant difference in ratings of guiltiness by gender (see Table 4.7). After controlling for the BIDR-16, RLAQ23 and BRS, this effect remained nonsignificant, $F(1, 165) = 2.27$, $p = .133$, $\eta^2 = .01$ (see Table 4.8 for means, adjusted means, and variances of guiltiness ratings).

Table 4.7

Means, Standard Deviations, and One-Way Analysis of Variance in Ratings of Guiltiness based on Gender

	N	*M*	*SD*	*F*(1, 192)	η^2
Gender				1.20	.01
Male	98	3.88	1.56		
Female	96	4.13	1.59		

Table 4.8

Means, Adjusted Means, and Variances of Ratings of Guiltiness Based on Gender After Controlling for the BIDR-16, RLAQ, and BRS

	M	*SD*	M_{adj}	*SE*
Male	3.75	1.51	3.74	.17
Female	4.08	1.57	4.09	.17

Hypothesis 3: Minority Participants Will Be More Lenient To Black Defendants Compared To White Defendants. A two-way ANOVA was run to determine if participant race and defendant race had an effect on ratings of guiltiness. Key assumptions were met, except for normality, as assessed by a significant Shapiro-Wilk test. However, visual inspection of histograms and normal Q-Q plots showed approximately normal distributions. Inconsistent with hypothesis 3, there was not a statistically significant difference in ratings of guiltiness by participant and defendant race (see Table 4.9). After controlling for the BIDR-16, RLAQ, and

BRS, this effect remained nonsignificant, $F(1, 163) = .19$, $p = .662$, $\eta^2 = .00$ (see Table 4.10 for means, adjusted means, and variances of ratings of guiltiness).

Table 4.9

Means, Standard Deviations, and Two-Way ANOVA Statistics for Ratings of Guiltiness by Participant and Defendant Race

	White Participants		Non-White Participants		ANOVA		
	M	*SD*	*M*	*SD*	Effect	*F* ratio (1, 190)	η^2
					DR	1.26	.01
Black	3.98	1.72	3.70	1.49	PR	.42	.00
White	4.11	1.52	4.09	1.56	DFxPR	.33	.00

Note. DR = defendant race; PR = participant race

Table 4.10

Means, Adjusted Means and Variances for Guiltiness Ratings with the BIDR-16, RLAQ23 and BRS as Covariates

	M	*SD*	*M*	*SE*
White participant				
Black defendant	3.98	1.68	3.99	.22
White defendant	4.00	1.51	4.03	.20
Non-White participant				
Black defendant	3.58	1.47	3.60	.30
White defendant	3.95	1.51	3.86	.26

Hypothesis 4: Attractive Opposite-Gender Defendants Will Be Perceived As Not Guilty More Often And Less Guilty Compared To All Other Defendants. A two-way ANOVA was run to determine if participant gender and defendant gender and attractiveness had an effect on ratings of guiltiness. Key assumptions were met, except for normality, as assessed by a significant Shapiro-Wilk test. However, visual inspection of histograms and normal Q-Q plots showed approximately normal distributions. Inconsistent with hypothesis four, there was not a significant difference in ratings of guiltiness by participant gender and defendant gender and attractiveness (see Table 4.11). After controlling for the BIDR-16 and BRS, this effect

remained nonsignificant, $F(1, 179) = 22$, $p = .639$, $\eta^2 = .00$ (see Table 4.12 for means, adjusted means and variances of guiltiness ratings).

Table 4.11

Means, Standard Deviations, and Two-Way ANOVA Statistics for Ratings of Guiltiness by Participant Gender and Defendant Gender and Attractiveness

	Unattractive		Attractive			ANOVA	
	M	*SD*	*M*	*SD*	Effect	*F* ratio (1, 190)	η^2
					A	.32	.00
Male	4.04	1.62	3.70	1.47	G	1.31	.01
Female	4.08	1.50	4.17	1.70	AxG	.91	.01

Note. A = attractiveness; G = Defendant gender

Table 4.12

Means, Adjusted Means and Variances for Confidence Ratings with the BIDR-16 and BRS as Covariates

	M	*SD*	*M*	*SE*
Unattractive				
Male	3.96	1.61	3.91	.23
Female	4.04	1.49	4.09	.24
Attractive				
Male	3.69	1.49	3.73	.24
Female	4.15	1.71	4.13	.23

Hypothesis 5: Attractive Opposite Gender Defendants Of Other Races Will Be Perceived As Not Guilty More Often And Less Guilty Compared To Attractive Same Gender Defendants Of Other Races. A three-way ANOVA was conducted to determine if attractiveness, gender, and race of the defendant and gender and race of the participants had an effect on guiltiness ratings. Key assumptions were met, except for normality, as assessed by a significant Shapiro-Wilk test. However, visual inspection of histograms and normal Q-Q plots showed approximately normal distributions. Inconsistent with hypothesis five, there was not a

significant difference in ratings of guiltiness by participant race and gender and defendant race, gender, and attractiveness (see Table 4.13).

Table 4.13

Means, Standard Deviations, and Two-Way ANOVA Statistics for Ratings of Guiltiness by Participant and Defendant Race and Gender

	White Participant		Non-White Participant		ANOVA	
Defendant	*M*	*SD*	*M*	*SD*	*F* ratio (1, 178)	η^2
Attractive						
Black Male	3.44	2.07	3.50	1.20	.01	.00
White Male	3.81	1.42	3.85	1.35	.00	.00
Black Female	4.43	1.79	3.43	1.27	1.84	.01
White Female	4.00	1.76	4.86	1.68	1.48	.01
Unattractive						
Black Male	3.65	1.42	3.75	2.25	.02	.00
White Male	4.64	1.15	4.20	1.93	.45	.00
Black Female	4.42	1.78	4.00	1.25	.37	.00
White Female	4.07	1.59	3.85	1.41	.14	.00

A three-way ANCOVA was run to determine if defendant attractiveness, gender, and race, and participant race and gender had a significant effect on ratings of guiltiness after controlling for the BIDR-16, RLAQ, and BRS. The BIDR-16 and RLAQ violated the assumption of linearity and were omitted from the analysis. All other key assumptions were met. After controlling for the BRS, there was not a significant difference in ratings of guiltiness based on defendant attractiveness, gender, and race, and participant race and gender (see Table 4.14).

Table 4.14

Three-Way ANCOVA for Guiltiness Ratings with the BRS as Covariate

	F ratio	*df* (174)	Partial η^2
Participant race	1.01	1	.01
Defendant race	2.35	3	.01
Gender and attractiveness	.67	1	.01
Participant race*defendant race	.74	3	.00
Participant race*gender and attractiveness	.27	3	.01
Defendant race*gender and attractiveness	.49	3	.01
Attractiveness* Gender and race * Participant race	1.03	3	.02

DISCUSSION

Guilt

This study examined whether defendant and participant gender and race, and defendant attractiveness influenced participant perceptions of guilt. The findings for the dichotomous guilty/not guilty verdict and the Likert-type ratings of guiltiness were nonsignificant based on these variables. Although the findings did not reach statistical significance, the pattern of results were mostly in alignment with hypotheses, but there were 60.5% White participants and 39.5% non-White participants, which could have affected the power. Additionally, small effects were found for the interaction of defendant and participant gender and defendant race and attractiveness.

Race

Non-White participants considered Black defendants to be less guilty than White defendants and less guilty than White participants considered Black defendants. However, White participants also considered Black defendants to be less guilty than White defendants. These findings are in alignment with Cothran et al. (2017), replicating and extending their findings by including participants of other minority groups and including female defendants. Race was made salient by the inclusion of images and identifying the defendant as a White or Black male or female. These results may indicate overcompensating for racial bias, or awareness of the systemic issues that Black people face, leading to leniency. Attending university has been found to be associated with increased culturally liberal values (Apfeld et al., 2022), which may increase students' exposure to social justice movements, such as a Black Lives Matter. These findings

make sense in the context of more liberal cultural values being associated with less Black-based leniency. Salerno et al. (2023) found that when liberal leaning mock-jurors had an opportunity to acquit Black defendants across two studies, they had similarly lenient sentencing practices. Conservative-leaning mock jurors gave harsher penalties to the second Black defendant, after 'proving' they were not racist by acquitting a Black defendant in the first study. They also found that more conservative mock jurors reported having more racial prejudice and less motivation to control those attitudes then did more liberal mock jurors. Furthermore, they found that more liberal mock jurors were more likely to convict White defendants than Black defendants, especially after first acquitting a Black defendant. Psychology is a field that is known for progressive attitudes (Redding, 2023), and liberal psychology students have been shown to outnumber conservative psychology students two to one (Maranto et al., 2023). Within this context, it makes sense that there may be a higher rate of race-based leniency for Black defendants in more recent mock-jury research with undergraduates.

Attractiveness and Gender

When comparing participant gender to attractiveness and gender of the defendants, participants considered attractive males to be the least guilty. This was consistent with our expectations, that opposite gender attractive people would be considered the least guilty. Additionally, attractive females were considered the most guilty, which was also expected as demonstrated in Agthe et al. (2016). These findings must be qualified in the context of not having male participants, which may have affected the non-significant findings.

Consistent with Cothran et al. (2017), White participants in this study considered attractive Black males to be the least guilty, followed by unattractive Black males and attractive White males. However, they considered unattractive White males to be the most guilty. These

findings are consistent with the supposition that university students' culturally liberal values lead them towards leniency with Black males, while the attractiveness bias also influences perceptions of guilt (Salerno et al., 2023; Rice et al., 2020). Attractive and unattractive Black females were considered the guiltiest, suggesting there is an opposite gender leniency effect, and a possible racial bias for same-gender defendants.

Exploratory Analyses

Confidence

Exploratory analyses were run to determine if participants' confidence in their decision-making, defendant credibility and victim credibility varied as a function of attractiveness, gender, and race while controlling for SDR, authoritarian legal attitudes and racial prejudice (see Appendix F). White participants' confidence in their decision-making was significantly higher when they had a transcript with a Black defendant compared to a White defendant (see Tables F.2 and F.3 and Figure F.1). They were also significantly more confident than non-White participants when they had a Black defendant. These findings make sense in the context of more liberal cultural values being associated with race-based leniency for Black defendants. However, there is also some evidence that when White mock jurors are presented with Black defendants, their attention to detail increases (Ewanation & Maeder, 2021). Ewanation and Maeder (2021) found that expert witness testimony was significantly more influential when pertaining to a Black defendant than a White defendant. Because the expert testimony in this case referred to a defendant who expressed remorse, it is possible it was used for decision making, and if so, the participants may have had more confidence in their understanding of the case and their decision making.

Defendant Credibility

There were no significant differences between scores for defendant credibility based on attractiveness, gender, and race; however, White participants considered unattractive Black female defendants to be significantly more credible than did non-White participants (see Table F.10 and Figure F.4). Although in this study these evaluations were race incongruent, Agthe et al. (2016) found that female participants evaluated unattractive females more positively than attractive females. It is possible this effect extended to other race targets in this study.

The exploratory analyses findings fit into the main findings as expected. There appears to be a race-based stigma that interacts with attractiveness for Black males and Black females. Although the findings did not penalize White females in the same way as Black females, Abwender and Hough's (2001) findings did indicate that female participants would be more likely to penalize unattractive women compared to attractive women, and perhaps this is especially salient when race is involved.

Implications

Studies such as this contributes to research on racial bias towards youth involved in the criminal justice. It also gives insight into how attractiveness and gender interact with racial bias to affect decision-making. The findings here do not represent the reality of the youth criminal justice system in Canada, which indicates mock jury research using undergraduate student biases may not be representative of community samples and decision-makers. Although, it is valuable to understand the implicit biases of undergraduate students, who may serve on juries. For researchers, this suggests that certain variables that affect mock juror decision-making differentially affect undergraduate versus jury-eligible community samples. Samples of convenience are a necessary and valuable part of research, and in some cases should be used for

preliminary investigations, which should be replicated with community samples to ensure the results are generalizable (Keller & Wiener, 2011).

Limitations and Future Directions

This study had several limitations. First, the examination of participant race effects on decision-making lacked power, due to the uneven samples between White and other-race participants. Additionally, lumping all of the minority participants together did not allow for culture-specific biases and attitudes to be shown. If possible, future studies should collect data from diverse backgrounds to ensure the nuances of race-based biases and attitudes are truly identified.

Another limitation of the study design was the exclusion of a First Nations defendant. Given the systemic inequities faced by Indigenous people in Canada, it is imperative that researchers investigate the implicit biases that perpetuate these inequities. Future studies should include First Nations defendants.

Other limitations of the study were the gender and race congruent defendant and victim pairings. Although this provides important information, gender and race incongruent pairings have been shown to affect perceptions of defendants (e.g., Maeder & Yamamoto, 2019; Pozzulo et al., 2009). Future research should include gender and race incongruent defendant and victim pairings.

Future studies should consider participant characteristics that may affect their decision-making, as well. Factors such as experiences of victimization or discrimination, political affiliation, religiosity, and personality traits may all affect perceptions of defendants and lead to biased decision-making.

A final limitation of the three studies is that attractiveness and gender were studied three times with a different third variable in each of the studies. Ideally, all five variables would have been examined together, as there may be significant interactions between race, crime type, and psychiatric diagnosis. The practicality of a study such as that may be difficult, given the large sample size that would be required, however, it would provide a more thorough examination of extralegal factors that influence mock-juror decision-making.

Conclusion

Study three aimed to identify the effects of attractiveness, participant and defendant gender, and participant and defendant race on mock-juror decision-making. Results were not significant for the dichotomous and Likert-type guilt ratings; however, the pattern of results was broadly consistent with hypotheses. This study partially replicated and extended Cothran et al.'s (2017) research on the effects of race and attractiveness bias on mock juror decision-making. Our study was an online study rather than in lab, we included female defendants, and considered participant gender as well. Although it is encouraging to have found leniency for Black males from White and other race participants, these preliminary findings should be replicated with other samples from the community that are gender and race balanced.

GENERAL DISCUSSION

These studies examined the effects of attractiveness, gender, crime type, psychiatric diagnosis, and race on participant decision-making using the mock juror paradigm. Despite evidence showing an effect of attractiveness on decision-making (e.g., Abwender & Hough, 2001; Esses & Webster, 1988; Putz et al., 2016; Rice et al., 2020), findings reported across the three studies indicated that attractiveness was only influential when paired with another variable. For example, attractiveness influenced ratings of guilt in relation to crime type – unattractive

defendants accused of assault were considered significantly less guilty than unattractive defendants accused of fraud. Attractiveness also influenced ratings of guilt in relation to psychiatric diagnosis. Attractive defendants with psychopathic traits were considered significantly more guilty than attractive defendants with schizophrenia. These findings were inconsistent with our expectations; however, in light of the expert witness testimony, perhaps they were consistent with Yang et al. (2019). It is possible participants believed the attractive male with psychopathic traits exploited his looks to manipulate others.

The attractiveness manipulation did not appear to have been as effective as had been intended, based on mean ratings of attractiveness for the samples and those from the standardized CFD normative data. However, because attractiveness did seem to interact with other variables, it is not tenable to claim that the attractiveness manipulation was wholly ineffective. Additionally, the photos were shared before and after the trial transcript. It is possible that the participants' perception of attractiveness was altered after reading the transcript and the description of the crime committed (Niimi & Goto, 2023).

These studies were not without their limitations. This study lacked ecological validity due to the use of undergraduate students (although many would be expected to have been eligible to serve on a jury); it did not include a community sample or a sample from a jury pool. Although some research has demonstrated undergraduate student samples are representative of community samples for ratings of guilt and sentencing severity (Bornstein et al., 2017), other research has shown leniency effects for undergraduates compared to community samples and poor comprehension in community samples (Keller & Wiener, 2011). Additionally, because recent mock jury studies with undergraduate students have shown a race-based leniency effect for minority defendants (Cothran et al., 2017; Fraser et al., 2023; Roberts & Maeder, 2023), studies like this help identify potential limitations of undergraduate student samples. Future research

should use comparison groups to identify where the disparity lies between the reality of the criminal justice system, which incarcerates visible minorities disproportionately, and mock juror attitudes. Because of these differences, undergraduate student research is an effective first step but should be followed up with community samples.

Another limitation of this study was absence of images in the adolescent age range in the Chicago Face Database for images of youth. To manage this, faces were selected from this standardized database for which normative data on youthfulness ("baby-facedness') indicated an approximately appropriate age range.

As above, having a female only sample was a notable shortcoming in all three studies, although not a planned or intended one. Research has shown gender-based differences in decision making, which a future study should investigate. Another limitation was that the studies were split into three. Several studies have examined the combined effects of race and psychiatric diagnosis (Abwender & Hough, 2001; Maeder et al., 2020; Perry et al., 2013). The null findings regarding race in study 3 suggest the sample of women in this study did not have implicit biases regarding guilt associated with attractiveness, gender and race. It is possible there may be race-based effects found when race is combined with psychiatric diagnosis or crime type. Future research should include all of the extralegal factors that may be present when youth come into contact with the criminal justice system.

Conclusion

The current set of studies aimed to identify the effects of attractiveness, gender, crime type, psychiatric diagnosis and race on mock-juror decision-making regarding guilt and perceptions of the defendant and victim. It was expected that attractive defendants and female defendants would be perceived as less guilty compared to all other defendants. In study one it

was also expected that attractive defendants who had committed fraud to be perceived as more guilty than all other defendants. Contrary to these expectations, attractiveness and gender in isolation did not influence guilt. Although attractiveness and crime type did influence ratings of guilt, attractive defendants accused of fraud were not considered guiltier than other defendants, but unattractive defendants were. This research shows evidence of the beauty bias as something that may affect youth involved with the criminal justice system. Additionally, although attractiveness did not affect guilt ratings, it did have an effect on defendant credibility. This indicates that perhaps attractiveness affects the perceptions of traits of people accused of a crime more strongly than it affects perceptions of people's behaviour.

In study two, it was expected that defendants with schizophrenia would be perceived as guilty less often and less guilty than the rest and defendants with psychopathic traits to be perceived as guilty more often and more guilty than the rest. Attractiveness and gender, in isolation, did not affect guilty verdicts. However, attractive defendants with a psychiatric diagnosis would be perceived as less guilty than unattractive defendants with a psychiatric diagnosis. Contrary to expectations, attractive defendants with psychopathic traits were considered the guiltiest. These findings indicate the beauty penalty may be especially salient when good looks can be used to manipulate others. The effect of psychiatric diagnosis was statistically significant in several of the exploratory analyses, and this remained true when attractiveness and psychiatric diagnosis were considered together; however, it was primarily the no diagnosis condition that produced significant effects. Participants' ratings of attractive defendants with no diagnosis showed that these defendants were considered the least likely to benefit from treatment, but also to have the lowest risk for future violence and to not be a threat to society. This suggests the positive traits that are associated with attractiveness remain salient when there are no other factors interfering with decision-making.

In study three, it was expected that non-White participants would be more lenient with Black defendants than White defendants would be, however, the results were not statistically significant, and White participants were most lenient with Black males. Although both White and non-White participants perceived the unattractive White male to be most guilty of the unattractive defendants, White participants also gave similar mean scores to Black females, attractiveness notwithstanding. Non-White participants considered the attractive White female to be the most guilty. These findings suggest that for White participants, there may be a gender bias that is made salient when race is a factor. For non-White participants, this bias may be dependent on attractiveness, as well. It is encouraging that all the participants in this study considered Black males to be the least guilty, as those youth are overrepresented in the criminal justice system. However, the findings must be considered in light of their non significance and should be replicated with equal race and gender sample sizes.

Taken together the findings of these three studies prompt further examination into the factors that influence mock-juror decision making, and the complexity of their interactions. These studies are an important contribution to the research on mock-juror decision-making, particularly as it pertains to biases against justice-involved youth.

REFERENCES

Abwender, D. A., & Hough, K. (2001). Interactive effects of characteristics of defendant and mock juror on U. S. participants' judgment and sentencing recommendations. *The Journal of Social Psychology, 141*(5), 603–615. https://doi.org/10.1080/00224540109600574

Agthe, M., Strobel, M., Spörrle, M., Pfundmair, M., & Maner, J. K. (2016). On the borders of harmful and helpful beauty biases: The biasing effects of physical attractiveness depend on sex and ethnicity. *Evolutionary Psychology, 14*(2), Article 1474704916653968. https://doi.org/10.1177/1474704916653968

Ahola, A. S., Hellström, Å., & Christianson, S. Å. (2010). Is justice really blind? Effects of crime descriptions, defendant gender and appearance, and legal practitioner gender on sentences and defendant evaluations in a mock trial. *Psychiatry, Psychology, and Law, 17*(2), 304–324. https://doi.org/10.1080/13218710903566896

American Psychiatric Association. (2022). *Diagnostic and statistical manual of mental disorders* (5th ed., text rev.). https://doi.org/10.1176/appi.books.9780890425787

Apfeld, B., Coman, E., Gerring, J., & Jessee, S. (2024). The impact of university attendance on partisanship. *Political Science Research and Methods, 12*(1), 45–58. doi:10.1017/psrm.2022.33

Arendt, F., & Northup, T. (2015). Effects of long-term exposure to news stereotypes on implicit and explicit attitudes. *International Journal of Communication*, *9*(1), 2370-2390.

Ashkar, P. J., & Kenny, D. T. (2008). Views from the inside: Young offenders' subjective experiences of incarceration (Review). *International Journal of Offender Therapy and Comparative Criminology 52*(5). https://doi.org/10.1177/0306624X08314181

Austin, R. L., & Allen, M. D. (2000). Racial disparity in arrest rates as an explanation of racial disparity in commitment to Pennsylvania's prisons. *The Journal of Research in Crime and Delinquency, 37*(2), 200–220. https://doi.org/10.1177/0022427800037002003

Bala, N. C., Carrington, P. J., & Roberts, J. V. (2009). Evaluating the Youth Criminal Justice Act after five years: A qualified success. *Canadian Journal of Criminology and Criminal Justice*, *51*(2), 131–167. https://doi.org/10.3138/cjccj.51.2.131

Beaver, K. M., Boccio, C., Smith, S., & Ferguson, C. J. (2019). Physical attractiveness and criminal justice processing: Results from a longitudinal sample of youth and young adults. *Psychiatry, Psychology and Law, 26*(4), 669-681. https://doi.org/10.1080/13218719.2019.1618750

Berryessa, C. M., Milner, L. C., Garrison, N. A., & Cho, M. K. (2015). Impact of psychiatric information on potential jurors in evaluating high-functioning autism spectrum disorder (hfASD). *Journal Of Mental Health Research In Intellectual Disabilities, 8*(3-4), 140–167. https://doi.org/10.1080/19315864.2015.1040176

Blais, J., & Forth, A. E. (2014). Potential labeling effects: influence of psychopathy diagnosis, defendant age, and defendant gender on mock jurors' decisions. *Psychology, Crime & Law*, *20*(2), 116–134. https://doi.org/10.1080/1068316X.2012.749473

Bleidorn, W., Hopwood, C. J., Back, M. D., Denissen, J. J. A., Hennecke, M., Hill, P. L., Jokela, M., Kandler, C., Lucas, R. E., Luhmann, M., Orth, U., Roberts, B. W., Wagner, J., Wrzus, C., & Zimmermann, J. (2021). Personality trait stability and change. *Personality Science*, *2*, 1-20. https://doi.org/10.5964/ps.6009

Boccaccini, M. T., Murrie, D. C., Clark, J. W., & Cornell, D. G. (2008). Describing, diagnosing, and naming psychopathy: how do youth psychopathy labels influence jurors? *Behavioral Sciences & the Law*, *26*(4), 487–510. https://doi.org/10.1002/bsl.821

Boduszek, D., Debowska, A., Sherretts, N., Willmott, D., Boulton, M., Kielkiewicz, K., Popiolek, K. & Hyland, P. (2019). Are prisoners more psychopathic than non-forensic populations? Profiling psychopathic traits among prisoners, community adults, university students, and adolescents. *Deviant Behavior, 42*(2), 232–244. https://doi.org/10.1080/01639625.2019.1665221

Bonner, R. L. (2006). Stressful segregation housing and psychosocial vulnerability in prison suicide ideators. *Suicide & Life-Threatening Behavior, 36*(2). https://doi.org/10.1521/suli.2006.36.2.250

Bornstein, B. H., Golding, J. M., Neuschatz, J., Kimbrough, C., Reed, K., Magyarics, C., & Luecht, K. (2017). Mock juror sampling issues in jury simulation research: A meta-analysis. *Law and Human Behavior, 41*(1), 13-28. http://dx.doi.org/10.1037/lhb0000223

Breheney, C., Groscup, J., & Galietta, M. (2007). Gender matters in the insanity defense. *Law & Psychology Review, 31*, 93-123.

Bruce, M., & Weinraub, D. (2023). Implicit gender bias in the clinical judgment of psychopathy and personality disorders among licensed psychologists in the USA. *Journal of Personality Assessment, ahead-of-print*(ahead-of-print), 1–7. https://doi.org/10.1080/00223891.2023.2178928

Canadian Charter of Rights and Freedoms, Part 1 of the Constitution Act, 1982, being Schedule B to the Canada Act 1982 (UK), 1982, c 11.

Cann, A., Siegfried, W. D. & Pearce, L. (1981). Forced attention to specific application qualifications: Impact on physical attractiveness and sex of applicant biases. Personnel Psychology, 24(1), 65-75. https://doi.org/10.1111/j.1744-6570.1981.tb02178.x

Cothran, D. L., Stepanova, E. V., & Barlow, K. R. (2017). Studying guilt perception in millennials: Unexpected effects of suspects' race and attractiveness. *Imagination, Cognition and Personality, 36*(4), 379–399. https://doi.org/10.1177/0276236617696718

Criminal Code, RSC 1985, c C – 46.

Curley, L. J., Munro, J. & Dror, I. E. (2022). Cognitive and human factors in legal layperson decision making: Sources of bias in juror decision making. *Medicine, Science and the Law, 62*(3):206-215. https://doi.org/10.1177/00258024221080655

Denomme, W. J., Curno, J., & Forth, A. (2020). Psychopathic traits, risk and protective factors, and attractiveness in forensic psychiatric patients: Their role in review board dispositions. *Journal of Forensic Psychology Research and Practice, 20*(3), 264–289. https://doi.org/10.1080/24732850.2020.1717904

Department of Justice. (n. d.). *Pre-trial detention* [Fact sheet].

Doerner, J. K., & Demuth, S. (2014). Gender and sentencing in the federal courts: Are women treated more leniently? *Criminal Justice Policy Review, 25*(2), 242–269. https://doi.org/10.1177/0887403412466877

Downs, A. C., & Lyons, P. M. (1991). Natural observations of the links between attractiveness and initial legal judgments. *Personality and Social Psychology Bulletin, 17*(5), 541-547. https://doi-org/10.1177/0146167291175009

Drerup, L. C., Croysdale, A., & Hoffmann, N. G. (2008). Patterns of behavioral health conditions among adolescents in a juvenile justice system. *Professional Psychology: Research and Practice*, *39*(2), 122–128. https://doi.org/10.1037/0735-7028.39.2.122

Dyck, H. L., Campbell, M. A., Schmidt, F., & Wershler, J. L. (2013). Youth psychopathic traits and their impact on long-term criminal offending trajectories. *Youth Violence and Juvenile Justice, 11*(3), 230-248. https://doi.org/10.1177/1541204012469414

Esses, V. M., & Webster, C. D. (1988). Physical attractiveness, dangerousness, and the Canadian Criminal Code. *Journal of Applied Social Psychology, 18*(12, Pt 2), 1017–1031. https://doi.org/10.1111/j.1559-1816.1988.tb01190.x

Farnworth, M., & Teske, R. (1995). Gender differences in felony court processing: three hypotheses of disparity. *Women and Criminal Justice, 6*(2), 23-44.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175-191.

Feng, Y., Chen, H., & He, L. (2019). Consumer responses to femvertising: A data-mining case of Dove's "campaign for real beauty" on YouTube. *Journal of Advertising, 48*(3), 292–301. https://doi.org/10.1080/00913367.2019.1602858

Franklin, C. A., & Fearn, N. E. (2008). Gender, race, and formal court decision-making outcomes: Chivalry/paternalism, conflict theory or gender conflict? *Journal of Criminal Justice, 36*(3), 279-290. https://doi.org/10.1016/j.jcrimjus.2008.04.009

Fraser, B. M., Pica, E., & Pozzulo, J. D. (2023). Mock-jurors' judgements in a sexual assault case: the influence of defendant race and occupational status, delayed reporting, and multiple allegations. *Journal of Interpersonal Violence, 38*(13-14), 7964-7989. https://doi-org/10.1177/08862605231153873

Freedman, J. L., Krismer, K., MacDonald, J. E., & Cunningham, J. A. (1994). Severity of penalty, seriousness of the charge, and mock jurors' verdicts. *Law and Human Behavior, 18*(2), 189-202. doi:https://doi.org/10.1007/BF01499015

Furnham, A. (2003). Belief in a just world: Research progress over the past decade. *Personality and Individual Differences, 34*(5), 795-817. https://doi.org/10.1016/S0191-8869(02)00072-7

Gabor, T. (2015). Costs of crime and criminal justice responses. *Ottawa: Public Safety Canada,* c2015. Publications.gc.ca/pub?id=9.803888&sl=0

Grier, S. A., & Deshpandé, R. (2001). Social dimensions of consumer distinctiveness: The influence of social status on group identity and advertising persuasion. *Journal of Marketing Research, 38*(2), 216–224. https://doi.org/10.1509/jmkr.38.2.216.18843

Hart, C. M., Ritchie, T. D., Hepper, E. G., & Gebauer, J. E. (2015). The Balanced Inventory of Desirable Responding Short Form (BIDR-16). *SAGE Open, 5*(4), 1-9. https://doi.org/10.1177/2158244015621113

Henderson, M. L., Cullen, F. T., Cao, L., Browning, S. L., & Kopache, R. (1997). The impact of race on perceptions of criminal injustice. *Journal of Criminal Justice, 25*(6), 447-462. https://dx.doi.org/10.1016/S0047-2352(97)00032-9

Hester, R., & Hartman, T. K. (2017). Conditional race disparities in criminal sentencing: A test of the liberation hypothesis from a non-guidelines state. *Journal of Quantitative Criminology, 33*(1), 77–100. https://doi.org/10.1007/s10940-016-9283-z

Hunter, E.A., Kluck, A.S., Ramon, A.E. Ruff, E. & Dario, J. (2021). The curvy ideal silhouette scale: Measuring cultural differences in the body shape ideals of young U.S. Women. *Sex Roles 84*, 238–251. https://doi.org/10.1007/s11199-020-01161-x

John Howard Society of Ontario. (2021). Experiences and outcomes of young people in Ontario's youth bail system. *UnEqual Justice*. https://johnhoward.on.ca/wp-content/uploads/2021/03/Youth-Bail-Highlights-Final.pdf

Jones, A. M., Jones, S., & Penrod, S. (2015). Examining legal authoritarianism in the impact of punishment severity on juror decisions. *Psychology, Crime & Law, 21*(10), 939-951.

Jones, W. H., Hansson, R. O., & Phillips, A. L. (1978). Physical attractiveness and judgments of psychopathology. *The Journal of Social Psychology, 105*(1), 79–84. https://doi.org/10.1080/00224545.1978.9924093

Kalkbrenner, M. T., James, C., & Pérez-Rojas, A. E. (2020). College students' awareness of mental disorders and resources: comparison across academic disciplines. *Journal of College Student Psychotherapy*, *36*(2), 113–134. https://doi.org/10.1080/87568225.2020.1791774

Keller, S. R., Wiener, R. L. (2011). What are we studying? Student jurors, community jurors, and construct validity. *Behavioral Sciences & The Law, 29*(3), 376-394. https://doi.org/10.1002/bsl.971

Kenny, A., Bizumic, B., &Griffiths, K. M. (2018). Prejudice Towards People with Mental Illness Scale [Database record]. Retrieved from PsycTESTS. http://dx.doi.org/10.1037/t70320-000

Kortright, K. (2019). Double stigma: How jurors perceive mentally ill defendants. [Unpublished master's thesis]. The University of Alabama

Kravitz, David A, Cutler, Brian L., & Brock, Petra. (1993). Reliability and validity of the original and Revised Legal Attitudes Questionnaire. *Law and Human Behavior, Vol 17*(6), 661-667. https://dx.doi.org/10.1007/BF01044688

Lambie, I., & Randell, I. (2013). The impact of incarceration on juvenile offenders. *Clinical Psychology Review, 33*(3). https://doi.org/10.1016/j.cpr.2013.01.007

Landy, D., & Aronson, E. (1969). The influence of the character of the criminal and his victim on the decisions of simulated jurors. *Journal of Experimental Social Psychology, 5*(2), 141–152. https://doi.org/10.1016/0022-1031(69)90043-2

Langlois J. H., Kalakanis L., Rubenstein A. J., Larson A., Hallamm M., Smoot M. 2000. Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychol. Bull*. 126, 390–423 10.1037/0033-2909.126.3.390

Lundrigan, S., Dhami, M. K., & Mueller-Johnson, K. (2016). Predicting verdicts using pre-trial attitudes and standard of proof. *Legal and Criminological Psychology., 21*(1), 95–110. https://doi.org/10.1111/lcrp.12043

Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago Face Database: A free stimulus set of faces and norming data. *Behavior Research Methods (Online), 47*(4), 1122. https://doi.org/10.3758/s13428-014-0532-5

Ma, Z. (2017). How the media cover mental illnesses: A review. *Health Education*, 117(1), 90-109. DOI:10.1108/HE-01-2016-0004

Maeder, E. M., McManus, L. A., Yamamoto, S., & McLaughlin, K. (2018). A test of gender-crime congruency on mock juror decision-making. *Cogent Psychology, 5*(1), 1461543–. https://doi.org/10.1080/23311908.2018.1461543

Maranto, R., Redding, R.E., Wai, J., Woessner, M. (2023). Does psychology's progressive ideology affect its undergraduates? A national test. In: Frisby, C.L., Redding, R.E., O'Donohue, W.T., Lilienfeld, S.O. (Eds.), *Ideological and political bias in psychology*. Springer, Cham. https://doi.org/10.1007/978-3-031-29148-7_14

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. https://doi.org/10.1111/j.1559-1816.1994.tb01552.x

McGowen, R. & King, G. D. (1982). Effects of authoritarian, anti-authoritarian, and egalitarian legal attitudes on mock juror and jury decisions. Psychological Reports, 51(3_suppl), 1067-1074.

Meaux, L. T., Cox, J., & Kopkin, M. R. (2018). Saving damsels, sentencing deviants and selective chivalry decisions: juror decision-making in an ambiguous assault case. *Psychiatry, Psychology, and Law, 25*(5), 724–736. https://doi.org/10.1080/13218719.2018.1474817

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. https://doi.org/10.1007/s10979-005-8122-9

Mossière, A., & Maeder, E. M. (2016). Juror decision making in not criminally responsible on account of mental disorder trials: Effects of defendant gender and mental illness type. *International Journal of Law and Psychiatry, 49*(Pt A), 47–54. https://doi.org/10.1016/j.ijlp.2016.05.008

Mustard, D. B. (2001). Racial, ethnic, and gender disparities in sentencing: Evidence from the U.S. federal courts. *The Journal of Law & Economics*, *44*(1), 285–314. https://doi.org/10.1086/320276

Narby, D. J., Cutler, B. L., & Moran, G. (1993). A meta-analysis of the association between authoritarianism and jurors' perceptions of defendant culpability. Journal of Applied Psychology, 78(1), 34-42. https://doi.org/10.1037/0021-9010.78.1.34

Niimi, R., & Goto, M. (2023). Good conduct makes your face attractive: The effect of personality perception on facial attractiveness judgments. PloS one, 18(2), e0281758. https://doi.org/10.1371/journal.pone.0281758

Perry, B. L., Neltner, M. & Allen, T. (2013). *A paradox of bias: Racial differences in forensic psychiatric diagnosis and determinations of criminal responsibility.* Springer Science + Business Media New York.

Pica, E., Pettalia, J., & Pozzulo, J. (2017). The influence of a defendant's chronological age, developmental age, and race on mock juror decision making. *Journal of Police and Criminal Psychology, 32*(1), 66–76. https://doi.org/10.1007/s11896-016-9201-1

Pichler, S., Kohli, C. & Granitz, N. (2021). DITTO for Gen Z: A framework for leveraging the uniqueness of the new generation. Business Horizons, 64(5), 599-610. https://doi.org/10.1016/j.bushor.2021.02.021

Pozzulo, J. D., Pettalia, J. L., Dempsey, J. L., & Gooden, Amanda. (2015). Juvenile offenders on trial: does alibi corroboration evidence and defendant age interact to influence jurors' perceptions and verdicts. Psychiatry, Psychology and Law, 22(2), 224-234.

Principe, C. P., & Langlois, J. H. (2011). Faces differing in attractiveness elicit corresponding affective responses. *Cognition and Emotion, 25*(1), 140–148. https://doi.org/10.1080/02699931003612098

Putz, Á., Palotai, R., Csertö, I., & Bereczkei, T. (2016). Beauty stereotypes in social norm enforcement. The effect of attractiveness on third-party punishment and reward. Personality and Individual Differences, 88, 230–235. https://doi.org/10.1016/j.paid.2015.09.025

Redding, R. E. (2023). Psychologists' politics. Ideological and political bias in psychology: Nature, scope, and solutions, 79-95.

Reich, C. M., Pegel, G. A., & Johnson, A. B. (2022). Are survivors of sexual assault blamed more than victims of other crimes? Journal of Interpersonal Violence, 37(19-20), NP18394-NP18416. https://doi.org/10.1177/08862605211037423

Rice, H., Murphy, C., Nolan, C., & Kelly, M. (2020). Measuring implicit attractiveness bias in the context of innocence and guilt evaluations. *International Journal of Psychology & Psychological Therapy, 20*(3), 273-285.

Roberts, A., & Maeder, E. (2023). The intersection of defendant gender and racialisation in a case of child neglect. Psychology, Crime & Law, 1–26. https://doi.org/10.1080/1068316X.2023.2222215

Rodriguez, S. F., Curry, T. R., & Lee, G. (2006). Gender differences in criminal sentencing: Do effects vary across violent, property, and drug offenses? *Social Science Quarterly, 87*(2), 318–339. https://doi.org/10.1111/j.1540-6237.2006.00383.x

Salerno, J. M., Kulak, K., Smalarz, L., Eerdmans, R. E., Lawrence, M. L., & Dao, T. (2023). The role of social desirability and establishing nonracist credentials on mock juror decisions about black defendants. Law and Human Behavior, 47(1), 100-118. doi:https://doi.org/10.1037/lhb0000496

Samuels-Wortley, K. (2019). Youthful discretion: Police selection bias in access to pre-charge diversion programs in Canada. *Race and Justice, 12*(2), 387–410. https://doi.org/10.1177/2153368719889093

Sanderson, C.A., Zanna, A.S. and Darley, J.M. (2000), Making the punishment fit the crime and the criminal: Attributions of dangerousness as a mediator of liability. *Journal of Applied Social Psychology, 30*: 1137-1159. https://doi.org/10.1111/j.1559-1816.2000.tb02514.x

Silcox J. (2022). Youth crime and depictions of youth crime in Canada: Are news depictions purely moral panic?. *Canadian Review of Sociology*, *59*(1), 96–114. https://doi.org/10.1111/cars.12370

Sommers, S. R., & Ellsworth, P. C. (2001). White juror bias: An investigation of prejudice against Black defendants in the American courtroom. *Psychology, Public Policy, and Law, 7*(1), 201–229. https://doi.org/10.1037/1076-8971.7.1.201

Statista Research Department. (2023, June 5th). Rate of youths incarcerated in provincial and territorial correctional services in Canada in fiscal years 2001 to 2022 (per 10, 000 youths) [Infographic]. Statista. https://www.statista.com/statistics/560952/rate-of-youths-in-provincial-and-territorial-correctional-services-canada/

Statistics Canada. (2022). A generational portrait of Canada's aging population from the 2021 census. https://www12.statcan.gc.ca/census-recensement/2021/as-sa/98-200-X/2021003/98-200-X2021003-eng.cfm

Statistics Canada. (2022). Adult and youth correctional statistics, 2020/2021. *The Daily.* https://www150.statcan.gc.ca/n1/daily-quotidien/220420/dq220420c-eng.htm

Statistics Canada. (2023). *Table 35-10-0006-02 Youth admissions to correctional services, by sex and age* [Data table]. https://doi.org/10.25318/3510000601-eng

Steffensmeier, D., & Demuth, S. (2006). Does gender modify the effects of race-ethnicity on criminal sanctioning? Sentences for male and female White, Black, and Hispanic defendants. *Journal of Quantitative Criminology, 22*(3), 241-261.

Stevenson, M. C., & Bottoms, B. L. (2009). Race shapes perceptions of juvenile offenders in criminal court. *Journal of Applied Social Psychology, 39*(7), 1660–1689. https://doi.org/10.1111/j.1559-1816.2009.00499.x

Sweeney, L. T., & Haney, C. (1992). The influence of race on sentencing: A meta-analytic review of experimental studies. *Behavioral Sciences & the Law, 10*(2), 179–195. https://doi.org/10.1002/bsl.2370100204

Taylor, Kaplan, T., Mulvey, P., & Miller, M. K. (2019). Perceptions of waived juvenile defendants across mental health diagnoses and demographic characteristics. *International Journal of Law and Psychiatry, 66*, 101474–101474. https://doi.org/10.1016/j.ijlp.2019.101474

Uhhmann, E. L., Brescoll, V. L., & Machery, E. (2010). The motives underlying stereotype-based discrimination against members of stigmatized groups. *Social Justice Research, 23*, 1–16. https://doi.org/10.1007/s11211-010-0110-7

Verbrugge, P. (2003). *Fetal alcohol spectrum disorder and the youth criminal justice system: A discussion paper.* Ottawa: Department of Justice Canada. Retrieved from https://www.justice.gc.ca/eng/rp-pr/cj-jp/yj-jj/rr03_yj6-rr03_jj6/rr03_yj6.pdf

Visher, C.A. (1983), Gender, police arrest decisions, and notions of chivalry. *Criminology*, 21: 5-28.

Voit, M., Weiß, M. & Hewig, J. The benefits of beauty – Individual differences in the pro-attractiveness bias in social decision making. Curr Psychol 42, 11388–11402 (2023). https://doi-org.ledproxy2.uwindsor.ca/10.1007/s12144-021-02366-3

Walker, S. C., & Herting, J. R. (2020). The impact of pretrial juvenile detention on 12-month recidivism: A matched comparison study. *Crime & Delinquency, 66*(13-14), 1865–1887. https://doi.org/10.1177/0011128720926115

Williams, M. & Holcomb, J. (2001). Racial disparity and death sentences in Ohio. *Journal of Criminal Justice*, *29*, 207-218.

Wilson, R. K., & Eckel, C. C. (2006). Judging a book by its cover: Beauty and expectations in the Trust Game. *Political Research Quarterly, 59*(2), 189–202. https://doi.org/10.1177/106591290605900202

World Health Organization. (2019). *International statistical classification of diseases and related health problems* (11th ed.). https://icd.who.int/

Wuensch, L., Castellow, W., & Moore, C. (1991). Effects of defendant attractiveness and type of crime on juridic judgment. *Journal of Social Behaviour and Personality, 6*(4), 713-724.

Yang, Q., Zhu, B., Zhang, Q., Wang, Y., Hu, R., Liu, S., & Sun, D. (2019). Effects of male defendants' attractiveness and trustworthiness on simulated judicial decisions in two different swindles. Frontiers in Psychology, 10, 2160–2160. https://doi.org/10.3389/fpsyg.2019.02160

Youth Criminal Justice Act, S.C., ch. 1 (2002) (Can.).

Young Offenders Act, R.S.C., ch. Y-1 (1985) (Can.).

Zhang, T. & Hoddenbagh, J. (2013). The costs of the youth criminal justice system 2010. *Department of Justice Canada,* J4-67/2013E-PDF. publications.gc.ca/pub?id=9.850749&sl=0

Zhao, N., Zhou, M., Shi, Y., & Zhang, J. (2015). Face attractiveness in building trust: Evidence from measurement of implicit and explicit responses. *Social Behavior and Personality, 43*(5), 855–866. https://doi.org/10.2224/sbp.2015.43.5.855

Zinger, A. & Forth, E. (1998). Psychopathy and Canadian criminal proceedings: The potential for human rights abuses. Canadian Journal of Criminology, 40(3), 237-276.

APPENDICES

Appendix A1: Short Transcript of a Court Case

Below is an amended version of the transcript used by Blais and Forth (2014). Permission has been granted to the researchers to amend and use this by the authors.

References

Blais, J., & Forth, A. E. (2014). Potential labeling effects: Influence of psychopathy diagnosis, defendant age, and defendant gender on mock jurors' decisions. *Psychology, Crime & Law, 20*(2), 116-134. doi:10.1080/1068316X.2012.749473

Court Summary

(a photo of the defendant will be displayed at the top of the screen for the entirety of the Court Summary in the Qualtrics survey)

PHOTO

Defendant: **David / Alice** Smith is a 16-year-old White **male/female.**

Victim: **George / Christine** Atwell is a 16-year-old White **male/female.**

Judge: In this alleged case of aggravated assault, **Mr./Ms. Smith** is 16 years old and has been charged with section **268 of the *Criminal Code,* aggravated assault/380 of the *Criminal Code,* fraud**. The Crown has the burden of proving that charge beyond a reasonable doubt. This means that, if you find the defendant is guilty, there cannot be any uncertainty or doubt about the evidence on which you have based your finding of guilt. It is your responsibility to listen to all the evidence and then to apply the law that I will give you at the end of the trial. The trial will begin with opening statements by the Crown and Defence attorneys. Following these statements, the Crown and Defence will present and question the parties involved, who will subsequently be cross-examined. You will also hear testimony from the court appointed clinical psychologist, Dr. Richards, who interviewed the defendant after **his/her** arrest. Please listen to the following arguments carefully. Following the testimonies, you will be asked to make a decision as to whether you find the defendant, **Mr./Ms. Smith**, guilty or not guilty of the charge.

The Crown makes their opening statements.
(ASSAULT)
Crown attorney: Mr./Ms. Smith is an acquaintance of the victim, **Mr./Ms. Atwell.** On October 15th, 2022, an altercation occurred between the defendant and the victim. **Mr./Ms. Atwell** and **Mr./Ms. Smith** were playing cards at **Mr./Ms. Atwell's** home. An argument broke out between them, and it is alleged that **Mr./Ms. Smith** pushed **Mr./Ms. Atwell** then took a pocketknife out of **his/her** waistband and stabbed the victim in the chest. **Mr./Ms. Atwell** was rushed to the hospital in an ambulance where he received several stitches following this incident. **Mr./Ms. Smith** is charged with aggravated assault, and it is your duty, as representatives of our community, to seek justice by finding **Mr./Ms. Smith** guilty.

OR

(FRAUD)
Crown attorney: Mr./Ms. Smith is an acquaintance of the victim, **Mr./Ms. Atwell.** On October 15th, 2022, **Mr./Ms. Atwell** and **Mr./Ms. Smith** were playing cards at **Mr./Ms. Atwell's** home.

An argument broke out between them, and it is alleged that **Mr./Ms. Smith** stole $1500 from **Mr./Ms. Atwell** by cheating during their card game. **Mr./Ms. Smith** is being charged with fraud, and it is your duty, as representatives of our community, to seek justice by finding **Mr./Ms. Smith** guilty.

The Defence makes their opening statement.
(ASSAULT)
Defence attorney: Mr./Ms. Smith is not guilty of this charge. Although the altercation between **Mr./Ms. Smith** and **Mr./Ms. Atwell** did occur, the details are as follows. **Mr./Ms. Atwell** and **Mr./Ms. Smith** were playing cards and **Mr./Ms. Atwell** accused **Mr./Ms. Smith** of cheating. **Mr./Ms. Smith** denied cheating and **Mr./Ms. Atwell** became angry and slapped **Mr./Ms. Smith** in the face. **Mr./Ms. Smith** pushed **Mr./Ms. Atwell** away but **Mr./Ms. Atwell** then began hitting and choking **Mr./Ms. Smith**. Fearing for **his/her** safety, **Mr./Ms. Smith** removed **his/her** pocketknife from **his/her** pocket and slashed at **Mr./Ms. Atwell**, striking **his/her** chest. **Mr./Ms. Smith** acted in self-defence. **He/She** is not guilty of this charge. I trust that you will consider all the information accordingly and find that my client is indeed innocent.

OR

(FRAUD)
Defence attorney: Mr./Ms. Smith is not guilty of this charge. Although an argument between **Mr./Ms. Smith** and **Mr./Ms. Atwell** did occur the details are as follows. **Mr./Ms. Atwell** and **Mr./Ms. Smith** were playing cards and **Mr./Ms. Smith** had won several hands equalling $500. **Mr./Ms. Atwell** accused **Mr./Ms. Smith** of cheating. **Mr./Ms. Smith** denied cheating and immediately left **Mr./Ms. Atwell's** home. **Mr./Ms. Smith** is not guilty of this charge. I trust that you will consider all the information accordingly and find that my client is indeed innocent.

The Crown calls the victim, **Mr./Ms. Atwell**.
Crown attorney: Please state your name for the Court?
Witness (victim): My name is George/Christine Atwell.
Crown attorney: And how do you know the defendant, **Mr./Ms. Smith?**
Witness (victim): We met in Grade 9. We've known each other for years. We hang out once in a while.
Crown attorney: Please tell the court what happened when **Mr./Ms. Smith** went to your home.

(ASSAULT)
Witness (victim): David/Alice came over after school at around 3:30 PM. **He/She** asked if I wanted to play cards, so we set up a game of poker. After **he/she** won a few hands in a row, I asked if **he/she** was cheating. **He/She** denied cheating and then became angry and pushed me hard enough to fall out of my chair. I got up and shouted at **him/her** to back off, and that's when I noticed a knife in **his/her** hand. I put my arms up and backed up, and **he/she** stabbed me in the chest with the knife.

OR

(FRAUD)

Witness (victim): David/Alice came over after school at around 3:30 PM. **He/She** asked if I wanted to play cards, so we set up a game of poker. After **he/she** won a bunch of hands in a row, I noticed **him/her** fumble a card, so **he/she** had six cards. **He/She** clearly tried to hide the dropped card, so I accused **him/her** of cheating which **he/she** denied.

Crown attorney: And what happened next?

(ASSAULT)
Witness (victim): I just remember that **he/she** ran out of my house, and I called my mom and told her I needed help. I guess I blacked out because the next thing I remember was waking up in an ambulance.

OR

(FRAUD)
Witness (victim): He/She ran out of my house with my money and blocked my number.

Crown attorney: Was there anyone else at home with you at the time?
Witness (victim): No. Both of my parents were at work.
Crown attorney: Thank you **Mr./Ms. Atwell**. I have no further questions, your honour.

The Defence cross-examines the victim, **Mr./Ms. Atwell**.
(ASSAULT)
Defence attorney: Hello Mr./Ms. Atwell.
Witness (victim): Hello.
Defence attorney: You said in your statement that **Mr./Ms. Smith** claimed that **he/she** had not been cheating?
Witness (victim): Yes, that's right.
Defence attorney: Isn't it true that you and **Mr./Ms. Smith** play cards often together?
Witness (victim): On occasion, yes, that's right.
Defence attorney: And has **Mr./Ms. Smith** ever cheated before in your card games?
Witness (victim): Not that I know of.
Defence attorney: Isn't it true that you started yelling at **Mr./Ms. Smith?**
Witness (victim): Well, yes. I guess I did. I was pissed that **he/she** was cheating and blatantly lying about it.
Defence attorney: I see. Isn't it true that you hit **Mr./Ms. Smith** in the face several times during the altercation?
Witness (victim): If I did, it was because I was defending myself after **he/she** pushed me.
Defence attorney: Did you choke Mr./Ms. Smith?
Witness (victim): If I did, it was only because **he/she** was going to hurt me first!
Defence attorney: No further questions. *The witness is excused.*

OR

(FRAUD)
Defence attorney: Hello Mr./Ms. Atwell.
Witness (victim): Hello.

Defence attorney: You said in your statement that **Mr./Ms. Smith** claimed that **he/she** had not been cheating?
Witness (victim): Yes, that's right.
Defence attorney: Isn't it true that you and **Mr./Ms. Smith** play cards often together?
Witness (victim): On occasion, yes, that's right.
Defence attorney: And has **Mr./Ms. Smith** ever cheated before in your card games?
Witness (victim): Not that I know of.
Defence attorney: The card that you say **Mr./Ms. Smith** dropped, was it a particularly good card that could have helped **his/her** hand?
Witness (victim): I didn't notice what the card was, I just saw that **he/she** had extra cards, and **he/she** was winning every hand. It wasn't hard to put two and two together.
Defence attorney: Isn't it true that you started yelling at **Mr./Ms. Smith?**
Witness (victim): Well, yes. I guess I did. I was pissed that **he/she** was cheating and blatantly lying about it.
Defence attorney: I see. Isn't it true that you threatened to hit **Mr./Ms. Smith** in the face?
Witness (victim): If I did, it was because **he/she** was stealing from me and lying about it.
Defence attorney: Did you try to grab the money from **Mr./Ms. Smith?**
Witness (victim): Yes, I tried to take my own money back.

The Defence calls the defendant, **Mr./Ms. Smith.**
Defence attorney: Please state your name for the Court.
Defendant: My name is David/Alice Smith
Defence attorney: Please tell the court what happened on October 15th.
Defendant: George/Christine invited me over after school to hang out. We've hung out before, so it was pretty normal.
Defence attorney: What happened next?
Defendant: When I got there, I suggested we play poker, which we've done before. We played a few hands, and I was on a lucky streak. After I won a **fourth/tenth** hand in a row, **George/Christine** asked me if I was cheating. I said I wasn't, and **he/she** got really mad and yelled that I was cheating.
Defence attorney: What happened next?

(ASSAULT)
Defendant: I yelled back that I wasn't and then **he/she** slapped my face. I pushed **him/her** away from me and was trying to get my things to leave. Then **he/she** pushed me back and started punching and hitting my head and face. I was trying to fight back, but I was stunned at what was happening.
Defence attorney: What did Mr./Ms. Atwell do next?
Defendant: He/She started choking me. I was so scared. I couldn't breathe. I grabbed my pocketknife and lashed out with it trying to hit anything to get **him/her** off of me.

OR

(FRAUD)
Defendant: I yelled back that I wasn't and then **he/she** started threatening me and waving the card I dropped in my face. I got scared and was trying to grab my things to leave. Then **he/she**

started trying to tear my winnings out of my hand. I just grabbed my stuff as fast as possible and ran out of there.

Defence attorney: Thank you Mr./Ms. Smith. No further questions.

The Crown cross-examines the defendant, **Mr./Ms. Smith.**

(ASSAULT)
Crown attorney: So, after you stabbed **Mr./Ms. Atwell** with the knife, what happened?
Defendant: What do you mean?
Crown attorney: Did you call the police?
Defendant: I was scared that **he/she** would come at me again, so I ran away.
Crown attorney: You left **him/her** there, and it didn't occur to you to call an ambulance?
Defendant: I honestly wasn't thinking. I was scared and just wanted to get away.
Crown attorney: Isn't it true that you told one of your friends that you hoped **Mr./Ms. Atwell** would die?
Defendant: I was so pissed off when I said that. People say things they don't mean all the time, but I didn't actually want **him/her** to die. **He's/She's** the one who could have killed me! I was defending myself! Why can't you see that?
Crown attorney: No further questions your honour. *The defendant is excused.*

OR

(FRAUD)
Crown attorney: So, right after you dropped the card, what happened?
Defendant: What do you mean?
Crown attorney: Did you try to hide it?
Defendant: Yes, of course. I didn't want **George/Christine** to see the cards in my hand.
Crown attorney: And how many cards did you have in your hand at this point?
Defendant: I had five cards in my hand! I wasn't cheating!
Crown attorney: When **George/Christine** accused you of cheating, did you try to show **him/her** that you had five cards, and not six?
Defendant: Well, no, but I was mad that **he/she** would accuse me like that.
Crown attorney: So, instead of showing proof that you weren't cheating, you just tried to grab the money and run?
Defendant: I needed that money and won it fair and square, and it was mine! **George/Christine** wasn't going to miss it anyway. **He/She** is loaded! And even if I had been cheating, **he/she** was still threatening me! Of course I was just trying to get out of there!
Crown attorney: No further questions your honour. *The defendant is excused.*

Testimony from the court appointed clinical psychologist, Dr. Richards.
Judge: Dr. Richards, you had the opportunity to assess the defendant, **Mr./Ms. Smith,** is that correct?

(ASSAULT)
Dr. Richards: Yes, that's correct. I spent several hours with **David/Alice** Smith. I found **David/Alice** to be an earnest young **man/woman**. **He/She** was forthright and candid, providing

careful and thorough answers. **He/She** was visibly upset about the incident and clearly showed empathy for **Mr./Ms. Atwell**. In my professional opinion, the defendant, **David/Alice** Smith, does not have any signs indicating a mental disorder and no diagnosis has been given.

OR

(FRAUD)
Dr. Richards: Yes, that's correct. I spent several hours with **David/Alice** Smith. I found **David/Alice** to be an earnest young **man/woman**. **He/She** was forthright and candid, providing careful and thorough answers. **He/She** was visibly upset about the incident. In my professional opinion, the defendant, **David/Alice** Smith, does not have any signs indicating a mental disorder and no diagnosis has been given.

Judge: Thank you, Dr. Richards.

The Judge provides the law and instructions for the jury.
Judge: Members of the jury, you have heard the testimony from both parties involved and the court appointed clinical psychologist. It is now my responsibility to provide you with the law. **Mr./Ms. Smith** is charged with the following:

268. AGGRAVATED ASSAULT
(1) Every one commits an aggravated assault who wounds, maims, disfigures or endangers the life of the complainant.
(2) Every one who commits an aggravated assault is guilty of an indictable offence and liable to imprisonment for a term not exceeding fourteen years.

Judge: Mr./Ms. Smith's defence is as follows:

> **34. SELF-DEFENCE AGAINST UNPROVOKED ASSAULT**
> (1) Every one who is unlawfully assaulted without having provoked the assault is justified in repelling force by force if the force he uses is not intended to cause death or grievous bodily harm and is no more than is necessary to enable him to defend himself. Extent of justification
> (2) Every one who is unlawfully assaulted and who causes death or grievous bodily harm in repelling the assault is justified if
> (*a*) he causes it under reasonable apprehension of death or grievous bodily harm from the violence with which the assault was originally made or with which the assailant pursues his purposes; and
> (*b*) he believes, on reasonable grounds, that he cannot otherwise preserve himself from death or grievous bodily harm.

OR

380. FRAUD
(1) Every one who, by deceit, falsehood or other fraudulent means, whether or not it is a false pretence within the meaning of this Act, defrauds the public or any person, whether ascertained or not, of any property, money or valuable security or any service,

(a) is guilty of an indictable offence and liable to a term of imprisonment not exceeding fourteen years, where the subject-matter of the offence is a testamentary instrument or the value of the subject-matter of the offence exceeds five thousand dollars; or
(b) is guilty
(i) of an indictable offence and is liable to imprisonment for a term not exceeding two years, or
(ii) of an offence punishable on summary conviction,

where the value of the subject-matter of the offence does not exceed five thousand dollars.

Judge: Please take into consideration all the information you have heard today, and do not let any biases you may have come into your decision-making process. Finally, remember that the charge must be proved beyond a reasonable doubt. Once again, this means that, if you find that the defendant is guilty, there cannot be any uncertainty or doubt about the evidence on which you have based your finding of guilt.

Appendix A2: Short Transcript of a Court Case

Below is an amended version of the transcript used by Blais and Forth (2014). Permission has been granted to the researchers to amend and use this by the authors.

References

Blais, J., & Forth, A. E. (2014). Potential labeling effects: Influence of psychopathy diagnosis, defendant age, and defendant gender on mock jurors' decisions. *Psychology, Crime & Law, 20*(2), 116-134. doi:10.1080/1068316X.2012.749473

Court Summary

(a photo of the defendant will be displayed at the top of the screen for the entirety of the Court Summary in the Qualtrics survey)

PHOTO

Defendant: **David / Alice** Smith is a 16-year-old White **male/female.**

Victim: **George / Christine** Atwell is a 16-year-old White **male/female.**

Judge: In this alleged case of aggravated assault, **Mr./Ms. Smith** is 16 years old and has been charged with section 268 of the *Criminal Code,* aggravated assault. The Crown has the burden of proving that charge beyond a reasonable doubt. This means that, if you find the defendant is guilty, there cannot be any uncertainty or doubt about the evidence on which you have based your finding of guilt. It is your responsibility to listen to all the evidence and then to apply the law that I will give you at the end of the trial. The trial will begin with opening statements by the Crown and Defence attorneys. Following these statements, the Crown and Defence will present and question the parties involved, who will subsequently be cross-examined. You will also hear testimony from the court appointed clinical psychologist, Dr. Richards, who interviewed the defendant after **his/her** arrest. Please listen to the following arguments carefully. Following the testimonies, you will be asked to make a decision as to whether you find the defendant, **Mr./Ms. Smith**, guilty or not guilty of the charge.

The Crown makes their opening statements.
Crown attorney: Mr./Ms. Smith is an acquaintance of the victim, **Mr./Ms. Atwell.** On October 15th, 2022, an altercation occurred between the defendant and the victim. **Mr./Ms. Atwell** and **Mr./Ms. Smith** were playing cards at **Mr./Ms. Atwell's** home. An argument broke out between them, and it is alleged that **Mr./Ms. Smith** pushed **Mr./Ms. Atwell** then took a pocketknife out of **his/her** waistband and stabbed the victim in the chest. **Mr./Ms. Atwell** was rushed to the hospital in an ambulance where he received several stitches following this incident. **Mr./Ms. Smith** is charged with aggravated assault, and it is your duty, as representatives of our community, to seek justice by finding **Mr./Ms. Smith** guilty.

The Defence makes their opening statement.
Defence attorney: Mr./Ms. Smith is not guilty of this charge. Although the altercation between **Mr./Ms. Smith** and **Mr./Ms. Atwell** did occur, the details are as follows. **Mr./Ms. Atwell** and **Mr./Ms. Smith** were playing cards and **Mr./Ms. Atwell** accused **Mr./Ms. Smith** of cheating. **Mr./Ms. Smith** denied cheating and **Mr./Ms. Atwell** became angry and slapped **Mr./Ms. Smith** in the face. **Mr./Ms. Smith** pushed **Mr./Ms. Atwell** away but **Mr./Ms. Atwell** then began hitting

and choking **Mr./Ms. Smith**. Fearing for **his/her** safety, **Mr./Ms. Smith** removed **his/her** pocketknife from **his/her** pocket and slashed at **Mr./Ms. Atwell**, striking **his/her** chest. **Mr./Ms. Smith** acted in self-defence. **He/She** is not guilty of this charge. I trust that you will consider all the information accordingly and find that my client is indeed innocent.

The Crown calls the victim, **Mr./Ms. Atwell**.
Crown attorney: Please state your name for the Court?
Witness (victim): My name is George/Christine Atwell.
Crown attorney: And how do you know the defendant, **Mr./Ms. Smith?**
Witness (victim): We met in Grade 9. We've known each other for years. We hang out once in a while.
Crown attorney: Please tell the court what happened when **Mr./Ms. Smith** went to your home.
Witness (victim): David/Alice came over after school at around 3:30 PM. **He/She** asked if I wanted to play cards, so we set up a game of poker. After **he/she** won a few hands in a row, I asked if **he/she** was cheating. **He/She** denied cheating and then became angry and pushed me hard enough to fall out of my chair. I got up and shouted at **him/her** to back off, and that's when I noticed a knife in **his/her** hand. I put my arms up and backed up, and **he/she** stabbed me in the chest with the knife.
Crown attorney: And what happened next?
Witness (victim): I just remember that **he/she** ran out of my house, and I called my mom and told her I needed help. I guess I blacked out because the next thing I remember was waking up in an ambulance.
Crown attorney: Was there anyone else at home with you at the time?
Witness (victim): No. Both of my parents were at work.
Crown attorney: Thank you **Mr./Ms. Atwell**. I have no further questions, your honour.

The Defence cross-examines the victim, **Mr./Ms. Atwell**.
Defence attorney: Hello Mr./Ms. Atwell.
Witness (victim): Hello.
Defence attorney: You said in your statement that **Mr./Ms. Smith** claimed that **he/she** had not been cheating?
Witness (victim): Yes, that's right.
Defence attorney: Isn't it true that you and **Mr./Ms. Smith** play cards often together?
Witness (victim): On occasion, yes, that's right.
Defence attorney: And has **Mr./Ms. Smith** ever cheated before in your card games?
Witness (victim): Not that I know of.
Defence attorney: Isn't it true that you started yelling at **Mr./Ms. Smith?**
Witness (victim): Well, yes. I guess I did. I was pissed that **he/she** was cheating and blatantly lying about it.
Defence attorney: I see. Isn't it true that you hit **Mr./Ms. Smith** in the face several times during the altercation?
Witness (victim): If I did, it was because I was defending myself after **he/she** pushed me.
Defence attorney: Did you choke Mr./Ms. Smith?
Witness (victim): If I did, it was only because **he/she** was going to hurt me first!
Defence attorney: No further questions. *The witness is excused.*

The Defence calls the defendant, **Mr./Ms. Smith.**

Defence attorney: Please state your name for the Court.
Defendant: My name is David/Alice Smith
Defence attorney: Please tell the court what happened on October 15th.
Defendant: George/Christine invited me over after school to hang out. We've hung out before, so it was pretty normal.
Defence attorney: What happened next?
Defendant: When I got there, I suggested we play poker, which we've done before. We played a few hands, and I was on a lucky streak. After I won a fourth hand in a row, **George/Christine** asked me if I was cheating. I said I wasn't, and **he/she** got really mad and yelled that I was cheating.
Defence attorney: What happened next?
Defendant: I yelled back that I wasn't and then **he/she** slapped my face. I pushed **him/her** away from me and was trying to get my things to leave. Then **he/she** pushed me back and started punching and hitting my head and face. I was trying to fight back, but I was stunned at what was happening.
Defence attorney: What did Mr./Ms. Atwell do next?
Defendant: He/She started choking me. I was so scared. I couldn't breathe. I grabbed my pocketknife and lashed out with it trying to hit anything to get **him/her** off of me.
Defence attorney: Thank you Mr./Ms. Smith. No further questions.

The Crown cross-examines the defendant, **Mr./Ms. Smith.**
Crown attorney: So, after you stabbed **Mr./Ms. Atwell** with the knife, what happened?
Defendant: What do you mean?
Crown attorney: Did you call the police?
Defendant: I was scared that **he/she** would come at me again, so I ran away.
Crown attorney: You left **him/her** there, and it didn't occur to you to call an ambulance?
Defendant: I honestly wasn't thinking. I was scared and just wanted to get away.
Crown attorney: Isn't it true that you told one of your friends that you hoped **Mr./Ms. Atwell** would die?
Defendant: I was so pissed off when I said that. People say things they don't mean all the time, but I didn't actually want **him/her** to die. **He's/She's** the one who could have killed me! I was defending myself! Why can't you see that?
Crown attorney: No further questions your honour. *The defendant is excused.*

Testimony from the court appointed clinical psychologist, Dr. Richards.
(SCHIZOPHRENIA)
Judge: Dr. Richards, you had the opportunity to assess the defendant, **Mr./Ms. Smith,** is that correct?
Dr. Richards: Yes, that's correct. I spent several hours with **David/Alice** Smith. I found **David/Alice** to be a distracted and agitated young **man/woman**. **He/She** seemed confused during the interview and appeared to be responding to questions I did not ask and had sensory experiences that seemed quite odd. **He/She** was not emotionally expressive, and sometimes would stare off only to respond again with disorganized speech that was difficult to follow. **He/she** expressed no remorse for **his/her** actions expressing that **he/she** had to protect **himself/herself** from **Mr./Ms. Atwell**. At the time of the interview, **David/Alice** expressed a belief that **Mr./Ms. Atwell** wanted to kill **David/Alice** and used the opportunity to hang out after school as a way to lure **him/her** to **Mr./Ms. Atwell's** home. **David/Alice** explained that it was

part of a plan to kill **him/her** and that **he/she** was not safe. In my professional opinion, the defendant, **David/Alice** Smith, is suffering from schizophrenia.

OR

(PSYCHOPATHIC TRAITS)
Judge: Dr. Richards, you had the opportunity to assess the defendant, **Mr./Ms. Smith,** is that correct?
Dr. Richards: Yes, that's correct. I spent several hours with **David/Alice** Smith. I found **David/Alice** to be a superficial and insincere young **man/woman**. **He/She** consistently tries to charm, con, and manipulate others to obtain what **he/she** wants. **He/She** has no empathy for **Mr./Ms. Atwell** and refuses to take responsibility for **his/her** actions. **He/She** explained that **Mr./Ms. Atwell** deserved **his/her** injuries for crossing **him/her**. In my professional opinion, the defendant, **David/Alice** Smith has psychopathic traits.

OR

(CONDUCT DISORDER)
Judge: Dr. Richards, you had the opportunity to assess the defendant, **Mr./Ms. Smith,** is that correct?
Dr. Richards: Yes, that's correct. I spent several hours with **David/Alice** Smith. I found **David/Alice** to be an aggressive and agitated young **man/woman**. **He/She** appeared bored with the interview and responded with anger to my questions. **He/She** either behaved aggressively or ignored me when **he/she** was asked direct questions about the altercation with **Mr./Ms. Atwell**. **David/Alice** has had recurring difficulties with the law and has a history of violating the rights and boundaries of others. **He/She** has no remorse for **Mr./Ms. Atwell**. In my professional opinion, the defendant, **David/Alice** Smith, has conduct disorder.

OR

(NO DIAGNOSIS)
Judge: Dr. Richards, you had the opportunity to assess the defendant, **Mr./Ms. Smith,** is that correct?
Dr. Richards: Yes, that's correct. I spent several hours with **David/Alice** Smith. I found **David/Alice** to be an earnest young **man/woman**. **He/She** was forthright and candid, providing careful and thorough answers. **He/She** was visibly upset about the incident and clearly showed empathy for **Mr./Ms. Atwell**. In my professional opinion, the defendant, **David/Alice** Smith, does not have any signs indicating a mental disorder and no diagnosis has been given.

Judge: Thank you, Dr. Richards.

The Judge provides the law and instructions for the jury.
Judge: Members of the jury, you have heard the testimony from both parties involved and the court appointed clinical psychologist. It is now my responsibility to provide you with the law. **Mr./Ms. Smith** is charged with the following:

268. AGGRAVATED ASSAULT
(1) Every one commits an aggravated assault who wounds, maims, disfigures or endangers the life of the complainant.
(2) Every one who commits an aggravated assault is guilty of an indictable offence and liable to imprisonment for a term not exceeding fourteen years.

Judge: Mr./Ms. Smith's defence is as follows:

34. SELF-DEFENCE AGAINST UNPROVOKED ASSAULT
(1) Every one who is unlawfully assaulted without having provoked the assault is justified in repelling force by force if the force he uses is not intended to cause death or grievous bodily harm and is no more than is necessary to enable him to defend himself. Extent of justification
(2) Every one who is unlawfully assaulted and who causes death or grievous bodily harm in repelling the assault is justified if
(*a*) he causes it under reasonable apprehension of death or grievous bodily harm from the violence with which the assault was originally made or with which the assailant pursues his purposes; and
(*b*) he believes, on reasonable grounds, that he cannot otherwise preserve himself from death or grievous bodily harm.

Judge: Please take into consideration all the information you have heard today, and do not let any biases you may have come into your decision-making process. Finally, remember that the charge must be proved beyond a reasonable doubt. Once again, this means that, if you find that the defendant is guilty, there cannot be any uncertainty or doubt about the evidence on which you have based your finding of guilt.

Appendix A3: Short Transcript of a Court Case

Below is an amended version of the transcript used by Blais and Forth (2014). Permission has been granted to the researchers to amend and use this by the authors.

References

Blais, J., & Forth, A. E. (2014). Potential labeling effects: Influence of psychopathy diagnosis, defendant age, and defendant gender on mock jurors' decisions. *Psychology, Crime & Law, 20*(2), 116-134. doi:10.1080/1068316X.2012.749473

Court Summary

(a photo of the defendant will be displayed at the top of the screen for the entirety of the Court Summary in the Qualtrics survey)

PHOTO






Defendant: **David / Alice** Smith is a 16-year-old **White/Black male/female.**

Victim: George / Christine Atwell is a 16-year-old White/Black male/female.

Judge: In this alleged case of aggravated assault, **Mr./Ms. Smith** is 16 years old and has been charged with section 268 of the *Criminal Code,* aggravated assault. The Crown has the burden of proving that charge beyond a reasonable doubt. This means that, if you find the defendant is guilty, there cannot be any uncertainty or doubt about the evidence on which you have based your finding of guilt. It is your responsibility to listen to all the evidence and then to apply the law that I will give you at the end of the trial. The trial will begin with opening statements by the Crown and Defence attorneys. Following these statements, the Crown and Defence will present and question the parties involved, who will subsequently be cross-examined. You will also hear testimony from the court appointed clinical psychologist, Dr. Richards, who interviewed the defendant after **his/her** arrest. Please listen to the following arguments carefully. Following the testimonies, you will be asked to make a decision as to whether you find the defendant, **Mr./Ms. Smith**, guilty or not guilty of the charge.

The Crown makes their opening statements.
Crown attorney: Mr./Ms. Smith is an acquaintance of the victim, **Mr./Ms. Atwell.** On October 15th, 2022, an altercation occurred between the defendant and the victim. **Mr./Ms. Atwell** and **Mr./Ms. Smith** were playing cards at **Mr./Ms. Atwell's** home. An argument broke out between them, and it is alleged that **Mr./Ms. Smith** pushed **Mr./Ms. Atwell** then took a pocketknife out of **his/her** waistband and stabbed the victim in the chest. **Mr./Ms. Atwell** was rushed to the hospital in an ambulance where he received several stitches following this incident. **Mr./Ms. Smith** is charged with aggravated assault, and it is your duty, as representatives of our community, to seek justice by finding **Mr./Ms. Smith** guilty.

The Defence makes their opening statement.
Defence attorney: Mr./Ms. Smith is not guilty of this charge. Although the altercation between **Mr./Ms. Smith** and **Mr./Ms. Atwell** did occur, the details are as follows. **Mr./Ms. Atwell** and **Mr./Ms. Smith** were playing cards and **Mr./Ms. Atwell** accused **Mr./Ms. Smith** of cheating. **Mr./Ms. Smith** denied cheating and **Mr./Ms. Atwell** became angry and slapped **Mr./Ms. Smith** in the face. **Mr./Ms. Smith** pushed **Mr./Ms. Atwell** away but **Mr./Ms. Atwell** then began hitting and choking **Mr./Ms. Smith**. Fearing for **his/her** safety, **Mr./Ms. Smith** removed **his/her** pocketknife from **his/her** pocket and slashed at **Mr./Ms. Atwell**, striking **his/her** chest. **Mr./Ms. Smith** acted in self-defence. **He/She** is not guilty of this charge. I trust that you will consider all the information accordingly and find that my client is indeed innocent.

The Crown calls the victim, **Mr./Ms. Atwell**.
Crown attorney: Please state your name for the Court?
Witness (victim): My name is George/Christine Atwell.
Crown attorney: And how do you know the defendant, **Mr./Ms. Smith?**
Witness (victim): We met in Grade 9. We've known each other for years. We hang out once in a while.
Crown attorney: Please tell the court what happened when **Mr./Ms. Smith** went to your home.
Witness (victim): David/Alice came over after school at around 3:30 PM. **He/She** asked if I wanted to play cards, so we set up a game of poker. After **he/she** won a few hands in a row, I asked if **he/she** was cheating. **He/She** denied cheating and then became angry and pushed me hard enough to fall out of my chair. I got up and shouted at **him/her** to back off, and that's when I noticed a knife in **his/her** hand. I put my arms up and backed up, and **he/she** stabbed me in the chest with the knife.

Crown attorney: And what happened next?
Witness (victim): I just remember that **he/she** ran out of my house, and I called my mom and told her I needed help. I guess I blacked out because the next thing I remember was waking up in an ambulance.
Crown attorney: Was there anyone else at home with you at the time?
Witness (victim): No. Both of my parents were at work.
Crown attorney: Thank you **Mr./Ms. Atwell**. I have no further questions, your honour.

The Defence cross-examines the victim, **Mr./Ms. Atwell**.
Defence attorney: Hello Mr./Ms. Atwell.
Witness (victim): Hello.
Defence attorney: You said in your statement that **Mr./Ms. Smith** claimed that **he/she** had not been cheating?
Witness (victim): Yes, that's right.
Defence attorney: Isn't it true that you and **Mr./Ms. Smith** play cards often together?
Witness (victim): On occasion, yes, that's right.
Defence attorney: And has **Mr./Ms. Smith** ever cheated before in your card games?
Witness (victim): Not that I know of.
Defence attorney: Isn't it true that you started yelling at **Mr./Ms. Smith?**
Witness (victim): Well, yes. I guess I did. I was pissed that **he/she** was cheating and blatantly lying about it.
Defence attorney: I see. Isn't it true that you hit **Mr./Ms. Smith** in the face several times during the altercation?
Witness (victim): If I did, it was because I was defending myself after **he/she** pushed me.
Defence attorney: Did you choke Mr./Ms. Smith?
Witness (victim): If I did, it was only because **he/she** was going to hurt me first!
Defence attorney: No further questions. *The witness is excused.*

The Defence calls the defendant, **Mr./Ms. Smith.**
Defence attorney: Please state your name for the Court.
Defendant: My name is David/Alice Smith
Defence attorney: Please tell the court what happened on October 15th.
Defendant: George/Christine invited me over after school to hang out. We've hung out before, so it was pretty normal.
Defence attorney: What happened next?
Defendant: When I got there, I suggested we play poker, which we've done before. We played a few hands, and I was on a lucky streak. After I won a fourth hand in a row, **George/Christine** asked me if I was cheating. I said I wasn't, and **he/she** got really mad and yelled that I was cheating.
Defence attorney: What happened next?
Defendant: I yelled back that I wasn't and then **he/she** slapped my face. I pushed **him/her** away from me and was trying to get my things to leave. Then **he/she** pushed me back and started punching and hitting my head and face. I was trying to fight back, but I was stunned at what was happening.
Defence attorney: What did Mr./Ms. Atwell do next?
Defendant: He/She started choking me. I was so scared. I couldn't breathe. I grabbed my pocketknife and lashed out with it trying to hit anything to get **him/her** off of me.

Defence attorney: Thank you Mr./Ms. Smith. No further questions.

The Crown cross-examines the defendant, **Mr./Ms. Smith.**
Crown attorney: So, after you stabbed **Mr./Ms. Atwell** with the knife, what happened?
Defendant: What do you mean?
Crown attorney: Did you call the police?
Defendant: I was scared that **he/she** would come at me again, so I ran away.
Crown attorney: You left **him/her** there, and it didn't occur to you to call an ambulance?
Defendant: I honestly wasn't thinking. I was scared and just wanted to get away.
Crown attorney: Isn't it true that you told one of your friends that you hoped **Mr./Ms. Atwell** would die?
Defendant: I was so pissed off when I said that. People say things they don't mean all the time, but I didn't actually want **him/her** to die. **He's/She's** the one who could have killed me! I was defending myself! Why can't you see that?
Crown attorney: No further questions your honour. *The defendant is excused.*

Testimony from the court appointed clinical psychologist, Dr. Richards.
Judge: Dr. Richards, you had the opportunity to assess the defendant, **Mr./Ms. Smith,** is that correct?
Dr. Richards: Yes, that's correct. I spent several hours with **David/Alice** Smith. I found **David/Alice** to be an earnest young **man/woman**. **He/She** was forthright and candid, providing careful and thorough answers. **He/She** was visibly upset about the incident and clearly showed empathy for **Mr./Ms. Atwell**. In my professional opinion, the defendant, **David/Alice** Smith, does not have any signs indicating a mental disorder and no diagnosis has been given.

Judge: Thank you, Dr. Richards.

The Judge provides the law and instructions for the jury.
Judge: Members of the jury, you have heard the testimony from both parties involved and the court appointed clinical psychologist. It is now my responsibility to provide you with the law. **Mr./Ms. Smith** is charged with the following:

268. AGGRAVATED ASSAULT
(1) Every one commits an aggravated assault who wounds, maims, disfigures or endangers the life of the complainant.
(2) Every one who commits an aggravated assault is guilty of an indictable offence and liable to imprisonment for a term not exceeding fourteen years.

Judge: Mr./Ms. Smith's defence is as follows:

34. SELF-DEFENCE AGAINST UNPROVOKED ASSAULT
(1) Every one who is unlawfully assaulted without having provoked the assault is justified in repelling force by force if the force he uses is not intended to cause death or grievous bodily harm and is no more than is necessary to enable him to defend himself. Extent of justification
(2) Every one who is unlawfully assaulted and who causes death or grievous bodily harm in repelling the assault is justified if

(*a*) he causes it under reasonable apprehension of death or grievous bodily harm from the violence with which the assault was originally made or with which the assailant pursues his purposes; and
(*b*) he believes, on reasonable grounds, that he cannot otherwise preserve himself from death or grievous bodily harm.

Judge: Please take into consideration all the information you have heard today, and do not let any biases you may have come into your decision-making process. Finally, remember that the charge must be proved beyond a reasonable doubt. Once again, this means that, if you find that the defendant is guilty, there cannot be any uncertainty or doubt about the evidence on which you have based your finding of guilt.

Appendix B: Transcript-Related Questions

It is very important that you carefully read the court summary. To demonstrate that you are familiar with the details, please answer the following questions about the case. Should you need to, you may scroll back through the court summary to ensure you are familiar with the details.

1. What crime is the defendant charged with?

2. What is the gender of the defendant?

3. What is the victim's name?

4. Where did the incident occur?

5. What is the defendant's name?

6. What is the defendant's age?

7. What is the victim's age?

Appendix C: Verdict Questions

Below are questions adapted, with permission, from Blais and Forth's (2014) study.

Select a verdict for the defendant based on the transcript of court proceeding that you just read.

1. In your opinion, is the defendant guilty or not-guilty?

   ☐ Yes

   ☐ No

Select the degree of guilt of the defendant. Select 1 if you believe that the defendant is not at all guilty, 7 if you believe the defendant is definitely guilty, or anywhere in between.

		1 Not at all guilty	2	3	4	5	6	7 Definitely guilty
2.	How guilty was the defendant?	1	2	3	4	5	6	7

Select your level of confidence with the verdict. Select 1 if you are not at all confident, 7 if you are very confident, or anywhere in between.

		1 Not at all confident	2	3	4	5	6	7 Very confident
3.	How confident are you with your verdict decision?	1	2	3	4	5	6	7

Read each statement and select the answer that best describes your opinion of the person in question. Select 1 if your answer is ‘not at all’, select 7 if your answer is ‘extremely’, or anywhere in between.

		1 Not at all	2	3	4	5	6	7 Extremely
4.	How credible was the testimony of the defendant?	1	2	3	4	5	6	7
5.	How credible was the testimony of the victim?	1	2	3	4	5	6	7
6.	How likely are you to recommend treatment for the defendant?	1	2	3	4	5	6	7
7.	How much would the defendant benefit from treatment?	1	2	3	4	5	6	7
8.	How useful would a treatment program aimed at reducing violence be for the defendant?	1	2	3	4	5	6	7
9.	How useful would any treatment program aimed at reducing criminal behaviour be for the defendant?	1	2	3	4	5	6	7
10.	How high is the defendant’s risk for future violence?	1	2	3	4	5	6	7

11.	How high is the defendant's risk for future criminal acts?	1	2	3	4	5	6	7
12.	How likely is it that the defendant poses a threat to society?	1	2	3	4	5	6	7

Now we want you to think about the physical appearance of the defendant and the victim and answer the following four questions.

		1 Very unattractive	2	3	4 Neutral: Neither unattractive nor attractive	5	6	7 Very attractive
13.	How physically attractive do you think the defendant is?	1	2	3	4	5	6	7

		1 Very untrustworthy	2	3	4 Neutral: Neither untrustworthy nor trustworthy	5	6	7 Very trustworthy

14.	How trustworthy do you think the defendant is?	1	2	3	4	5	6	7

Appendix D: Exploratory Analyses for Study 1

Confidence

An ANCOVA was run to determine the effect of attractiveness on confidence ratings after controlling for the BIDR-16 and RLAQ. Key assumptions for ANCOVA were met. After adjustment for socially desirable responding and authoritarian legal attitudes, attractiveness did not have a significant effect on confidence ratings, $F(1, 158) = .07$, $p = .799$, $\eta^2 = .00$.

An ANCOVA was run to determine the effect of gender on confidence ratings after controlling for the BIDR-16 and RLAQ. Key assumptions for ANCOVA were met, except for normality of the standardized residuals for males ($p = .039$). Visual inspection of histograms revealed an approximately normal distribution. After adjustment for socially desirable responding and authoritarian legal attitudes, gender did not have a significant effect on confidence ratings, $F(1, 158) = 3.04$, $p = .083$, $\eta^2 = .19$.

An ANCOVA was run to determine the effect of crime type on confidence ratings after controlling for the BIDR-16 and RLAQ. Key assumptions for ANCOVA were met, except for normality of the standardized residuals overall, and for the assault group. Visual inspection of histograms and normal Q-Q plots revealed approximately normal distributions. After adjustment for socially desirable responding and authoritarian legal attitudes, crime type did not have a significant effect on confidence ratings, $F(1, 158) = .09$, $p = .771$, $\eta^2 = .00$.

A two-way ANCOVA was run to determine if attractiveness and crime type had an effect on confidence ratings after controlling for the BIDR-16 and RLAQ. There was one outlier, as assessed by one studentized residual exceeding ±3 standard deviations. The

analysis was run both with and without this case, and the outlier had a significant effect on the analysis; therefore, the outlier was removed from the analysis. Homogeneity of variances was not met, as assessed by Levene's statistic, and Huber-White robust standard errors were used. The remaining key assumptions were met. There was a statistically significant two-way interaction between attractiveness and crime type after controlling for socially desirable responding and authoritarian legal attitudes, $t(155) = 2.00$, $p = .048$, partial $\eta^2 = .03$ (see Tables D.1 and D.2 and Figure D.1). An analysis of simple main effects with a Bonferroni adjustment was conducted, which showed no statistically significant differences between the different levels of attractiveness and crime type.

Table D.1

Analysis of Covariance Results of Confidence Ratings Based on Attractiveness and Crime Type when Controlling for Socially Desirable Responding and Authoritarian Legal Attitudes

	$F(1, 155)$	Partial η^2	Robust *SE*
Attractiveness	.43	.00	.29
Crime Type	.03	.00	.30
Attractiveness * Crime Type	4.07**	.03	.40**

**p < .05

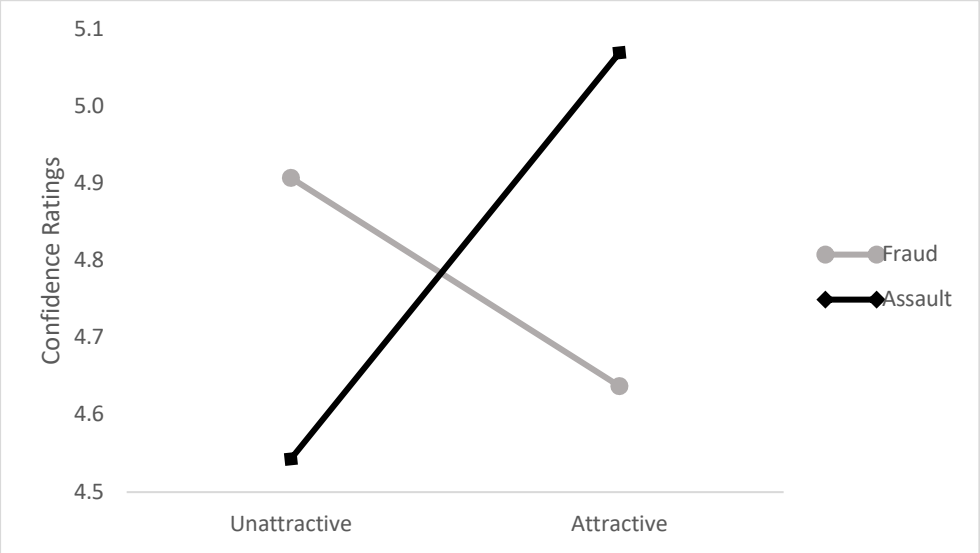
Table D.2

Means, Adjusted Means, Standard Deviations, and Standard Errors for Confidence for the Attractiveness and Crime Type Groups

	Unattractive		Attractive	
Crime Type	Fraud	Assault	Fraud	Assault
M	4.89	4.56	4.65	5.06
SD	.96	1.42	1.42	1.12
M_{adj}	4.91	4.54	4.64	5.07
SE	.19	.19	.20	.22

Figure D.1

Two-Way Interaction of Attractiveness and Crime Type after Controlling for the BIDR-16 and RLAQ



A three-way ANCOVA was conducted to examine the effects of attractiveness, gender, and crime type on confidence, after controlling for socially desirable responding (BIDR-16), and authoritarian legal attitudes (RLAQ). Homogeneity of variance was violated; therefore, Huber-White's robust standard errors were interpreted. There was one outlier, as assessed by standardized residuals exceeding ± 3 standard deviations. After running the analysis both with and without the outlier, the results did not change significantly; therefore, it was kept in the analysis. Other key assumptions were met. Means, adjusted means, standard deviations and standard errors are presented in Table D.3. There was not a statistically significant three-way interaction between attractiveness, gender, and crime type on ratings of confidence while controlling for socially desirable responding and authoritarian legal attitudes, $t(152) = -.10$, $p = .921$, partial $\eta^2 = .00$ (see Table D.4).

Table D.3

Means, Adjusted Means, Standard Deviations, and Standard Errors for Ratings of

Confidence by Attractiveness, Gender, and Crime Type while Controlling for Socially Desirable Responding and Authoritarian Legal Attitudes

		M	SD	M_{adj}	SE
Unattractive					
Fraud	Male	5.20	.89	5.22	.28
	Female	4.64	.95	4.66	.25
Assault	Male	4.55	1.40	4.57	.28
	Female	4.57	1.47	4.52	.27
Attractive					
Fraud	Male	4.91	1.51	4.95	.27
	Female	4.33	1.28	4.25	.31
Assault	Male	5.00	1.61	5.00	.30
	Female	4.88	.89	4.90	.32

Table D.4

Analysis of Covariance Results of Confidence Ratings Based on Attractiveness, Crime Type, and Gender when Controlling for the BIDR-16 and RLAQ

	$F(1, 152)$	Partial η^2	Robust *SE*
Attractiveness	.03	.00	.38
Crime Type	.02	.00	.39
Gender	3.02	.02	.45
Attractiveness*Gender	.06	.00	.63
Attractiveness*Crime Type	3.40	.02	.53
Gender*Crime Type	1.81	.01	.65
Attractiveness*Crime Type*Gender	.01	.00	.83

Defendant Credibility

An ANCOVA was run to determine the effect of attractiveness on defendant credibility ratings after controlling for the BIDR-16) and RLAQ. Key assumptions were met, except for normality of residuals between groups, as assessed by a significant Shapiro-Wilk test of normality. Visual inspection of histograms and normal Q-Q plots revealed approximately normal distributions. After controlling for socially desirable responding and authoritarian legal attitudes, attractiveness had a statistically significant effect on defendant credibility ratings, $F(1, 158) = 10.51$, $p = .001$, $\eta^2 = .06$. Attractive defendants were

considered significantly less credible than unattractive defendants (Mean difference = .56, 95% *CI* =.22, .90, *p* = .001) (see Table D.5).

Table D.5

Means, Adjusted Means, Standard Deviations, and Standard Errors for Defendant Credibility based on Attractiveness with the BIDR-16 and RLAQ as Covariates

	Unadjusted		Adjusted	
	M	SD	M	SE
Unattractive	4.51	1.03	4.51	.12
Attractive	3.96	1.18	3.96	.13

An ANCOVA was run to determine the effect of gender on defendant credibility ratings after controlling for the BIDR-16 and RLAQ. Key assumptions were met, except for normality of residuals between groups, as assessed by a significant Shapiro-Wilk test of normality. Visual inspection of histograms and normal Q-Q plots revealed approximately normal distributions. After controlling for socially desirable responding and authoritarian legal attitudes, gender did not have a statistically significant effect on defendant credibility ratings, $F(1, 158) = 1.39$, $p = .240$, $\eta^2 = .01$.

An ANCOVA was run to determine the effect of crime type on defendant credibility ratings after controlling for the BIDR-16 and RLAQ. Key assumptions were met, except for normality of residuals, as assessed by significant Shapiro-Wilk tests of normality. Visual inspection of histograms and normal Q-Q plots revealed approximately normal distributions. After controlling for socially desirable responding and authoritarian legal attitudes, crime type had a statistically significant effect on defendant credibility ratings, $F(1, 158) = 3.35$, $p = .013$, $\eta^2 = .04$. Defendants accused of assault were considered significantly more credible than those accused of fraud (Mean difference = .44, 95% *CI* =.09, .78, *p* = .013)(see Table D.6).

Table D.6

Means, Adjusted Means, Standard Deviations, and Standard Errors for Defendant Credibility based on Attractiveness with the BIDR-16 and RLAQ as Covariates

	M	SD	M	SE
Fraud	4.05	1.01	4.05	.12
Assault	4.49	1.21	4.49	.13

A two-way ANCOVA was performed to assess the effect of crime type and attractiveness on defendant credibility ratings while controlling for the BIDR-16 and RLAQ. Key assumptions were met, except for normality as assessed by a significant Shapiro-Wilk test. Examination of histograms and normal Q-Q plots revealed approximately normal distributions. There was not a significant interaction of attractiveness and crime type on defendant credibility ratings, $F(1, 156) = 1.13$, $p = .289$, partial $\eta^2 = .01$. There was a statistically significant main effect of crime type ($F(1, 156) = 6.67$, $p = .011$, $\eta^2 = .04$), such that those accused of fraud were considered significantly less credible than those accused of assault (Mean difference = .44, 95% *CI* =.10, .77, $p = .011$) . There was also a main effect of attractiveness ($F(1, 156) = 9.98$, $p = .002$, $\eta^2 = .06$), such that those who were unattractive were considered significantly more credible than those who were attractive (Mean difference = .54, 95% *CI* = .20, .87, $p = .002$). There were also significant pairwise comparisons between crime type for the attractive defendants, $F(1, 156) = 7.00$, p .015, partial $\eta^2 = .04$. Those attractive defendants accused of assault were considered significantly more credible than those accused of fraud (Mean difference = .62, 95% *CI* = .12, 1.11, $p = .015$). Additionally, those who were unattractive and accused of fraud were considered significantly more credible than those who were attractive and accused of fraud (Mean difference = .72, 95% *CI* =.26, 1.18, $p = .003$)(see Table D.7 for means, adjusted means and variances).

Table D.7

Means, Adjusted Means, Standard Deviations, and Standard Errors for Defendant Credibility for Attractiveness and Crime Type

	Unattractive		Attractive	
Crime Type	Fraud	Assault	Fraud	Assault
M	4.38	4.65	3.68	4.29
SD	.91	1.13	1.00	1.29
M_{adj}	4.39	4.65	3.67	4.29
SE	.16	.16	.17	.18

A three-way ANCOVA was conducted to determine the effect of attractiveness, gender, and crime type on defendant credibility ratings while controlling for the BIDR-16 and RLAQ. Key assumptions were met. There was not a statistically significant interaction between the independent variables after controlling for socially desirable responding and authoritarian legal attitudes, $F(1, 152) = 2.62$, $p = .108$, partial $\eta^2 = .02$. After controlling for the BIDR-16 and RLAQ, there were main effects of crime type and attractiveness (see Table D.8). Pairwise comparisons with a Bonferroni adjustment showed that those accused of assault were considered significantly more credible than those accused of fraud (Mean difference = .42, 95% *CI* = .08, .75, $p = .015$). Unattractive defendants were considered significantly more credible than attractive defendants (Mean difference = .46, 95% *CI* = .22, .89, $p = .001$). Credibility ratings were higher for attractive defendants accused of assault compared to those accused of fraud (Mean difference = .59, 95% *CI* = .10, 1.09, $p = .019$), and for unattractive defendants accused of fraud compared to attractive defendants accused of fraud (Mean difference = .73, 95% *CI* = .27, 1.20, $p = .002$). Males accused of assault were considered significantly more credible than those accused of fraud (Mean difference = .51, 95% *CI* = .04, 1.00, $p = .034$). Unattractive males were considered significantly more credible than their attractive counterparts (Mean difference = .51, 95% *CI* =.04, 1.00, $p = .034$), and unattractive females were significantly more credible than attractive females

(Mean difference = .60, 95% *CI* = .13, 1.07, *p* = .014). Unattractive males accused of fraud were considered significantly more credible than attractive males accused of fraud (Mean difference = .97, 95% CI= .31, 1.62, *p* = .004). Unattractive females accused of assault were considered significantly more credible than attractive females accused of assault (Mean difference = .70, 95% *CI* = .01, 1.39, *p* = .047). Means, standard deviations, adjusted means, and standard errors can be seen in Table D.9.

Table D.8

Three-Way ANCOVA for Defendant Credibility with the BIDR-16 and RLAQ as Covariates

	F ratio (1, 152)	η^2
Attractiveness	10.79**	.07
Crime type	6.00*	.04
Gender	2.57	.02
Attractiveness*Crime type	1.10	.01
Attractiveness*Gender	.07	.00
Crime type*Gender	.33	.00
Attractiveness*Crime type*Gender	2.62	.02

*$p < .05$, **$p < .01$

Table D.9

Means, Adjusted Means, Standard Deviations, and Standard Errors for Defendant Credibility by Attractiveness, Gender, and Crime Type while Controlling for Socially Desirable Responding and Authoritarian Legal Attitudes

		M	*SD*	M_{adj}	*SE*
Unattractive					
Fraud	Male	4.60	.94	4.62	.24
	Female	4.20	.87	4.21	.21
Assault	Male	4.65	1.39	4.67	.24
	Female	4.65	.89	4.62	.22
Attractive					
Fraud	Male	3.68	1.00	3.65	.23
	Female	3.67	1.03	3.71	.26
Assault	Male	4.61	1.20	4.62	.25
	Female	3.94	1.34	3.92	.27

Victim Credibility

Three one-way ANCOVAs were conducted to determine the effects of attractiveness, gender, and crime type on victim credibility ratings while controlling for the BIDR-16 and RLAQ. Key assumptions were met for the attractiveness analysis, except for the assumption of normality. Visual inspection of histograms and normal Q-Q plots showed approximately normal distributions. For the gender and crime type analyses, the BIDR-16 was not linearly related to victim credibility and was omitted from the analysis.

After controlling for the BIDR-16 and the RLAQ, attractiveness did not have a significant effect on victim credibility, $F(1, 158) = .27$, $p = .605$, partial $\eta^2 = .00$. After controlling for the RLAQ, gender did not have a significant effect on victim credibility ratings, $F(1, 161) = .58$, $p = .446$, partial $\eta^2 = .00$. After controlling for the RLAQ, crime type did have a significant effect on victim credibility ratings, $F(1, 161) = 7.56$, $p = .007$, partial $\eta^2 = .05$. Post-hoc analysis was performed with a Bonferroni adjustment. Victim credibility ratings were statistically significantly higher when defendants were accused of fraud compared to assault (Mean difference = .49, 95% *CI* = .14, .83, $p = .007$) (see Table D.10).

Table D.10

Adjusted and Unadjusted Means and Variability in Victim Credibility Ratings by Crime Type with the BIDR-16 and RLAQ as Covariates

	N	*M*	*SD*	M_{adj}	*SE*
Fraud	86	4.48	1.04	4.48	.12
Assault	78	3.99	1.24	3.99	.13

A two-way ANCOVA was conducted to determine the effect of attractiveness and crime type on victim credibility ratings while controlling for the BIDR-16 and RLAQ. All key assumptions were met. There was a statistically significant interaction between attractiveness and crime type after controlling for the BIDR-16 and RLAQ, $F(1, 156) = 6.65$,

p = .011, partial η^2 = .04. An analysis of simple main effects for attractiveness and crime type was performed with a Bonferroni adjustment. The effect of attractiveness in the assault condition was statistically significantly different than in the fraud condition (F(1, 156) = 4.79, p = .03, partial η^2 = .03), such that means were significantly higher for defendant credibility for attractive versus unattractive defendants (Mean difference = .56, 95% CI = .06, 1.06, p = .030). The effect of crime type was statistically significantly different in the unattractive condition (F(1, 156) = 13.76, p < .001, partial η^2 = .08), such that defendant credibility was significantly higher when unattractive defendants were accused of fraud versus assault (Mean difference = .88, 95% CI =.41, 1.35, p < .001)(see Table D.11 and Figures D.2 and D.3).

Table D.11

Means, Adjusted Means, Standard Deviations and Standard Errors for Victim Credibility by Attractiveness and Crime Type with BIDR-16 and RLAQ as Covariates

	M	SD	M_{adj}	SE
Unattractive				
Fraud	4.62	.91	4.63	.17
Assault	3.77	1.17	3.75	.17
Attractive				
Fraud	4.30	1.16	4.29	.18
Assault	4.29	1.29	4.31	.19

Figure D.2

Estimated Marginal Means of Victim Credibility by Attractiveness and Crime Type with the BIDR-16 and RLAQ as Covariates

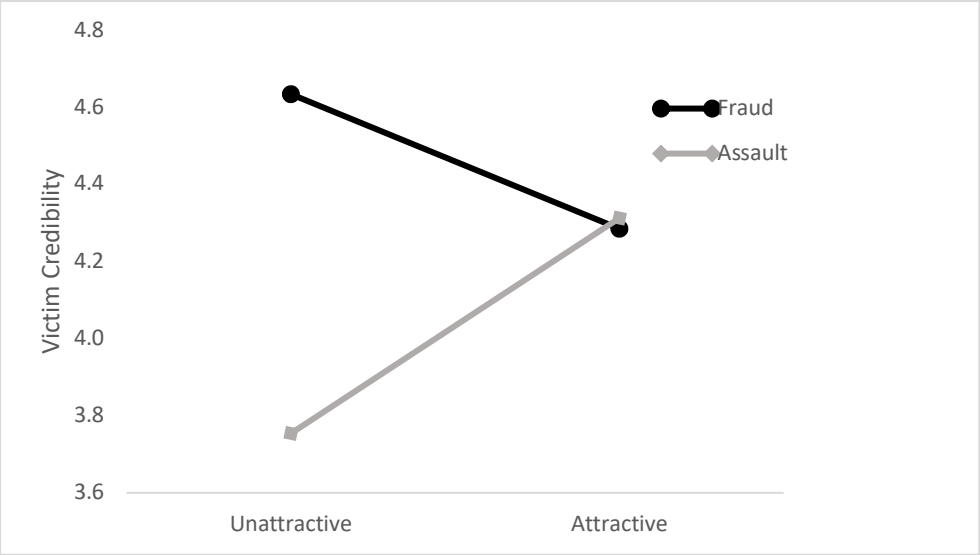


Figure D.3

Estimated Marginal Means of Victim Credibility by Crime Type and Attractiveness with the BIDR-16 and RLAQ as Covariates

A three-way ANCOVA was run to determine the effect of attractiveness, gender, and crime type on victim credibility ratings while controlling for the BIDR-16 and RLAQ. Key assumptions were met. There was not a statistically significant three-way interaction of attractiveness, gender, and crime type on victim credibility ratings after controlling for the

BIDR-16 and RLAQ, $F(1, 152) = .25$, $p = .620$, partial $\eta^2 = .00$). There was a significant two-way interaction between attractiveness and crime type when controlling for the BIDR-16 and RLAQ, $F(1, 152) = 6.16$, $p = .014$, partial $\eta^2 = .04$. Pairwise comparisons using a Bonferroni adjustment indicated ratings of victim credibility were significant when attractive defendants were accused of assault, compared to unattractive defendants, $F(1, 152) = 4.32$, $p = .039$, partial $\eta^2 = .03$ (Mean difference = .53, 95% *CI* = .03, 1.04, $p = .039$). When unattractive defendants were accused of fraud, victim credibility was significantly higher than when they were accused of assault, $F(1, 152) = 13.26$, $p < .001$, partial $\eta^2 = .08$ (Mean difference = .87, 95% *CI* = .40, 1.34, $p < .001$). When female defendants were accused of fraud, victim credibility was significantly higher than when they were accused of assault, $F(1, 152) = 7.15$, $p = .008$, partial $\eta^2 = .05$ (Mean difference = .67, 95% *CI* =.18, 1.17, $p = .008$). Victim credibility was also significantly higher when attractive versus unattractive males were accused of assault, $F(1, 152) = 5.21$, $p = .024$, partial $\eta^2 = .03$ (Mean difference = .82, 95% *CI* =.11, 1.54, $p = .024$). Victim credibility ratings were significantly higher when unattractive males were accused of fraud compared to assault, $F(1, 152) = 4.13$, $p = .044$, partial $\eta^2 = .03$ (Mean difference = .71, 95% *CI* =.02, 1.41, $p = .044$), and when unattractive females were accused of fraud compared to assault, $F(1, 152) = 10.06$, $p = .002$, partial $\eta^2 = .06$ (Mean difference = 1.02, 95% *CI* =.39, 1.66, $p = .002$)(see Table D.12 and Figures D.4 and D.5).

Table D.12

Means, Adjusted Means, Standard Deviations, and Standard Errors for Victim Credibility by Attractiveness, Gender, and Crime Type while Controlling for Socially Desirable Responding and Authoritarian Legal Attitudes

		M	*SD*	M_{adj}	*SE*
Unattractive					
Fraud	Male	4.55	.89	4.55	.25

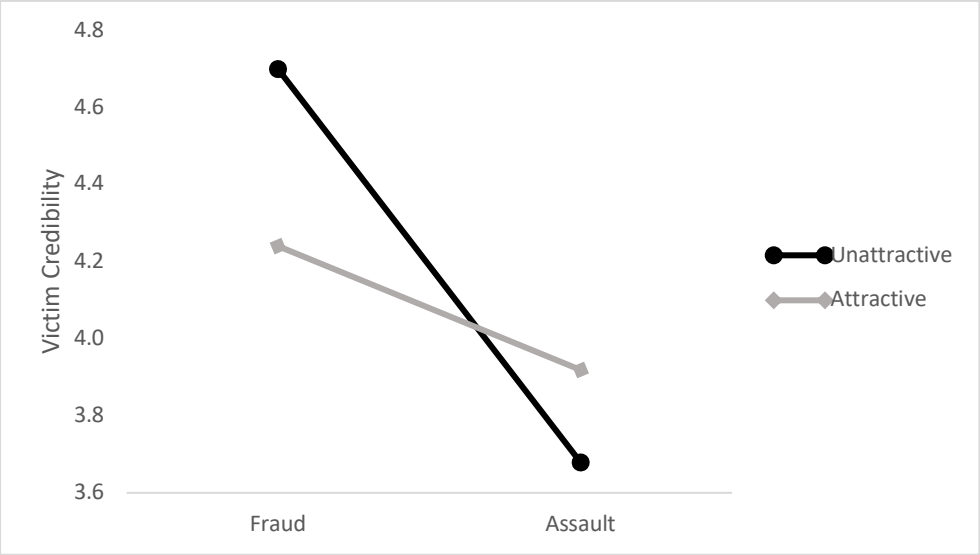
	Female	4.68	.95	4.70	.22
Assault	Male	3.85	1.39	3.84	.25
	Female	3.70	.97	3.68	.23
Attractive					
Fraud	Male	4.23	1.31	4.32	.24
	Female	4.39	.98	4.24	.27
Assault	Male	4.67	1.03	4.66	.26
	Female	3.88	1.45	3.92	.28

Figure D.4

Estimated Marginal Means of Victim Credibility of Males by Attractiveness and Crime Type with the BIDR-16 and RLAQ as Covariates

Figure D.5

Estimated Marginal Means of Victim Credibility of Females by Crime Type and Attractiveness with the BIDR-16 and RLAQ as Covariates


4.8
4.6
4.4
4.2
4.0
3.8
3.6
Victim Credibility
Fraud
Assault
Unattractive
Attractive

Appendix E: Exploratory Analyses for Study 2

Confidence

Three ANCOVAs were run to test the effect of attractiveness, gender, and psychiatric diagnosis on confidence ratings after controlling for the BIDR-16, RLAQ, and PPMI. The BIDR-16 and RLAQ were not linearly related to confidence ratings and were omitted from the analyses.

For the attractiveness analysis, key assumptions were met, except for two outliers and normality for five of the groups. Visual inspection of histograms revealed approximately normal distributions. The analysis was run with and without the outliers without significant changes in the results; therefore, the outliers were kept in the analysis. After controlling for the PPMI confidence ratings based on attractiveness did not significantly differ between transcript conditions, $F(1, 199) = .30$, $p = .584$, partial $\eta^2 = .00$.

For the gender analysis, key assumptions were met, except for normality for six of the groups. Visual inspection of histograms, and skewness and kurtosis z-scores within ±1.96 standard deviations suggested the data were approximately normally distributed. After controlling for the PPMI confidence ratings based on gender did not significantly differ between transcript conditions, $F(1, 199) = 2.17$, $p = .142$, partial $\eta^2 = .01$.

For the psychiatric diagnosis analysis, key assumptions were met, except for normality for six of the groups. Visual inspection of histograms, and skewness and kurtosis z-scores within ±1.96 standard deviations suggested the data were approximately normally distributed. After controlling for the PPMI confidence ratings based on psychiatric diagnosis did not significantly differ between transcript conditions, $F(3, 197) = 2.07$, $p = .105$, partial $\eta^2 = .03$.

A two-way ANCOVA was conducted to determine the effect of attractiveness and psychiatric diagnosis on confidence ratings while controlling for the BIDR-16, RLAQ, and PPMI. The BIDR-16 was not linearly related to confidence ratings at each level of the independent variables and was omitted from the analysis. Key assumptions were met, except for two outliers, which had an effect on the outcome and were removed from the analysis. After controlling for the RLAQ and PPMI, there was a significant effect of attractiveness and psychiatric diagnosis on confidence ratings, $F(3, 182) = 4.50$, $p = .004$, partial $\eta^2 = .07$. Means, adjusted means, and variances can be seen in Table E.1. An analysis of simple main effects with a Bonferroni adjustment indicated significant differences in the conduct disorder group at different levels of attractiveness, $F(1, 182) = 12.01$, $p < .001$, partial $\eta^2 = .07$. Participant confidence ratings were significantly lower when defendants with conduct disorder were attractive compared to unattractive (Mean difference = 1.26, 95% *CI* =.57, 1.96, $p < .001$) (see Figure E.1). There were also significant differences in the attractive group at different levels of psychiatric diagnosis, $F(3, 182) = 6.31$, $p < .001$, partial $\eta^2 = .09$. Participant confidence ratings were significantly lower when attractive defendants had conduct disorder compared to psychopathic traits (Mean difference = 1.44, 95% *CI* =.53, 2.35, $p < .001$), schizophrenia (Mean difference = 1.10, 95% *CI* =.09, 2.11, $p = .024$), and no diagnosis (Mean difference = 1.02, 95% *CI* =.05, 2.00, $p = .034$)(see Figure E.2).

Table E.1

Means, Adjusted Means and Variances of Confidence Ratings Based on Attractiveness and Psychiatric Diagnosis While Controlling for the RLAQ and PPMI

	M	*SD*	M_{adj}	*SE*
Attractive				
Conduct Disorder	4.00	1.00	3.98	.25
Psychopathic Traits	5.43	1.14	5.42	.23
Schizophrenia	5.05	1.18	5.08	.28
No Diagnosis	5.00	1.18	5.00	.26

Unattractive				
Conduct Disorder	5.24	1.39	5.24	.24
Psychopathic Traits	4.96	1.15	4.96	.24
Schizophrenia	5.22	1.31	5.24	.23
No Diagnosis	5.00	1.24	4.99	.25

Figure E.1

Estimated Marginal Means of Confidence Ratings by Attractiveness and Psychiatric Diagnosis

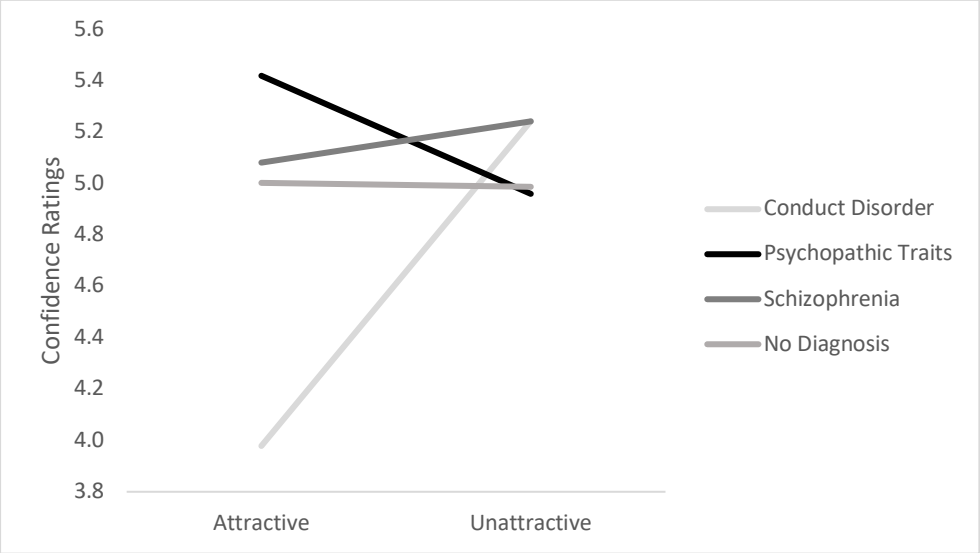
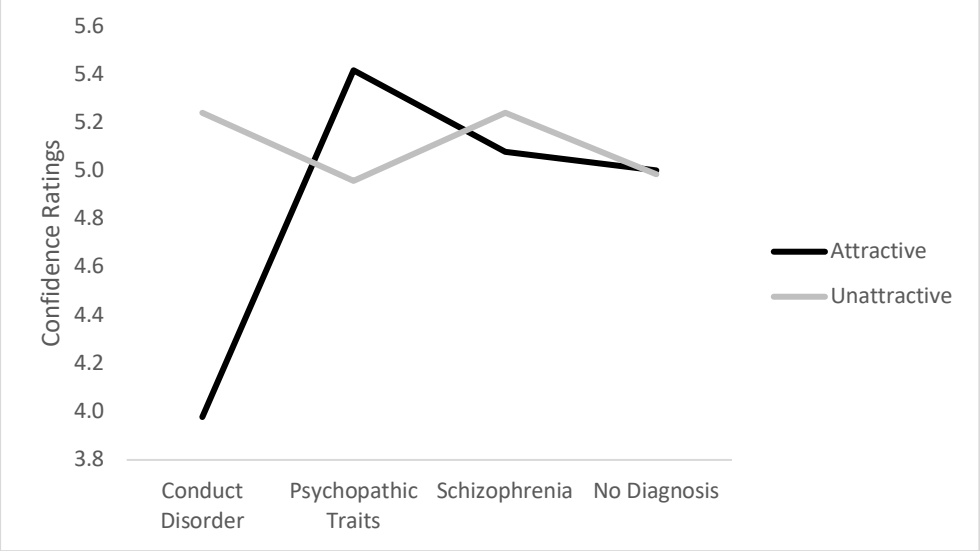


Figure E.2

Estimated Marginal Means of Confidence by Psychiatric Diagnosis and Attractiveness



A three-way ANCOVA was conducted to examine the effects of attractiveness, gender, and psychiatric diagnosis on confidence ratings after controlling for the BIDR-16, RLAQ, and PPMI. Key assumptions were met, except for two outliers, which had an effect on the outcome and were removed from the analysis. After controlling for the BIDR-16, RLAQ and PPMI, there was not a significant interaction between attractiveness, gender, and psychiatric diagnosis on ratings of confidence, $F(3, 166) = 1.62$, $p = .186$, partial $\eta^2 = .03$. There were significant interactions between gender and psychiatric diagnosis and attractiveness and psychiatric diagnosis (see Table E.2). See Table E.3 for means, adjusted means, and variances.

Table E.2

Three-Way ANCOVA for Confidence Ratings with the BIDR-16, RLAQ, and PPMI as Covariates

	F ratio	df	η^2
Attractiveness	2.83	1, 166	.02
Psychiatric Diagnosis	2.95*	3, 166	.05
Gender	1.14	1, 166	.01
Attractiveness * Psychiatric Diagnosis	3.92*	3, 166	.07
Attractiveness*Gender	.14	1, 166	.00
Psychiatric Diagnosis *Gender	5.88**	3, 166	.10
Attractiveness*Gender* Psychiatric Diagnosis	1.62	3, 166	.03

*$p < .05$, **$p < .001$

Table E.3

Adjusted and Unadjusted Means for Confidence Ratings Based on Attractiveness, Gender, and Psychiatric Diagnosis with the BIDR-16, RLAQ, and PPMI as Covariates

	N	M	SD	M_{adj}	SE
Attractive					
Female					
Conduct Disorder	11	4.09	1.22	4.08	.36
Psychopathic Traits	14	5.07	1.21	5.04	.31
Schizophrenia	9	4.67	1.00	4.67	.39
No Diagnosis	12	5.17	1.47	5.17	.34
Male					
Conduct Disorder	12	3.92	.79	3.88	.34

Psychopathic Traits	13	5.77	1.01	5.75	.32
Schizophrenia	10	5.40	1.26	5.47	.37
No Diagnosis	7	5.00	.58	4.90	.45
Unattractive					
Female					
Conduct Disorder	11	5.82	.98	5.85	.36
Psychopathic Traits	17	4.35	.86	4.41	.29
Schizophrenia	12	5.00	1.13	5.04	.34
No Diagnosis	10	5.10	1.45	5.10	.37
Male					
Conduct Disorder	13	4.69	1.55	4.72	.33
Psychopathic Traits	9	6.11	.60	6.11	.39
Schizophrenia	13	5.23	1.48	5.21	.32
No Diagnosis	12	4.92	1.16	4.87	.34

An analysis of simple main effects with a Bonferroni adjustment indicated significant differences in the conduct disorder group at different levels of attractiveness, $F(1, 166) = 14.57$, $p < .001$, partial $\eta^2 = .08$. Participants were more confident when defendants with conduct disorder were unattractive compared to attractive (Mean difference = 1.31, 95% *CI* =.63, 1.98, $p < .001$). There was a significant difference in the psychopathic traits group at different levels of gender, $F(1, 166) = 13.34$, $p < .001$, partial $\eta^2 = .07$. Confidence ratings were significantly higher when male defendants had psychopathic traits compared to female defendants (Mean difference = 1.21, 95% *CI* =.55, 1.86, $p < .001$).

There were also significant differences in the attractive group at different levels of psychiatric diagnosis, $F(3, 166) = 6.54$, $p < .001$, partial $\eta^2 = .11$. Participants were less confident when attractive defendants had conduct disorder compared to psychopathic traits (Mean difference = 1.41, 95% *CI* =.53, 2.30, $p < .001$), schizophrenia (Mean difference = 1.09, 95% *CI* =.11, 2.07, $p = .020$), or no diagnosis (Mean difference = 1.05, 95% *CI* =.07, 2.04, $p = .029$).

There were significant differences for male defendants at different levels of psychiatric diagnosis, $F(3, 166) = 8.04$, $p < .001$, partial $\eta^2 = .13$. Confidence ratings were

significantly higher when males had psychopathic traits compared to conduct disorder (Mean difference = 1.63, 95% *CI* =.71, 2.55, $p < .001$) or no diagnosis (Mean difference = 1.04, 95% *CI* =.04, 2.05, $p < .037$), and when they had schizophrenia compared to conduct disorder (Mean difference = 1.04, 95% *CI* =.13, 1.94, $p < .015$).

There were significant differences in confidence ratings for females with conduct disorder at different levels of attractiveness, $F(1, 166) = 12.48$, $p < .001$, partial $\eta^2 = .07$ (see Figure E.3). They also differed significantly for unattractive defendants with conduct disorder ($F(1, 166) = 5.41$, $p = .021$, partial $\eta^2 = .03$) and psychopathic traits ($F(1, 166) = 12.43$, $p < .001$, partial $\eta^2 = .07$) at different levels of gender. Confidence ratings were significantly higher for unattractive females with conduct disorder compared to their male counterparts (Mean difference = 1.14, 95% *CI* =.17, 2.10, $p = .021$), and for unattractive males with psychopathic traits compared to their female counterparts (Mean difference = 1.70, 95% *CI* =.75, 2.66, $p < .001$) (see Figure E.4). Confidence ratings also differed significantly for attractive males at different levels of psychiatric diagnosis, and unattractive males and females at different levels of psychiatric diagnosis. Confidence ratings were lower when attractive males had conduct disorder compared to psychopathic traits (Mean difference = 1.86, 95% =.62, 3.11, $p < .001$) and schizophrenia (Mean difference = 1.59, 95% *CI* =.24, 2.93, $p < .001$). Ratings were higher when unattractive females had conduct disorder compared to psychopathic traits (Mean difference = 1.44, 95% *CI* =.24, 2.65, $p = .010$), and when unattractive males had psychopathic traits compared to conduct disorder (Mean difference = 1.40, 95% *CI* =.05, 2.75, $p = .039$) (see Figure E.5).

Figure E.3

Estimated Marginal Means of Confidence for Female Defendants

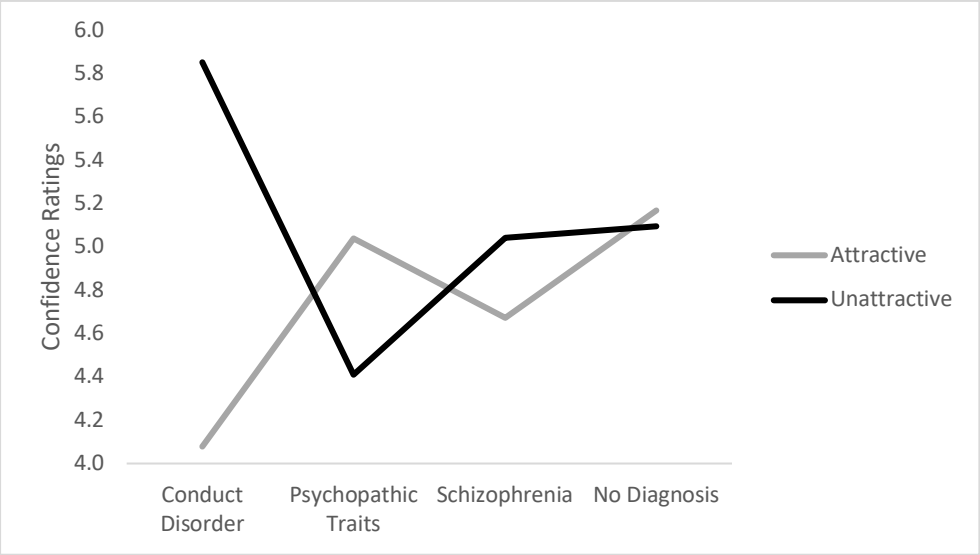


Figure E.4

Estimated Marginal Means of Confidence for Unattractive Defendants

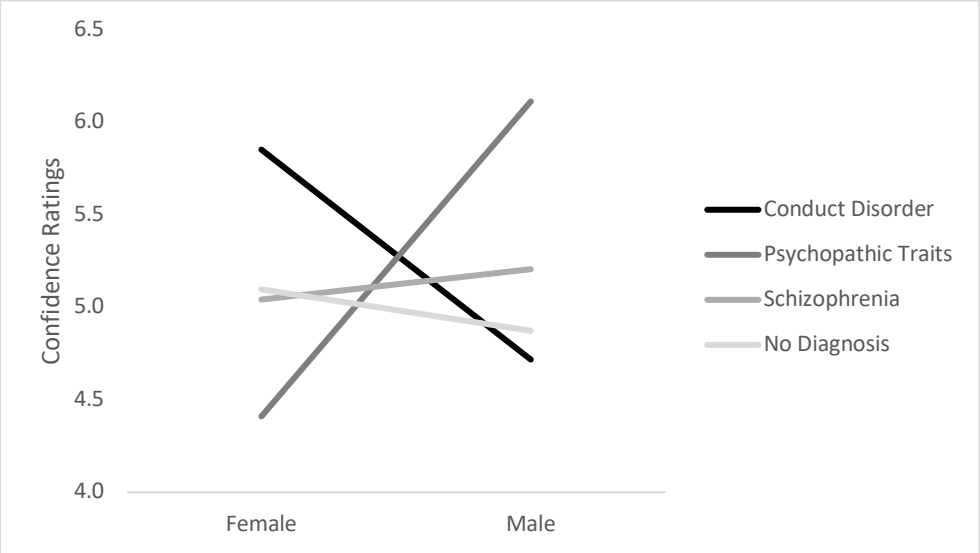
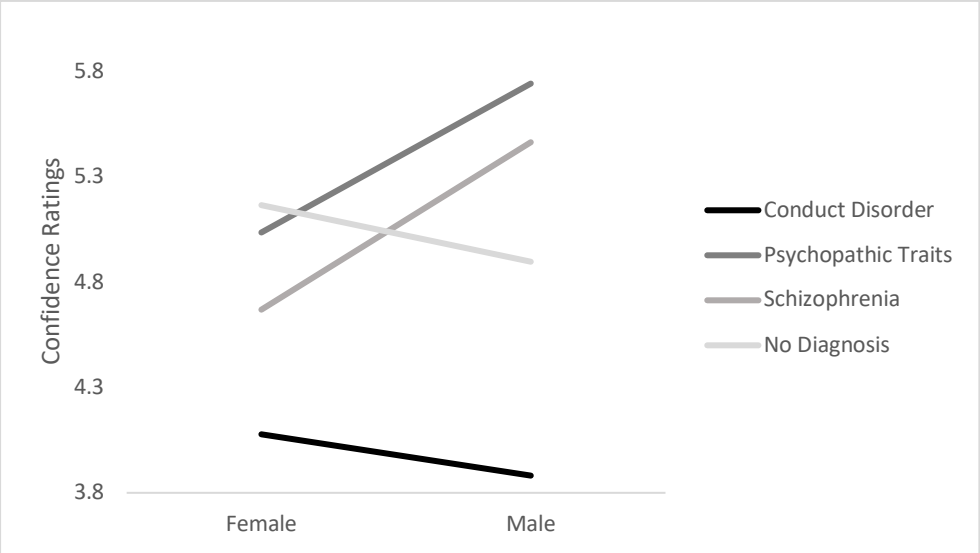


Figure E.5

Estimated Marginal Means of Confidence for Attractive Defendants



Defendant Credibility

Three one-way ANCOVAs were run to determine the effect of attractiveness, gender, and psychiatric diagnosis on defendant credibility ratings after controlling for socially desirable responding (BIDR-16), authoritarian legal attitudes (RLAQ), and mental illness prejudice (PPMI). For the attractiveness analysis, key assumptions were met. After controlling for the BIDR-16, RLAQ, and PPMI, attractiveness did not have an effect on defendant credibility ratings, $F(1, 182) = 1.07$, $p = .303$, partial $\eta^2 = .01$.

For the gender analysis, key assumptions were met. After controlling for the BIDR-16, RLAQ, and PPMI, gender did not have an effect on defendant credibility ratings, $F(1, 182) = 2.65$, $p = .105$, partial $\eta^2 = .01$.

For the psychiatric diagnosis analysis, homogeneity of variances was violated, and a square root transformation was applied, which resolved the heterogeneity of variances. Normality was violated for two of the groups, however, visual inspection of histograms showed approximately normal distributions. All other key assumptions were met. After controlling for the BIDR-16, RLAQ, and PPMI, psychiatric diagnosis had a statistically

significant effect on defendant credibility ratings, $F(3, 180) = 5.78$, $p < .001$, partial $\eta^2 = .09$. Defendant credibility was significantly higher for defendants with no diagnosis ($M = 4.57$, $SD = 1.17$), than for those with psychopathic traits ($M = 3.58$, $SD = .85$; Mean difference = .95, 95% CI =.34, 1.57, $p < .001$).

A two-way ANCOVA was run to test the effect of attractiveness and psychiatric diagnosis on ratings of defendant credibility while controlling for the BIDR-16, RLAQ, and PPMI. Key assumptions were met, except for normality for one of the groups. Skewness and kurtosis z-scores were within ±1.96 standard deviations suggesting the data were approximately normally distributed. There was not a statistically significant interaction of attractiveness and psychiatric diagnosis on defendant credibility ratings, $F(3, 176) = .93$, $p = .429$, partial $\eta^2 = .02$. There was a statistically significant main effect of psychiatric diagnosis, $F(3, 176) = 5.54$, $p = .001$, $\eta^2 = .09$. Those with no diagnosis were considered significantly more credible than those with psychopathic traits (Mean difference = .93, 95% CI =.32, 1.55, $p < .001$). See Table E.4 for means, adjusted means, and variances.

Table E.4

Means, Adjusted Means, Standard Deviations and Standard Errors for Defendant Credibility by Attractiveness and Psychiatric Diagnosis with BIDR-16, RLAQ and PPMI as Covariates

	M	*SD*	M_{adj}	*SE*
Attractive				
Conduct Disorder	4.13	.87	4.14	.23
Psychopathic Traits	3.59	1.42	3.57	.21
Schizophrenia	3.95	1.31	3.92	.26
No Diagnosis	4.37	1.16	4.33	.26
Unattractive				
Conduct Disorder	3.83	.82	3.87	.23
Psychopathic Traits	3.58	.90	3.63	.22
Schizophrenia	4.35	1.20	4.32	.22
No Diagnosis	4.74	1.18	4.73	.23

A three-way ANCOVA was conducted to test the effect of attractiveness, gender, and psychiatric diagnosis on defendant credibility ratings while controlling for the BIDR-16, RLAQ, and PPMI. See Table E.5 for means, adjusted means, and variances. There was not a significant three-way interaction of attractiveness, gender, and psychiatric diagnosis on ratings of defendant credibility, $F(3, 168) = .26$, $p = .851$, partial $\eta^2 = .01$. There were no significant two-way interactions, however, there was a main effect of psychiatric diagnosis (see Table E.6).

Table E.5

Adjusted and Unadjusted Means for Defendant Credibility Ratings Based on Attractiveness, Gender, and Psychiatric Diagnosis with the BIDR-16, RLAQ, and PPMI as Covariates

	N	*M*	*SD*	M_{adj}	*SE*
Attractive					
Female					
Conduct Disorder	11	4.18	.87	4.22	.34
Psychopathic Traits	14	3.86	1.23	3.81	.30
Schizophrenia	9	4.11	1.36	4.01	.37
No Diagnosis	12	4.58	1.08	4.52	.33
Male					
Conduct Disorder	12	4.08	.90	4.07	.33
Psychopathic Traits	13	3.31	1.60	3.29	.31
Schizophrenia	10	3.80	1.32	3.85	.35
No Diagnosis	7	4.00	1.29	4.00	.43
Unattractive					
Female					
Conduct Disorder	11	3.64	.92	3.79	.34
Psychopathic Traits	17	3.71	.85	3.80	.27
Schizophrenia	13	4.23	1.24	4.26	.31
No Diagnosis	11	5.18	1.08	5.24	.34
Male					
Conduct Disorder	13	4.00	.71	3.95	.31
Psychopathic Traits	9	3.33	1.00	3.33	.37
Schizophrenia	13	4.46	1.20	4.38	.31
No Diagnosis	12	4.33	1.15	4.26	.32

Table E.6

Three-Way ANCOVA for Defendant Credibility Ratings with the BIDR-16, RLAQ, and PPMI as Covariates

	F ratio	*df*	Partial η^2
Attractiveness	.85	1, 168	.01
Psychiatric Diagnosis	5.54*	3, 168	.09
Gender	3.36	1, 168	.02
Attractiveness * Psychiatric Diagnosis	1.10	3, 168	.02
Attractiveness*Gender	.02	1, 168	.00
Psychiatric Diagnosis *Gender	1.15	3, 168	.02
Attractiveness*Gender* Psychiatric Diagnosis	.26	3, 168	.01

*$p < .01$

Although the interaction was not significant, pairwise comparisons with a Bonferroni adjustment showed some significant differences. Those with no diagnosis were considered significantly more credible than those with psychopathic traits (Mean difference = .95, 95% *CI* =.32, 1.58, $p < .001$). Credibility ratings were significantly different for unattractive defendants at different levels of psychiatric diagnosis, $F(3, 168) = 5.04$, $p = .002$, partial $\eta^2 = .08$. For unattractive defendants, those with no diagnosis were considered significantly more credible than those with conduct disorder (Mean difference = .88, 95% *CI* =.01, 1.75, $p = .046$) or psychopathic traits (Mean difference = 1.18, 95% *CI* =.31, 2.06, $p = .002$). Ratings were also significantly different for defendants with no diagnosis at different levels of gender, $F(1, 168) = 4.41$, $p = .037$, partial $\eta^2 = .03$. Females with no diagnosis were considered significantly more credible than their male counterparts (Mean difference = .75, 95% *CI* =.05, 1.45, $p = .037$). For unattractive defendants with no diagnosis, credibility ratings were significantly different between females and males, $F(1, 168) = 4.38$, $p = .038$, partial $\eta^2 = .03$. Females were significantly more credible than males (Mean difference = .98, 95% *CI* =.06, 1.90, $p = .038$). Finally, unattractive females had significantly different credibility ratings at different levels of psychiatric diagnosis, $F(3, 168) = 4.39$, $p = .005$, partial $\eta^2 = .07$. Those with no diagnosis were considered significantly more credible than those with conduct

disorder (Mean difference = 1.45, 95% *CI* =.18, 2.72, p = .017) or psychopathic traits (Mean difference = 1.44, 95% *CI* =.28, 2.60, p = .006).

Victim Credibility

Three one-way ANCOVAs were run to determine the effect of attractiveness, gender, and psychiatric diagnosis on defendant credibility ratings after controlling for the BIDR-16, RLAQ, and PPMI. The BIDR-16 and RLAQ did not have linear relationships with the dependent variable and were omitted from the analyses. For the attractiveness analysis, key assumptions were met except for normality. Visual inspection of histograms and normal Q-Q plots showed approximately normal distributions. After controlling for the PPMI, attractiveness did not have a significant effect on victim credibility ratings, $F(1, 198) = 3.58$, p = .060, partial η^2 = .02.

For the gender analysis, key assumptions were met except for normality. Visual inspection of histograms and normal Q-Q plots showed approximately normal distributions. After controlling for the PPMI, gender did not have a significant effect on victim credibility ratings, $F(1, 198) = .001$, p = .976, partial η^2 = .00.

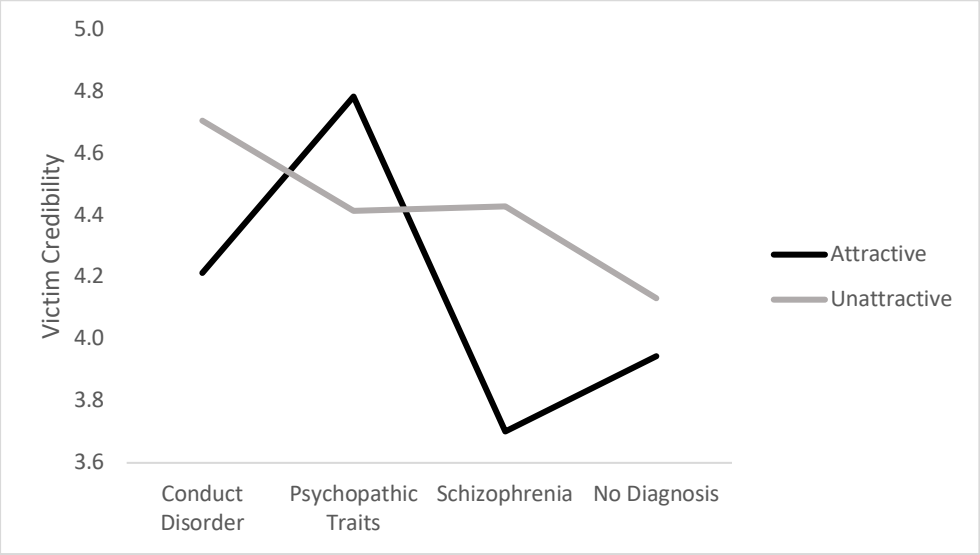
For the psychiatric diagnosis analysis, key assumptions were met except for normality of three groups. Visual inspection of histograms and normal Q-Q plots showed approximately normal distributions. After controlling for the PPMI, psychiatric diagnosis did not have a significant effect on victim credibility ratings, $F(1, 198) = 2.56$, p = .056, partial η^2 = .04.

A two-way ANCOVA was conducted to determine the effect of attractiveness and psychiatric diagnosis on victim credibility ratings after controlling for the BIDR-16, RLAQ and PPMI. Key assumptions were met except for normality of three groups. Visual inspection of histograms showed approximately normal distributions, and skewness and kurtosis z-

scores were within ± 3 z-scores, suggesting the data were approximately normally distributed. After controlling for the BIDR-16, RLAQ, and PPMI, there was not a significant interaction between attractiveness and psychiatric diagnosis on victim credibility ratings, $F(3, 175) = 1.61$, $p = .188$, partial $\eta^2 = .03$. However, pairwise comparisons with a Bonferroni adjustment showed a significant difference for attractive defendants at different levels of psychiatric diagnosis ($F(3, 175) = 2.94$, $p = .034$, partial $\eta^2 = .05$). When comparing attractive defendants with psychopathic traits ($M = 4.78$, $SD = 1.34$) to attractive defendants with schizophrenia ($M = 3.68$, $SD = 1.38$), victims were considered significantly more credible than when defendants had schizophrenia (Mean difference =1.08, 95% CI =.03, 2.13, $p = .039$) (see Figure E.6).

Figure E.6

Estimated Marginal Means of Victim Credibility by Attractiveness and Psychiatric Diagnosis



A three-way ANCOVA was conducted to evaluate the interaction effect of attractiveness, gender, and psychiatric diagnosis on victim credibility ratings after controlling for the BIDR-16, RLAQ, and PPMI. Key assumptions were met, except for normality for three groups. Visual inspection of histograms, and skewness and kurtosis z-values indicated

approximately normal distributions. After controlling for the BIDR-16, RLAQ, and PPMI, there was not a significant interaction effect of attractiveness, gender, and psychiatric diagnosis on victim credibility ratings, $F(3, 167) = .59$, $p = .626$, partial $\eta^2 = .01$. An analysis of main effects was also not significant for any of the independent variables (see Table E.7).

Table E.7

Three-Way ANCOVA for Victim Credibility Ratings with the BIDR-16, RLAQ, and PPMI as Covariates

	F ratio	df	η^2
Attractiveness	1.96	1, 167	.01
Psychiatric Diagnosis	2.30	3, 167	.04
Gender	.15	1, 167	.00
Attractiveness * Psychiatric Diagnosis	1.43	3, 167	.03
Attractiveness*Gender	.93	1, 167	.01
Psychiatric Diagnosis *Gender	.43	3, 167	.01
Attractiveness*Gender* Psychiatric Diagnosis	.59	3, 167	.01

Likelihood of Recommending Treatment

Three one-way ANCOVAs were run to determine the effect of attractiveness, gender, and psychiatric diagnosis on the likelihood of recommending treatment after controlling for the BIDR-16, RLAQ, and PPMI. The BIDR-16 and RLAQ did not have linear relationship with attractiveness, and the RLAQ did not have a linear relationship with psychiatric diagnosis and were omitted from those analyses. For all of the analyses, normality was violated. Visual inspection of histograms and normal Q-Q plots showed approximately normal distributions. After controlling for the PPMI, attractiveness did not have a significant effect on likelihood of recommending treatment ratings, $F(1, 199) = 2.26$, $p = .134$, partial $\eta^2 = .01$. After controlling for the BIDR-16, RLAQ and PPMI, gender did not have a significant effect on likelihood of recommending treatment ratings, $F(1, 182) = 1.76$, $p = .286$, partial $\eta^2 = .01$.

For the psychiatric diagnosis analysis, homogeneity of variances was violated. A Log10 transformation was applied which resolved the heterogeneity of variances. After controlling for the BIDR-16 and PPMI, psychiatric diagnosis had a significant effect on the likelihood of recommending treatment, $F(3, 189) = 9.91$, $p < .001$, partial $\eta^2 = .14$. See means, adjusted means, and variances in Table E.8. Pairwise comparisons with a Bonferroni adjustment showed that participants were significantly less likely to recommend treatment to defendants with no diagnosis compared to those with conduct disorder (Mean difference = 1.46, 95% *CI* =.64, 2.27, $p < .001$), psychopathic traits (Mean difference = 1.69, 95% *CI* =.90, 2.49, $p < .001$) or schizophrenia (Mean difference = 1.20, 95% *CI* =.36, 2.03, $p = .004$).

Table E.8

Means, Adjusted Means, Standard Deviations, and Standard Errors for Likelihood of Recommending Treatment by Psychiatric Diagnosis with the BIDR-16 and PPMI as Covariates

	M	*SD*	M_{adj}	*SE*
Conduct Disorder	5.53	1.45	5.50	.21
Psychopathic Traits	5.76	1.16	5.74	.20
Schizophrenia	5.17	1.64	5.24	.22
No Diagnosis	4.04	1.83	4.04	.22

A two-way ANCOVA was run to test the interaction of attractiveness and psychiatric diagnosis on the likelihood of recommending treatment after controlling for the BIDR-16, RLAQ, and PPMI. Homogeneity of variances was violated and a Log10 transformation was applied, which resolved the heterogeneity of variances. Normality was violated for two of the groups. Visual inspection of histograms and skewness and kurtosis z-scores indicated the distributions were approximately normal. After controlling for the BIDR-16, RLAQ, and PPMI, there was not a significant interaction between attractiveness and psychiatric diagnosis on the likelihood of recommending treatment, $F(3, 176) = 1.68$, $p = .173$, partial $\eta^2 = .03$. An analysis of main effects showed a significant effect of psychiatric diagnosis (see Table E.9).

Pairwise comparisons with a Bonferroni adjustment showed that participants were significantly less likely to recommend treatment to defendants with no diagnosis than those with conduct disorder (Mean difference = 1.45, 95% *CI* =.61, 2.29, $p < .001$), psychopathic traits (Mean difference = 1.68, 95% *CI* =.87, 2.50, $p < .001$), or schizophrenia (Mean difference = 1.17, 95% *CI* =.31, 2.03, $p = .009$). There were also significant differences in ratings between psychiatric diagnoses for attractive defendants, $F(3, 176) = 7.84$, $p < 001$, partial $\eta^2 = .12$. Attractive defendants with no diagnosis were significantly less likely to receive a treatment recommendation than those with conduct disorder (Mean difference = 1.67, 95% *CI* =.44, 2.89, $p = .010$) and psychopathic traits (Mean difference = 2.38, 95% *CI* =1.20, 3.56, $p < .001$). See Table E.10 for means, adjusted means, and variances.

Table E.9

Two-Way ANCOVA for Likelihood of Recommending Treatment with the BIDR-16, RLAQ, and PPMI as Covariates

	F ratio	*df*	η^2
Attractiveness	.95	1, 176	.01
Psychiatric Diagnosis	8.78*	3, 176	.13
Attractiveness * Psychiatric Diagnosis	1.68	3, 176	.03

*$p < .001$

Table E.10

Means, Adjusted Means, Standard Deviations, and Standard Errors for Likelihood of Recommending Treatment by Attractiveness and Psychiatric Diagnosis with the BIDR-16, RLAQ and PPMI as Covariates

	M	*SD*	M_{adj}	*SE*
Attractive				
Conduct Disorder	5.26	1.76	5.20	.31
Psychopathic Traits	5.89	1.25	5.92	.28
Schizophrenia	5.00	1.56	5.09	.34
No Diagnosis	3.53	1.84	3.54	.34
Unattractive				
Conduct Disorder	5.79	1.02	5.76	.30
Psychopathic Traits	5.58	1.06	5.51	.29
Schizophrenia	5.27	1.73	5.31	.29

No Diagnosis	4.52	1.70	4.52	.31

A three-way ANCOVA was conducted to evaluate the interaction effect of attractiveness, gender, and psychiatric diagnosis on the likelihood of recommending treatment after controlling for the BIDR-16, RLAQ, and PPMI. Homogeneity of variances was violated, and a Log10 transformation was applied, which resolved the violation. The assumption of normality was violated for three groups. Visual inspection of histograms, and skewness and kurtosis z-values indicated approximately normal distributions. After controlling for the BIDR-16, RLAQ, and PPMI, there was a significant interaction effect of attractiveness, gender, and psychiatric diagnosis, $F(3, 177) = 3.12$, $p = .027$, partial $\eta^2 = .05$. See Table E.11 for means, adjusted means, and variances. Pairwise comparisons with a Bonferroni correction showed, a significant effect of attractiveness for males with no diagnosis, $F(1, 77) = 6.40$, $p = .012$, partial $\eta^2 = .04$. Those who were attractive were significantly less likely to receive a treatment recommendation (Mean difference = 2.04, 95% *CI* =.74, 3.35, $p = .012$). There was also a significant effect of psychiatric diagnosis and gender for attractive defendants, ($F(1, 177) = 4.73$, $p = .031$, partial $\eta^2 = .03$) and unattractive defendants ($F(1, 177) = 4.03$, $p = .46$, partial $\eta^2 = .02$). Attractive females with schizophrenia were significantly less likely to receive a treatment recommendation than their male counterparts (Mean difference = 1.30, 95% *CI*=.02, 2.59, $p = .031$). Unattractive females with psychopathic traits were significantly less likely to receive a treatment recommendation than their male counterparts (Mean difference = 1.01, 95% *CI* =.16, 2.18, $p = .046$). Attractive males with no diagnosis were also significantly less likely to be recommended treatment than those with a psychiatric diagnosis, $F(3, 177) = 8.70$, $p < .001$, partial $\eta^2 = .13$. Those with no diagnosis had lower mean scores than those with conduct disorder (Mean difference = 2.10,

95% *CI* =.39, 3.82, *p* = .007), psychopathic traits (Mean difference = 3.36, 95% *CI* =1.62, 5.09, *p* < .001), and schizophrenia (Mean difference = 3.00, 95% *CI* =1.16, 4.84, *p* < .001).

Table E.11

Adjusted and Unadjusted Means for Treatment Recommendation Based on Attractiveness, Gender, and Psychiatric Diagnosis with the BIDR-16, RLAQ, and PPMI as Covariates

	N	*M*	*SD*	M_{adj}	*SE*
Attractive					
Female					
Conduct Disorder	11	5.82	1.08	5.71	.44
Psychopathic Traits	14	5.64	1.22	5.69	.39
Schizophrenia	10	4.30	1.49	4.49	.46
No Diagnosis	14	3.86	2.11	3.94	.39
Male					
Conduct Disorder	14	4.86	2.07	4.90	.39
Psychopathic Traits	13	6.15	1.28	6.15	.40
Schizophrenia	10	5.80	1.23	5.80	.46
No Diagnosis	8	2.88	1.25	2.80	.51
Unattractive					
Female					
Conduct Disorder	11	5.73	.90	5.45	.44
Psychopathic Traits	18	5.33	1.08	5.22	.34
Schizophrenia	13	5.54	1.61	5.49	.40
No Diagnosis	12	4.33	1.78	4.19	.42
Male					
Conduct Disorder	13	5.85	1.14	6.00	.40
Psychopathic Traits	9	6.22	.83	6.23	.48
Schizophrenia	13	5.00	1.87	5.13	.40
No Diagnosis	12	4.75	1.60	4.84	.42

Treatment Benefit

Three one-way ANCOVAs were run to test the effect of attractiveness, gender, and psychiatric diagnosis on treatment benefit ratings while controlling for the BIDR-16, RLAQ, and PPMI. Key assumptions were met for the attractiveness and gender analyses, except for the assumption of normality. Visual inspection of histograms and normal Q-Q plots showed approximately normal distributions. Homogeneity of variances was violated for the

psychiatric diagnosis analysis, and a Log10 transformation was applied, which resolved the heterogeneity of variances. All other key assumptions were met.

There was not a significant effect of attractiveness on the benefit of treatment after controlling for the BIDR-16, RLAQ, and PPMI, $F(1, 182) = 1.96$, $p = .163$, partial $\eta^2 = .01$. There was also not a significant effect of gender, $F(1, 1, 182) = 1.13$, $p = .289$, partial $\eta^2 = .01$. There was a significant effect of psychiatric diagnosis on the benefit of treatment, $F(3, 180) = 3.93$, $p = .010$, partial $\eta^2 = .06$. Those with no diagnosis were considered significantly less likely to benefit from treatment compared to those with conduct disorder (Mean difference = 1.05, 95% *CI* =.23, 1.88, $p = .019$) and schizophrenia (Mean difference = 1.04, 95% *CI* =.21, 1.88, $p = .025$) (see Table E.12).

Table E.12

Means, Adjusted Means, Standard Deviations, and Standard Errors for Treatment Benefit by Psychiatric Diagnosis with the BIDR-16, RLAQ, and PPMI as Covariates

	M	*SD*	M_{adj}	*SE*
Conduct Disorder	5.43	1.43	5.38	.21
Psychopathic Traits	5.36	1.18	5.34	.20
Schizophrenia	5.31	1.50	5.38	.22
No Diagnosis	4.33	1.86	4.33	.23

A two-way ANCOVA was conducted to determine the effect of attractiveness and psychiatric diagnosis on treatment benefit ratings after controlling for the BIDR-16, RLAQ, and PPMI. Homogeneity of variances was violated and a Log10 transformation was applied, which resolved the heterogeneity of variances. Overall residuals were not normally distributed; however, visual inspection of histograms showed an approximately normal distribution. All other key assumptions were met. After controlling for the covariates, there was not a significant interaction of attractiveness and psychiatric diagnosis on treatment benefit ratings, $F(3, 176) = 2.57$, $p = .056$, partial $\eta^2 = .04$. There was a significant main

effect of psychiatric diagnosis (see Table E.13). Pairwise comparisons with a Bonferroni correction showed that those with no diagnosis were considered significantly less likely to benefit from treatment than those with conduct disorder (Mean difference = 1.10, 95% *CI* =.29, 1.91, $p = .011$), psychopathic traits (Mean difference = 1.07, 95% *CI* =.28, 1.85, $p = .037$), or schizophrenia (Mean difference = 1.10, 95% *CI* =.27, 1.92, $p = .015$). Pairwise comparisons also showed that attractive defendants with no diagnosis were significantly less likely to benefit from treatment than unattractive defendants with no diagnosis, $F(1, 176) = 5.27$, $p = .023$, partial $\eta^2 = .03$ (Mean difference = 1.18, 95% *CI* =.31, 2.05, $p = .023$). Attractive defendants with no diagnosis were also significantly less likely to benefit from treatment than attractive defendants with psychopathic traits (Mean difference = 1.93, 95% *CI* =.79, 3.07, $p = .002$) or schizophrenia (Mean difference = 1.63, 95% *CI* =.39, 2.87, $p = .018$), $F(3, 176) = 4.96$, $p = .002$, partial $\eta^2 = .08$.

Table E.13

Two-Way ANCOVA for Treatment Benefit with the BIDR-16, RLAQ, and PPMI as Covariates

	F ratio	*df*	η^2
Attractiveness	2.77	1, 176	.02
Psychiatric Diagnosis	4.42*	3, 176	.07
Attractiveness * Psychiatric Diagnosis	2.57	3, 176	.04

*$p < .01$

A three-way ANCOVA was run to determine the effect of attractiveness, gender, and psychiatric diagnosis on treatment benefit ratings while controlling for the BIDR-16, RLAQ, and PPMI. The assumption of linearity was violated for the BIDR-16 and RLAQ which were omitted from the analysis. Homogeneity of variances was violated and a Log10 transformation was applied which resolved it. All other assumptions were met. There was a significant interaction between attractiveness, gender, and psychiatric diagnosis while controlling for the PPMI (see Table E.14). See Table E.15 for means, adjusted means, and

variances. Pairwise comparisons with a Bonferroni correction showed significantly lower treatment benefit for attractive females with schizophrenia (Mean difference = 1.40, p = .022, CI = .25, 2.55) compared to unattractive females with schizophrenia, $F(1, 185) = 5.300$, p = .022, partial η^2 = .03. Attractive males with conduct disorder (Mean difference = 1.48, p = .008, CI = .44, 2.53) were considered significantly less likely to benefit from treatment than their unattractive counterparts, $F(1, 185) = 7.16$, p = .008, partial η^2 = .04. Attractive males with no diagnosis (Mean difference = 2.08, p = .004, CI = .86, 3.30) were also considered less likely to benefit from treatment than their unattractive counterparts, $F(1, 185) = 8.50$, p = .004, partial η^2 = .04. Treatment benefit scores were lower for attractive females with schizophrenia (Mean difference = 1.60, p = .009, CI =.38, 2.81) compared to their male counterparts, $F(1, 185) = 6.93$, p = .009, partial η^2 = .04. Scores were also lower for attractive males with no diagnosis than those with psychopathic traits (Mean difference = 2.73, p < .001, CI = 1.10, 4.36), or schizophrenia (Mean difference = 3.25, p < .001, CI = 1.50, 5.00), $F(3, 185) = 8.29$, p < .001, partial η^2 = .12.

Table E.14

Three-Way ANCOVA for Treatment Benefit with the PPMI as Covariate

	F ratio	df	η^2
Attractiveness	5.23*	1, 185	.03
Psychiatric Diagnosis	7, 49***	3, 185	.11
Gender	.89	1, 185	.01
Attractiveness * Psychiatric Diagnosis	1.89	3, 185	.03
Attractiveness*Gender	.10	1, 185	.00
Psychiatric Diagnosis *Gender	1.07	3, 185	.02
Attractiveness*Gender* Psychiatric Diagnosis	4.68**	3, 185	.07

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

Table E.15

Adjusted and Unadjusted Means for Treatment Benefit Based on Attractiveness, Gender, and Psychiatric Diagnosis with the PPMI as Covariate

	N	*M*	*SD*	M_{adj}	*SE*
Attractive					
Female					
Conduct Disorder	11	5.55	1.13	5.40	.42
Psychopathic Traits	14	5.36	.74	5.35	.37
Schizophrenia	10	4.40	1.58	4.56	.44
No Diagnosis	16	4.06	1.88	4.15	.35
Male					
Conduct Disorder	14	4.57	1.83	4.58	.37
Psychopathic Traits	14	5.64	1.34	5.64	.37
Schizophrenia	10	6.10	.74	6.16	.44
No Diagnosis	8	3.00	1.20	2.91	.49
Unattractive					
Female					
Conduct Disorder	12	5.58	1.08	5.45	.40
Psychopathic Traits	18	5.06	1.47	5.00	.32
Schizophrenia	13	6.00	.71	5.96	.38
No Diagnosis	12	4.58	1.88	4.47	.40
Male					
Conduct Disorder	13	5.92	1.04	6.07	.38
Psychopathic Traits	9	5.44	1.13	5.46	.46
Schizophrenia	15	5.07	1.79	5.17	.36
No Diagnosis	13	5.00	1.58	4.99	.38

Treatment Aimed at Reducing Violence

Three one-way ANCOVAs were run to test the effect of attractiveness, gender, and psychiatric diagnosis on the effectiveness of a treatment aimed at reducing violence while controlling for the BIDR-16, RLAQ, and PPMI. The RLAQ violated the assumption of linearity for the attractiveness and psychiatric diagnosis analyses and was omitted from those. The assumption of normality was violated for all of the analyses. Visual inspection of histograms and normal Q-Q plots and skewness and kurtosis z-scores indicated approximately normal distributions. Homogeneity of variances was violated for the

psychiatric diagnosis analysis, and a Log10 transformation was applied, which resolved the heterogeneity of variances. All other key assumptions were met.

After controlling for the BIDR-16, PPMI, and RLAQ attractiveness and gender did not have significant effects on the effectiveness of a treatment aimed at reducing violence (see Table E.16). Psychiatric diagnosis did have a significant effect on ratings of the effectiveness of a treatment aimed at reducing violence. Pairwise comparisons with a Bonferroni correction indicated that participants considered defendants with no diagnosis to be less likely to benefit from a treatment aimed at reducing violence than those with conduct disorder (Mean difference = 1.02, $p = .050$, $CI = .18, 1.86$) or psychopathic traits (Mean difference = 1.08, $p = .017$, $CI = .19, 1.52$).

Table E.16

One-Way ANCOVAs for Effectiveness of a Treatment Aimed at Reducing Violence by Attractiveness, Gender, and Psychiatric Diagnosis with the BIDR-16, RLAQ, and PPMI as Covariates

	M	*SD*	M_{adj}	*SE*	*F* ratio	*df*	Partial η^2
Attractive	5.09	1.66	5.10	.16	.23	1, 191	.00
Unattractive	5.23	1.53	5.21	.16			
Female	5.02	1.58	4.99	.16	2.54	1, 182	.01
Male	5.33	1.55	5.36	.17			
Conduct disorder	5.49	1.33	5.47	.22	4.99*	3, 189	.07
Psychopathic traits	5.54	1.30	5.52	.21			
Schizophrenia	5.07	1.72	5.11	.23			
No diagnosis	4.46	1.81	4.45	.23			

*$p < .01$

A two-way ANCOVA was run to determine if attractiveness and psychiatric diagnosis had an interaction effect on ratings of the effectiveness of a treatment aimed at reducing violence while controlling for the PPMI. The BIDR-16 and RLAQ violated the assumption of linearity and were omitted. A Log10 transformation was applied due to a violation of

homogeneity of variances and all other key assumptions were met. After controlling for the PPMI, there was not a significant interaction of attractiveness and psychiatric diagnosis on treatment effectiveness ratings (see Table E.17). There was a significant main effect of psychiatric diagnosis. Pairwise comparisons with a Bonferroni correction showed that those with no diagnosis were considered significantly less likely to benefit from treatment than those with conduct disorder (Mean difference = 1.05, $p = .03$, $CI = .23, 1.88$) or psychopathic traits (Mean difference = 1.09, $p = .011$, $CI = .28, 1.89$).

Table E.17

Two-Way ANCOVA for the Effectiveness of a Treatment Aimed at Reducing Violence with the PPMI as Covariate

	F ratio	*df*	η^2
Attractiveness	.39	1, 193	.00
Psychiatric Diagnosis	4.15*	3, 193	.06
Attractiveness * Psychiatric Diagnosis	.40	3, 193	.01

*$p < .01$

A three-way ANCOVA was conducted with the PPMI as covariate. The BIDR-16 and RLAQ were omitted due to violations of linearity. Levene's statistic was significant, therefore, a Log10 transformation was applied which resolved the heterogeneity of variances. All other key assumptions were met. When controlling for the PPMI, attractiveness, gender, and psychiatric diagnosis did not have a significant effect on ratings of the effectiveness of a treatment aimed at reducing violence (see Table E.18). Means, adjusted means, and their variances can be found in Table E.19. Pairwise comparisons with a Bonferroni correction indicated a significant two-way interaction between attractiveness and gender for males, $F(1, 185) = 4.15$, $p = .043$, partial $\eta^2 = .02$. Attractive males were considered significantly less likely to benefit from a treatment aimed at reducing violence than unattractive males (Mean difference = .63, $p = .043$, $CI = .00, 1.25$).

Table E.18

Three-Way ANOVA Statistics for Treatment Aimed at Reducing Violence

	F ratio	*df*	Partial η^2
Attractiveness	.83	1, 185	.00
Psychiatric Diagnosis	5.53**	3, 185	.08
Gender	2.48	1, 185	.01
Attractiveness * Psychiatric Diagnosis	.48	3, 185	.01
Attractiveness*Gender	4.15*	1, 185	.02
Psychiatric Diagnosis *Gender	.97	3, 185	.02
Attractiveness*Gender* Psychiatric Diagnosis	1.88	3, 185	.03

*$p < .05$, **$p < .01$

Table E.19

Adjusted, Unadjusted Means and Variances for a Treatment Aimed at Reducing Violence with the PPMI as Covariate

	N	*M*	*SD*	M_{adj}	*SE*
Attractive					
Female					
Conduct Disorder	11	5.64	1.21	5.51	.46
Psychopathic Traits	14	5.50	1.22	5.50	.40
Schizophrenia	10	4.60	1.84	4.74	.48
No Diagnosis	16	4.44	2.06	4.51	.38
Male					
Conduct Disorder	14	4.86	1.83	4.87	.40
Psychopathic Traits	14	5.57	1.34	5.57	.40
Schizophrenia	10	5.90	1.20	5.95	.48
No Diagnosis	8	3.50	1.07	3.42	.54
Unattractive					
Female					
Conduct Disorder	12	5.42	1.00	5.30	.44
Psychopathic Traits	18	5.22	1.48	5.17	.36
Schizophrenia	13	5.00	1.73	4.96	.42
No Diagnosis	12	4.25	1.76	4.15	.44
Male					
Conduct Disorder	13	6.00	.82	6.13	.42
Psychopathic Traits	9	5.89	1.27	5.90	.50
Schizophrenia	15	5.13	1.92	5.22	.39
No Diagnosis	13	5.08	1.61	5.06	.42

Risk for Future Violence

Three one-way ANCOVAs were attempted to evaluate the effect of attractiveness, gender, and psychiatric diagnosis on ratings of risk for future violence while controlling for the BIDR-16, RLAQ, and PPMI. The RLAQ was not linearly related to any of the dependent variables. The attractiveness analysis was not run due to violations of the assumptions of ANCOVA. The assumption of normality was violated for gender and psychiatric diagnosis; however, visual inspection of histograms showed approximately normal distributions. There was not a significant effect of gender on ratings of future risk when controlling for the BIDR-16 and PPMI, $F(1, 191) = .51$, $p = .476$, partial $\eta^2 = .00$. Due to a significant Levene's test for the psychiatric diagnosis analysis, Huber-White robust standard errors are reported (see Table E.20). There was a significant effect of psychiatric diagnosis, $F(3, 191) = 6.44$, $p < .001$, partial $\eta^2 = .09$. Those with no diagnosis were considered at significantly lower risk than those with conduct disorder (Mean difference = .99, $p = .007$, *CI* = .19, 1.79) or psychopathic traits (Mean difference = 1.23, $p < .001$, *CI* = .45, 2.01).

Table E.20

Means, Adjusted Means, and Variances for Risk of Violence by Psychiatric Diagnosis with the BIDR-16 and PPMI as Cov

	M	*SD*	M_{adj}	*SE*	Robust *SE*
Conduct Disorder	5.02	1.42	5.02	.21	.33
Psychopathic Traits	5.26	1.33	5.25	.20	.32
Schizophrenia	4.70	1.26	4.70	.22	.33
No Diagnosis	4.02	1.76	4.03	.22	.32

A two-way ANCOVA evaluated the effects of attractiveness and psychiatric diagnosis while controlling for the BIDR-16, RLAQ, and PPMI. Key assumptions were met. See Table E.21 for means, adjusted means, and variances. There was not a significant interaction between attractiveness and psychiatric diagnosis on violence risk ratings (see Table E.22).

There was a significant main effect of psychiatric diagnosis. Pairwise comparisons with a Bonferroni correction showed that attractive defendants with psychopathic traits were considered at significantly higher risk than those with no diagnosis (Mean difference = 1.83, $p < .001$, $CI = .68, 2.99$). Defendants with psychopathic traits were considered at significantly higher risk when they were attractive compared to unattractive (Mean difference = .82, $p = .043$, $CI = .03, 1.61$).

Table E.21

Means, Adjusted Means, Standard Deviations, and Standard Errors for Risk of Future Violence by Attractiveness and Psychiatric Diagnosis with the BIDR-16, RLAQ and PPMI as Covariates

	M	*SD*	M_{adj}	*SE*
Attractive				
Conduct Disorder	5.00	1.31	5.00	.30
Psychopathic Traits	5.63	1.24	5.64	.28
Schizophrenia	4.53	1.31	4.52	.33
No Diagnosis	3.79	1.78	3.81	.33
Unattractive				
Conduct Disorder	5.13	1.57	5.11	.30
Psychopathic Traits	4.85	1.35	4.82	.29
Schizophrenia	4.77	1.24	4.78	.28
No Diagnosis	4.43	1.67	4.44	.30

Table E.22

Two-Way ANCOVA for Future Violence Risk with the BIDR-16, RLAQ, and PPMI as Covariates

	F ratio	*df*	η^2
Attractiveness	.041	1, 176	.00
Psychiatric Diagnosis	5.23*	3, 176	.08
Attractiveness * Psychiatric Diagnosis	2.18	3, 176	.04

*$p < .01$

A three-way ANCOVA was attempted to evaluate the interaction effect of attractiveness, gender, and psychiatric diagnosis while controlling for the BIDR-16, RLAQ, and PPMI. However, due to the assumption of linearity being violated for all of the

covariates, a three-way ANOVA was conducted, which was not significant (see Table E.23). There were also no significant two-way interactions.

Table E.23

Three-Way ANOVA Statistics for Risk of Future Violence

	F ratio	*df*	Partial η^2
Attractiveness	.14	1, 196	.00
Psychiatric Diagnosis	5.75*	3, 196	.08
Gender	.08	1, 196	.00
Attractiveness * Psychiatric Diagnosis	2.53	3, 196	.04
Attractiveness*Gender	.30	1, 196	.00
Psychiatric Diagnosis *Gender	2.09	3, 196	.03
Attractiveness*Gender* Psychiatric Diagnosis	1.53	3, 196	.02

*$p < .001$

Threat to Society

Three one-way ANCOVAs were conducted to determine the effects of attractiveness, gender, and crime type on threat to society ratings while controlling for the BIDR-16, RLAQ, and PPMI. Linearity was not established for the BIDR-16 and RLAQ and were dropped from the analyses. Levene's test was significant for the gender analysis, and a Log10 transformation was applied which resolved the heterogeneity of variances. After controlling for the PPMI, attractiveness and gender were not significant predictors of threat to society ratings (see Table E.24). Psychiatric diagnosis did have a statistically significant effect on threat to society ratings. Those with no diagnosis were considered to be at a significantly lower threat to society when compared to those with conduct disorder (Mean difference = 1.07, $p = .004$, $CI = .24, 1.90$) or psychopathic traits (Mean difference = 1.04, $p = .005$, $CI = .23, 1.85$).

Table E.24

One-Way ANCOVAs for Threat to Society by Attractiveness, Gender, and Psychiatric Diagnosis with the PPMI as Covariate

	M	*SD*	M_{adj}	*SE*	*F* ratio	*df*	Partial η^2
Attractive	4.19	1.67	4.18	.16	.02	1, 199	.00
Unattractive	4.21	1.54	4.21	.16			
Female	4.09	1.44	4.10	.16	.14	1, 199	.00
Male	4.31	1.76	4.30	.16			
Conduct disorder	4.58	1.46	4.59	.22	5.26*	3, 197	.07
Psychopathic traits	4.55	1.46	4.55	.21			
Schizophrenia	4.10	1.57	4.08	.23			
No diagnosis	3.51	4.70	3.52	.22			

*$p < .01$

A two-way ANCOVA was run with attractiveness and psychiatric diagnosis while controlling for the PPMI. Key assumptions were met, except for normality for one of the groups. Visual inspection of histograms revealed approximately normal distributions. There was a significant interaction effect of attractiveness and psychiatric diagnosis on threat to society ratings (see Table E.25). Pairwise comparisons with a Bonferroni correction indicated significant differences at different levels of attractiveness of psychopathic traits, $F(1, 193) = 5.74$, $p = .018$, partial $\eta^2 = .03$. Those who were attractive were considered a significantly higher threat to society than those who were unattractive (Mean difference = .99, $p = .018$, *CI* = .18, 1.80). For attractive defendants, there were also significant differences at different levels of psychiatric diagnosis, $F(3, 193) = 6.05$, $p < .001$, partial $\eta^2 = .09$. Those with psychopathic traits were considered a significantly higher threat than those with schizophrenia (Mean difference = 1.27, $p = .03$, *CI* = .08, 2.47) or no diagnosis (Mean difference = 1.71, $p < .001$, *CI* = .48, 2.85) (see Figures E.7 and E.8).

Table E.25

Two-Way ANCOVA for Threat to Society by Attractiveness and Psychiatric Diagnosis with the PPMI as Covariate

	F ratio	*df*	Partial η^2

Attractiveness	.16	1, 193	.00
Psychiatric diagnosis	5.45**	3, 193	.08
Attractiveness * psychiatric diagnosis	2.95*	3, 193	.04

*$p < .05$, **$p < .01$

Figure E.7

Estimated Marginal Means of Threat to Society Scores by Attractiveness and Psychiatric Diagnosis

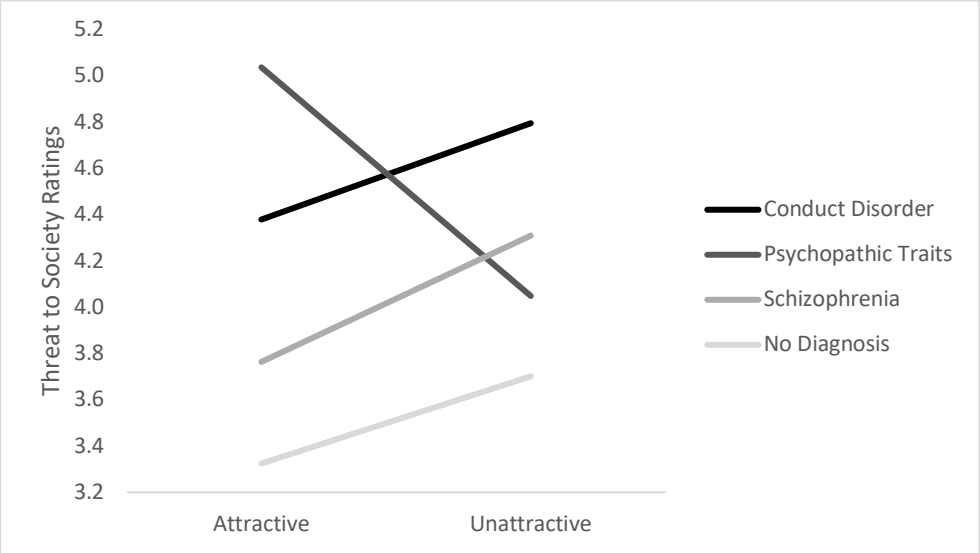
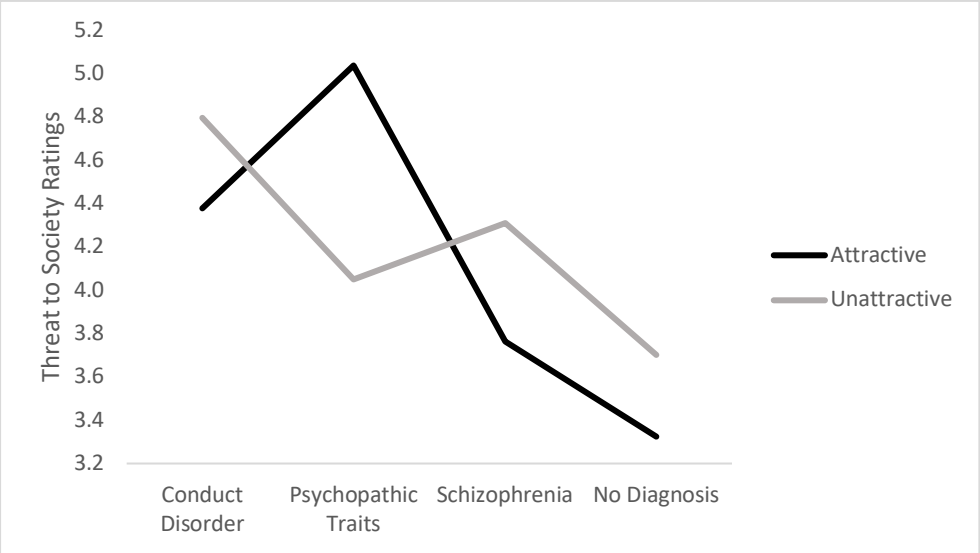


Figure E.8

Estimated Marginal Means of Threat to Society Scores by Psychiatric Diagnosis and Attractiveness



A three-way ANCOVA was attempted to evaluate the interaction effect of attractiveness, gender, and psychiatric diagnosis while controlling for the BIDR-16, RLAQ, and PPMI. However, due to the assumption of linearity being violated for all of the covariates, a three-way ANOVA was conducted. Normality was violated for four of the groups. Examination of skewness and kurtosis z-scores indicated the data were approximately normally distributed. Huber-White robust standard errors were used due to heterogeneity of variances. There was not a significant interaction between attractiveness, gender, and psychiatric diagnosis. There was a significant two-way interaction between attractiveness and psychiatric diagnosis (see Table E.26). Pairwise comparisons with a Bonferroni correction showed that attractive males with psychopathic traits (Mean difference = 1.33, p = .038, CI = .08, 2.59) were considered a significantly higher threat to society than unattractive males, $F(1, 196) = 4.37$, p = .038, partial η^2 = .02. Attractive females with psychopathic traits (Mean difference = 1.57, p = .035, CI = .07, 3.06) were considered a significantly higher threat to society than attractive females with no diagnosis, $F(3, 196) = 3.06$, p = .029, partial η^2 = .05.

Table E.26

Three-Way ANOVA Statistics for Threat to Society

	F ratio	*df*	Partial η^2
Attractiveness	.00	1, 196	.00
Psychiatric Diagnosis	4.54**	3, 196	.07
Gender	.14	1, 196	.00
Attractiveness * Psychiatric Diagnosis	3.20*	3, 196	.05
Attractiveness*Gender	1.50	1, 196	.01
Psychiatric Diagnosis *Gender	.46	3, 196	.01
Attractiveness*Gender* Psychiatric Diagnosis	.70	3, 196	.01

*$p < .05$, **$p < .01$

Appendix F: Exploratory Analyses for Study 3

Confidence

Two one-way ANCOVAs were conducted to evaluate mean differences in participants' confidence in their decision-making based on attractiveness and gender, while controlling for the BIDR-16, RLAQ, and BRS. The RLAQ and BIDR-16 lacked linearity for attractiveness and gender, respectively and were omitted from the analyses. All other key assumptions were met except for normality for both analyses. Visual inspection of histograms showed approximately normal distributions. After controlling for the covariates, there was not a statistically significant difference in ratings of confidence by attractiveness. There was also not a statistically significant difference in ratings of confidence by gender (see Table F.1).

Table F.1

Means, Adjusted Means, Variances and One-Way Analyses of Covariance for Confidence Ratings

	M	*SD*	*M*	*SE*	*F*	*df*	Partial η^2
Attractiveness					.78	1, 180	.00
Unattractive	4.77	1.53	4.78	.15			
Attractive	4.60	1.40	4.59	.16			
Gender					.51	1, 172	.00
Male	4.57	1.40	4.57	.16			
Female	4.73	1.50	4.73	.15			

A two-way ANCOVA was conducted to assess mean differences in confidence ratings by participant race and defendant race while controlling for the BIDR-16, RLAQ, and BRS. The RLAQ was omitted due to violations of linearity. Key assumptions were met, except for normality, which was assessed by a significant Shapiro-Wilk test. Visual inspection of histograms and normal Q-Q plots showed approximately normal distributions.

There was a significant interaction between defendant race and participant race (see Table F.2 and Figure F.1). An analysis of simple main effects was run with Bonferroni adjustment. There was a statistically significant difference in mean confidence scores for White participants, such that their confidence in their decision-making was higher when they had a transcript with a Black defendant (Mean difference = .58, p = .035, 95% CI = .04, 1.11) versus a White defendant, $F(1, 178) = 4.53$, p = .035, partial η^2 = .03. There was also an effect of defendant race, such that when participants had a transcript with a Black defendant, White participants (Mean difference = .77, p = .024, 95% CI = .10, 1.43) were significantly more confident than non-White participants, $F(1, 178) = .5.21$, p = .024, partial η^2 = .03 (see Table F.3 for means, adjusted means, and variances).

Table F.2

Two-Way ANCOVA for Confidence Ratings with the BIDR-16 and BRS as Covariates

	F ratio (1, 178)	η^2
Defendant race	.21	.00
Participant race	1.63	.01
Defendant race * participant race	4.54*	.03

*$p < .05$

Figure F.1

Estimated Marginal Means of Confidence Ratings by Defendant and Participant Race

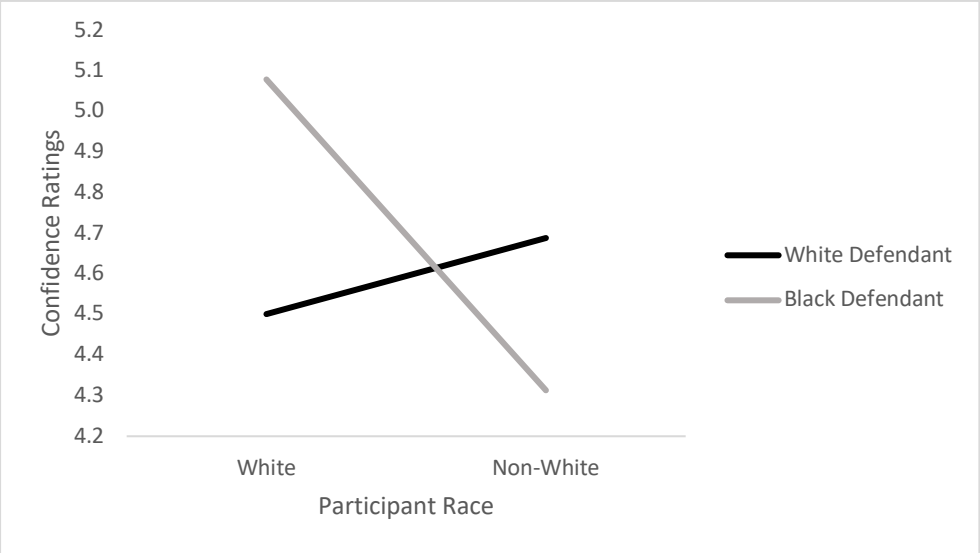


Table F.3

Means, Adjusted Means and Variances for Confidence Ratings with the BIDR-16 and BRS as Covariates

	M	*SD*	*M*	*SE*
White participant				
Black defendant	5.07	1.48	5.08	.20
White defendant	4.48	1.42	4.50	.19
Non-White participant				
Black defendant	4.34	1.54	4.31	.27
White defendant	4.70	1.38	4.69	.23

A two-way ANCOVA was conducted to determine if defendant gender and attractiveness and participant gender had an effect on confidence ratings while controlling for the BIDR-16, RLAQ, and BRS. The BIDR-16 was omitted from the analysis due to nonlinearity. All other key assumptions were met, except for normality. Visual inspection of histograms showed approximately normal distributions. When controlling for the RLAQ and BRS, there was not a significant interaction between defendant attractiveness and gender and participant gender or significant main effects (see Table F.4).

Table F.4

Two-Way ANCOVA for Confidence Ratings with the BIDR-16 and BRS as Covariates

	F ratio (1, 170)	η^2
Attractiveness	.83	.01
Gender	.56	.00
Attractiveness * gender	2.25	.01

A three-way ANCOVA was run to determine the effect of participant race and gender and defendant race, gender, and attractiveness on confidence ratings after controlling for the BIDR-16, RLAQ, and BRS. The BIDR-16 and RLAQ were omitted from the analysis due to violations of linearity. All other assumptions were met except for normality for one of the groups. Visual inspection of histograms showed approximately normal distributions. After controlling for the BRS, there were not significant interaction or main effects on confidence ratings (see Table F.5). Although there was not a significant interaction, pairwise comparisons with a Bonferroni correction indicated that White participants were significantly more confident in their ratings than non-White participants when defendants were unattractive Black females (Mean difference = 1.37, $p = .041$, 95% *CI* = .06, 2.68) or attractive Black males (Mean difference = 1.44, $p = .043$, 95% *CI* = .05, 2.83) (see Table F.6 and Figures F.2 and F.3).

Table F.5

Three-Way ANCOVA for Confidence Ratings with the BRS as Covariate

	F ratio	*df* (173)	Partial η^2
Attractiveness	.54	1	.00
Gender and race	.93	3	.02
Participant race	1.07	1	.01
Attractiveness*Gender and race	.62	3	.01
Attractiveness* Participant race	.35	1	.00
Gender and race* Participant race	2.00	3	.03
Attractiveness* Gender and race * Participant race	1.67	3	.03

Table F.6

Means, Adjusted Means, and Variances of Confidence Ratings with the BRS as Covariate

	M	*SD*	*M*	*SE*
Unattractive				
Black Female				
White participant	5.25	1.76	5.24	.42
Non-White participant	3.88	1.89	3.88	.51
White Female				
White participant	4.71	1.59	4.71	.39
Non-White participant	4.54	1.45	4.54	.40
Black Male				
White participant	4.95	1.19	4.95	.32
Non-White participant	5.43	1.51	5.43	.55
White Male				
White participant	4.36	1.34	4.36	.39
Non-White participant	5.00	1.56	5.01	.46
Attractive				
Black Female				
White participant	4.86	1.66	4.86	.39
Non-White participant	4.29	.76	4.28	.55
White Female				
White participant	4.72	1.53	4.72	.34
Non-White participant	5.00	1.00	5.01	.55
Black Male				
White participant	5.44	1.51	5.44	.48
Non-White participant	4.00	1.41	4.00	.51
White Male				
White participant	4.13	1.15	4.12	.36
Non-White participant	4.38	1.39	4.39	.40

Figure F.2

Estimated Marginal Means of Confidence for Black Female Defendants

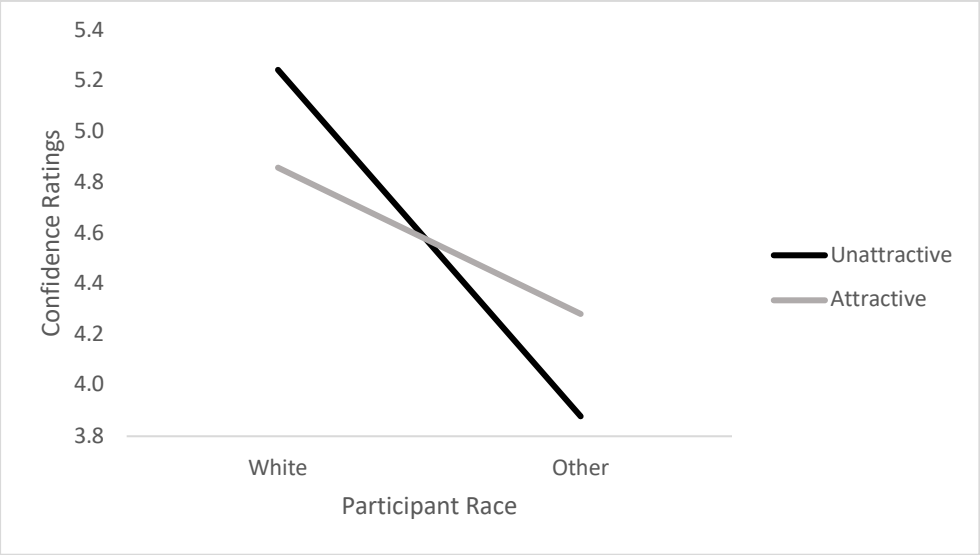
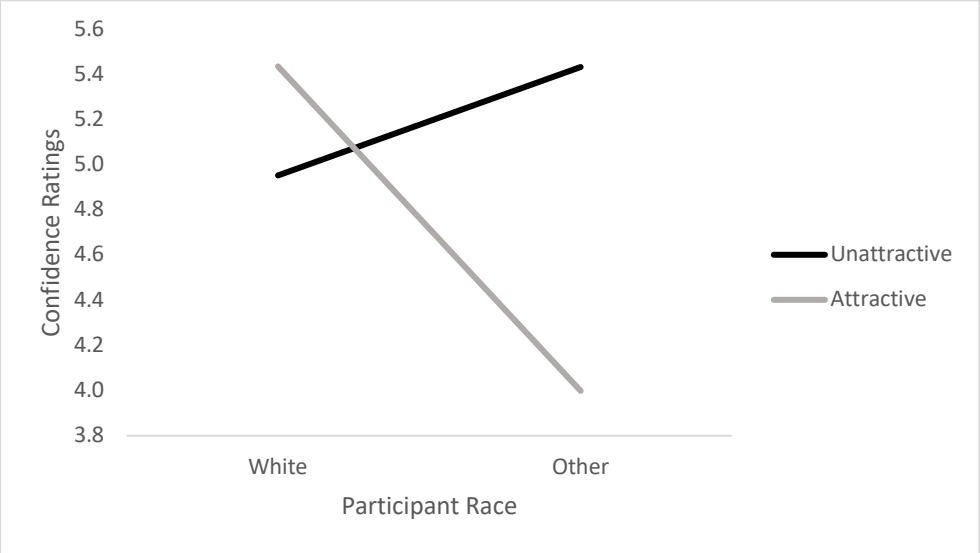


Figure F.3

Estimated Marginal Means of Confidence for Black Male Defendants



Defendant Credibility

Two one-way ANCOVAs were conducted to determine if attractiveness and gender had an effect on defendant credibility ratings when controlling for the BIDR-16, RLAQ, and BRS. All key assumptions were met. There were not significant effects of attractiveness or gender on defendant credibility ratings when controlling for the covariates (see Table F.7).

Table F.7

Means, Variances, and One-Way ANCOVAs of Defendant Credibility Ratings by Attractiveness and Gender with the BIDR-16, RLAQ, and BRS as Covariates

	M	*SD*	*M*	*SE*	*F* ratio (1, 166)	Partial η^2
Unattractive	4.47	1.14	4.49	.13	.02	.00
Attractive	4.48	1.19	4.46	.13		
Male	4.44	1.25	4.46	.13	.02	.00
Female	4.51	1.08	4.49	.12		

A two-way ANCOVA was conducted to determine the effect of participant and defendant race on defendant credibility ratings while controlling for the BIDR-16, RLAQ, and BRS. The RLAQ was omitted from the analysis due to violations of linearity. There was one outlier, and the analysis was run both with it and without it. Overall normality of the residuals improved without the outlier, but it did not significantly change the outcome of the analysis. Visual inspection of histograms showed an approximately normal distribution. Other key assumptions were met. There was not a significant interaction of participant and defendant race while controlling for the covariates, $F(1, 179) = .25$, $p = .616$, partial $\eta^2 = .00$ (see Table F.8 for means, adjusted means, and variances of defendant credibility).

Table F.8

Means, Adjusted Means and Variances for Defendant Credibility Ratings with the BIDR-16 and BRS as Covariates

	M	*SD*	*M*	*SE*
White participant				
Black defendant	4.56	1.30	4.57	.16
White defendant	4.47	1.10	4.47	.15
Non-White participant				
Black defendant	4.41	.95	4.38	.22
White defendant	4.44	1.27	4.46	.18

A two-way ANCOVA was conducted to determine if defendant attractiveness and gender and participant gender had an effect on defendant credibility ratings while controlling

for the BIDR-16, RLAQ, and BRS. Key assumptions were met. There was not a significant interaction or main effects of defendant attractiveness and gender and participant gender on defendant credibility ratings (see Table F.9).

Table F.9

Two-Way ANCOVA for Defendant Credibility Ratings with the BIDR-16, RLAQ and BRS as Covariates

	F ratio (1, 164)	Partial η^2
Attractiveness	.01	.00
Gender	.03	.00
Attractiveness*Gender	3.42	.02

A three-way ANCOVA was conducted to determine if participant race and gender and defendant attractiveness, race and gender had an effect on defendant credibility ratings while controlling for the BIDR-16, RLAQ, and BRS. The BIDR-16 and RLAQ were omitted due to violations of linearity. Normality was violated for one group; however, visual inspection of histograms showed an approximately normal distribution. When controlling for the BRS, there was not a significant interaction between participant and defendant race and gender and defendant attractiveness (see Table F.10). However, pairwise comparisons with a Bonferroni correction indicated that White participants (Mean difference = 1.19, $p = .027$, 95% *CI* = .13, 2.24) considered unattractive Black female defendants to be significantly more credible than did non-White participants, $F(1, 174) = 4.95$, $p = .027$, partial $\eta^2 = .03$ (see Figure F.4). White participants also considered unattractive Black female defendants (Mean difference = .92, $p = .048$, 95% *CI* = .01, 1.82) to be significantly more credible than attractive Black female defendants, $F(1, 174) = 3.97$, $p = .048$, partial $\eta^2 = .02$ (see Table F.11 and Figure F.5).

Table F.10

Three-Way ANCOVA for Defendant Credibility Ratings with the BRS as Covariate

	F ratio	*df* (174)	Partial η^2
Attractiveness	.05	1	.00
Gender and race	.42	3	.01
Participant race	.25	1	.00
Attractiveness*Gender and race	1.08	3	.02
Attractiveness* Participant race	.21	1	.00
Gender and race* Participant race	1.45	3	.02
Attractiveness* Gender and race * Participant race	1.42	3	.02

Figure F.4

Estimated Marginal Means of Defendant Credibility for Black Female Defendants with the BRS as Covariate

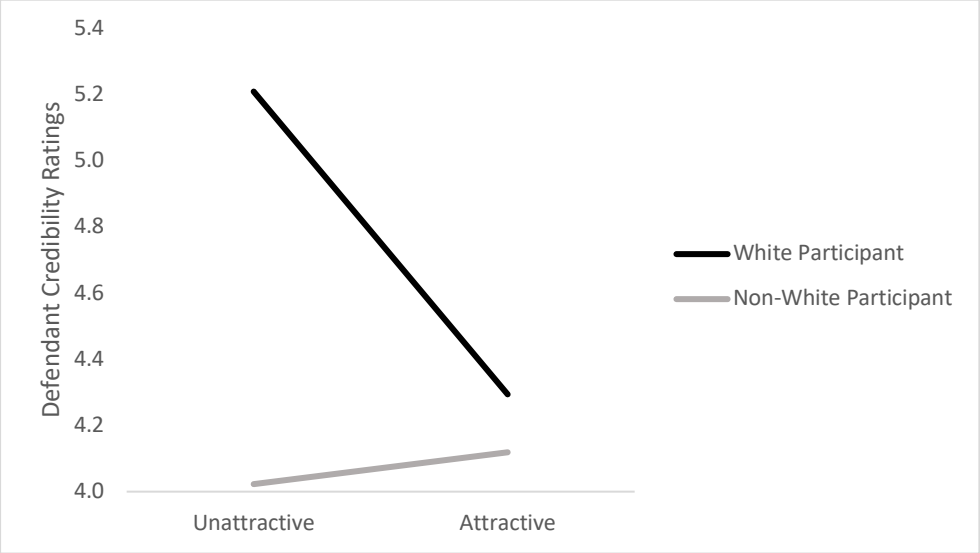


Table F.11

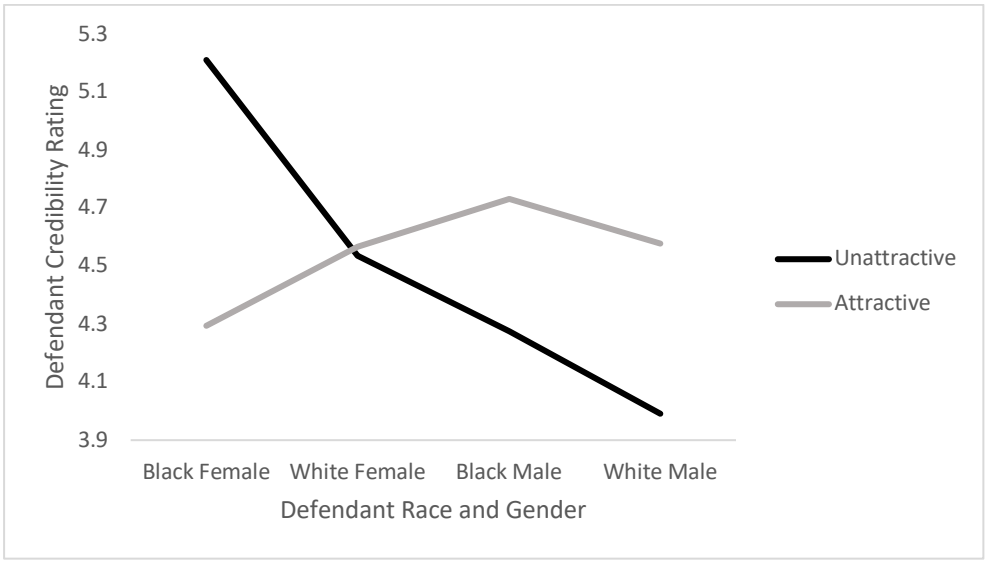
Means, Adjusted Means, and Variances of Defendant Credibility Ratings with the BRS as Covariate

	M	*SD*	*M*	*SE*
White participant				
Black Female				
Unattractive	5.25	1.22	5.21	.34
Attractive	4.29	1.27	4.29	.31
White Female				
Unattractive	4.57	.76	4.54	.31

Attractive	4.56	1.15	4.57	.27
Black Male				
Unattractive	4.25	1.29	4.27	.26
Attractive	4.78	1.30	4.73	.39
White Male				
Unattractive	4.00	1.11	3.99	.31
Attractive	4.63	1.26	4.58	.29
Non-White participant				
Black Female				
Unattractive	4.00	.93	4.02	.41
Attractive	4.14	.69	4.12	.44
White Female				
Unattractive	4.69	1.03	4.69	.32
Attractive	4.00	1.20	4.05	.41
Black Male				
Unattractive	4.57	.98	4.61	.44
Attractive	5.00	.93	4.99	.41
White Male				
Unattractive	4.60	1.51	4.65	.37
Attractive	4.31	1.38	4.33	.32

Figure F.5

Estimated Marginal Means of Defendant Credibility for White Participants with the BRS as Covariate



Victim Credibility

Two one-way ANCOVAs were conducted to evaluate mean differences in participants' victim credibility ratings based on attractiveness and gender, while controlling for the BIDR-16, RLAQ, and BRS. All key assumptions were met. There was not a significant effect of attractiveness or gender on victim credibility ratings when controlling for the BIDR-16, RLAQ, and BRS (see Table F.12).

Table F.12

Means, Variances and One-Way ANCOVAs of Victim Credibility Ratings by Attractiveness and Gender with the BIDR-16, RLAQ, and BRS as Covariates

	M	*SD*	*M*	*SE*	*F* ratio (1, 166)	Partial η^2
Unattractive	4.09	1.28	4.08	.15	.10	.00
Attractive	4.13	1.40	4.14	.14		
Male	4.00	1.23	3.99	.15	1.31	.01
Female	4.22	1.43	4.23	.14		

A two-way ANCOVA was conducted to determine the effect of participant and defendant race on victim credibility while controlling for the BIDR-16, RLAQ, and BRS. The BIDR-16 and RLAQ were omitted from the analysis due to violations of linearity. Due to a significant Levene's statistic, Huber-White robust standard errors are reported. All other key assumptions were met, except for normality which was shown to be approximately normally distributed via visual inspection of histograms. After controlling for the BRS, defendant and participant race did not have significant interaction or main effects on ratings of victim credibility (see Table F.13).

Table F.13

Two-Way ANCOVA for Victim Credibility Ratings with the BRS as Covariate

	F ratio (1, 186)	Partial η^2	Robust *SE*
Defendant race	.05	.00	.30
Participant race	.57	.00	.28

Defendant race * participant race	.25	.00	.40

A two-way ANCOVA was conducted to determine if defendant attractiveness and gender and participant gender had an effect on victim credibility ratings while controlling for the BIDR-16, RLAQ, and BRS. The RLAQ was omitted from the analysis due to violations of linearity. Key assumptions were met. After controlling for the BIDR-16 and BRS, there was not a significant interaction between defendant attractiveness and gender and participant gender, nor were there significant main effects (see Tables F.14 and F.15).

Table F.14

Two-Way ANCOVA for Victim Credibility Ratings with the BIDR-16 and BRS as Covariates

	F ratio (1, 164)	Partial η^2
Attractiveness	.00	.00
Gender	1.00	.01
Attractiveness*Gender	.04	.00

Table F.15

Means, Adjusted Means and Variances for Victim Credibility Ratings with the BIDR-16 and BRS as Covariates

	M	*SD*	*M*	*SE*
Unattractive				
Male	4.08	1.24	4.02	.20
Female	4.22	1.36	4.26	.20
Attractive				
Male	4.02	1.31	4.06	.20
Female	4.23	1.44	4.22	.20

A three-way ANCOVA was conducted to determine if participant race and gender and defendant attractiveness, race and gender had an effect on victim credibility ratings while controlling for the BIDR-16, RLAQ, and BRS. The BIDR-16 and RLAQ were omitted due to violations of linearity. All other key assumptions were met. When controlling for the BRS, there was not a significant interaction effect between participant race and gender, and

defendant race, gender, and attractiveness on ratings of victim credibility, nor were there any significant main effects (see Tables F.16 and F.17).

Table F.16

Three-Way ANCOVA for Victim Credibility Ratings with the BRS as Covariate

	F ratio	*df* (174)	Partial η^2
Attractiveness	.12	1	.00
Gender and race	1.00	3	.02
Participant race	.35	1	.00
Attractiveness*Gender and race	1.12	3	.02
Attractiveness* Participant race	.41	1	.00
Gender and race* Participant race	.56	3	.01
Attractiveness* Gender and race * Participant race	.56	3	.01

Table F.17

Means, Adjusted Means, and Variances of Victim Credibility Ratings with the BRS as Covariate

	M	*SD*	*M*	*SE*
White participant				
Black Female				
Unattractive	4.92	1.44	4.99	.38
Attractive	4.29	1.49	4.27	.35
White Female				
Unattractive	4.07	1.21	4.13	.36
Attractive	4.22	1.40	4.20	.31
Black Male				
Unattractive	4.05	1.28	4.01	.30
Attractive	3.33	1.50	3.42	.44
White Male				
Unattractive	4.00	.88	4.02	.35
Attractive	4.44	1.15	4.52	.33
Non-White participant				
Black Female				
Unattractive	4.13	.99	4.08	.47
Attractive	4.29	.95	4.33	.50
White Female				
Unattractive	3.69	1.44	3.69	.37
Attractive	4.25	2.12	4.17	.47
Black Male				
Unattractive	4.43	1.13	4.36	.50
Attractive	3.75	1.16	3.76	.47

White Male				
Unattractive	4.10	1.66	4.00	.42
Attractive	4.15	1.28	4.11	.37

VITA AUCTORIS

NAME:	Jennifer A. Ranjit
PLACE OF BIRTH:	Windsor, ON
YEAR OF BIRTH:	1984
EDUCATION:	Sandwich Secondary School LaSalle, ON 1999-2003 University of Windsor, B.A. Windsor, ON 2018-2021 University of Windsor, M.A. Windsor, ON 2021-present