Studying Guilt Perception in Millennials: Unexpected Effects of Suspects’ Race and Attractiveness

D. Lisa Cothran*
Alabama State University, USA

Elena V. Stepanova*
The University of Southern Mississippi, USA

K. Raquel Barlow
Vanderbilt University, USA

Authors’ Note

*Contributions of the first and second authors are equal, and their names are arranged in alphabetical order.

Correspondence concerning this article should be addressed to D. Lisa Cothran, Department of Psychology, Alabama State University, 915 S. Jackson St., Montgomery, AL 36104. Email: dcothran@alasu.edu and dee.lisa.cothran@gmail.com

Imagination, Cognition and Personality, in press
Abstract

The present study explored mock jurors’ guilt judgments with a 2 (jurors’ race: Black vs. White) x 2 (suspects’ race: Black vs. White) x 2 (suspects’ attractiveness: high vs. low) design in a group of Millennials ($N = 331$). Black jurors were more lenient; all jurors were more lenient toward Black suspects; and White jurors were less lenient toward Black unattractive suspects.

The current study contributes the following novel findings to the literature: documentation of a possible Black experimenter effect in mock jurors; an interaction among suspects’ race, suspects’ attractiveness and jurors’ race suggesting that racial bias exhibited by White jurors may be masking itself as an “unattractiveness” bias; and additive empathy by Black jurors toward persons who fall within more than one underprivileged group.

Keywords: race, attractiveness, jurors’ bias, mock-juror characteristics, defendant characteristics
Studying Guilt Perception in Millennials: Unexpected Effects of Suspects’ Race and Attractiveness

Bob Dylan (1975) wrote in his song *Hurricane*, “How can the life of such a man/ Be in the palm of a fool’s hand.../ Put in a prison cell he could-a been/ The champion of the world” (Sporer & Goodman-Delahunty, 2009, p. 379). This song was written about the wrongful conviction in 1966 of an African American male, Ruben “The Hurricane” Carter, who became a symbol of racial injustice. Such instances, which remain frequent (about 4% in capital punishment cases, e.g., Gross, O’Brien, Hu, & Kennedy, 2014), embody and exemplify ongoing disparities in the criminal justice system. The current research examines two factors contributing to disparate treatment in the United States’ judicial system, specifically, racial bias and physical attractiveness bias in juridical decision making.

As humans, we rely on stereotypes as cognitive shortcuts (Allport, 1954; Kunda, 1999). Problems arise however, when the stereotypes are incorrect and lead to harmful decisions and behaviors (e.g., unfair sentencing practices). Literature on the criminal justice system has examined whether juries, judges, and parole boards are influenced by suspects’ unrelated characteristics such as race (for review see, Sommers, 2007; Sommers & Ellsworth, 2009) or attractiveness (e.g., Abwender & Hough, 2001; DeSantis & Kayson, 1997; Patry; 2008) and how characteristics of jurors (e.g., race) interact with suspects’ characteristics (e.g., race) to affect decision-making (e.g., Bernard, 1979; Bradbury & Williams, 2012; Cohn, Bucallo, & Pride, 2009; Shaw & Skolnick, 1995; Skolnik & Shaw, 1997; Sommers & Ellsworth, 2000). Institutional bias exists at all levels of the justice system. From the juvenile arena to the adult arena and for both males and females, African Americans face disparities and inequities (Cothran, 2011; Sickmund & Puzzanchera, 2014). Jurors often develop impressions of
defendants’ guilt early in a trial, and these impressions are carried forward into the jury room and likely are expressed in the deliberation of verdicts (Mazzella & Feingold, 1994). Below we will review how perception of guilt, likelihood of conviction, and length of sentencing are affected by (a) the race of defendants and jurors and (b) the attractiveness of defendants in the context of the criminal justice system.

Race of defendants and jurors

Racial bias in White jurors. Previous extensive research on the effects of race of defendants on decision making of jurors in the criminal justice system has revealed inconsistent findings and almost exclusively focused on White jurors’ perceptions of Black versus White defendants (for review, see Sommers, 2007). Most of this work has shown that White jurors exhibit race bias toward Black defendants. For example, two recent meta-analyses indicate that White jurors are more biased toward Black (compared to White defendants) in verdict decisions and sentencing decisions (Sweeney & Haney, 1992; Mitchell, Haw, Pfeifer, & Meissner; 2005). The pattern of these results parallels the higher conviction and incarceration rates of Blacks (compared to Whites) in America. For example, according to the U.S. Census (2010), Blacks’ rate of incarceration is 5 times higher than Whites’ rate. Brewer (2004) suggests that Blacks historically have been executed for less serious crimes than Whites and their death sentences were carried out with fewer legal appeals. Research conducted in the criminal justice system also has shown that Black suspects with more stereotypically Afrocentric features (e.g. large nose, thick lips, dark skin) receive longer sentences compared to suspects who have less stereotypically Afrocentric features (Blair, Judd, & Chapleau, 2004; Blair et al., 2002; Eberhardt, Davis, Purdie-Vaughns, & Johnson, 2006; Goff, Eberhardt, Williams, & Jackson, 2008). Jackson (2006) also provides an excellent survey of the pervasive, negative depiction of Black
masculinity in popular press and social media. Such inequalities are rooted in outgroup bias, consistent with the social identity theory (Tajfel & Turner, 1985) postulating that people categorize others as belonging to an ingroup or an outgroup, identify with their ingroup and engage in social comparisons between the ingroup and the outgroup, and then favor the ingroup members to enhance self-image. In the dynamics of juries and defendants, such ougroup bias is explained by the similarity-leniency hypothesis: when jurors categorize a defendant as in-group rather than an outgroup member, jurors are more lenient toward him or her (Kerr, Hymes, Anderson, & Weathers, 1995).

**Black criminal stereotypes.** Racial inequalities in the criminal justice system also stem from certain stereotypes that the American public and law enforcement have about Blacks. Stereotypes have been historically conceptualized as culturally-learned, distorted images (Lippmann, 1922) of social groups. While stereotypes have been traditionally considered as consciously-held beliefs, subject to deliberate control (Fazio, Jackson, Dunton, & Williams, 1995), newer dual-process models postulate that stereotypes can be explicit (consciously held beliefs) and implicit (associations that are not easily assessable) (e.g., Smith & DeCoster, 2000). Stereotypes are commonly automatically activated and applied upon the mere presence of a member of a social group, sometimes without awareness (Devine, 1989). Throughout the history of the United States Blacks have been consistently stereotyped as criminals; dating back to the enslavement of Africans in the United States, Blacks have had their reputation besieged and besmirched by society’s perception that they are predisposed towards criminality (Kennedy, 1997). While Whites commit the majority of crime (Federal Bureau of Investigation, 2013), the common belief is that Blacks commit the majority of crime (Gillens, 1996; for review see Welch, 2007). Stereotyping research (Neiman, 1994) has shown that violent was the sixth most frequent
stereotype ascribed to Black men, and this crime-related descriptor was not listed among the stereotypes of the other groups (Black women, White women, and White men). Encountering some form of criminal punishment from the justice system has become something of an expectation for many young, urban Black men (Welch, 2007). Stereotyping of Black (compared to White) suspects as more criminal, violent, dangerous and less amenable to rehabilitation (for review, Spohn, 2000) leads to discriminatory behavior by judges and jurors and takes into account the race of the suspect.

(Absence of) Racial bias in White jurors. While many studies report racial bias in White jurors toward Black defendants, these findings are not uniform. Some studies report no difference in treatment of Black and White defendants by White jurors (Abwender & Hough, 2001; Dean, Wayne, Mack, & Thomas, 2000; McGuire & Bermant, 1977; Skolnick & Shaw, 1997; Voss et al., 1996). A meta-analysis by Mazella and Feingold (1994) indicates that neither Black nor White jurors exhibit racial bias in verdict or sentencing decisions. In fact, some studies report White jurors give harsher judgments to White (compared to Black) defendants (Blair, Judd, Chapleau, 2004; McGowen & King, 1982; Poulson, 1990; Shaw & Skolnick, 1995). In an attempt to reconcile these inconsistent findings, Sommers and Ellsworth (2000; 2001) proposed and empirically tested a hypothesis that “race salience” changes White jurors’ decision to convict a defendant. They found that White jurors are usually more lenient to White (compared to Black) defendants when the crime in question is not racially charged (e.g., when both the victim and the defendant are in the same racial category); however, when the crime is racially charged (e.g., when the victim is White and the defendant is Black), the defendants’ race did not impact White jurors’ decisions. Sommers and Ellsworth explained these findings by
using the aversive racism theoretical framework (Gaertner & Dovidio, 1986) suggesting that jurors are more cognizant that their verdict could appear racist in the “race salience” condition.

**Racial bias in Black jurors.** While the majority of studies address racial bias in White jurors, several studies investigated the dynamics between Black jurors and defendants from various racial groups. Bernard (1979) reported that Black jurors are more likely than White jurors to acquit a defendant regardless of their race. Foley and Chamblin (1982) found no differences in conviction rates of Black and White defendants in a rape case for Black mock jurors.

Outgroup bias among Black mock jurors has been found in several studies: Black jurors were less likely to convict Black (compared to White) defendants (Bradbury & Williams, 2013) and rendered harsher judgments for White (compared to Black) defendants in a simulated crimes resembling the infamous O. J. Simpson case (Skolnick & Shaw, 1997), an assault case (Sommers & Ellsworth, 2000) and a vehicular case (Abwender & Hough, 2001). Black jurors also considered a White defendant more culpable than a Black defendant in a simulated rape case trial (Ugwuegbu, 1979) and in a simulated sexual harassment case (Wuensch et al., 2002). In sum, there are only a few studies that investigated racial biases in Black versus White jurors and they did not produce uniform findings. One of the goals of this work is to extend current literature on comparing Black and White mock jurors’ guilt perceptions of Black and White defendants.

**Attractiveness of the defendants**

One of the most frequently studied characteristic of defendants that influences both perception of guilt and severity of sentencing is attractiveness. The attractiveness-leniency effect (ALE) occurs when more attractive suspects are treated more leniently than less attractive suspects (for review, see Mazzella & Feingold, 1994).
Suspects’ appearance influences jurors’ decisions despite physical attractiveness having been deemed an extralegal/legally irrelevant factor in jurors’ considerations and decisions (Mazzella & Feingold, 1994). The effect of physical attractiveness in day-to-day life is that beauty belies goodness (Dion, Berscheid, & Walster, 1972); attractiveness bias is present as early as infancy. Langlois, Roggman, and Reiser-Danner (1990) found that even twelve-month-old infants showed more positive reactions, greater attentiveness and longer play time with attractive (compared to unattractive) dolls and masks. This influence of physical attractiveness on social behavior at such a young age has led some researchers to hypothesize that one’s preference for attractive faces may even be innate instead of culturally learned. Whatever the basis, this preference is exercised throughout one’s life whether it is the grade school teacher giving more attractive (compared to less attractive) students more positive attention, or a judge and jurors giving more attractive (compared to less attractive) suspects greater leniency. Just as prosecutors may exploit racial stereotypes to gain convictions, defense attorneys may exploit attractiveness stereotypes to gain lesser charges, lighter punishments or acquittals. For example, Wiseman (1998) conducted an experiment with BBC Television to illustrate the bias. On the show viewers were shown evidence about a burglary. The independent variables were the suspects’ physical traits. Such that, the suspect either had stereotypical criminal traits (unattractive face, crooked nose, small eyes) or baby-faced traits (attractive, symmetrical face with large blue eyes). With a sample size of 64,000 people phoning in their verdict, 41 percent judged the stereotypical criminal guilty as compared to the 31 percent that found the attractive criminal guilty.

Mazzella and Feingold (1994) contend that physically unattractive suspects are treated more harshly because “jurors may unconsciously find personal characteristics relevant, and be more likely to assume a suspect is guilty when the suspect possesses characteristics that are more
commonly observed in criminals” (p. 1316). Desantis and Kayson (1997) propose another explanation for the attractiveness bias. They suggest that some jurors may perceive physically attractive suspects as more capable of remorse and less likely to make crime a habit. Notwithstanding, another explanation is the possibility that jurors may perceive more attractive suspects as having comparatively more to lose if imprisoned.

While the ALE has been demonstrated and replicated multiple times, it does not hold up for all types of crimes (e.g., swindle, Sigall & Ostrove, 1975; negligent homicide, see Mazzella & Feingold, 1994). For example, Desantis and Kayson (1997) manipulated attractiveness, race and sex of the defendant and race of the jurors in a fictitious burglary case. They found that there were no interactions among those factors. Abwender and Hough (2001) examined the effects of defendants’ attractiveness, race and gender and jurors’ race and gender on judgment of guilt and sentencing using a simulated vehicular homicide case. Race of jurors and race of defendants did not interact with attractiveness of defendants. Since studies simultaneously examining the effects of defendants’ and jurors’ race and defendants’ race and attractiveness are scarce, the second goal of our study is to investigate the combined and interactive effects of jurors’ race and defendants’ race and attractiveness on guilt perception in mock jurors.

The current research

Although the prejudice against Blacks and the prejudice against people who are unattractive are quite different, each has a measurable impact on decision making in the justice system. The current study is designed to extend and reconcile previous findings by testing the effects of mock jurors’ race, defendants’ race, and defendants’ attractiveness on mock jurors’ guilt perceptions.
Shaked-Schroer, Costanzo and Marcus-Newhall (2008) note that racial bias in sentencing has been investigated using a variety of research methods such as “post-trial interviews, government analyses of sentencing trends, experimental research, and survey methods” (p. 604). Mazzella and Feingold (1994) note that for the past two decades research on juridical decision-making and the influence of physical attractiveness have employed the “mock jury paradigm…[which] ‘simulate[s]’ jurors independently [reviewing] materials (vignettes, audiotapes, videotapes) that describe a criminal case” (p. 1316). Mitchell et al. (2005) also note the frequent and standard use of this paradigm in this area of study. Each method has its advantages and disadvantages but a majority of those listed rely on an array of resources not easily acquired such as interviews from trials or longitudinal sentencing data. For this study the mock-jury survey method is ideal because of its simple construction and exceptional ability to be distributed far and fast.

The current study is designed as a conceptual replication of previous work. In the past several years, the field of psychology has begun an initiative encouraging the replication of previously published work (Makel, Plucker, & Hegarty, 2012), including conceptual replication. As an example, one journal, Social Psychology, devoted an entire issue to the effort of replication in the area (Nosek & Lakens, 2014). As the previous findings of the previous studies on the effects of jurors’ and defendants’ race are mixed, and the previous findings on the interactive effects of defendants’ attractiveness and race are scarce, the current study aims to conceptually replicate previous work. The current study also is an extension of previous work reviewed in that we explore race bias in sentencing in a group of Millennials (born after 1980s, Pew Research Center, 2014). People in the millennial generation are frequently described as despising racism (Cohen, 2011).
The current study examines extralegal factors, racial and attractiveness bias, in evaluations of suspects’ guilt among a group of Millennial mock jurors. The design is a 2 (mock jurors’ race: Black vs. White) x 2 (suspects’ race: Black vs. White) x 2 (suspects’ level of attractiveness: high vs. low) between-subjects design.

As the majority of studies reviewed found White jurors’ bias toward Black suspects, Prediction 1 is that White mock jurors will rate Black (compared to White) suspects as more guilty. Analogously, more studies support the outgroup bias on the part of Black jurors, so Prediction 2 is that Black jurors will rate White (compared to Black) suspects as more guilty. However, we would like to note that any pattern of results including an interactive effect between suspects’ race and mock jurors’ race would suggest that race of suspects is weighted differently by jurors of different races.

Based upon attractiveness research reviewed, Prediction 3 is that all mock jurors will rate unattractive (compared to attractive) suspects as more guilty. While we did not formulate a specific hypothesis regarding an interaction among the race of the jurors, race of the suspects, and attractiveness of suspects, as prior studies investigating all these factors did not detect one (e.g., Abwender & Hough, 2001; Desantis & Kayson, 1997), we were interested in exploring all potential additive and interactive effects of these factors.

**Method**

**Participants**

A total of 363 university undergraduates participated in this study. Regarding race, the participants self-identified as Black (n = 165; 45.7%), White (n = 170; 47.1%), or ‘other’ (n = 25; 6.7%). Note that some participants had missing data, so only 331 participants were included.
in the final analyses. For the final sample, 161 participants (48.6%) identified as Black and 170 participants (51.4%) identified as White. The mean age was 21.7 years ($SD = 4.5$ years).

Regarding sex, there were 97 men (29.3%) and 232 women (70.1%). Two people left this item blank. The overwhelming majority of the sample ($n = 326; 98.5$%) had never participated on a jury, yet all of them were over 18 (one of the legal qualifiers for jury service in the US, see United States Courts, 2015).

**Materials**

**Suspects’ Photographs.** In order to validate that the suspects’ photos clearly represented the intended levels of attractiveness, two convenience samples of university undergraduates enrolled in psychology classes provided normative data for the pictorial stimuli. A series of male facial photographs (i.e., ‘headshots’) were gathered from the internet through Google image search. They were independently assessed by the authors who selected the two most physically attractive men’s pictures and the two least physically attractive men’s pictures. A convenience sample of university undergraduates (Sample One) enrolled in psychology classes provided normative data for these photographs. The majority of Sample One ($N = 34$) was comprised of White (79.4%) and Black (8.8%) female (52.9%) first year undergraduate (79.4%) university students. The average age of the participants was 19.8 years ($SD = 3.8$). Photographs were shown to Sample One and participants were asked to rate the attractiveness of each photo on a scale of 0 (*not at all physically attractive*) to 5 (*very physically attractive*). A 2 (race: Black vs. White) x 2 (attractiveness: attractive vs. unattractive) Repeated Measures ANOVA suggested that ratings were significantly impacted by the attractiveness of the photographs, $F (1, 32) = 96.39, p < .001, \eta^2_p = .751$. Participants rated the attractive faces as more attractive than the unattractive faces ($ps < .001$). Close examination of the average ratings of the pictures in each group revealed that,
while the ‘unattractive’ males’ pictures were rated quite low (white: $x = .91$, $sd = .95$; black: $x = .64$, $sd = 1.06$), the ‘attractive’ males’ pictures were also rated quite low (white: $x = 2.91$, $sd = 1.31$; black: $x = 1.82$, $sd = 1.23$). On the basis of that data, we selected two unattractive photos (one Black and one White) for the study.

So, in an effort to strengthen the attractiveness manipulation, the researchers collected additional 'headshot' pictures of attractive Black and White males from the internet. The majority of Sample Two ($N = 29$) was comprised of White (51.7%) and Black (34.5%) female (79.3%) university students. The average age of these participants was 23 years ($SD = 4.06$). Photos were shown to the participants in Sample Two and they were asked to rate the attractiveness of each photo on a scale of 0 (not at all physically attractive) to 5 (very physically attractive). On the basis of that data, the two pictures (one Black and one White) with the highest average attractiveness ratings in the respective racial categories were selected and included in the current study (white: $x = 3.93$, $sd = .75$; black: $x = 3.31$, $sd = 1.29$). There were no differences in attractiveness ratings between participants of various racial groups for both samples. Given that there is a very high agreement for facial attractiveness judgments/preferences among raters of different ages and cultural backgrounds (for review, see Langlois et al., 2000), we were not concerned about using two different samples for pre-testing. All photos were matched on age. The participants’ norming data from Samples 1 and 2 reflect the following: We utilized 4 pictures (two Black and two White) that participants perceived to be clearly attractive ($n = 2$) and clearly unattractive ($n = 2$). Pictures were copied and pasted on mugshot background using Photoshop.

**Crime Report.** One self-created crime report was based upon a fictitious home invasion, burglary and assault report provided by an active-duty Police Officer. Information on
the report included the suspect’s race and prior record, a photo of the offender, and a description of the crime. There were four versions of this crime report; all four versions were identical except for references to the suspects’ race (i.e., Black vs. White) and a picture of the suspect (i.e., attractive vs. unattractive). The manipulations for race included direct mention of the suspect’s race in the crime report in multiple places, and a mug shot that was 5 inches x 7 inches in size. The manipulation for attractiveness was the same mug shot photo. We specifically chose male suspects to increase ecological validity of our work, as most of the violent crimes are committed by men (see Federal Bureau of Investigation, 2012, 2014). Suspects’ prior record was included for the same reason, as many offenders are repeat, according to the Bureau of Justice Statistics (2014). See Figure 1 for mug shots used in the study.

Questionnaire. Attached to the crime report were survey questions assessing participants’ demographic information, and evaluations of the suspect’s guilt of four crimes (i.e., theft, home invasion, aggravated robbery, aggravated assault) on a scale of 0 (not at all guilty) to 3 (definitely guilty).

Procedure

Researchers (Black females) visited several undergraduate classes in psychology, sociology, and criminal justice and organizations to collect convenience samples’ data. Participants were randomly assigned to one of the four experimental conditions (i.e., the Black attractive suspect, the Black unattractive suspect, the White attractive suspect, and the White unattractive suspect). Following a brief overview of the experiment, participants were given consent forms, experimental materials, and a sheet defining the possible crimes. Participants were instructed not to discuss the surveys until all were returned to the researcher. Participants were encouraged to think about the case and their decisions carefully and at length before
making their final decisions. Extra credit was offered by some of the professors to the students for participating.

**Results**

A 2 (mock jurors’ race: Black vs. White) x 2 (suspects’ race: Black vs. White) x 2 (suspects’ level of attractiveness: high vs. low) MANOVA was performed. The dependent variables were guilt judgments for four crimes (i.e., theft, home invasion, aggravated robbery and aggravated assault). Box’s test of equality of covariance matrices was statistically significant; Levene’s test of equality of error variance was not statistically significant for any dependent variables. Only significant main effects and interactions are reported below. Regarding guilt judgments, multivariate Fs indicated a significant main effect for suspects’ race, Roy’s Largest Root = 0.63, $F\ (4, 320) = 5.06, p < .001, \eta_p^2 = .06$; a marginally significant effect for participants’ race, Roy’s Largest Root = 0.25, $F\ (4, 320) = 1.97, p = .098, \eta_p^2 = .02$; as well as a marginally significant 2-way interaction among mock jurors’ race and suspects’ level of attractiveness, Roy’s Largest Root = 0.26, $F\ (4, 320) = 2.06, p = .086, \eta_p^2 = .03$; which was qualified by a 3-way interaction among mock jurors’ race, suspects’ race, and suspects’ level of attractiveness, Roy’s Largest Root = 0.53, $F\ (4, 320) = 4.27, p = .002, \eta_p^2 = .05$.

**Main effects**

Univariate analyses showed that there were main effects of mock jurors’ race on perceptions of guilt for three of the four crimes: theft, $F\ (1, 323) = 4.74, p < .05, \eta_p^2 = .014$; home invasion, $F\ (1, 323) = 4.45, p < .05, \eta_p^2 = .014$; and aggravated robbery, $F\ (1, 323) = 4.97, p < .05, \eta_p^2 = .015$. Black (compared to White) mock jurors rated the suspects as less likely to be guilty of theft, home invasion, and aggravated robbery. The trend for the crime, aggravated
assault, was in the same direction but did not reach statistical significance \((p > .10)\). See Figure 2 for means and standard errors.

There were main effects of suspects’ race on guilt judgments for all four crimes: theft, \(F(1, 323) = 10.55, p < .01, \eta_p^2 = .032\); home invasion, \(F(1, 323) = 5.56, p < .05, \eta_p^2 = .017\); aggravated robbery, \(F(1, 323) = 13.18, p < .001, \eta_p^2 = .039\); and aggravated assault, \(F(1, 323) = 15.92, p < .001, \eta_p^2 = .047\). Mock jurors rated Black (compared to White) suspects as less likely to be guilty of theft, home invasion, aggravated robbery, and aggravated assault. See Figure 3 for means and standard errors.

**Interactions**

There were statistically significant interactions between mock jurors’ race and suspects’ level of attractiveness on perception of guilt for all four crimes: theft, \(F(1, 323) = 4.37, p < .05, \eta_p^2 = .013\); home invasion, \(F(1, 323) = 7.33, p < .01, \eta_p^2 = .022\); aggravated robbery, \(F(1, 323) = 4.86, p < .05, \eta_p^2 = .015\); and aggravated assault, \(F(1, 323) = 4.31, p < .05, \eta_p^2 = .013\). However, these two-way interactions were further qualified by three-way interactions among mock jurors’ race, suspects’ race, and suspects’ level of attractiveness for all four crimes: theft, \(F(1, 323) = 9.19, p < .01, \eta_p^2 = .028\); home invasion, \(F(1, 323) = 3.92, p < .05, \eta_p^2 = .012\); aggravated robbery, \(F(1, 323) = 13.67, p < .001, \eta_p^2 = .041\); and aggravated assault, \(F(1, 323) = 5.84, p < .05, \eta_p^2 = .018\).

To ease presentation of results for these interactions, we calculated a composite crime guilt index, summing participants’ ratings for all four different types of crimes. For the Mock Jurors’ Race x Suspects’ Attractiveness interaction, follow-up comparisons indicated that White participants \((M = 9.68, SE = .23)\) rated unattractive suspects as more likely to be guilty than
RACE AND ATTRACTIVENESS

Black jurors ($M = 8.32, SE = .32$); but there were no differences in ratings of guilt of attractive suspects for White ($M = 9.27, SE = .28$) participants and Black ($M = 9.39, SE = .23$).

For the Mock Jurors’ Race x Suspects’ Race x Suspects’ Attractiveness interaction, White participants (compared to Black participants) rated Black unattractive suspects as more likely to be guilty. There were no differences in White versus Black mock jurors’ ratings for the rest of the suspects (i.e., attractive Black suspect, unattractive White suspect and attractive White suspect) (see Figure 4). Each of the differences reported was statistically significant at the .05 level or greater.

**Discussion**

While at a first glance, the results we report here are somewhat surprising, this study both partially replicates and extends earlier research on the complex effects of race bias and attractiveness bias on mock jury guilt decisions. Contrary to Prediction 1, White mock jurors did not exhibit an outgroup bias. Sommers and Ellsworth (2000; 2001, see also Cohn, Bucolo, Pride & Sommers, 2009) found that when race is salient, that is, when the crime in question is racially charged, race did not have a statistically significant impact on White jurors’ decisions. Sommers and Ellsworth (2009) argue that “race salience” in Sommers and Ellsworth’s initial and subsequent studies is primarily defined as highlighting racial issues at trial/presenting racially charged incidents; making race salient during a trial causes White jurors to be concerned about introducing race bias and subsequent self-correction not to appear biased. Yet, a question remains whether “race salience” is a potential mechanism for the effect described, as we do not define it in the same way as Sommers and Ellsworth did. In the present study, racial salience consisted of stating the suspects’ race in the crime report, displaying mug shots of the suspects that clearly conveyed the suspects’ race, and the experimenters’ presence.
Specifically, the current data collection team was comprised solely of Black women. The experimenter effect brought about by the presence of an all-Black female group likely made race *glaringly salient* for the participants. This *glaring salience*, in turn, may have prompted mock jurors to answer in a manner that is “politically correct” (i.e., in a manner that is inconsistent with race stereotypes). Perhaps White mock jurors in our study overcorrected for expected racial bias. A search of the literature did not detect previous studies that document a Black experimenter effect in the mock juror literature; the current study may be the first. The same effects also could have contributed to the leniency of Black jurors to Black suspects above and beyond in-group favorability.

Partially consistent with Prediction 2, Black mock jurors were more lenient toward all suspects, and all jurors were more lenient toward Black (compared to White) suspects. Consistent with previous research (Bernard, 1979), we anticipated and found a stronger in-group bias effect for Black (compared to White) mock jurors. While Black mock jurors’ leniency toward Black suspects has been discussed in the literature (and is consistent with our Prediction 2), our study’s findings are novel because *both* White and Black jurors simultaneously exhibited greater leniency to Black subjects and are likely to be driven by different mechanisms (i.e., glaring race salience for White jurors and in-group favorability effect for Black jurors).

Contrary to Prediction 3, while unattractive suspects were not perceived as ‘more likely to be guilty’, White jurors were more likely to rate Black unattractive suspects to be guilty. The absence of a main effect for suspect attractiveness, although unanticipated, is consistent with some of the previous work (Blair, Judd, & Chapleau, 2004; Downs & Lyons, 1991; Patry, 2008; Sporer & Goodman-Delhaunty, 2009). Downs and Lyons (1991) found no effect of attractiveness on sentences for felony crimes. After controlling for criminal record, Blair, Judd
and Chapleau (2004) found no effect of attractiveness on sentence length. Patry (2008) examined the effect of deliberation on guilt judgments for attractive (compared to unattractive) defendants. Findings were that mock jurors who did deliberate in a group and at length were more likely to find guilty the more attractive (compared to the less attractive) defendant. Unfortunately, Patry (2008) confounded deliberation in a group with the length of time that mock jurors were permitted to make their guilt decisions. The current study partially replicates and expounds upon Patry’s (2008) findings. We replicate Patry’s findings that, when jurors deliberate at length and are not hasty, they are less susceptible to the physical attractiveness stereotype bias; we extend Patry’s findings by clarifying that the effect may be achieved whether the deliberation occurs in an individual setting. Our participants were encouraged to think about the case and their decisions carefully and at length before making their final decisions; results support the notion that deliberation, not necessarily group discussion, may dissipate the attractiveness effect on guilt judgments.

More interestingly, we obtained two interactions involving attractiveness. White jurors (compared to Black jurors) rated Black unattractive suspects as more likely to be guilty. While we did not observe the expected pattern of racial bias in White mock jurors, this pattern of results might signify that race bias may be masked as unattractiveness bias in White mock jurors. Notwithstanding, Black mock jurors who were exposed to the Black unattractive suspects, compared to Black mock jurors in the other three conditions (i.e., attractive black suspect, unattractive white suspect and attractive white suspect) and White mock jurors in all four conditions, rated the Black unattractive suspect as least likely to be guilty. These findings might demonstrate the additive empathy that Blacks may exhibit toward members of their own as well as other marginalized groups who are discriminated against by the larger society.
In the current study, a criminal suspect who fell within two underprivileged groups received greater sympathy and leniency only from Black jurors. There is experimental evidence that Black (compared to White) jurors are more distrustful of the judicial process, are more concerned that a jury may have rendered a wrong verdict, and are more likely than White jurors to identify with the defendant’s family/perceive it as similar to their own family (Bowers, Sandys, & Brewer, 2004). These findings explain why Black participants might be more lenient overall and especially lenient toward members of historically marginalized or persecuted groups (i.e., Black suspects and especially unattractive Black suspects). At the same time, the absence of such leniency in White jurors can be interpreted as potential bias.

We are cautiously optimistic about our pattern of findings because of the following: (a) the congruency between crime type and race, and (b) the implications for societal change. The crimes discussed in the report are considered “stereotypically Black crimes” for which Black suspects are usually considered more likely to be guilty and receive harsher sentencing (e.g., Gordon et al., 1988; Jones & Kaplan, 2003; Sunnafrank & Fontes, 1983). It is somewhat encouraging not to find this pattern in the current results. While the pattern of results including attractiveness might indicate bias or lack of sensitivity toward low-status marginalized groups (unattractive Black suspects) for White jurors, the absence of out-group bias among White jurors may be indicative of positive societal changes. The overwhelming majority of this sample is in the millennial generation – a group known for its bold embrace of diversity, inclusion and more liberal racial attitudes (e.g., American National Election Studies, Stanford University, and University of Michigan, 2016). Black criminal suspects received greater sympathy and leniency in the current study. It is possible that, if selected for a jury, members of this group (millennial college students and alumni) may not be as susceptible to the prejudices of older generations in
America that are plagued with racial biases. If selected for a jury, members of this group might be lenient toward members of historically marginalized or persecuted groups.

This study also highlights the importance of diversity at all levels of the justice system. It is the case that Blacks are overrepresented among those who are arrested, on probation and imprisoned in the United States. Blacks are largely absent, and in some cases absent altogether, from other areas of the justice system. Just as it is likely that a Black experimenter may have been an anomaly and racial prime for some in the current study, it also is likely for any (or all) of the following to be anomalies and racial primes for some in real world situations: a Black investigator, a Black prosecutor, a Black defense attorney, a Black juror, a Black jury foreperson, a Black judge, etc. While the experimenter effect and resulting racial prime in this instance led to greater leniency (on paper), a fair and impartial process (in the actual court system) remains the goal in the American justice system.

There are several potential limitations to this study. We only assessed the perception of guilt. Other measures, such as likelihood of conviction and length of sentencing might serve as better or more complete indicators of biases. We did not manipulate experimenter race or gender or recruit participants of various age groups to fully assess the experimenter effect or the leniency exhibited by the millennial generation. Our race and attractiveness manipulation relied on a limited number of photographs of male subjects (although these photographs were carefully normed on attractiveness). Future research could expand the number of photographs and include pictures of female subjects thus enabling the manipulation of defendants’ gender.

We acknowledge that these limitations should be addressed in future studies. The next step will be to replicate and extend the current research in other age groups and groups of millennials that are more gender-balanced in different regions of the United States. Additionally,
future research also could vary experimenter race and gender to further explore their effects on racial salience and mock jurors’ responses. Photographs of female suspects could also be used to investigate the complexity of race and attractiveness effects. Future research should also seek to extend the examination of race and attractiveness in other areas of the justice system such as arrest and arraignment.
References


doi:10.1073/pnas.13064171


legal attitudes on mock juror and jury decisions. *Psychological Reports, 51,* 1067–1074. doi:10.2466/pr0.1982.51.3f.1067


Footnotes


2. Sometimes effects for defendants’ attractiveness differ depending upon whether the juror is same- vs. opposite-sex (Abwender & Hough, 2001). To explore the impact of mock jurors’ race and sex and the impact of suspects’ race and level of attractiveness on participants’ guilt judgments, a 2 (mock jurors’ race: Black vs. White) x 2 (mock jurors’ sex: woman vs. man) x 2 (suspects’ race: Black vs. White) x 2 (suspects’ level of attractiveness: high vs. low) MANOVA was initially performed. There were no main effects or interactions including the mock jurors’ sex factor, so we dropped this variable from the analysis.
Figure 1. Stimuli used. Note: From the left upper corner, clockwise: Black attractive, Black unattractive, White unattractive, White attractive suspects.
Figure 2. Black versus White mock jurors’ Mean (SE) guilt judgments.
Figure 3. Mock jurors’ Mean (SE) guilt judgments of Black versus White suspects.
Figure 4. Mock jurors’ Mean (SE) composite crime guilt index as a function of mock jurors’ race, suspects’ race, and suspects’ attractiveness.