Psychology, Public Policy, and Law
2013, Vol. 19, No. 4, 466–478

1076-8971/13/$12.00 DOI: 10.1037/a0034736

Criminal Stereotypes in the Courtroom: Facial Tattoos Affect Guilt and Punishment Differently

Friederike Funk and Alexander Todorov
Princeton University

Three studies using between-subjects designs examined the effect of facial tattoos on two stages of a courtroom trial. The presence of a facial tattoo affected judgments of guilt, but it did not lead to higher punishment ratings. This was the case for different types of crime varying in offense seriousness and for faces varying in perceived trustworthiness. The effect on guilt was fully mediated by the perceived criminal appearance of the tattooed defendant. These findings are the first that systematically address the question whether facial tattoos can bias legal outcomes. They further suggest that the psychological mechanisms by which an activated criminal stereotype influences legal judgments can differ for two stages of a trial. Policy implications are discussed.

Keywords: stereotypes, punishment, tattoos, trustworthiness, guilt

Person perceptions from faces are formed automatically and rapidly (Todorov, Pakrashi, & Oosterhof, 2009). It takes people about 40 ms of exposure to a stranger's face to come to a consistent judgment of the stranger's character (Bar, Neta, & Linz, 2006). Longer exposure time to the face does not lead to a change in judgment but only to increased confidence about it (Willis & Todorov, 2006). These fast and automatic character inferences have been found to impact decisions in the courtroom; for instance, untrustworthy-looking faces require less evidence for a guilty verdict than do trustworthy-looking faces (Porter, ten Brinke, & Gustaw, 2010). Moreover, babyfaced or attractive offenders are often judged more leniently than mature-looking or unattractive offenders (Sigall & Ostrove, 1975; Zebrowitz & McDonald, 1991). Facial appearance also activates stereotypes that, in turn, influence legal decisions. Afrocentric features, for instance, have been found to affect sentencing decisions (Blair, Judd, & Chapleau, 2004; Eberhardt, Davies, Purdie-Vaughns, & Johnson, 2006), because people use these features to infer traits that are stereotypically associated with African Americans (Blair, Judd, & Fallman, 2004; Blair, Judd, Sadler, & Jenkins, 2002).

This article focuses on the effect of tattoos on legal decisions. The prevalence of tattoos is high among the general population: 24% of Americans between the ages of 18 and 50 are tattooed (Laumann & Derick, 2006). Yet despite their prevalence, tattoos are still mostly associated with negative characteristics. Perceivers judge tattooed people as more creative but less attractive, less fashionable, less athletic, less intelligent, less caring, and less religious than nontattooed people (see Degelman & Price, 2002; Hawkes, Senn, & Thorn, 2004; Resenhoeft, Villa, & Wiseman, 2008). Tattoos are also stereotypically associated with lower social class and middle class but not with upper social class (DeMello, 2000).

In addition to these negative associations, there is a *criminal* stereotype of tattoos in particular. People have clear ideas about how criminals differ in their physical appearance from noncriminals, and tattoos are the main physical characteristic that is linked to criminals (MacLin & Herrera, 2006). The perceptions of association between tattoos and delinquent behavior can already be found in children at the age of six (Durkin & Houghton, 2000).

Given these associations, studying tattoos in a legal context is highly relevant. Estimates of the incidence of tattoos among criminal populations range between 15% and 32% (Manuel & Retzlaff, 2002; Palermo, 2004); some even estimate that over 50% of all the male prisoners are tattooed (DeMello, 1993). Existing research explores the special meaning of prison-themed tattoos compared with other tattoos (see, e.g., DeMello, 1993) and has examined behavioral and psychopathological differences between prison-themed tattooed and otherwise tattooed prisoners or between tattooed and nontattooed prisoners (see, e.g., Lozano, Morgan, Murray, & Varghese, 2011; Manuel & Retzlaff, 2002; Palermo, 2004). Yet, to the best of our knowledge, no systematic research has studied the effect of tattoos on sentencing decisions. We aim to fill this gap to inform policymakers how to deal with tattooed defendants in the courtroom.

The present research focuses particularly on the subcategory of *facial* tattoos and its impact on legal decisions. Although it is easy to cover tattoos on arms and legs with clothes, facial tattoos require more effort. In some legal cases, facial tattoos get covered with make-up during trial (see, e.g., Schwartz, 2010), but in others they remain exposed (see, e.g., Casarez, 2009).

We conducted three studies to examine the effect of facial tattoos on decisions in the courtroom. In brief, we hypothesized that the presence of a facial tattoo would activate a criminal stereotype. In Study 1, we examined whether such a criminal stereotype would affect decisions on guilt. In Study 2, we looked at the effects on recommended punishment severity after guilt was

Friederike Funk and Alexander Todorov, Department of Psychology, Princeton University.

We thank Deborah Prentice, Andrew Conway, Mike North, and Pam Mueller for their comments on previous drafts of this article.

Correspondence concerning this article should be addressed to Friederike Funk, Department of Psychology, Princeton University, Princeton, NJ 08540. E-mail: ffunk@princeton.edu

presumed. In Study 3, we assessed both guilt and punishment ratings in the same research design.

Study 1

Study 1 examined the effect of facial tattoos and their criminal stereotypes on guilt judgments. Previous research has shown that once a stereotype is activated, people make dispositional and stable attributions about the causes of a person's behavior and neglect situational or external factors that might have led to the specific behavior (see, e.g., Bodenhausen & Wyer, 1985; Duncan, 1976; Gordon, 1990). Therefore, people interpret any behavior that is congruent with the stereotype as more intentional than incongruent behaviors. We thus hypothesized that an activated criminal stereotype would reduce doubt about whether the defendant actually committed the crime. In addition, when confronted with stereotypes, people use a confirmation strategy to validate their stereotype-congruent expectations instead of using a diagnostic information-processing strategy that would objectively test their expectations (see, e.g., Darley & Gross, 1983; Kelley, 1950; Snyder & Cantor, 1979). This suggests that in stereotypical situations people need less information to accept their preferred hypothesis—that is, to confirm their expectancy.

Applied to a legal proceeding, we therefore assumed that even with weak evidence, people's threshold to hand down a guilty verdict would be reduced in the presence of facial tattoos. In addition, we wanted to study whether such an effect would be affected by circumstances of the depicted offense or by the general facial appearance of the defendant, as further described below.

Potential Effect of Different Types of Crimes

Previous research on stereotypes has found that people show a high level of consensus about what certain kinds of criminals look like (Bull & Green, 1980; Reed & Reed, 1973; Yarmey, 1993). Faces that are congruent with the facial stereotype of a particular offense are more likely to be found guilty (Dumas & Testé, 2006; Macrae & Shepherd, 1989; Shoemaker, South, & Lowe, 1973). Thus, there can be an interaction between characteristics of a crime and stereotypes about the offender.

Since tattoos are stereotypically associated with lower social class and middle class—rather than upper class (DeMello, 2000)—we wanted to test whether the effect of a facial tattoo in the courtroom would depend on the type of crime committed. It might be more stereotypical for a tattooed offender to commit a blue-collar crime than a white-collar crime. We therefore decided to use two different types of crimes in our studies. We examined whether legal judgments for an offender who committed a stereotypical, blue-collar crime (e.g., assault) would be affected more strongly by the presence of facial tattoos than would be decisions for an offender who committed a white-collar crime such as tax fraud.

Different Kinds of Faces

Finally, we sought to rule out the possibility that a potential facial tattoo effect would simply be driven by pure valence judgments. Previous research has already shown that attractive offenders can receive less severe sentences (Sigall & Ostrove, 1975) and that baby-faced offenders can be more likely to win cases (Zebrowitz & McDonald, 1991). Because tattoos are associated with many negative characteristics (Degelman & Price, 2002; Hawkes et al., 2004; Resenhoeft et al., 2008), we wanted to make sure that any potential effect would be caused by an activated *criminal* stereotype, which includes not only negative connotations but also criminal information.

We addressed this concern in two ways: First, we operationalized criminal stereotype activation as a combination of criminal appearance (including negative valence) and likelihood to reoffend (judgment about criminal domain). Second, we manipulated whether the tattooed or nontattooed face appeared trustworthy or untrustworthy. Trustworthiness judgments have been shown to reliably act as a proxy for face valence evaluation (Oosterhof & Todorov, 2008). If our hypothesized effects on guilt and punishment were just driven by valence, a facial tattoo should not matter for a face that looks untrustworthy anyway. We expected that a tattoo would matter for either type of face because we assume that it is the *criminal* stereotype that is driving the effects of a facial tattoo on judgments in the courtroom rather than its negative valence only.

Pilot study for experimental face stimuli. Two faces from the Karolinska Directed Emotional Faces set (Lundqvist, Flykt, & Ohman, 1998) with neutral emotional expression and direct eye gaze were used that had been found to be either trustworthy or untrustworthy looking (Oosterhof & Todorov, 2008; supplemental material). First, we digitally adjusted the hair of one face to make it more similar to the other face so that any effect would not be caused by the different hair style. In a subsequent step, we digitally pasted a real facial tattoo onto each face, inspired by an actual criminal case (Schwartz, 2010; see Figure A1 in the Appendix).

To validate the trustworthy versus untrustworthy perception obtained in previous studies and to test whether the tattooed versions were perceived to look more criminal than the nontattooed versions, we conducted a pilot study in which participants only rated the criminal appearance of a face without reading any crime scenario.

One-hundred workers were recruited on Amazon Mechanical Turk (referred to as Mturk throughout the text; for its use in social science research see Buhrmester, Kwang, & Gosling, 2011; Paolacci, Chandler, & Ipeirotis, 2010) and were paid $0.10 for their participation. The survey took participants about 1 min to complete. Prior to conducting the analyses, data from 8 subjects were removed because they were not from the United States or because they did not indicate correctly in the manipulation checks whether the person depicted had a tattoo. The final sample consisted of 92 participants (different users than in Studies 1–3, 55 male, age range from 18–61, $M = 29.7$, $SD = 10.1$).

Participants were randomly assigned to one of the four faces (i.e., either the trustworthy or the untrustworthy not tattooed or tattooed face) and were asked to rate the extent to which they thought the depicted person looks criminal, dangerous, aggressive, trustworthy [reverse coded], and honest [reverse coded] (order randomized; 1 = *not at all*, 7 = *very much*, Cronbach's $\alpha = .87$). We take these items as a scale to measure the strength of criminal stereotype activation by the defendant's criminal appearance. We will refer to the average of these items as "criminal appearance scale" throughout the article.

To make sure that the digital adjustment of hair length did not affect trustworthiness ratings from previous studies (Oosterhof & Todorov, 2008), we first looked at the nontattooed faces only ($N = 46$). Analysis of variance (ANOVA) results confirmed that participants rated the trustworthy-looking face as more trustworthy

($M = 3.92$, $SD = 1.10$) than the untrustworthy-looking face ($M = 2.95$, $SD = 1.09$), $F(1, 44) = 8.9$, $p < .01$, $\eta_p^2 = .17$.

Next, we looked at the effect of a facial tattoo on the criminal appearance scale. When the face had a facial tattoo, participants rated its appearance as more criminal looking ($M = 5.42$, $SD = 0.88$) than when it did not have a tattoo ($M = 4.20$, $SD = 1.11$), $F(1, 88) = 35.8$, $p < .001$, $\eta_p^2 = .29$. An untrustworthy-looking face was also rated as more criminal looking ($M = 5.06$, $SD = 0.90$) than a trustworthy-looking face ($M = 4.56$, $SD = 1.36$), $F(1, 88) = 4.8$, $p = .03$, $\eta_p^2 = .05$. In addition, there was a significant interaction effect of the type of face and the presence of a facial tattoo, indicating that the facial tattoo affected criminal appearance ratings more for the trustworthy face (tattoo $M = 5.40$, $SD = 1.05$; no tattoo $M = 3.79$, $SD = 1.14$) than for the untrustworthy face (tattoo $M = 5.44$, $SD = 0.70$; no tattoo $M = 4.64$, $SD = 0.92$), Face × Tattoo, $F(1, 88) = 4.0$, $p = .05$, $\eta_p^2 = .04$. These means indicated that once both the trustworthy and untrustworthy face had a facial tattoo, they received about similar criminal appearance ratings. Important to note for the purpose of this research, however, post hoc tests indicated for both faces that appearance was rated as more criminal looking when the faces were tattooed compared to when they were not tattooed, ps $< .01$. That is, the presence of a facial tattoo added information to the face independent of whether it originally looked trustworthy or untrustworthy. Summing up, the pilot study confirmed that the faces were appropriate to manipulate face valence (with the type of face) and criminal appearance (with the presence of a facial tattoo).

Method

Participants. Participants ($N = 320$) were recruited on Mturk and were paid $0.50 for their participation. Thirty-four participants were excluded from the analyses, either because they did not pass instructional manipulation checks testing for careful reading (e.g., asking participants to leave certain items blank; see Oppenheimer, Meyvis, & Davidenko, 2009) or because they were not from the United States. Furthermore, we excluded all of the participants from the tattoo conditions who indicated in the manipulation checks that the face did not have a tattoo or who mentioned in the open-ended comments that they thought the tattoo was a scar or that is was not real. The final sample consisted of 286 Mturk users (175 male, 55 [19%] tattooed; age range from 18–64, age $M = 28.4$, $SD = 9.2$). Both sex and whether participants had a tattoo were equally distributed across conditions: sex, $\chi^2(7, 286) = 4.3$, $p = .75$; tattooed, $\chi^2(7, 286) = 3.3$, $p = .85$. Both variables did not affect our findings and will not be discussed further.

Design. The design of Study 1 was a 2 (tattoo–no tattoo) × 2 (type of crime: assault–tax fraud) × 2 (type of face: untrustworthy–trustworthy) between-subjects design. The survey was administered online and took participants about 5 min to complete.

Materials and procedure. Participants were asked to imagine they were at a court hearing. They were introduced to Jack who was accused of either assault or tax fraud. Participants saw a picture of Jack (trustworthy or untrustworthy version) that either showed a facial tattoo or not. Below the picture, participants read a short scenario about the crime of which Jack was accused. The wording of the scenarios was as follows:

> *Tax fraud:* In the course of the last few years, Jack has underpaid his taxes by $18,000. The prosecutor assumes that Jack deliberately underreported his earnings and overreported his expenses to pay fewer taxes. Jack himself says that he didn't intend to do so but that he misinterpreted how to report his taxes.
>
> *Assault:* Jack was at a bar. He bumped into another man while going to get a drink, causing the other man to spill his drink. Jack states that this happened accidentally. The other man began to scream obscenities at Jack. According to the prosecutor, Jack then punched the man, breaking his jaw. Jack himself says that the other man tried to beat him up but lost balance and fell, knocking his jaw on the counter.

Participants were first asked to rate how likely it is that Jack is guilty on a 7-point Likert scale (1 = *very unlikely*, 7 = *very likely*). Subsequently, they were asked what they would choose if they had to decide between *guilty* and *not guilty*, using a forced-choice response option. Next, participants saw Jack's face again on a subsequent page and were asked to what extent they thought Jack looks criminal, dangerous, aggressive, trustworthy [reverse coded], and honest [reverse coded] to assess criminal appearance (order randomized; all 7-point Likert scales range from *not at all* to *very much*; Cronbach's $\alpha = .84$). To avoid demand effects, we also included filler items (e.g., to what extent Jack looks friendly, similar to them, and attractive).

On the next page, participants rated the likelihood that Jack would reoffend in the future assuming that he is guilty (1 = *very unlikely* to 7 = *very likely*). Although we assumed stereotype activation to be a mediator for the effects of a facial tattoo on guilt decisions, we measured the defendant's criminal appearance and likelihood to reoffend after the dependent variable to avoid demand effects—that is, we did not want participants to think about the criminal appearance before they had to report their legal decisions, assuming that the assessment of guilt would neither affect the perception nor the self-report about the perception of the defendant's criminal appearance.

Lastly, participants indicated demographic information and answered manipulation checks about Jack's appearance, such as whether Jack had a tattoo. Participants were also asked to indicate whether they have a tattoo (yes, no). At the end of the study, participants were debriefed and thanked and provided the option to leave feedback in a textbox.

Results and Interpretation

Guilt ratings. For each continuous dependent variable, we conducted full-factorial ANOVAs with the three experimental manipulations as independent variables, allowing for two-way and three-way interactions (unless otherwise noted, these interactions were not significant). On the continuous guilt scale, we found the predicted trend, such that participants indicated that it was more likely that Jack is guilty when he had a facial tattoo (tattoo $M = 4.82$, $SD = 1.13$) than when he did not have a facial tattoo (no tattoo $M = 4.60$, $SD = 1.04$), but this difference was only marginally significant, $F(1, 278) = 3.4$, $p = .09$. There were no significant effects for type of crime or type of face on continuous ratings of guilt: type of crime, $F(1, 278) = 0.4$, $p = .51$; face, $F(1, 278) = 0.0$, $p = .99$.

For the dichotomous decision about guilt, we used binary logistic regression analyses with categorical dummy variables as predictors. When forced to decide between guilty or not guilty, participants were significantly more likely to indicate that Jack was guilty when he was tattooed than when he was not: tattoo, $\beta =$

.53, Wald = 4.8, odd's ratio (*OR*) = 1.70, $p = .03$. More precisely, 76 (53%) participants in the no tattoo condition decided that Jack was guilty and 68 (47%) decided that he was not guilty, whereas in the tattoo condition, 93 participants (65%) decided that Jack was guilty and 49 participants (35%) decided that he was not guilty. Thus, the relative risk for Jack to be judged guilty was 1.2 times higher when he had a facial tattoo compared with when he did not have a facial tattoo. There were no significant effects for type of crime or type of face on dichotomous guilt decisions: type of crime, $\beta = .06$, Wald = .06, *OR* = 1.06, $p = .81$; face, $\beta = .19$, Wald = .61, *OR* = 1.21, $p = .44$. Taken together, these findings suggest that a facial tattoo affected dichotomous decisions about guilt. This effect was independent of the type of crime and independent of the type of face.

Criminal stereotype activation. Three-way ANOVA results for the criminal appearance scale showed that when Jack had a tattoo, participants rated his appearance as more criminal looking (tattoo $M = 4.89$, $SD = 0.99$) than when he did not have a tattoo (no tattoo $M = 4.03$, $SD = 0.91$), $F(1, 278) = 59.5$, $p < .001$, $\eta_p^2 = .18$. Similarly, when Jack had an untrustworthy-looking face, his look was rated as more criminal looking (untrustworthy face $M = 4.63$, $SD = 0.95$) than when he had a trustworthy-looking face (trustworthy face $M = 4.29$, $SD = 1.10$), $F(1, 278) = 9.6$, $p < .01$, $\eta_p^2 = .03$. For both faces, post hoc tests revealed a significant difference if Jack was tattooed or not ($ps < .001$). That is, a facial tattoo led to higher ratings of criminal appearance independent of whether the face already looked untrustworthy or not. There was no significant effect of type of crime, $F(1, 278) = 1.6$, $p = .21$, and none of the interaction terms was significant.

Results also showed that when Jack had a tattoo, participants rated him as being more likely to reoffend (tattoo $M = 4.69$, $SD = 1.34$) than when he did not have a tattoo (no tattoo $M = 4.06$, $SD = 1.48$), $F(1, 278) = 15.2$, $p < .001$, $\eta_p^2 = .05$. There was no effect of type of face on the likelihood to reoffend, $F(1, 278) = 0.34$, $p = .56$. It is interesting that there was a significant effect of type of crime indicating that when Jack was accused of assault, he was also perceived to be more likely to reoffend (assault $M = 4.71$, $SD = 1.30$) than when Jack was accused of tax fraud (tax fraud $M = 4.05$, $SD = 1.50$), $F(1, 278) = 16.6$, $p < .001$, $\eta_p^2 = .06$. None of the interactions was significant.

Summing up, a tattoo led to higher criminal appearance ratings for both trustworthy- and untrustworthy-looking faces. Moreover, it was only the tattoo but not the kind of face that resulted in a higher estimated likelihood of reoffending. It therefore seems that the effect of a facial tattoo is not only about negative valence but also about criminal stereotypes.

Mediation analysis. To test whether the activated criminal stereotype statistically mediated the effect of the presence of a facial tattoo on dichotomous judgments of guilt, a mediation analysis (collapsed across both types of faces and both types of crimes) was performed following the three steps suggested by Baron and Kenny (1986). We did not include likelihood to reoffend as a statistical mediator because its bivariate correlation with dichotomous guilt judgments was not significantly different from zero.

In a first step, as reported above, logistic regression analyses showed that there was a direct effect of the presence of a facial tattoo on guilt judgments (unstandardized $c = .53$, $p = .03$). In a second step, separate regression analyses revealed that the presence of a facial tattoo affected criminal appearance ratings (unstandardized $a = .86$, $p < .001$). In a third step, we conducted logistic regression analyses with both the independent variable and the presumed mediator as predictors. Results indicated that the direct effect of a facial tattoo on dichotomous guilt judgments became nonsignificant once the criminal appearance scale was entered as a mediator (unstandardized $c' = -.05$, $p = .86$), and only criminal appearance predicted guilt judgments (unstandardized $b = .74$, $p < .001$). Following Preacher and Hayes' (2004) suggestion to establish mediation, results further revealed that the indirect effect was significant (unstandardized $ab = .64$, 99% confidence interval [0.24, 1.03]); see Figure 1 for standard errors and other details. Thus, perceived criminal appearance fully mediated the effect of a facial tattoo on guilt judgments.

We also tested whether attractiveness ratings that were assessed as a filler item could explain our findings. Attractiveness ratings significantly differed when Jack was tattooed (tattoo, $M = 2.41$, $SD = 1.27$) compared to when he was not tattooed (no tattoo, $M = 2.99$, $SD = 1.39$), $F(1, 278) = 14.5$, $p < .001$, $\eta_p^2 = .05$. Attractiveness ratings further correlated with the criminal appearance scale, $r(284) = -.41$, $p < .001$, and predicted dichotomous decisions of guilt: attractive, $\beta = -.24$, Wald = 6.9, *OR* = .79, $p = .01$ (Nagelkerke's $R^2 = .03$). Statistically, attractiveness ratings fully mediated the effect of a facial tattoo on dichotomous guilt judgments; that is, once attractiveness ratings were added as mediator (attractive, $\beta = -.21$, Wald = 4.9, *OR* = .82, $p = .03$), the effect of a facial tattoo on guilt judgments became nonsignificant (tattoo, $\beta = .42$, Wald = 2.8, *OR* = 1.52, $p = .09$). However, the explained variance of this model (Nagelkerke's $R^2 = .05$) was

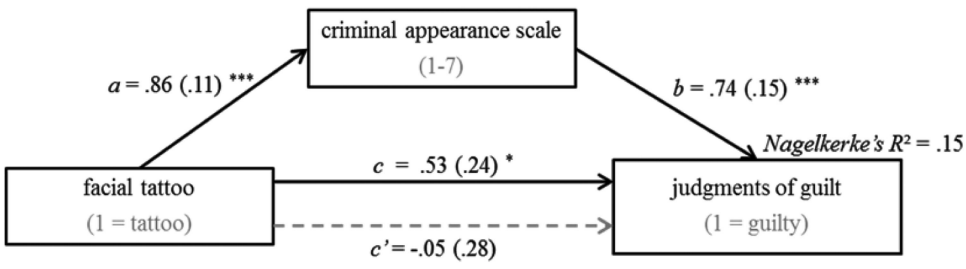

Figure 1. Perceived criminal appearance of the defendant fully mediates the effect a facial tattoo on dichotomous judgments of guilt in Study 1, unstandardized path coefficients (standard errors), $N = 286$. *** $p < .001$. * $p < .05$.

remarkably lower than the mediation model with criminal appearance as mediator. When the criminal appearance scale was entered as a third variable to the mediation model (i.e., testing the mediation depicted in Figure 1 and statistically controlling for attractiveness ratings), attractiveness also did not predict guilt judgments beyond the effect of criminal appearance: attractive, $\beta = -.04$, Wald = .16, $OR = .96$, $p = .69$. Only criminal appearance remained a significant predictor of guilt decisions (criminal, $\beta = .72$, Wald = 21.4, $OR = 2.06$, $p < .001$), and the presence of a facial tattoo became nonsignificant (tattoo, $\beta = -.06$, Wald = .04, $OR = .95$, $p = .84$; Nagelkerke's $R^2 = .15$). These results demonstrate that criminal appearance ratings are a better explanation for the effect of a facial tattoo on guilt judgments than attractiveness ratings.

Discussion

The results suggest that a facial tattoo activated a criminal stereotype. For different kinds of faces, people rated a tattooed face as looking more criminal and thought that the tattooed offender was more likely to reoffend. As hypothesized, a facial tattoo affected participants' judgments of guilt. This effect was mediated by criminal stereotype activation, particularly by the perceived criminal appearance of the defendant.

We did not find a stereotype–congruency effect (see, e.g., Shoemaker et al., 1973). That is, the effect of a tattoo was the same for both tax fraud and assault. This does not necessarily mean that stereotype–congruency effects do not exist for tattoos. Instead, it is possible that we did not include the most stereotypical crime for a tattooed offender, if such a stereotype exists. Future research is needed to identify the characteristics of a crime that are stereotypically associated with tattoos. Alternatively, it is possible that tattoos are linked to criminality in general but are not stereotypically associated with certain types of crime. For the sake of the present research, the failure of finding a stereotype–congruency effect shows that a facial tattoo can affect guilt judgments for a variety of crimes.

Although previous research has found that defendants can be put at a disadvantage when they look untrustworthy (see, e.g., Porter et al., 2010), we did not find this effect for our experimental manipulation of trustworthy versus untrustworthy faces (it occurred only for the experimental manipulation of tattooed vs. nontattooed faces). It is possible that the extent of manipulated trustworthiness was not large enough to replicate previous studies on untrustworthy-looking faces that require less evidence to be judged guilty. Nevertheless, this study suggests that it is important to study the psychology of criminal face perception and not only the psychology of the perception of trustworthiness. The presence of a facial tattoo affected guilt decisions even when the defendant was perceived to look untrustworthy anyway. Thus, our results show that a facial tattoo seems to matter for different kinds of faces.

One limitation of Study 1 was that participants did not have the possibility to indicate how severely they would like to punish the defendant. It is possible that the effect we found does not reflect a true psychological link between criminal appearance and guilt judgments but that participants only used the guilt item to express a higher desire to punish. To study the link of facial tattoos and punishment, we conducted a second study that looked at various punishment ratings instead of guilt judgments and a third study that looked at both guilt and punishment ratings in the same research design.

Study 2

In a court setting after guilt is established, an appropriate magnitude of punishment must be set. Previous research studies have found, as noted above, that once a stereotype is activated, people interpret stereotype-congruent behavior as more intentional (Bodenhausen & Wyer, 1985; Duncan, 1976; Gordon, 1990). When a person commits an offense intentionally, it is perceived to be more severe (Darley & Pittman, 2003). Perceived offense severity (i.e., moral wrongness and harmfulness), in turn, is a strong predictor of recommended punishment severity (see, e.g., Carlsmith, 2006; Carlsmith, Darley, & Robinson, 2002; Darley, 2009; Keller, Oswald, Stucki, & Gollwitzer, 2010). Thus, we wanted to test whether the presence of a facial tattoo would lead to harsher punishment ratings.

As in Study 1, we wanted to examine whether the type of crime might have an effect on the link between criminal appearance and punishment ratings. Previous research findings suggest that stereotypes do not always affect decisions on punishment severity the same way. For instance, cultural stereotypes only lead to harsher sentences when the offense is related to the stereotype (Bodenhausen & Wyer, 1985). Similarly, racial stereotypes about Black persons can lead to more severe judgments when the crime is congruent with the Black stereotype (e.g., negligent homicide or a blue-collar crime), but lead to less severe judgments when the crime is incongruent with the stereotype (e.g., fraud or a white-collar crime; see Gordon, 1990; Jones & Kaplan, 2003; Mazzella & Feingold, 1994). Therefore, we included again two types of crimes. As in Study 1, we also used two types of faces.

Method

Participants. Participants ($N = 251$) were recruited on Mturk and paid $0.80 for their participation. Twenty-two participants were excluded for failing instructional manipulation checks (Oppenheimer et al., 2009) or for being from outside the United States. Also, participants from the tattoo conditions who indicated in the manipulation checks that Jack did not have a tattoo were excluded. The final sample consisted of 229 individuals (different users than in Study 1, 122 male, 64 [28%] tattooed, age range from 18–66, age $M = 30.3$, $SD = 10.7$). Again, sex and whether participants had a tattoo were equally distributed across conditions and did not affect our findings: sex, $\chi^2(7, 229) = 2.4$, $p = .93$; tattooed, $\chi^2(7, 229) = 5.3$, $p = .63$.

Design. The design of Study 2 was again a 2 (tattoo–no tattoo) × 2 (type of crime: assault–tax fraud) × 2 (type of face: untrustworthy–trustworthy) between-subjects design. The survey was administered online and took participants about 5 min to complete.

Materials and procedure. Participants were asked to imagine being at a court hearing. They were introduced to Jack, who was accused of either assault or tax fraud. Participants saw the same pictures of Jack (trustworthy or untrustworthy version) as in Study 1, which showed either a facial tattoo or no tattoo. Below the picture of Jack, participants read about the crime of which Jack was accused. Both scenarios indicated that Jack is guilty. The wordings of the scenarios were as follows:

> *Tax fraud:* Jack deliberately underreported his earnings and overreported his expenses to pay fewer taxes. In the course of the last years, he has avoided paying $18,000 in taxes by doing this. Jack is accused of tax fraud.

Assault: Jack was at a bar. He bumped into another man while going to get a drink, causing the other man to spill his drink. The other man began to scream obscenities at him. Jack punched the man, breaking his jaw. Jack is accused of assault for an altercation causing minor bodily injury.

In general, the items were similar to Study 1. This time, however, we assessed several measures related to punishment instead of guilt. First, participants rated separately how harmful and morally wrongful the offense was (both 7-point Likert scales ranging from 1 = *not harmful/wrongful at all* to 7 = *extremely harmful/wrongful*) to assess perceived offense seriousness. Next, participants indicated the strength of their desire to punish Jack (1 = *very weak*, 7 = *very strong*). On an abstract punishment scale, they were asked how severe they thought the punishment should be (1 = *very mild,* 9 = *very severe*). At first sight, such an abstract punishment scale might not seem applicable to a court room setting, yet previous studies have found that people's abstract punishment ratings are closely related to people's concrete punishment judgments (see, e.g., Aharoni & Fridlund, 2012; Carlsmith et al., 2002). Nevertheless, participants were also asked to indicate a concrete punishment in an open-ended text field. They could indicate the kind of punishment they would recommend as well as its magnitude (e.g., if they indicated a prison sentence, they could write down the concrete number of years, months, or both).

On the subsequent page, participants saw Jack's face again and were asked to indicate the extent to which Jack looked criminal, dangerous, aggressive, trustworthy, and honest (order randomized) on a 7-point scale. Like before, these items formed a reliable scale measuring the strength of criminal stereotype activation by the defendant's criminal appearance (Cronbach's $\alpha = .85$). To avoid demand effects, we included the same filler items as in Study 1. Next, participants were asked about the likelihood that Jack will reoffend in the future (doing the same crime or a different one) using a 7-point Likert scale ranging from 1 (*very unlikely*) to 7 (*very likely*). Lastly, participants filled out the same demographic background items and manipulation check items that were also used in Study 1, and indicated whether they themselves have a tattoo (yes, no).

Results and Interpretation

Punishment ratings. Similar to Study 1, we conducted full-factorial ANOVAs for all of the continuous dependent variables with the three experimental manipulations as independent variables, allowing for two-way and three-way interactions.

First, we checked whether the presence of a facial tattoo affected the two indicators of offense seriousness. Wrongness and harmfulness ratings were correlated, $r(227) = .37$, $p < .001$, but we conducted separate ANOVAs. There was no effect of the presence of a facial tattoo on wrongness ratings, $F(1, 221) = 0.13$, $p = .72$, and no main effects on harmfulness ratings, $F(1, 221) = 0.30$, $p = .59$. There was only a significant interaction effect of type of crime and the presence of a facial tattoo on harmfulness ratings, Tattoo × Type of Crime, $F(1, 221) = 4.2$, $p = .04$, $\eta_p^2 = .02$, indicating a trend that the presence of a facial tattoo made assault appear more harmful (tattoo, $M = 5.04$; no tattoo, $M = 4.63$) and tax fraud less harmful (tattoo, $M = 3.97$; no tattoo $M = 4.26$), but post hoc tests indicated that these differences were not significant, $ps > .05$. Type of crime had a significant main effect on both wrongness and harmfulness ratings. Assault was perceived to be more harmful ($M = 4.82$, $SD = 1.14$) than tax fraud ($M = 4.11$, $SD = 1.48$), $F(1, 221) = 17.7$, $p < .001$, $\eta_p^2 = .07$; and tax fraud was perceived to be more wrongful ($M = 5.41$, $SD = 1.34$) than assault ($M = 4.64$, $SD = 1.34$), $F(1, 221) = 18.4$, $p < .001$, $\eta_p^2 = .08$. Lastly, type of face had a significant effect on both judgments. The trustworthy face received higher ratings of offense harmfulness ($M = 4.64$, $SD = 1.27$) than the untrustworthy face ($M = 4.29$, $SD = 1.44$), $F(1, 221) = 4.8$, $p = .03$, $\eta_p^2 = .02$. The same was the case for ratings of moral wrongness (trustworthy face, $M = 5.23$, $SD = 1.20$; untrustworthy face, $M = 4.83$, $SD = 1.53$), $F(1, 221) = 5.2$, $p = .02$, $\eta_p^2 = .02$. None of the other interaction terms were significant. Summing up, the measures were sensitive enough to detect differences in participants' perceptions of offense seriousness, but participants' ratings on these items were not affected by the presence of a facial tattoo.

Overall, there was no effect of facial tattoo presence on any of the punishment measures. There were no significant effects of any of the experimental manipulations on the desire to punish ($M = 3.97$, $SD = 1.55$), tattoo, $F(1, 221) = 0.0$, $p = .90$; type of crime, $F(1, 221) = 1.8$, $p = .18$; face, $F(1, 221) = 2.4$, $p = .12$. The same was the case for the abstract punishment severity on a 9-point scale ($M = 4.71$, $SD = 1.90$), tattoo, $F(1, 221) = 0.0$, $p = .96$; type of crime, $F(1, 221) = 2.2$, $p = .14$; face $F(1, 221) = 1.6$, $p = .21$. None of the interaction terms was significant.

We also assessed participants' concrete punishment severity. The majority of participants recommended concrete prison sentences in years or months, but some participants also indicated other forms of punishment, such as "community service," "fines or restitution," or a combination of imprisonment and other forms of punishment. Only the data from participants who recommended a concrete prison sentence as punishment ($N = 169$) were used to conduct parametric analyses. Because concrete punishment ratings usually show a huge variance and are highly skewed (which was also the case in this study, $M = 19.4$ months, $SD = 23.1$ months), these values were log-transformed using the natural logarithm ($M = 2.47$, $SD = 1.07$). Neither the presence of a facial tattoo nor the manipulation of face trustworthiness had an effect on these concrete punishment ratings: tattoo, $F(1, 161) = 0.3$, $p = .62$; face, $F(1, 161) = 0.0$, $p = .91$. Only type of crime had a significant effect on concrete punishment ratings, such that tax fraud ($M = 2.91$, $SE = 0.10$) garnered longer sentences than assault ($M = 1.98$, $SE = 0.11$), $F(1, 161) = 36.3$, $p < .001$, $\eta_p^2 = .19$. As reported above, tax fraud was perceived to be significantly more wrongful than assault. Although these findings are not central to the main purpose of this article, they are in line with previous studies showing that people sometimes rely more on the wrongness of an offense than on the harm caused when they are asked to assign punishment (Alter, Kernochan, & Darley, 2007; Carlsmith, 2006).

To check whether the null findings that we obtained for the presence of a facial tattoo on any of the punishment measures might have been caused by a lack of power, we conducted post hoc power analyses about the sensitivity of our experimental design (using the software G*Power; see Faul, Erdfelder, Lang, & Buchner, 2007). Given our sample size ($N = 229$) and number of experimental groups ($k = 8$) and given certain power ($1\text{-}\beta \geq .80$) and significance levels ($\alpha \leq .05$), the results indicated that our experimental design would have been sensitive enough to detect a small to medium effect size (Cohen's $f = 0.186$; $\eta_p^2 = .034$). Thus, lack of power seems to be an unlikely explanation for the findings.

Criminal stereotype activation. Again, participants rated Jack's appearance as more criminal-looking with a tattoo ($M = 5.01$, $SD = 1.12$) than without a tattoo ($M = 3.98$, $SD = 1.04$),

$F(1, 221) = 54.9$, $p < .001$, $\eta_p^2 = .20$. Similarly, Jack's appearance was rated to look more criminal when participants saw the untrustworthy-looking face ($M = 4.73$, $SD = 1.11$) than when participants saw the trustworthy-looking face ($M = 4.22$, $SD = 1.23$), $F(1, 221) = 12.7$, $p < .001$, $\eta_p^2 = .05$. There was a significant interaction of tattoo and type of face, showing that the difference of Jack being tattooed or not was greater for the trustworthy face (tattoo, $M = 4.94$; no tattoo, $M = 3.56$) than for the untrustworthy face (tattoo, $M = 5.08$; no tattoo, $M = 4.39$), Tattoo × Face, $F(1, 221) = 6.0$, $p = .02$, $\eta_p^2 = .03$. Yet for both faces, post hoc tests revealed a significant difference if Jack was tattooed or not ($ps < .001$), again indicating that the criminal appearance of a facial tattoo is relevant for different types of faces. There was no significant effect of type of crime on criminal appearance ratings. The interaction of tattoo and type of crime was not significant.

Replicating Study 1, when Jack had a tattoo, participants rated him as being more likely to reoffend ($M = 4.60$, $SD = 1.34$) than when he did not have a tattoo ($M = 3.82$, $SD = 1.46$), $F(1, 220) = 17.6$, $p < .001$, $\eta_p^2 = .07$. The effect of type of face was only marginally significant, $F(1, 220) = 3.3$, $p = .07$. There was no significant effect of type of crime, $F(1, 220) = 0.1$, $p = .71$. None of the interaction terms was significant.

Although a facial tattoo activated a criminal stereotype, there were no differences in recommended punishment severity for tattooed versus nontattooed offenders. Therefore, no mediation analysis was performed.

Discussion

The presence of a facial tattoo did not affect participants' perceptions of offense seriousness. Also, and in contrast to judgments of guilt, the presence of a facial tattoo did not lead to differences in punishment ratings. But similar to Study 1, facial tattoos activated a criminal stereotype. Participants rated a tattooed face as looking more criminal and thought that the tattooed offender was more likely to reoffend. As in the pilot study, there was a significant interaction of tattoo and type of face on criminal appearance ratings (i.e., a trustworthy face showed a higher difference in criminal appearance ratings than an untrustworthy face when it was tattooed vs. nontattooed). Nevertheless, both the trustworthy- and the untrustworthy-looking faces showed a significant difference for when they were tattooed versus nontattooed, indicating that a facial tattoo added to the information about facial trustworthiness. Study 1 did not find such an interaction effect. Future research findings will show whether an interaction effect can be replicated more often. It is important to note that, however, independent of whether there was an interaction effect or not, in each study the patterns of results showed that a facial tattoo significantly increased perceptions of criminal appearance for both types of faces.

As in Study 1, we did not find any stereotype-congruency effect for the punishment measures; that is, there were no differential effects for different types of crimes on sentencing decisions. It is possible that this lack of effect was caused by the choice of scenarios, as discussed above. Alternatively, it is also possible that it was caused by the nature of the criminal stereotype that a tattoo activates. Previous findings on stereotype-congruency effects have highlighted the match between stereotypical information and the characteristics of the crime (see, e.g., Bodenhausen & Wyer, 1985, showing that a cultural stereotype about laziness makes an offense punished more harshly when it is presumably caused by laziness). It is possible that the nature of the criminal stereotype that a tattoo activates only carries categorical information like "this person is a criminal." Once people have formed the opinion that the defendant did the crime of which he is accused, the characteristics of the crime would be irrelevant to the dichotomous nature of such a stereotype. And therefore once the decision of guilt has been made, a facial tattoo would not affect punishment ratings.

A limitation of the designs we used for Studies 1 and 2 is that participants were only asked to either make judgments about guilt or recommend punishment. Another potential limitation of the two studies is that the crimes were not sufficiently serious. It is possible that a facial tattoo would show differential effects on punishment severity when the offense was more serious. Previous studies have found that untrustworthy-looking faces require less evidence to obtain a guilty verdict than trustworthy-looking faces (Porter et al., 2010) but only when the offense was severe. To closer examine the effect of offense seriousness, in Study 3 we therefore experimentally varied the seriousness of the crime scenario.

Study 3

Results from Studies 1 and 2 suggested that a facial tattoo activates a criminal stereotype, and that this activated stereotype affects guilt decisions but does not have a direct effect on several kinds of punishment ratings. Study 3 was designed to study both guilt decisions and punishment ratings in the same research design.

Because null effects on punishment ratings (as obtained in Study 2) are difficult to interpret, we included a manipulation on offense seriousness as an independent variable. Because neither Study 1 nor Study 2 found a stereotype–congruency effect, we decided to only use one type of offense in Study 3, the assault scenario.

Building on findings from Studies 1 and 2, we predicted that a facial tattoo would once again activate a criminal stereotype and that this criminal stereotype in turn would affect judgments of guilt but not ratings of punishment severity. We predicted that recommended punishment severity would be affected by offense seriousness.

Method

Participants. The study was conducted online. Participants ($N = 206$) were recruited on Mturk and paid $0.50 for their participation. Using manipulation check and country confirmation procedures similar to Studies 1 and 2, 37 participants were excluded from the analyses. The final sample consisted of 169 Mturk users (98 male, 40 [24%] tattooed; age range from 18–66; age $M = 30.2$, $SD = 10.5$; all individuals had not participated in any of the previous studies). As in Studies 1 and 2, both sex and whether participants had a tattoo were equally distributed across conditions and again these variables did not affect our findings: sex, $\chi^2(7, 169) = 5.5$, $p = .60$; tattooed, $\chi^2(7, 169) = 6.2$, $p = .51$.

Design. The design of Study 3 was a 2 (tattoo–no tattoo) × 2 (offense seriousness: low–high) × 2 (type of face: untrustworthy–trustworthy) between-subjects design.

Materials and procedure. Participants saw a picture of Jack and were asked to imagine they were at a court hearing and that Jack has been accused of assault for an altercation causing minor or severe bodily injury. The wording of the assault scenarios [low seriousness/high seriousness] was as follows:

> Jack was at a bar. He bumped into another man while going to get a drink, causing the other man to spill his drink. Jack states that this happened accidentally. The other man began to scream obscenities at Jack. According to the prosecutor, Jack then punched the man, [causing some bruises which will probably take two weeks to heal/breaking his jaw which will probably take five months to heal]. Jack himself says that the other man tried to beat him up but lost balance and fell, knocking his jaw on the counter.

First, as in Study 1, participants indicated on a 7-point Likert scale how likely it is that Jack is guilty, followed by a forced-choice question asking for a dichotomous guilt judgment. Next, as in Study 2, participants indicated on a 9-point Likert scale how severe they think the punishment should be assuming Jack is guilty (again ranging from 1 = *very mild* to 9 = *very severe*). Participants also filled out an open-ended textbox that asked them to specify the concrete punishment they thought of.

On the subsequent page, participants saw Jack's face again and rated his criminal appearance (Cronbach's $\alpha = .88$) and the likelihood to reoffend in the same way as in Studies 1 and 2. At the end, we used the same demographic items and manipulation checks that were already assessed in Studies 1 and 2.

Results and Interpretation

Guilt ratings. Full-factorial ANOVAs for the continuous dependent variables showed that ratings of guilt likelihood were significantly higher when Jack had a tattoo ($M = 4.98$, $SD = 1.14$) than when he did not have a tattoo ($M = 4.40$, $SD = 1.10$), $F(1, 161) = 10.8$, $p = .001$, $\eta_p^2 = .06$. Neither type of face nor offense seriousness affected continuous guilt ratings: face, $F(1, 161) = 1.71$, $p = .19$; offense seriousness, $F(1, 161) = 1.98$, $p = .21$.

Similar to findings obtained in Study 1 using binary logistic regression analyses, participants were also significantly more likely to indicate on the dichotomous scale that Jack was guilty when he was tattooed than when he was not: tattoo, $\beta = .63$, Wald = 3.9, $OR = 1.89$, $p < .05$. In the no tattoo condition, 45 participants (52%) decided that Jack was guilty and 41 participants (48%) decided that he was not guilty, whereas in the tattoo condition, 56 participants (67.5%) decided that Jack was guilty and 27 participants (32.5%) decided that he was not. That is, when Jack had a facial tattoo his relative risk to be judged guilty was 1.3 times higher than when Jack did not have a facial tattoo. Neither the type of face nor the seriousness of the offense were significant predictors.

Punishment ratings. Similar to findings obtained in Study 2, there was no significant effect of a facial tattoo (or of an untrustworthy-looking face) on ratings of abstract punishment severity, $F(1, 161) = 2.64$, $p = .11$. Only offense seriousness significantly affected recommended punishment severity, $F(1, 161) = 39.2$, $p < .001$, $\eta_p^2 = .20$, indicating that Jack received higher punishment ratings for being accused of a more serious assault ($M = 4.55$, $SD = 1.74$) compared with a less serious assault ($M = 2.98$, $SD = 1.55$).

Unlike Study 2, we did not perform ANOVAs for concrete punishment ratings. Across the eight experimental conditions only 49 participants (29%) wrote "jail" as punishment, and only 24 of these participants indicated a concrete sentence length. The other concrete punishment suggestions that participants mentioned were (either alone or in combination) community service (40%), probation (20%), fines (28%), paying medical bills (14%), and anger management classes (11%).

Criminal stereotype activation. For both types of faces and across both levels of offense seriousness, participants rated Jack's criminal appearance significantly higher when participants saw Jack tattooed ($M = 5.06$, $SD = 1.07$) compared with nontattooed ($M = 3.95$, $SD = 1.01$), $F(1, 161) = 52.4$, $p < .001$, $\eta_p^2 = .25$. Participants also rated Jack's appearance as more criminal when Jack was untrustworthy looking ($M = 4.84$, $SD = 1.06$) compared with when he was trustworthy looking ($M = 4.15$, $SD = 1.20$), $F(1, 161) = 19.8$, $p < .001$, $\eta_p^2 = .11$. There were no significant effects for offense seriousness. As in Study 1, none of the two-way or three-way interactions were significant.

In addition, Jack was rated significantly more likely to reoffend when he had a facial tattoo ($M = 5.11$, $SD = 1.02$) than when he did not have a facial tattoo ($M = 4.06$, $SD = 1.32$), $F(1, 161) = 33.0$, $p < .001$, $\eta_p^2 = .17$. There were no significant differences for the two kinds of faces (nor for the two levels of offense seriousness), $F(1, 161) = 3.09$, $p = .08$.

Path analysis. To test whether criminal appearance mediated the effect of facial tattoos on guilt ratings, mediation analyses were performed, including offense seriousness as an additional independent variable and punishment severity as an additional dependent variable. Similar to Study 1, we focused on criminal appearance ratings as a mediator and did not include likelihood to reoffend as a potential second mediator. Ratings on Jack's likelihood to reoffend bivariately correlated with continuous guilt ratings, $r(167) = .21$, $p < .01$, but did not remain a significant predictor once criminal appearance was also entered into the model.

The analyses yielded similar results for both dichotomous and continuous guilt judgments. For the sake of interpretability, we discuss the results for continuous guilt ratings this time. Similar to Study 1, the effect of a tattoo on guilt ratings was fully mediated by ratings on the defendant's criminal appearance (see Figure 2). That is, the direct effect of the presence of a facial tattoo on guilt ratings ($c = .25$, $p = .001$) became nonsignificant when criminal appearance was entered as a mediator ($c' = .01$, $p = .86$; $b = .50$, $p < .001$). Moreover, the indirect effect of a facial tattoo on guilt ratings through perceived criminal appearance was significant ($ab = .24$, bootstrap [$N = 1{,}000$] 99% CI for the indirect effect [0.11, 0.37]).[1] Punishment ratings were affected by the manipulation of offense seriousness and by participants' guilt ratings but were not affected by the presence of a facial tattoo (see also ANOVA results above) or by the criminal appearance ratings.

[1] Logistic regressions with dichotomous guilt judgments would lead to the same conclusions that are discussed in the main text concerning continuous guilt judgments. As we reported, logistic regression analyses revealed a significant effect of the presence of a facial tattoo on dichotomous guilt judgments (unstandardized $c = .63$, $p < .05$). The presence of a facial tattoo also affected criminal appearance ratings (unstandardized $a = 1.11$, $p < .001$). Logistic regression analyses indicated that the direct effect of a facial tattoo on dichotomous guilt judgments became nonsignificant once the criminal appearance scale was entered as a mediator (unstandardized $c' = -.28$, $p = .47$), and only criminal appearance significantly predicted dichotomous guilt judgments (unstandardized $b = .96$, $p < .001$). In addition, the indirect effect of facial tattoos on dichotomous guilt judgments via criminal appearance ratings was significant (unstandardized $ab = 1.06$, 99% CI [0.38; 1.74]).

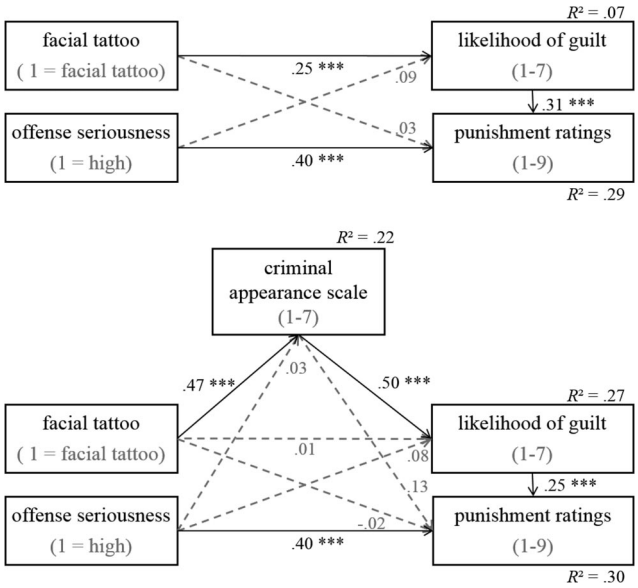

Figure 2. Differential direct effects of a facial tattoo and offense seriousness on guilt ratings and punishment ratings in Study 3 (above), mediated by the perceived criminal appearance of the defendant (below); standardized path coefficients ($N = 169$). *** $p \leq .001$, the dashed paths are all nonsignificant; $ps > .07$.

Criminal appearance ratings correlated negatively with attractiveness ratings that were assessed as filler items, $r(167) = -.49$, $p < .001$. To test the possibility that the defendant's rated attractiveness could impact the effect of facial tattoos on guilt ratings, we wanted to run the same mediation analysis with attractiveness as mediator, but there was no significant correlation between attractiveness ratings and continuous guilt ratings, $r(167) = -.13$, $p = .09$. When attractiveness was entered as a mediator nonetheless, the direct effect of a facial tattoo on guilt ratings remained significant, suggesting that the effect of a facial tattoo on guilt ratings is not only due to negative valence.

Discussion

Replicating Studies 1 and 2, Study 3 found that the presence of a facial tattoo affected guilt but not punishment ratings. Again, a facial tattoo activated a criminal stereotype. In addition, the effect of a facial tattoo on guilt ratings was fully mediated by the perceived criminal appearance of the defendant.

We included a manipulation of offense seriousness for two reasons. First, we wanted to test for the possibility that a facial tattoo might have differential effects on guilt or punishment severity for different levels of offense seriousness. Results from Study 3 did not replicate previous findings showing that untrustworthy faces require less evidence for a guilty verdict only when offense seriousness is high (Porter et al., 2010) nor did Study 3 find differential effects of a facial tattoo for different levels of offense seriousness. It is possible that we did not find such an effect because the range of seriousness we used was not wide enough. The scenarios used by Porter et al. included cases that were more severe, such as murder. Future research should continue to examine the effect of tattoos for different levels of offense seriousness. For the sake of this research, these findings suggest that the effects found for facial tattoos are generalizable across different levels of offense seriousness.

Second, we included a manipulation of offense seriousness to avoid obtaining complete null effects of all our independent variables on punishment ratings like in Study 2. In Study 3, offense seriousness affected punishment ratings but facial tattoos did not. That we found an effect for offense seriousness on abstract punishment severity means that the scales we used were sensitive enough to detect psychological differences in desired punishment severity. Therefore the null findings for the effect of facial tattoos on abstract punishment ratings found in Studies 2 and 3 were not artificially caused by properties of the punishment scale.

General Discussion

Previous research has shown that character attributions drawn from the physical appearance of a defendant and related stereotypes can influence legal decisions (see, e.g., Bodenhausen & Wyer, 1985; Eberhardt et al., 2006; Zebrowitz & McDonald, 1991). We examined whether facial tattoos would show similar effects. Building on previous research findings, we hypothesized that a facial tattoo would activate a criminal stereotype (see, e.g., MacLin & Herrera, 2006). We expected that the defendant would be rated to appear more criminal and to be more likely to reoffend in the future. In three experimental studies, a facial tattoo showed this effect.

In addition, we aimed at examining whether there were different effects of a tattoo on two stages of a trial; that is, making judgments of guilt versus assigning punishment. Results showed that the presence of a facial tattoo and its activated criminal stereotype influenced decisions of guilt, such that a tattooed defendant was more likely to be found guilty than a nontattooed defendant. However, tattooed and nontattooed defendants received the same amount of punishment once guilt was certain. Participants in our study neither indicated a higher desire to punish a tattooed defendant nor did they actually assign harsher punishment.

Each of the effects was independent of the type of crime (blue-collar or white-collar crime), independent of the type of face we used (trustworthy or untrustworthy looking), and unaffected by the level of offense severity. This indicates that the findings are generalizable across different crimes and different types of faces.

These findings add to the general literature on stereotypes and their potential effects on judgments in the courtroom and show the importance of disentangling measures of guilt from sentence severity. Criminal stereotypes might influence legal decisions differently depending on the stage of the trial. It is important to note that although we did not find a direct effect of a tattoo on punishment ratings, this does not imply that a facial tattoo does not affect the second stage of a trial at all. Only people who are judged guilty receive a punishment (see also Jones & Kaplan, 2003, on this matter). Yet, we think it is important to disentangle the two stages when one aims to understand the psychological basis of the criminal stereotype effect on legal outcomes.

Although previous research has found that stereotypes in general can influence sentence length (see, e.g., Bodenhausen & Wyer, 1985; Mazzella & Feingold, 1994; Sigall & Ostrove, 1975), facial tattoos did not show such an effect. Presumably, this divergence is caused by the nature of the tattoo stereotype. The content of the criminal stereotype that tattoos activate is very unspecific

and covers a broad range of negative characteristics: People with a facial tattoo are perceived as deviant from the norm (i.e., they are perceived to look less trustworthy, less honest, more dangerous, and more aggressive). This characteristic does not add information to the nature of the crime itself. Specific stereotypes, on the other hand, such as cultural stereotypes about laziness (see Bodenhausen & Wyer, 1985), affect punishment ratings more directly, for instance by making the defendant appear more blameworthy for what he or she did (especially when the offense is related to laziness) and thereby letting the offense itself seem more serious (Darley & Pittman, 2003). The general information that a facial tattoo seems to convey is that the person is more likely to commit a crime. Such a criminal stereotype might matter for any kind of crime. In addition, it only affects judgments of guilt but not sentence length like other stereotypes do.

Limitations and Future Directions

Our findings are limited by the fact that procedures of actual trials differ from 5-min studies. Although actual trials present evidence in different ways and have more features that influence legal decisions, we think that our results can speak to the question whether facial tattoos can potentially bias these decisions.

We used scales that many psychological studies on punishment motives have used before. Nevertheless, future studies should also use punishment scales that show a higher ecological validity and show a concrete range of possible sentences for a particular offense to rule out possible effects of a facial tattoo on punishment ratings.

In addition, these studies were juror studies in which participants were asked individually about their opinion. In a jury setting, it is possible that facial tattoos have different effects on guilt judgments. Future studies should examine whether the bias of facial tattoos can be prevented by group deliberation.

Another limitation of our studies is that we studied only the effect of a prison-themed tattoo (barbed wire). Although this kind of tattoo is likely to occur in court settings and therefore is very important to study (Casarez, 2009; Schwartz, 2010), its effects might not be generalizable to any kind of tattoo. Character inferences have been shown to be affected by the specific image that is depicted by a tattoo (a dragon leads to more negative ratings than a dolphin; see Resenhoeft et al., 2008). Previous studies also found that the bigger a tattoo, the more negatively it is perceived (Hawkes et al., 2004). Moreover, it is possible that the location of a facial tattoo leads to higher ratings of criminality than other kinds of tattoos. Future studies should therefore manipulate both type of tattoo and location to examine whether the findings were caused by criminal imagery or the facial tattoo itself and whether the criminal stereotype applies to any kind of tattoo or whether it is special to facial tattoos.

It is interesting that we did not find any differences in ratings between participants who indicated to have a tattoo (and most of them probably did not have a facial tattoo) and participants who indicated not to have a tattoo. Assuming that tattooed people are less biased toward tattoos, this lack of a difference could mean that the effect of a facial tattoo is distinct from a general tattoo effect and should not be generalized to all kinds of tattoos.

Future research should also vary the tattooed persons' sex and ethnicity to study whether female defendants elicit the same criminal stereotype as male tattooed defendants and whether a tattoo leads to similar changes of character ratings for Black or Hispanic defendants. Previous research suggests that crime stereotypes can differ for different ethnicities, such that Hispanic Americans compared to European Americans are perceived to be more likely to commit assault, for instance, whereas Black Americans compared to European Americans are perceived to be more likely to commit blue-collar crimes like robbery or burglary (Bodenhausen & Wyer, 1985; Gordon, 1990; Esqueda, 1997). In our studies, we only used pictures of European Americans, but it is possible that there are differential tattoo stereotypes for different ethnic groups that could lead to different effects on guilt or punishment ratings for defendants with different ethnicities.

It is also unclear what exactly people perceive when they see a facial tattoo. Tattoos are a form of visual communication that—independent of their original meaning—might be perceived as a symbol that the tattooed belongs to a different subculture and that the tattooed wants to be different from the mainstream (Kosut, 2000). We did not assess what people think about the origin of the facial tattoo. Some participants might have interpreted the facial tattoo as a sign that the defendant does not adhere to the general values of society, for instance, or that he got the tattoo to look tougher or because he belongs to a criminal gang.

Future research should further look at the mechanism for why a facial tattoo and its criminal stereotype affect guilt decisions. Is it because perceivers attribute the offense to stable internal dispositions rather than external circumstances (cf. Gordon, 1990, concerning the link of offender race and stereotypical crimes)? If that were the case, it would be even more surprising that a facial tattoo did not affect punishment severity, because intentionally committed crimes could be perceived as more severe and could lead to an increased desire to punish (Darley & Pittman, 2003). Alternatively, facial tattoos might affect guilt decisions because perceivers use a confirmation strategy to validate their stereotype-congruent information and therefore need less information to confirm their expectancy. Thus, they may not objectively process information regarding the crime (e.g., Jones & Kaplan, 2003). Although this research established that facial tattoos affect guilt decisions, many questions about the complete mechanism still remain to be explored.

Policy Implications

The present research findings show that facial tattoos can be linked to unfair prejudice, as they activate a criminal stereotype and increase the likelihood of guilty verdicts. This danger of prejudice seems to apply to tattooed and nontattooed jurors alike because tattooed participants and nontattooed participants were equally affected by the presence of facial tattoos in all of our studies.

These findings suggest that the impact of facial tattoos should be taken seriously by policymakers. It is already required by the law to allow prisoners to wear street clothes at court, for instance, instead of prison garb to prevent biases and receive a fair trial. Yet as of now, the court has the discretion to decide about the display of the defendant's tattoos on a case-by-case basis, and there is no widely shared standard on how the court should deal with tattoos in the courtroom (see Lozar, 2012, for a discussion of sample cases). Our research findings suggest that if guilt is uncertain and the defendant wishes to cover visible tattoos, the court should grant

the right to do so. Only in cases where the defendant pleads guilty and is sentenced immediately, our findings suggest that a facial tattoo would not bias legal outcomes.

There are several challenges that need to be addressed by policymakers. First, even if the court lets a defendant cover facial tattoos, should the court also be responsible to help the defendant do so? How should the court deal with the case of a defendant who wishes to cover facial tattoos but does not have the means to pay for it, for instance? On the one hand, one could argue that the self-infliction of a facial tattoo comes with certain risks and that the defendant should be aware that there are certain stereotypes associated with tattoos. It could be argued that if a defendant wishes to hide tattoos with make-up, the defendant needs to cover the expenses for a make-up artist. On the other hand, it is important that stereotypes and their biasing effects in the courtroom are taken seriously, even if the feature that activates the stereotype was self-inflicted. Nobody would deny that babyfaced, attractive, and untrustworthy-looking defendants all have the right to a fair trial. If facial tattoos bias the outcome of a trial, one could argue, it is not the defendant's fault that criminal stereotypes continue to be linked to tattoos despite their considerable occurrence among the general population.

Another challenge that policymakers need to address is where to draw the line. If facial tattoos count as biasing and are allowed to be covered and future research identifies other facial characteristics that potentially bias the jury and that can be changed, should there always be new policies that specify how to deal with these facial features? If facial tattoos can be covered, how should the court deal with bushy eyebrows that may make the eyes look sinister, for instance? Should administrative costs and the effort for the court to deal with these policies be limited, or should the criminal justice system do whatever it takes to grant a fair trial?

Lastly, criminal stereotypes are also linked to characteristics that could not be physically changed at court, even if the defendant would like to do so, such as afrocentric facial features (Blair, Judd, & Chapleau, 2004; Eberhardt et al., 2006). How can the court be equally fair to defendants encountering all kinds of unfair prejudice linked to facial appearance?

Independent of how these questions are answered, our findings highlight the need for jurors and judges to be educated about potentially biasing effects of criminal stereotypes and facial tattoos. Whether such an education would be sufficient as a possible alternative to the use of make-up remains to be tested in future research studies.

Conclusion

Despite limitations, the present studies clearly found that a facial tattoo activated a criminal stereotype and that a criminal stereotype can affect decisions on guilt. These results indicate that more research on tattooed defendants is needed. Previous research on tattoos in relation to the justice system has mostly focused on tattooed prisoners. This line of research shows, for instance, that inmates who have tattoos are at greater risk for recidivism (Lozano et al., 2011). This article emphasizes the importance of studying tattooed defendants as well. Because the effects of facial tattoos on guilt judgments found in these studies document the danger of unfair prejudice linked to facial tattoos, people with facial tattoos might already be at a greater risk to be judged guilty and to go to prison in the first place when the defendant's guilt is ambiguous. We hope that our research findings will raise awareness for the potentially biasing effects of facial tattoos on courtroom decisions and stimulate discussions about effective and uniform ways to prevent those biases.

References

Aharoni, E., & Fridlund, A. J. (2012). Punishment without reason: Isolating retribution in lay punishment of criminal offenders. *Psychology, Public Policy, and Law, 18,* 599–625. doi:10.1037/a0025821

Alter, A. L., Kernochan, J., & Darley, J. M. (2007). Transgression wrongfulness outweighs its harmfulness as a determinant of sentence severity. *Law and Human Behavior, 31,* 319–335. doi:10.1007/s10979-006-9060-x

Bar, M., Neta, M., & Linz, H. (2006). Very first impressions. *Emotion, 6,* 269–278. doi:10.1037/1528-3542.6.2.269

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51,* 1173–1182. doi:10.1037/0022-3514.51.6.1173

Blair, I. V., Judd, C. M., & Chapleau, K. M. (2004). The influence of afrocentric facial features in criminal sentencing. *Psychological Science, 15,* 674–679. doi:10.1111/j.0956-7976.2004.00739.x

Blair, I. V., Judd, C. M., & Fallman, J. L. (2004). The automaticity of race and afrocentric facial features in social judgments. *Journal of Personality and Social Psychology, 87,* 763–778. doi:10.1037/0022-3514.87.6.763

Blair, I. V., Judd, C. M., Sadler, M. S., & Jenkins, C. (2002). The role of afrocentric features in person perception: Judging by features and categories. *Journal of Personality and Social Psychology, 83,* 5–25. doi:10.1037/0022-3514.83.1.5

Bodenhausen, G. V., & Wyer, R. S. (1985). Effects of stereotypes in decision making and information-processing strategies. *Journal of Personality and Social Psychology, 48,* 267–282. doi:10.1037/0022-3514.48.2.267

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's mechanical turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science, 6,* 3–5. doi:10.1177/1745691610393980

Bull, R. H., & Green, J. (1980). The relationship between physical appearance and criminality. *Medicine, Science, and the Law, 20,* 79–83. doi:10.1177/002580248002000202

Carlsmith, K. M. (2006). The roles of retribution and utility in determining punishment. *Journal of Experimental Social Psychology, 42,* 437–451. doi:10.1016/j.jesp.2005.06.007

Carlsmith, K. M., Darley, J. M., & Robinson, P. H. (2002). Why do we punish? Deterrence and just deserts as motives for punishment. *Journal of Personality and Social Psychology, 83,* 284–299. doi:10.1037/0022-3514.83.2.284

Casarez, J. (2009, February 18). Defendant's tattoos become an issue at trial. *CNN News*. Retrieved from http://insession.blogs.cnn.com/2009/02/18/defendants-tattoos-reporters-notebook/

Darley, J. M. (2009). Morality in the law: The psychological foundations of citizens' desires to punish transgressions. *Annual Review of Law and Social Science, 5,* 1–23. doi:10.1146/annurev.lawsocsci.4.110707.172335

Darley, J. M., & Gross, P. H. (1983). A hypothesis-confirming bias in labeling effects. *Journal of Personality and Social Psychology, 44,* 20–33. doi:10.1037/0022-3514.44.1.20

Darley, J. M., & Pittman, T. S. (2003). The psychology of compensatory and retributive justice. *Personality and Social Psychology Review, 7,* 324–336. doi:10.1207/S15327957PSPR0704_05

Degelman, D., & Price, N. D. (2002). Tattoos and ratings of personal characteristics. *Psychological Reports, 90,* 507–514. doi:10.2466/pr0.2002.90.2.507

DeMello, M. (1993). The convict body: Tattooing among male American prisoners. *Anthropology Today, 9,* 10–13. doi:10.2307/2783218

DeMello, M. (2000). *Bodies of inscription: A cultural history of the modern tattoo community*. Durham, NC: Duke University Press.

Dumas, R., & Testé, B. (2006). The influence of criminal facial stereotypes on juridic judgments. *Swiss Journal of Psychology/Schweizerische Zeitschrift für Psychologie/Revue Suisse de Psychologie, 65,* 237–244.

Duncan, B. L. (1976). Differential social perception and attribution if intergroup violence: Testing the lower limits of stereotyping of Blacks. *Journal of Personality and Social Psychology, 34,* 590–598. doi:10.1037/0022-3514.34.4.590

Durkin, K., & Houghton, S. (2000). Children' and adolescents' stereotypes of tattooed people as delinquent. *Legal and Criminological Psychology, 5,* 153–164. doi:10.1348/135532500168065

Eberhardt, J. L., Davies, P. G., Purdie-Vaughns, V. J., & Johnson, S. L. (2006). Looking deathworthy. *Psychological Science, 17,* 383–386. doi:10.1111/j.1467-9280.2006.01716.x

Esqueda, C. W. (1997). European American students' perceptions of crimes committed by five racial groups. *Journal of Applied Social Psychology, 27,* 1406–1420. doi:10.1111/j.1559-1816.1997.tb01605.x

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. doi:10.3758/BF03193146

Gordon, R. A. (1990). Attributions for blue-collar and white-collar crime: The effects of subject and defendant race on simulated juror decisions. *Journal of Applied Social Psychology, 20,* 971–983. doi:10.1111/j.1559-1816.1990.tb00385.x

Hawkes, D., Senn, C., & Thorn, C. (2004). Factors that influence attitudes toward women with tattoos. *Sex Roles, 50,* 593–604. doi:10.1023/B:SERS.0000027564.83353.06

Jones, C. S., & Kaplan, M. F. (2003). The effects of racially stereotypical crimes on juror decision-making and information-processing strategies. *Basic and Applied Social Psychology, 25,* 1–13. doi:10.1207/S15324834BASP2501_1

Keller, L., Oswald, M., Stucki, I., & Gollwitzer, M. (2010). A closer look at an eye for an eye: Laypersons' punishment decisions are primarily driven by retributive motives. *Social Justice Research, 23,* 99–116. doi:10.1007/s11211-010-0113-4

Kelley, H. H. (1950). The warm-cold variable in first impressions of persons. *Journal of Personality, 18,* 431–439. doi:10.1111/j.1467-6494.1950.tb01260.x

Kosut, M. (2000). Tattoo narratives: The intersection of the body, self-identity and society. *Visual Sociology, 15,* 79–100. doi:10.1080/14725860008583817

Laumann, A. E., & Derick, A. J. (2006). Tattoos and body piercings in the United States: A national data set. *Journal of the American Academy of Dermatology, 55,* 413–421. doi:10.1016/j.jaad.2006.03.026

Lozano, A. T. R., Morgan, R. D., Murray, D. D., & Varghese, F. (2011). Prison tattoos as a reflection of the criminal lifestyle. *International Journal of Offender Therapy and Comparative Criminology, 55,* 509–529. doi:10.1177/0306624X10370829

Lozar, R. (2012, January). Tattoos as evidence. *California Lawyer*. Retrieved from http://www.callawyer.com/Clstory.cfm?eid=919809

Lundqvist, D., Flykt, A., & Ohman, A. (1998). *Karolinska directed emotional faces*. Stockholm, Sweden: Psychology Section, Department of Clinical Neuroscience, Karolinska Institute.

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

Macrae, C. N., & Shepherd, J. W. (1989). Do criminal stereotypes mediate juridic judgements? *British Journal of Social Psychology, 28,* 189–191. doi:10.1111/j.2044-8309.1989.tb00860.x

Manuel, L., & Retzlaff, P. D. (2002). Psychopathology and tattooing among prisoners. *International Journal of Offender Therapy and Comparative Criminology, 46,* 522–531. doi:10.1177/030662402236738

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,* 1315–1338. doi:10.1111/j.1559-1816.1994.tb01552.x

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences of the United States of America, 105,* 11087–11092. doi:10.1073/pnas.0805664105

Oppenheimer, D. M., Meyvis, T., & Davidenko, N. (2009). Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology, 45,* 867–872. doi:10.1016/j.jesp.2009.03.009

Palermo, G. B. (2004). Tattooing and tattooed criminals. *Journal of Forensic Psychology Practice, 4,* 1–25. doi:10.1300/J158v04n01_01

Paolacci, G., Chandler, J., & Ipeirotis, P. G. (2010). Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making, 5,* 411–419.

Porter, S., ten Brinke, L., & Gustaw, C. (2010). Dangerous decisions: The impact of first impressions of trustworthiness on the evaluation of legal evidence and defendant culpability. *Psychology, Crime & Law, 16,* 477–491. doi:10.1080/10683160902926141

Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, & Computers, 36,* 717–731. doi:10.3758/BF03206553

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

Resenhoeft, A., Villa, J., & Wiseman, D. (2008). Tattoos can harm perceptions: A study and suggestions. *Journal of American College Health, 56,* 593–596. doi:10.3200/JACH.56.5.593-596

Schwartz, J. (2010, December 5). Extreme makeover: Criminal court ed. *The New York Times*. Retrieved from http://www.nytimes.com/2010/12/06/us/06tattoo.html

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

Sigall, H., & Ostrove, N. (1975). Beautiful but dangerous: Effects of offender attractiveness and nature of the crime on juridic judgment. *Journal of Personality and Social Psychology, 31,* 410–414. doi:10.1037/h0076472

Snyder, M., & Cantor, N. (1979). Testing hypotheses about other people: The use of historical knowledge. *Journal of Experimental Social Psychology, 15,* 330–342. doi:10.1016/0022-1031(79)90042-8

Todorov, A., Pakrashi, M., & Oosterhof, N. N. (2009). Evaluating faces on trustworthiness after minimal time exposure. *Social Cognition, 27,* 813–833. doi:10.1521/soco.2009.27.6.813

Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after a 100-ms exposure to a face. *Psychological Science, 17,* 592–598. doi:10.1111/j.1467-9280.2006.01750.x

Yarmey, A. D. (1993). Stereotypes and recognition memory for faces and voices of good guys and bad guys. *Applied Cognitive Psychology, 7,* 419–431. doi:10.1002/acp.2350070505

Zebrowitz, L. A., & McDonald, S. M. (1991). The impact of litigants' baby-facedness and attractiveness on adjudications in small claims courts. *Law and Human Behavior, 15,* 603–623. doi:10.1007/BF01065855

(*Appendix follows*)

Appendix

Experimental Face Stimuli

Figure A1. Pictures from the Karolinska Directed Emotional Faces set (Lundqvist et al., 1998). To the left: AM66NES with digitally adjusted hair (rated as less trustworthy). To the right: AM38NES (rated as more trustworthy).

Received December 29, 2012
Revision received September 19, 2013
Accepted September 19, 2013 ■