

Dangerous decisions: the impact of first impressions of trustworthiness on the evaluation of legal evidence and defendant culpability

Stephen Porter^{a*}, Leanne ten Brinke^a and Chantal Gustaw^b

^a*Department of Psychology, University of British Columbia Okanagan, 3333 University Way, Kelowna, British Columbia, V1V 1V7 Canada;* ^b*Department of Psychology, Dalhousie University, 1355 Oxford Street, Halifax, Nova Scotia, B3H 4J1 Canada*

(Received 19 August 2008; final version received 18 March 2009)

There is little support for the long-standing assumption that judges and jurors can accurately assess credibility. According to Dangerous Decisions Theory (DDT; Porter & ten Brinke, *Legal and Criminological Psychology*, 14, 119–134, 2009), intuitive evaluations of trustworthiness based on the face may strongly bias the interpretation of subsequent information about a target. In a courtroom setting, the assessment of evidence provided by or concerning a defendant may be fundamentally flawed if its interpretation is influenced by an initial, spontaneous assessment of trustworthiness. In an empirical test of DDT, participants were presented with two vignettes describing major or minor crimes, accompanied by a photograph of the supposed defendant, previously rated as highly trustworthy or untrustworthy in appearance. Participants evaluated culpability following the presentation of evidence in each case. Participants required less evidence to arrive at a guilty verdict and were more confident in this decision for untrustworthy-appearing defendants. The current evidence supports DDT and has implications for legal decision-making practices.

Keywords: facial trustworthiness; intuition; bias; legal decision-making; credibility assessment

Introduction

Judges and jurors are faced with the task of weighing the available evidence against a defendant following arguments presented by adversarial parties. In some cases, their decisions come down to a ‘credibility contest’ as observed by Judge Josephson in the Air India mass murder case (*R v. Malik & Bagri*, 2005): ‘... the determination of guilt devolves to the weighing of the credibility of a number of witnesses who testified in these proceedings’. Thus, particularly in cases lacking physical evidence, the ability of legal decision-makers to make accurate determinations of credibility through an unbiased, objective evaluation of evidence for or against the defendant’s guilt is essential. While our justice system acknowledges that decisions are not infallible, relying on the principle of reasonable doubt, there is good reason to question the validity of credibility assessment in the courtroom. In a Canadian inquiry into the wrongful murder conviction of Thomas Sophonow, Justice Cory observed that witnesses commonly use deception, and many are ‘smooth and convincing liars’ who

*Corresponding author. Email: stephen.porter@ubc.ca

have fooled judges and juries (Wilson, 2003). As highlighted by a recent string of notorious wrongful convictions (in Canada, these include Donald Marshall Jr, David Milgaard, and Guy Paul Morin), concern over credibility assessment in the courtroom is warranted given the enormous individual and societal consequences that can result from incorrect decisions.

Seeking the 'ring of truth'

An examination of case law illuminates the judiciary's dominant attitude that credibility assessment is a straightforward matter, best guided by simple common sense. In *R. v. Marquard* (1993), the Supreme Court of Canada concluded that: 'Credibility is a matter within the competence of lay people. Ordinary people draw conclusions about whether someone is lying or telling the truth on a daily basis' (p. 248). Further, in the Supreme Court case *R. v. Francois* (1994), Justice McLachlin stated: 'In the end, the jury must decide whether it believes the witness's story in whole or in part. That determination turns . . . on the demeanour of the witness and the common sense of the jury'. Despite the general attitude that people are able to make accurate determinations of credibility based on common sense, empirical research suggests that deception detection is a flawed process with errors occurring in approximately 45% of all assessments (e.g. Bond & DePaulo, 2006). While one might predict that professionals who detect deception on a daily basis (e.g. judges, police officers, etc.) would outperform laypersons, they too typically perform around the level of chance, or even below chance, in judging the credibility of speakers (Ekman & O'Sullivan, 1991; Porter, Woodworth, & Birt, 2000; see Porter & ten Brinke, 2009; Vrij, 2008). However, judges and laypersons alike tend to be highly confident in their assessments despite the fallibility of their judgements (e.g. Porter & ten Brinke, 2009; Vrij, 2008). While it is not possible to know how often mistakes concerning credibility occur in the courtroom, numerous wrongful convictions contest the belief that such assessments and the evaluation of evidence in general are matters of mere common sense. As such, the common sense argument advocated by the high courts, and the confidence that is placed in these decisions, may be considered the first sign of 'danger' in courtroom assessments of credibility.

Dangerous decisions in the courtroom

Previous research has attributed faulty credibility assessments, in part, to a reliance on false, stereotypical beliefs about deceptive behaviour and overconfidence in one's ability to detect deception, leading to a biased 'tunnel vision' view of the facts (e.g. Meissner & Kassin, 2002; Porter, McCabe, Woodworth & Peace, 2007). In addition to these problems, the Dangerous Decisions Theory (DDT) (Porter & ten Brinke, 2009) suggests that instantaneous impressions of trustworthiness based on facial appearance may play a major role in the assessment of credibility and ensuing decisions about the target. DDT suggests that motivated assessments of guilt or innocence may be influenced by facial appearance biases to such an extent that the ensuing decisions are largely irrational (Kahneman & Tversky, 1982). The human face provides a canvas from which we scrutinize and 'read' the faces of strangers to make evaluations of their state (emotions and intentions) and trait characteristics (e.g. Martelli, Majib, & Pelli, 2005; Porter & ten Brinke, 2009). The development of

rapid assessments of trustworthiness, and their ability to shape later decision-making, likely originated in the distant evolutionary past, allowing our ancestors to assess the best course of action for survival (e.g. Livingstone-Smith, 2004). Confronted by an individual exhibiting an angry facial expression, one might wisely escape a dangerous encounter. In fact, emotional expressions are highly influential in determinations of mock witness testimony credibility, even more influential than the content of the testimony (Kaufmann, Drevland, Wessel, Overskeid, & Magnussen, 2003). However, the evolutionary development of deception and emotional concealment (which we think occurred later than rapid trustworthiness or dangerousness assessments in human evolution) has complicated the assessment of interpersonal trustworthiness (e.g. O'Sullivan, 2003), making assessments of trustworthiness more difficult and prone to error (Ekman, 1992; Leach, Talwar, Lee, Bala, & Lindsay, 2004; Porter & ten Brinke, 2008).

Despite people's well-adapted ability to control certain aspects of their facial presentation, the analysis of emotional facial expressions appears to play a critical role in credibility assessment in the courtroom (*R. v. B. (KG)*, 1993). Thus, judges and jurors view the face of a witness as a rich source of presumably relevant information. Information about an individual's character gleaned from the face seems to be gathered extremely quickly upon seeing an individual for the first time. Willis and Todorov (2006) examined the effect of exposure time to the face on trait inferences, including evaluations of trustworthiness. Results suggested that individuals infer the trustworthiness of others almost instantaneously upon seeing the face, after only 100 ms of exposure. Further, longer exposure time increased levels of confidence without changing judgements. Thus, their findings indicate that assessments of trustworthiness occur instantaneously and are enduring in nature.

Despite highlighting the speed of this process, the paradigm designed by Willis and Todorov (2006) did not offer insights into the accuracy of such assessments, as the actual trustworthiness of target faces was unknown. Although some research suggests that particular attributes can be correctly inferred from physical appearance (e.g. Roney, Hanson, Durante, & Maestriperi, 2006; Shevlin, Walker, Davies, Banyard, & Lewis, 2003), the assessment of trustworthiness does not appear to be among them. Porter, England, Juodis and ten Brinke (2008) presented participants with a group of faces consisting of America's Most Wanted criminals and humanitarian award (e.g. Nobel Peace Prize) winners (simply asking them to evaluate 'strangers' faces). Although from common lore one might assume that the most and least trustworthy members of society could be discriminated by looking at their faces, accuracy levels were only slightly above chance, with errors occurring in over 40% of assessments. Thus, assessments of trustworthiness based on the face alone, while instantaneous, are highly fallible. It is a similar assessment that may strike judges and jurors as valid intuition upon viewing a defendant, complainant, or witness for the first time (Porter & ten Brinke, 2009).

Other facial features correlated with trustworthiness and perceived honesty include 'babyfacedness', symmetry, and attractiveness (Bull, 2006; Bull & Vine, 2003; Lundqvist, Flykt, & Ohman, 1998; Todorov, 2008; Zebrowitz, Voinescu, & Collins, 1996). For example, baby-faced individuals receive more lenient judicial outcomes than mature-faced individuals. Reflecting the impact of attractiveness on impression formation and the 'what is beautiful is good' phenomenon (Dion, Berscheid, &

Walster, 1972), attractive defendants are more likely to be found not guilty, dealt shorter sentences, and considered less dangerous than their unattractive counterparts (e.g. Bull & Rumsey, 1988; Downs & Lyons, 1991; Esses & Webster, 1988). Similarly, particular faces are viewed as being congruent with certain criminal offences. Thus, there are some faces that people agree 'look like' that of a rapist, armed robber, or murderer (Bull & McAlpine, 1998; Dumas & Testé, 2006). Further, research suggests that if a defendant has a face that is considered to be consistent with the charged offence, they are more likely to be convicted than an individual with an 'incongruent' face (Macrae & Shepherd, 1989; Shoemaker, South, & Lowe, 1973). Recent work indicates that this effect occurs regardless of the strength of the prosecution's evidence such that defendants whose faces are congruent with the offence are more likely to be found guilty even when evidence against them is weak (Dumas & Testé, 2006). Although the impact of these facial attributions (most notably attractiveness) in courtroom decisions has been well established, we believe that evaluations of trustworthiness may play a central role in biased legal decision-making, given the evolutionary roots of friend or foe judgements and the appropriate (but not necessarily accurate) application of such evaluations to potentially dangerous criminals.

The initial intuitive evaluation of trustworthiness proposed in DDT is thought to influence subsequent inferences concerning the defendant (or other witness) by making decision-making about him/her increasingly irrational, leading to such biased decisions (Kahneman & Tversky, 1982). Although judges and jury members no doubt strive to maintain objectivity, they are not immune from normal human biases and may not be aware of the influences upon their decision-making (Kaufmann et al., 2003). Further, judges are susceptible to critical thinking errors and a reliance on false stereotypes that may be cited to justify their initial, unreliable assessment (e.g. Granhag & Strömwall, 2004; Vrij & Mann, 2004; Vrij, 2004, 2008). For example, while judges and laypersons alike adhere to the belief that deceptive individuals avert their gaze, empirical research suggests that no such relationship exists (e.g. DePaulo et al., 2003; Porter & ten Brinke, in press; Vrij, 2008). DDT suggests that the adherence to initial impressions and subsequent justification with erroneous beliefs will result in a non-critical, 'tunnel vision' assimilation of potentially ambiguous and even contradictory evidence concerning the defendant resembling a confirmation bias described by Carretta and Moreland (1982). Ask and Granhag (2007) found that criminal investigators have a tendency to be more sceptical about evidence that runs counter to their original perception than evidence consistent with that perception. Further, Porter et al. (2007) found that an increased level of motivation exacerbated tunnel vision in a deception detection task, leading to decreased accuracy. Thus, Porter and ten Brinke (2009) proposed that judges and jury members are likely to subconsciously prefer evidence confirming their initial assessment of trustworthiness, slanting ambiguous evidence in favour of this assessment, and creating tunnel vision in their decision-making. Particularly when assessing evidence in the case of a severe crime where the consequences of a faulty decision are the greatest, the strong motivation of judges and juries, coupled with the complexity of the task, may lead to exacerbated tunnel vision and over-confidence in the initial assessments.

The current study

The current study was an empirical test of DDT, and investigated how participants assimilate information about an individual's alleged involvement in a crime with their initial impression based on a photograph of the accused person's face. Participants were presented with two vignettes describing criminal acts of varying severity (severe or petty) each paired with a photo of the defendant previously rated as highly trustworthy or untrustworthy. Assessments of guilt/innocence and ratings of confidence were reported after the presentation of each piece of increasingly incriminating and exonerating evidence. It was predicted that participants would interpret ambiguous information to further support and rationalize their initial assessment of the individual's trustworthiness. Further, it was predicted that this effect would be exacerbated for severe crimes. The increased motivation felt by the judge or jurors during a murder trial, given the weighty consequences of a faulty decision (i.e. incarcerating an innocent individual or allowing a murderer to escape justice) was expected to exacerbate tunnel vision, confirming initial impressions.

Specifically, if a participant initially concluded (instantaneously or intuitively) that the defendant was untrustworthy based on his physical appearance, it was expected that ambiguous information would be interpreted as indicative of guilt. As such, fewer pieces of ambiguous evidence would need to be considered prior to reaching a decision of guilt. Additionally, it was hypothesized that participants would be highly confident in their guilty verdict for untrustworthy defendants and would be less likely to overturn their verdict in the face of powerful exonerating evidence, particularly for severe crimes where motivation and resulting tunnel vision was expected to be exacerbated.

Pilot study

A pilot study was conducted in order to rate the photographs and created evidence, allowing for selection of trustworthy and untrustworthy defendant faces as well as increasingly incriminating evidence sets for each of four (two petty, two severe) crime vignettes.

Method

Participants

Participants in the pilot study were 27 individuals (15 females, 12 males) recruited from the student population of Dalhousie University.

Apparatus

Twenty photographs of Caucasian males were chosen from the Karolinska Set of Images (Lundqvist et al., 1998) and were rated on various attributes by participants in the pilot study. The photographed individuals were actors between 20 and 30 years old, exhibiting neutral expressions. Pilot participants rated attractiveness, babyfacedness, symmetry, kindness, and trustworthiness on a 1 (not at all) to 7 (highly) scale for each face. Pilot participants also rated the severity of four crime vignettes describing a robbery resulting in murder, a sexually motivated murder, a car theft and a fraud, as

well as the strength of evidence pieces (e.g. eyewitness, alibi, confession, DNA, etc.) associated with each scenario on seven-point scales. The determination of severity was based on whether the crimes were summary (punishable by a sentence of 2 years less a day) or indictable (felony in the US) offences, and was confirmed with pilot study ratings. For crime severity, a rating of one denoted a very petty crime while seven indicated an extremely serious, major crime. Similarly, when rating evidence pieces, a one indicated a highly ambiguous piece of evidence not indicative of guilt, whereas a seven indicated that the evidence was highly incriminating and suggestive of guilt. Last, a subset of exonerating evidence associated with each vignette was also scored on a 1 (not at all exonerating) to 7 (highly exonerating) scale.

Procedure

Participants were presented with four hypothetical criminal cases (suggested to be actual cases), and a set of at least 24 pieces of evidence corresponding to each crime. They were asked to read each case, and then on seven-point scales (i) rate the severity of each crime and (ii) evaluate each piece of corresponding evidence. After all evidence was rated, participants were presented with a series of 20 photographs of male actors and were asked to rate each photograph (as described above) and their confidence in each such rating (also on seven-point scales).

Results

Pilot study ratings

Based on average trait inferences made by pilot study participants, the most ($M = 4.70$, $SD = 0.99$) and least trustworthy ($M = 2.63$, $SD = 1.45$) individuals were chosen from the rated images. A paired samples t -test indicated that the photographs were rated as significantly different from one another on trustworthiness, $t(26) = 6.99$, $p < 0.01$. Other attributions based on the face, including attractiveness, babyfacedness, symmetry and kindness also were rated as significantly different ($p < 0.05$) and, as such, were examined for use as covariates in the primary study analyses. Crime vignette severity also was examined; a repeated measures ANOVA followed by Bonferroni corrected *post hoc* tests confirmed that the severe crimes ($M = 6.60$, $SD = 0.63$) were indeed rated as significantly more severe than the petty crimes ($M = 4.21$, $SD = 0.92$), $F(1,25) = 183.4$, $p < 0.01$. Further, the two vignettes within each severity condition were equivalent in terms of severity ($ps > 0.05$).

In order to ensure that evidence sets accompanying each crime vignette were similar, mean scores of the level of incrimination for each piece of evidence were calculated and evidence was rank ordered. The five most ambiguous pieces of evidence (defined as pieces with an average rating below four) corresponding to each vignette were chosen for use in the main study. Next, five progressively incriminating pieces were chosen with mean ratings between four and seven for each vignette. Paired samples t -tests confirmed that the increasingly incriminating evidence was rated as significantly more incriminating than ambiguous evidence for each vignette ($ps < 0.01$). Further, ratings of ambiguous and incriminating evidence were similar across petty and severe crime vignettes ($ps > 0.05$). The final, exonerating piece of evidence for each vignette was chosen on the basis of being rated as the strongest

piece of exonerating information. Ratings were similar for exonerating evidence across vignettes, $F(1,26) = 1.28, p > 0.05$.

Main study

Method

Participants

Eighty participants from the student population of Dalhousie University were recruited for the main study with an offer of class credit points in return for participation. Participants were predominantly female (12 male, 68 female) and had a mean age of 20.05 years ($SD = 3.36$).

Apparatus

The ratings obtained from the pilot study led to the selection of photographs and evidence for the primary study. The most trustworthy and untrustworthy faces (highest and lowest mean ratings) were selected, and a smaller evidence set was established for each crime. Final evidence sets consisted of five pieces of ambiguous, followed by five pieces of progressively incriminating and one piece of exonerating evidence. The two severe crime vignettes described a robbery resulting in murder and a sexually motivated murder, while the petty crime vignettes described a car theft and a fraud. Although the crime description and corresponding evidence was different across cases within each severity condition, both were equivalent on pilot ratings of ambiguity/incrimination. See the Appendix for crime vignettes and associated evidence of one petty (car theft) and one severe (robbery resulting in murder) case.

In addition to providing verdicts and measures of confidence, main study participants completed a series of personality questionnaires, which served as distracter tasks during the experiment.

Procedure

The severity of crime vignettes (severe or petty) served as a between-subjects variable such that participants were either presented with two severe or two petty crime descriptions. With facial trustworthiness as a within-subjects variable, each crime vignette was accompanied by an image of the trustworthy or untrustworthy looking defendant. As such, this study utilized a 2 (severity) \times 2 (trustworthiness) mixed design.

Half of the participants ($n = 40$) were randomly assigned to complete seven-point ratings of perceived attractiveness, babyfacedness, symmetry, and kindness on the two selected faces prior to vignette presentation. This served to confirm the pilot study finding that these two individuals differed in perceived trustworthiness, allowed an examination of the impact of making explicit ratings on subsequent decision-making, and allowed for an examination of these facial traits as potential covariates in subsequent analyses. After these photographs were rated, participants completed a distracter personality questionnaire. Subsequently, the participant was presented with the first of two counterbalanced crime vignettes. Participants assigned

to the severe crime condition were presented with vignettes describing (i) a robbery resulting in murder and (ii) a sexually motivated murder. Petty crime condition participants were presented with (i) a fraud and (ii) a car theft. The defendants' photographs were counterbalanced to the case file such that trustworthy and untrustworthy images were associated with each crime vignette equally. After reading each vignette, participants were presented with evidence associated with the crime including five ambiguous, five increasingly incriminating, and one exonerating piece of evidence – in that order. After each piece of evidence was presented, the participants, based on what they knew up to that point, were instructed to indicate their verdict (guilty or not guilty beyond a reasonable doubt) and rate their confidence in that decision from one to seven. After participants provided their final verdict, they were administered another distracter personality questionnaire. Next, the remaining photograph and corresponding crime scenario were presented, and the participant repeated the above procedure. At completion of the study, all participants completed a short demographic questionnaire, and were fully debriefed.

Results

To ensure that the selected photographs were rated similarly on trustworthiness in the main and pilot studies, a paired samples *t*-test was conducted. Assessments of trustworthiness by those in the rating condition replicated the results of the pilot, and the two photographed individuals again were shown to be significantly different from one another, $t(39) = 2.82, p < 0.01$.

Latency prior to guilty verdict

A series of bivariate correlations was conducted to determine whether perceived attractiveness, babyfacedness, symmetry, and/or kindness were related to the latency to reach a decision of guilt. All correlations were nonsignificant ($ps > 0.05$). Subsequently, none of the above variables were treated as covariates in the following analyses.

A 2 (crime severity) \times 2 (trustworthiness) mixed ANOVA was conducted to examine the number of evidentiary items required before participants declared a guilty verdict.¹ There was a main effect of severity, $F(1,78) = 25.43, p < 0.001$, such that participants required more evidence in order to render a guilty verdict for a severe crime ($M = 5.05, SD = 2.85$) compared to a petty crime ($M = 2.43, SD = 1.65$). The main effect of trustworthiness was not significant, $F(1,78) = 2.81, p > 0.05$, however, a severity \times trustworthiness interaction emerged, $F(1,78) = 5.17, p < 0.05$. As illustrated in Figure 1, untrustworthy faces accused of a severe crime required less evidence ($M = 4.23, SD = 3.34$) for a guilty verdict than trustworthy faces accused of the same crime ($M = 5.88, SD = 3.19$), $t(39) = 3.26, p < 0.01$. However, there was no difference in latency to decisions of guilt for trustworthy ($M = 2.30, SD = 2.54$) and untrustworthy ($M = 2.55, SD = 2.80$) defendants accused of petty crimes, $p > 0.05$. Follow-up *t*-tests further revealed that the effect of severity was consistent over both levels of defendant trustworthiness, wherein additional evidence was always necessary to convict an individual of a severe crime, $ps < 0.05$.

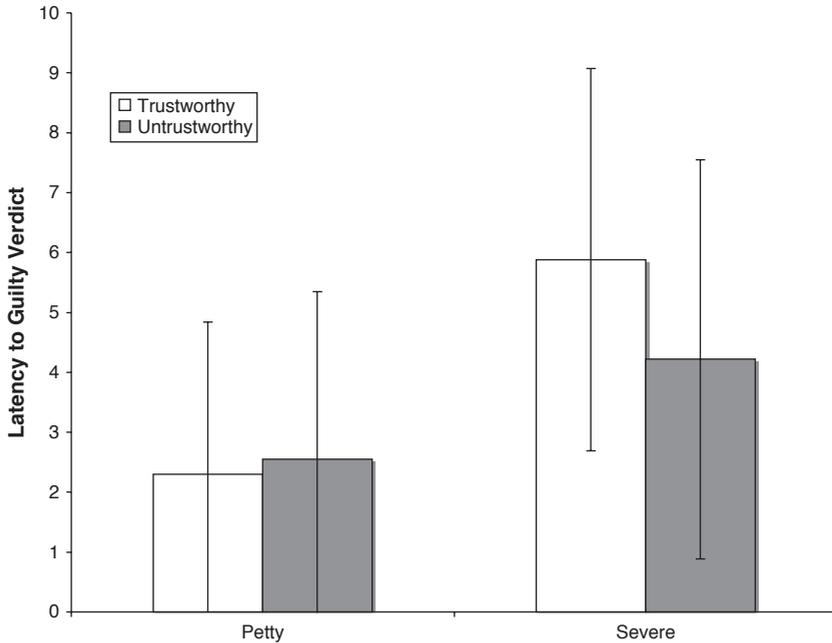


Figure 1. Latency, reflected by number of pieces of evidence prior to a guilty verdict, as a function of crime severity and defendant trustworthiness. Error bars represent standard deviation.

Confidence in verdicts

A 2 (crime severity) \times 2 (trustworthiness) mixed ANOVA² was conducted to examine participants' confidence in their (guilty) verdicts following the final piece of incriminating evidence. As such, only those participants who returned a guilty verdict after evidence piece 10 were included in this analysis (severe: $n=30$; petty: $n=29$). The analysis revealed a significant interaction, $F(1,57)=9.44$, $p<0.01$ (see Figure 2), but no significant main effects, $ps>0.05$. Follow-up t -tests revealed that participants were similarly convinced in the guilt of trustworthy individuals in petty ($M=5.90$, $SD=1.59$) and severe crimes ($M=5.47$, $SD=1.36$), $p>0.05$. However, participants were significantly more confident that untrustworthy defendants had committed a severe crime ($M=6.13$, $SD=0.86$) compared to a petty crime ($M=5.30$, $SD=1.75$), $t(58)=2.35$, $p<0.05$. Further, participants were more confident that untrustworthy defendants ($M=6.13$, $SD=0.86$) were guilty of a severe crime than trustworthy defendants ($M=5.47$, $SD=1.36$), $t(29)=2.61$, $p<0.05$, while no such difference emerged for confidence associated with petty crime verdicts, $p>0.05$.

Interpretation of exonerating evidence

The interpretation of exonerating evidence was examined after excluding participants who returned a not guilty verdict on evidence piece 10 for either vignette they examined, because they were already convinced of the target's innocence. This resulted in $n=13$ eligible participants in the severe and $n=17$ in the petty crime

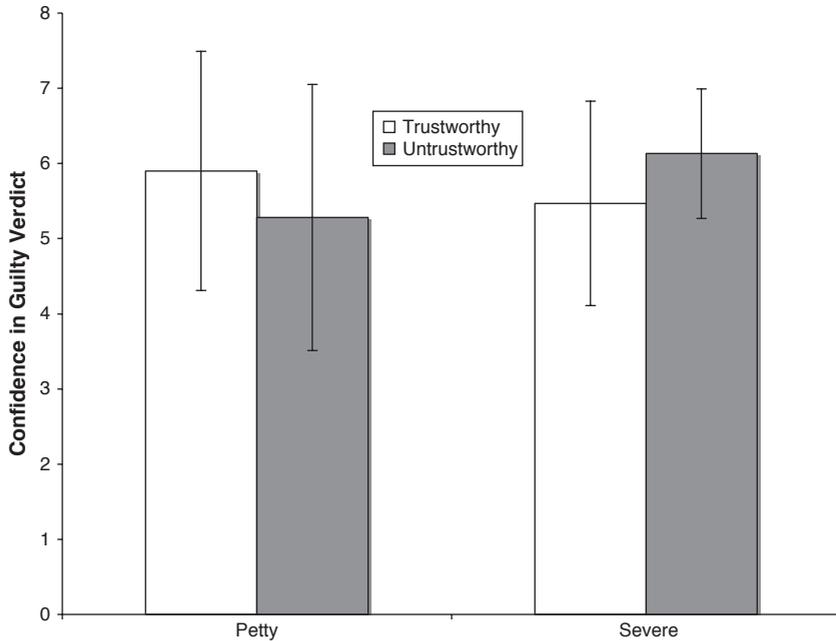


Figure 2. Confidence in guilty verdict, rated on a 1–7 scale, as a function of crime severity and defendant trustworthiness. Error bars represent standard deviation.

conditions. A series of bivariate correlations revealed only a significant relationship between the interpretation of exonerating evidence and ratings of attractiveness for the trustworthy-looking defendant, $r = -0.40$, $p < 0.05$. To statistically control for the effect of attractiveness across levels of trustworthiness, separate regression analyses were conducted for the interpretation of exonerating evidence for trustworthy and untrustworthy defendants wherein respective attractiveness ratings served as predictor variables. Residuals for each regression analysis were saved, and served as measures of the dependent variables in subsequent analyses. A 2 (severity) \times 2 (trustworthiness) ANOVA, using the aforementioned residuals as dependent variables, did not reveal significant main effects or interaction, $ps > 0.05$.

Discussion

The predictions based on DDT were supported in this study, suggesting that this model may elucidate the process of biased legal decision-making that can lead to wrongful convictions. As predicted by DDT, less evidence was necessary to convict an untrustworthy looking defendant of (the same) crime compared to a perceived trustworthy person, particularly in cases of murder (severe crimes). Further supporting DDT, untrustworthy defendants were considered guilty based on an average of fewer than five pieces of ‘ambiguous’ evidence, while more incriminating evidence was necessary to convict the trustworthy defendant of the same (severe) crime. On the other hand, there was no such difference in evidence necessary to convict trustworthy and untrustworthy individuals of minor crimes. Thus, the suggestion that increased crime severity might increase motivation and exacerbate

the effect of initial impressions appears to be supported and, at least in a mock jury study context, the presentation of severe crimes is necessary to reveal dangerous decision-making based on initial impressions.

DDT also suggested that the effect of trustworthiness impressions will be reflected in confidence for verdicts. In accordance with this tenet, confidence levels were highest when participants were delivering a guilty verdict to an untrustworthy looking individual accused of murder. Thus, it appears that perceived trustworthiness not only affects verdicts in general but also jurors' confidence in the types of crimes that an individual could have committed. This finding is in line with legal decision-making research examining face/offence congruency, wherein participants were more likely to find a photographed individual guilty if they 'looked like' the type of individual who would engage in this criminal activity (Dumas & Testé, 2006).

The effect of exonerating evidence on verdict decision-making also was examined. Although the effect of facial trustworthiness was not significant, it should be noted (for future research) that the hypotheses were supported by the pattern of means. Specifically, participants were less likely to accept the innocence of untrustworthy defendants after powerful exonerating evidence – particularly for severe crime vignettes. In general, participants appeared to have been unable to overcome their initial impression of the untrustworthiness of the accused, and interpreted evidence in such a way as to reinforce their initial impression. Future research should further examine the impact of facial trustworthiness on the assimilation of, not only incriminating, but also subsequently presented exonerating evidence.

The current study provides support for DDT and suggests that this model could have implications for understanding wrongful convictions, and perhaps for improving legal decision-making. The impact of trustworthiness, discovered here, was found to have considerable implications for decisions of defendant culpability over and above that associated with the well-documented 'what is beautiful is good' (and honest) phenomenon (Downs & Lyons, 1991). Given that the impact of attractiveness (among other potentially related facial features) was considered as a covariate during statistical analyses, it appears that trustworthiness holds unique biasing effects in the courtroom. Future studies might expand the present design to better understand the effect of trustworthiness assessments on information assimilation, and subsequent decision-making. To fully explicate the impact of trustworthiness biases, this study could be extended by including a third case file in each severity condition with no associated photograph. Further, the current study might be modified to include photographs of truly wrongly convicted individuals to examine if the individual's untrustworthy physical appearance may have been a factor in their conviction. DDT also suggests more complex interactions between intuitive facial assessments and subsequent legal decision-making. Further research may examine the impact of the victim or complainant's perceived trustworthiness on assimilation of evidence against the defendant. Additionally, the trustworthiness of the source of information may also play a role in legal decision-making. For example, information concerning a defendant from a source perceived as trustworthy may be favoured in the mind of a judge or juror over information derived from similarly knowledgeable but untrustworthy-looking witnesses.

The results of this study could have important implications for the justice system. Particularly for severe crimes such as murder – transgressions that also carry the

greatest punishments – defendants who are considered untrustworthy based on physical characteristics may be vulnerable to faulty decision-making and wrongful convictions, even in the face of powerful exonerating evidence. Thus, it is important to acknowledge and consider strategies to combat these biases to increase the fairness of the system. Given that judges consider an unobstructed view of the defendant to be essential to credibility assessment (*R. v. B. (KG), 1993*), perhaps the best approach to combat the misinterpretation of defendant appearance might be judicial education. The efficacy of illustrating potential biases as a tool to overcome their effects should be the focus of future research to determine if more dramatic measures are necessary.

In reaching a final verdict, judges and jurors rely on more than the evidence presented before them. A reliance on gut instinct, careful examinations of an individual's demeanour, particularly the face, are commonly cited components of credibility assessments that are actively encouraged within the court (e.g. see Porter & ten Brinke, 2009). The current study found that assessments of a defendant's trustworthiness, based on facial characteristics, can be so potent as to taint the interpretation of subsequent evidence even when the impact of other known biases (i.e. attractiveness) were considered. Those who have the misfortune of appearing untrustworthy in the eyes of a judge or juror are at a disadvantage in our justice system; jurors may be more likely to determine their guilt before hearing all the evidence, and may be unlikely to alter their perceptions in the face of exonerating evidence. This study calls for further research and a greater appreciation and understanding of the major impact that biases can have on legal decision-making.

Acknowledgements

This project was supported by the Social Sciences and Humanities Research Council of Canada (SSHRC) and the Natural Sciences and Engineering Research Council of Canada (NSERC) through grants awarded to the first author. Graduate scholarships from NSERC and Killam Trusts awarded to the third author also contributed to the successful completion of this project.

Notes

1. Counterbalancing was successful, as the order of faces had no effect on confidence ratings or latency to a guilty verdict, $ps > 0.05$.
2. A series of bivariate correlations did not reveal any significant ($ps > 0.05$) relationships between ratings of perceived facial characteristics and confidence in verdicts; therefore, no covariates were included in subsequent analyses.

References

- Ask, K., & Granhag, P.A. (2007). Motivational bias in criminal investigators' judgments of witness reliability. *Journal of Applied Social Psychology, 37*, 561–591.
- Bond, C.F., & DePaulo, B.M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review, 10*, 214–234.
- Bull, R. (2006). Detecting lies and deceit: The psychology of lying and the implications for professional practice. *Journal of Community & Applied Social Psychology, 16*, 166–167.
- Bull, R., & McAlpine, S. (1998). Facial appearance and criminality. In A. Memon, A. Vrij, & R. Bull (Eds.), *Psychology and law. Truthfulness, accuracy and credibility* (pp. 59–76). London: McGraw-Hill.

- Bull, R., & Rumsey, N. (1988). *The social psychology of facial appearance*. Berlin: Springer.
- Bull, R., & Vine, M. (2003). *Attractive people tell the truth: Can you believe it?* Poster presented at the Annual Conference of the European Association of Psychology and Law, Edinburgh, UK.
- Carretta, T.R., & Moreland, R.L. (1982). Nixon and Watergate. *Personality and Social Psychology Bulletin*, 8, 446–453.
- DePaulo, B.M., Lindsay, J.J., Malone, B.E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin*, 129, 74–118.
- Dion, K.K., Berscheid, E., & Walster, E. (1972). What is beautiful is good. *Journal of Personality and Social Psychology*, 24, 285–290.
- Downs, A.C., & Lyons, P.M. (1991). Natural observations of the links between attractiveness and initial legal judgments. *Personality and Social Psychology Bulletin*, 17, 541–547.
- Dumas, R., & Testé, B. (2006). The influence of criminal facial stereotypes on juridic judgments. *Swiss Journal of Psychology*, 65, 237–244.
- Ekman, P. (1992). *Telling lies: Clues to deceit in the marketplace, politics, and marriage*. New York: W.W. Norton.
- Ekman, P., & O'Sullivan, M. (1991). Who can catch a liar? *American Psychologist*, 46, 913–920.
- Esses, V.M., & Webster, C.D. (1988). Physical attractiveness, dangerousness, and the Canadian Criminal Code. *Journal of Applied Social Psychology*, 18, 1017–1031.
- Granhag, P.A., & Strömwall, L.A. (Eds.). (2004). *The detection of deception in forensic contexts*. Cambridge: Cambridge University Press.
- Kahneman, D., & Tversky, A. (1982). The psychology of preferences. *Scientific American*, 246, 160–173.
- Kaufmann, G., Drevland, G.C.B., Wessel, E., Overskeid, G., & Magnussen, S. (2003). The importance of being earnest: Displayed emotions and witness credibility. *Applied Cognitive Psychology*, 17, 21–34.
- Leach, A.-M., Talwar, V., Lee, K., Bala, N., & Lindsay, R.C.L. (2004). 'Intuitive' lie detection and children's deception by law enforcement officials and university students. *Law and Human Behavior*, 28, 661–685.
- Livingstone-Smith, D. (2004). *Why we lie: The evolutionary roots of deception and the unconscious mind*. St Martin's Griffin: New York.
- Lundqvist, D., Flykt, A., & Ohman, A. (1998). *The Karolinska Directed Emotional Faces*. Stockholm: Karolinska Institute, Psychology Section, Department of Clinical Neuroscience.
- Macrae, C.N., & Shepherd, J.W. (1989). Do criminal stereotypes mediate juridic judgments? *British Journal of Social Psychology*, 28, 189–191.
- Martelli, M., Majib, J.M., & Pelli, D.G. (2005). Are faces processed like words? A diagnostic test for recognition by parts. *Journal of Vision*, 5, 58–70.
- Meissner, C.A., & Kassin, S.M. (2002). 'He's guilty!': Investigator bias in judgments of truth and deception. *Law and Human Behavior*, 26, 469–480.
- O'Sullivan, M. (2003). The fundamental attribution error in detecting deception: The boy-who-cried-wolf effect. *Personality and Social Psychology Bulletin*, 29, 1316–1327.
- Porter, S., & ten Brinke, L. (2008). Reading between the lies: Identifying concealed and falsified emotions in universal facial expressions. *Psychological Science*, 19, 508–514.
- Porter, S., & ten Brinke, L. (in press). The truth about lies: What works in detecting high-stakes deception? Special Issue of *Legal and Criminological Psychology*.
- Porter, S., & ten Brinke, L. (2009). Dangerous decisions: A theoretical framework for understanding how judges assess credibility in the courtroom. *Legal and Criminological Psychology*, 14, 119–134.
- Porter, S., Woodworth, M., & Birt, A.R. (2000). Truth, lies, and videotape: An investigation of the ability of federal parole officers to detect deception. *Law and Human Behavior*, 24, 643–658.
- Porter, S., McCabe, S., Woodworth, M., & Peace, K.A. (2007). Genius is 1% inspiration and 99% perspiration? ... or is it? An investigation of the effects of motivation and feedback on deception detection. *Legal and Criminological Psychology*, 12, 297–309.

- Porter, S., England, L., Juodis, M., & ten Brinke, L. (2008). Is the face a window to the soul? Investigation of the validity of intuitive judgments of the trustworthiness of human faces. *Canadian Journal of Behavioural Science*, *40*, 171–177.
- R. v. B. (K.G.), [1993] 1 S.C.R. 740.
- R. v. Francois [1994] 2 S.C.R. 827.
- R. v. Malik & Bagri, 2005 BCSC 350.
- R. v. Marquard [1993] 4 S.C.R. 223.
- Roney, J.R., Hanson, K.N., Durante, K.M., & Maestripieri, D. (2006). Reading men's faces: women's mate attractiveness judgments track men's testosterone and interest in infants. *Proceedings of the Royal Society of London B*, *273*, 2169–2175.
- Shevlin, M., Walker, S., Davies, M., Banyard, P., & Lewis, C.A. (2003). Can you judge a book by its cover? Evidence of self-stranger agreement on personality at zero acquaintance. *Personality and Individual Differences*, *35*, 1373–1383.
- Shoemaker, D.J., South, D.R., & Lowe, J. (1973). Facial stereotypes of deviants and judgments of guilt or innocence. *Social Forces*, *51*, 427–433.
- Todorov, A. (2008). Evaluating faces on trustworthiness: An extension of systems for recognition of emotions signaling approach/avoidance behaviors. *Annals of the New York Academy of Sciences*, *1124*, 208–224.
- Vrij, A. (2004). Why professionals fail to catch liars and how they can improve. *Legal and Criminological Psychology*, *9*, 159–181.
- Vrij, A. (2008). *Detecting lies and deceit: Pitfalls and opportunities*. Chichester, UK: Wiley.
- Vrij, A., & Mann, S. (2004). Detecting deception: The benefit of looking at a combination of behavioral, auditory and speech content related cues in a systematic manner. *Group Decision and Negotiation*, *13*, 61–79.
- Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after a 100-ms exposure to a face. *Psychological Science*, *17*, 592–598.
- Wilson, P.J. (2003). *Wrongful convictions: Lessons learned from the Sophonow Public Inquiry*. Canadian Police College.
- Zebrowitz, L.A., Voinescu, L., & Collins, M.A. (1996). 'Wide-eyed' and 'crooked-faced': Determinants of perceived and real honesty across the life span. *Personality and Social Psychology Bulletin*, *22*, 1258–1269.

Appendix

Petty crime vignette: car theft

In October 2005 a 46-year-old male reported that his car had been stolen. The man stated that he had stayed overnight at the home of a friend after a dinner party. He reported that his car had been parked at the curb next to the driveway, and that he was sure he had locked his doors. He stated that at approximately 3:30 am he was awoken by a noise outside, and when he got up to look out the window became aware that his car was missing. He woke up his friends, as many people had spent the night, and phoned the police. Upon interviewing the other people who had spent the night, two of the partygoers reported noticing a man outside the house the night before. They indicated that he had walked by the house a couple of times, but they assumed he was just out for a walk, as it was a nice night. They described him as white, average weight, and average height. When police looked into the matter further, they discovered that within a five block radius there were approximately seven reports of stolen cars over the past 11 months. This struck the investigators as odd, since the particular area was a suburban Toronto area known to have a relatively low crime rate. All the reports indicated that the victims had been awoken between 2:00 am and 3:30 am to find their cars stolen. In five of the reports there was mention of an unknown man being seen outside the houses the same evenings that the cars were stolen. Within 2 months of the cars being stolen, they were discovered in a downtown Toronto alley, stripped of all valuable parts. The alleys the cars were discovered in were all within approximately three blocks of a particular garage. The accused was arrested and charged with car theft. The accused was arrested and charged with Taking a Motor Vehicle without Consent [CC: 335 (1)].

Evidence:

- (1) The accused is relatively new to the neighbourhood, having moved three times in the past year.
- (2) The accused matches the physical description (white, average height and weight).
- (3) The accused is very interested in cars.
- (4) The accused was very short on cash at the time of the crime.
- (5) The accused says on the night of the theft he was home alone.
- (6) The accused works in the garage near the alley where the cars were found.
- (7) The accused does not own a car. A friend testified that once when he had locked his keys in his car, the accused was able to pick his lock in under a minute.
- (8) The accused tried to run from the police when they attempted to arrest him.
- (9) The garage where the accused works has been investigating previously for being a suspected 'chop shop'.
- (10) Parts from the previously stolen cars were found in the garage where the accused works.
- (11) A teenager arrested in Hamilton admitted to having stolen the car.

Severe crime vignette: robbery resulting in murder

In August 1998, a convenience store was robbed in a Toronto suburb. At the time the only people in the store were the clerk behind the counter, and one customer who was shopping in the back of the store, who hid when the two suspects entered the store. The store had video surveillance, but the footage was extremely grainy, and it was unable to provide investigators with any identifying information. One of the suspects remained on watch at the door of the store, while the other threatened the store clerk with what appeared to be a large kitchen knife. After the store clerk handed over \$140 from the cash register along with several cartons of cigarettes, the suspect demanded more money. When the clerk failed to provide the additional money, the suspect stabbed the clerk once in the chest, and then ran out with his accomplice. The other witness called 911, but was only able to describe the suspects as white, young, and average height and weight. Unfortunately, emergency medical teams were unable to save the store clerk. The police later recovered the murder weapon from a nearby dumpster. The accused was arrested that same night, and was charged with robbery (CC: 343), and murder in commission of an offence (CC: 230).

Evidence:

- (1) The accused admits to using marijuana.
- (2) The accused is from a Toronto suburb.
- (3) The accused smokes cigarettes.
- (4) The accused has older friends who are unemployed.
- (5) The accused is unemployed.
- (6) The footprint of the accused was found in the convenience store.
- (7) Upon apprehension, a stocking mask was found in the pocket of the accused.
- (8) Acquaintances of the accused testified that he had previously mentioned robbing a store as a source of revenue.
- (9) Another suspect in the robbery has identified the accused as his partner.
- (10) When police showed the knife to the mother of the accused, she identified it as one that she owned.
- (11) A young man came forward and confessed to the robbery. The other man he identified as his partner subsequently confessed.