Trust and distrust: the perception of trustworthiness of faces in psychopathic and non-psychopathic offenders

R.A. Richell \textsuperscript{a,1}, D.G.V. Mitchell \textsuperscript{a,b,2}, K.S. Peschardt \textsuperscript{a,b,2}, J.S. Winston \textsuperscript{c}, A. Leonard \textsuperscript{d}, R.J. Dolan \textsuperscript{c}, R.J.R. Blair \textsuperscript{a,*}

\textsuperscript{a} Affective Cognitive Neuroscience, Mood and Anxiety Disorders Program, National Institute of Mental Health, National Institutes of Health, Department of Health and Human Services, 15K North Drive, Room 206, MSC 2670, Bethesda, MD 20892-2670, USA

\textsuperscript{b} Institute of Cognitive Neuroscience and Department of Psychology, University College, London, UK

\textsuperscript{c} Wellcome Department of Imaging Neuroscience, 12 Queen Square, London, WC1N 3BG, UK

\textsuperscript{d} HMP Wormwood Scrubs, P.O. Box 757, Du Cane Road, London, W12 0AE, UK

Received 10 October 2003; received in revised form 28 August 2004; accepted 15 November 2004
Available online 26 January 2005

Abstract

Psychopathy is a disorder linked to impairments in social cognition. Individuals with psychopathy and comparison individuals, as defined by the Hare Psychopathy Checklist Revised (Hare, 1991), completed a task requiring complex social judgments. Participants viewed a series of male faces, and made judgments concerning how trustworthy they considered the person pictured to be and the degree to which each face was displaying specific emotions. Judgments of trustworthiness did not differ between the individuals with psychopathy and the comparison individuals. Trustworthiness judgments of the faces negatively correlated with, in particular, ratings of anger in the faces for both groups of participants. The data are discussed with...
reference to the neural systems, including the amygdala, considered to be involved in the mediation of task performance and also neuro-cognitive models of psychopathy.

Published by Elsevier Ltd.

Keywords: Social cognition; Trustworthiness; Psychopathy; Amygdala

1. Introduction

There are two disorders clearly linked to impairments in social cognition: psychopathy and autism. Psychopathy is a developmental disorder, characterized by a core impairment in social behavior, that is identified using the Hare Psychopathy Checklist-Revised [PCL-R] (Hare, 1991). The PCL-R is a clinically based rating scale, and has been shown to be a reliable method of measuring psychopathy in forensic settings. Factor analyses of behaviors rated on the PCL-R reveal a two-factor structure (Hare, 1991) relating to an emotion dysfunction component and an antisocial component. Autism is a severe developmental disorder where there is marked neuro-cognitive impairment. It is described by the American Psychiatric Association’s Diagnostic and Statistical Manual (DSM-IV, American Psychiatric Association, 1994) as ‘the presence of markedly abnormal or impaired development in social interaction and communication and a markedly restricted repertoire of activities and interests’ (p. 66).

At the neural level both psychopathy and autism have been linked to impairments in amygdala functioning (Baron-Cohen et al., 2000; Blair, 2001; Patrick, 1994). With respect to psychopathy, we have argued that the functional consequence of this amygdala dysfunction in individuals with psychopathy is impairment in the formation of stimulus–reinforcement associations, particularly the formation of stimulus–punishment associations (Blair, 2001, 2004). Certainly, individuals with psychopathy have been found to present with reduced amygdaloid volume (Tiihonen et al., 2000), and reduced amygdala activation during emotional memory (Kiehl et al., 2001) and aversive conditioning (Veit et al., 2002) tasks, relative to comparison individuals. Similarly to amygdala-lesioned patients, individuals with psychopathy have also shown impaired aversive conditioning (Lykken, 1957), reduced augmentation of the startle response to threat primes (Patrick, Bradley, & Lang, 1993), and impaired recognition of fearful facial and vocal expressions (Blair, Colledge, Murray, & Mitchell, 2001; Blair et al., 2002).

As regards autism, neuro-imaging techniques have revealed an increase in grey matter volume in the amygdala/peri-amygdaloid nucleus in a group of high functioning autistic individuals (Abell et al., 1999). However, functionally, there is no reason to believe that individuals with autism are impaired in the formation of stimulus–reinforcement associations. Indeed, individuals with autism are at increased risk of presenting with anxiety (Gillott, Furniss, & Walter, 2001; Rumsey, Rapoport, & Scerey, 1985), rather than at decreased risk as seen in individuals with psychopathy (Frick, Lilienfeld, Ellis, Loney, & Silverthorn, 1999; Patrick, 1994; Verona, Patrick, & Joiner, 2001).

In contrast, there may be impairment in aspects of social cognition involving the amygdala in individuals with autism for tasks that require social judgments from facial stimuli. Thus, patients with autism are impaired on both the Reading the Mind in the Eyes task and the face trustworthiness judgment tasks (Adolphs, Sears, & Piven, 2001; Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001). In the Reading the Mind in the Eyes task, the participant makes a forced choice
regarding the mental or emotional state of a character on the basis of an image of their eyes alone. Performance on this task has been shown to recruit the amygdala (Baron-Cohen et al., 1999) and is impaired following amygdala lesions (Adolphs, Baron-Cohen, & Tranel, 2002) and in individuals with autism (Baron-Cohen et al., 1999). In the trustworthiness judgment task, the participant judges a person’s trustworthiness on the basis of a picture of their face. Performance on this task has been shown to recruit the amygdala (Winston, Strange, O’Doherty, & Dolan, 2002), is impaired following amygdala lesions (Adolphs, Tranel, & Damasio, 1998) and in individuals with autism (Adolphs et al., 2001).

Despite the extensive evidence for amygdala dysfunction in individuals with psychopathy (Blair, 2004), it is unclear whether individuals with psychopathy are impaired in these specific aspects of social cognition thought to involve the amygdala mentioned above. Individuals with psychopathy are impaired in the recognition of fearful expressions (Blair et al., 2001, 2002). However, we have argued elsewhere that fearful expressions act as unconditioned aversive stimuli (Blair, 2003). Thus, the impaired processing of these expressions by individuals with psychopathy is consistent with their proposed functional impairment related to stimulus–punishment associations. In contrast, ascribing a mental state to an individual on the basis of information from their eyes, as is necessary in the Reading the Eyes task, cannot easily be related to stimulus–reinforcement processing. Interestingly, individuals with psychopathy show no impairment on this task (Richell et al., 2003).

In the trustworthiness task, the participant judges a person’s trustworthiness on the basis of a picture of their face. Presumably, the participant judges on the basis of negative valence cues generated by specific aspects of the viewed face. Certainly, as noted above and consistent with this position, the role of the amygdala appears to be important in trustworthiness judgments (Adolphs et al., 1998; Winston et al., 2002). This might predict that individuals with psychopathy should be impaired on this task. Alternatively, the task may be highly dependent on face processing skills. Individuals with autism may fail this task due to their known dysfunction in crucial face processing regions such as fusiform gyrus and superior temporal sulcus (Schultz et al., 2003). There are no reasons to believe that individuals with psychopathy are impaired in more general face processing capacities. This would thus predict that individuals with psychopathy would not show impairment on this task. The current study tests these contrasting predictions.

2. Methods

2.1. Participants

From a pool of 200 men incarcerated in a forensic institution in the London area, two groups of male participants were identified who satisfied the criteria for the individuals with psychopathy ($n = 19$) and comparison ($n = 19$) groups. The sample was made up of 29 Caucasian, 1 Asian, and 8 Afro-Caribbean participants; 1 Asian and 2 Afro-Caribbean participants were in the comparison group, and 6 Afro-Caribbean participants were in the psychopathic group. Files were pre-screened to exclude individuals whose psychiatric reports revealed a diagnosis for psychosis, organic brain damage, or neurological disorder.

In accordance with the literature and the guidelines of the PCL-R (Hare, 1991), individuals with a score of 30 or above on the PCL-R were assigned to the psychopathic group, whilst those with a
score of 20 or less were assigned to the control group. The ages of the participants ranged from 22 to 56 years, with a mean of 34.5 years (SD = 10.1). The mean ages of the psychopathic and comparison groups were 37.2 (SD = 8.73) and 31.8 (SD = 10.9) years, respectively.

The Raven’s Advanced Progressive Matrices, Set I (Raven, 1976) was administered to provide an estimate of intelligence. The Raven’s Scores of the participants ranged from 4 to 12, with a mean of 8.13 (SD = 1.82). The mean Raven’s Scores of the psychopathic and comparison groups were 8.32 (SD = 1.89) and 7.95 (SD = 1.78), respectively. Participant characteristics are given in Table 1. The two groups did not differ significantly in terms of age or Raven’s score. Written consent was obtained from each inmate who participated in the study, and all were informed that they were free to withdraw from the study at any time.

2.2. Psychopathy checklist-revised (PCL-R; Hare, 1991)

The PCL-R consists of 20 behavioral items that are scored on the basis of a file review and, where possible, a semi-structured interview. Each item has a maximum possible score of 2, and the maximum total score is therefore 40. The PCL-R has been shown to be a valid and reliable method for assessing psychopathy (Hare, 1991).

Inmates were scored by two independent raters. Inter-rater reliability for the PCL-R total score, assessed by Spearman Rank correlation, was high ($r_{ranks} = 0.88, p < 0.001$). The agreement between the two raters for diagnostic group (psychopathic versus comparison) was 100%. The mean PCL-R scores of the psychopathic and comparison groups are given in Table 1.

2.3. Trustworthiness task

Grayscale frontal images of 120 Caucasian male faces were used as stimuli. These stimuli were identical to those used by Winston et al. (2002). Gaze direction of all stimuli was directly forward, and no overt emotions were expressed. Participants viewed each stimulus separately, in a random order, and made a judgment as to how trustworthy they considered the person in the photograph to be. Judgments were made on a 7-point rating scale, running from 1 (extremely untrustworthy) to 7 (extremely trustworthy). The task was self-paced.

A subset of the participants ($n = 11$ comparison individuals and $n = 11$ individuals with psychopathy) also rated the degree of emotional expression that they perceived within each face after completing the trustworthiness ratings. These participants rated each face, in turn, on each of the following basic emotions: anger, fear, happy and sad. Faces were presented in a random order.

<table>
<thead>
<tr>
<th>Group</th>
<th>PCL-R</th>
<th>Age</th>
<th>RAVENS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Factor 1</td>
<td>Factor 2</td>
<td>Total</td>
</tr>
<tr>
<td>Comparison ($n = 19$)</td>
<td>2.11 (1.92)</td>
<td>3.72 (2.81)</td>
<td>7.05 (4.51)</td>
</tr>
<tr>
<td>Psychopathic ($n = 19$)</td>
<td>11.2* (1.87)</td>
<td>15.6* (1.17)</td>
<td>32.2* (2.00)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>31.8 (10.9)</td>
</tr>
</tbody>
</table>

Values are mean (SD).
* $p < 0.001$, significantly different from comparison group.
Ratings were made on a 7-point rating scale, running from 1 (no emotion) to 7 (extremely angry/fearful/happy/sad for each respective emotion). Again this part of the task was self-paced.

2.4. Analysis

The 120 stimuli were ranked based on their mean trustworthiness rating from pilot data (Winston et al., 2002). The stimuli were then divided into three categories by rank score, i.e., the 40 most untrustworthy stimuli, the 40 median stimuli, and the 40 most trustworthy stimuli. The mean rating for each stimulus category was computed for all participants. This data was analysed using repeated measures ANOVA, with rank category (most untrustworthy; median; most trustworthy) as a within-subjects factor and group (psychopathic; comparison) as a between-subjects factor. In addition, each participant’s ratings were correlated with the mean ratings from the pilot study (Winston et al., 2002) to produce individual Spearman rho values. The mean rho value for each group was compared using an unpaired t-test.

For the subset of participants who rated the emotional expression of the faces, the mean trustworthiness scores were correlated with the emotion ratings within each subject group. Multiple regression analyses were conducted to determine the ability of the emotional expression ratings to predict trustworthiness score.

2.5. Procedure

All participants were tested individually, in a quiet room attached to the wing. The trustworthiness task was administered as part of a larger neurocognitive test battery. Written instructions were given to each participant prior to commencing the test, and use of the rating scale was verbally clarified.

3. Results

Mean ratings of trustworthiness for the three-rank stimulus categories are presented in Table 2. Repeated measures ANOVA revealed a main effect of stimulus category ($F(2,72) = 83.1$, $p < 0.001$). Ratings for the three-rank categories did not differ significantly between groups (rank category × group interaction, $F(2,72) = 2.16$, $p = 0.12$).

The trustworthiness ratings of the participants were correlated with the ratings obtained from the pilot data. The majority of ratings of both the control and psychopathic groups correlated positively, and significantly, with ratings from the pilot study (89% of control and 74% of

<table>
<thead>
<tr>
<th>Group</th>
<th>Least trustworthy</th>
<th>Median</th>
<th>Most trustworthy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comparison ($n = 19$)</td>
<td>3.20 (0.43)</td>
<td>3.87 (0.46)</td>
<td>4.45 (0.67)</td>
</tr>
<tr>
<td>Psychopathic ($n = 19$)</td>
<td>3.31 (0.97)</td>
<td>3.78 (0.93)</td>
<td>4.21 (0.88)</td>
</tr>
</tbody>
</table>

The least trustworthy, median and most trustworthy categories each contain 40 stimuli. Values are mean (SD).
individuals with psychopathy). The mean correlations (rho values) for the comparison group and the individuals with psychopathy were 0.34 and 0.30, respectively. There was no difference between the mean group correlations ($t = 0.60$, $p = 0.56$).

Although the stimuli did not express overt emotion, in the subset of participants who additionally rated the emotional expression of the faces, mean trustworthiness ratings were significantly negatively correlated with ratings of anger, fear and sad in both the psychopathic and comparison groups ($p < 0.001$ for all; see Table 3). Trustworthiness ratings were significantly positively correlated with ratings of happy in both groups ($p < 0.001$ for both).

As the trustworthiness judgments correlated with the ratings of every emotion, it appeared that we were identifying a relationship between trustworthiness and a more general affective judgment for the expression. For this reason, we entered the emotion expression ratings into a multiple regression analysis, with trustworthiness as the dependent variable. This should identify which expression(s) was/were driving the trustworthiness ratings. The model was significant in both the comparison individuals ($R = 0.61$, $F = 17.4$, $p < 0.001$) and the individuals with psychopathy ($R = 0.64$, $F = 20.2$, $p < 0.001$), and explained approximately 38% and 42% of the variance in trustworthiness scores in these groups, respectively. Examination of the factors in the model revealed that ratings of anger were the only significant correlate of trustworthiness score for both the controls ($t = -6.0$, $p < 0.001$) and the individuals with psychopathy ($t = -5.05$, $p < 0.001$), after the other emotions were partialed out.

<table>
<thead>
<tr>
<th>Group</th>
<th>Mean score</th>
<th>Anger</th>
<th>Fear</th>
<th>Happy</th>
<th>Sad</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comparison</td>
<td></td>
<td>-0.57*</td>
<td>-0.24*</td>
<td>0.43*</td>
<td>-0.42*</td>
</tr>
<tr>
<td>Psychopathic</td>
<td>-0.65*</td>
<td>-0.33*</td>
<td>0.54*</td>
<td>-0.45*</td>
<td></td>
</tr>
</tbody>
</table>

* $p < 0.005$.

4. Discussion

This study investigated the performance of individuals with psychopathy and comparison individuals on judging trustworthiness in faces. The results indicated that the two groups did not differ in their evaluation of trustworthiness. In addition, a multiple regression investigating the relative predictive power of the expression ratings indicated that ratings of trustworthiness were uniquely related to ratings of anger.

The current results interestingly extend previous findings suggesting a double dissociation in the form of the social cognition impairment shown by individuals with psychopathy and individuals with autism. While both disorders have been linked to deficits in empathy, it has been suggested that the empathy deficit in psychopathy relates to impairments in emotional processing, whilst in autism the deficit relates to an impairment in Theory of Mind (Blair, 2002). In line with this, individuals with psychopathy have been found to present with decreased responsiveness to specific emotional expressions but intact Theory of Mind; i.e., the ability to represent the mental states
of others (Blair, 1999; Blair et al., 1996, Blair, Jones, Clark, & Smith, 1997, 2001, 2002; Richell et al., 2003; Widom, 1978). In contrast, individuals with autism present with unimpaired facial affect recognition, if the control group is matched on verbal mental age, but profound Theory of Mind impairment (Adolphs et al., 2001; Baron-Cohen, Leslie, & Frith, 1985, 2001; Baron-Cohen, Wheelwright, & Jolliffe, 1997; Happe, 1994). Previous work has suggested that individuals with autism perform poorly on facial trustworthy tasks similar to the current paradigm (Adolphs et al., 2001). However, in the current study we observed no significant impairment in facial trustworthiness judgments in the individuals with psychopathy.

We do not believe that the trustworthiness task indexes the ability to represent the mental states of others (i.e., Theory of Mind). Theory of Mind refers to the ability to represent the mental states of others; i.e., their thoughts, desires, beliefs, intentions and knowledge (Leslie, 1987; Premack & Woodruff, 1978). It could be argued that trustworthiness is a mental state, but such an argument weakens the tight theoretical nature of the Theory of Mind construct, at least as it is represented by Leslie (e.g., Leslie, 1987). Trustworthiness judgments do not have the prepositional status that Leslie has argued underpin Theory of Mind computations. Moreover, by this relaxed criteria for mental state it becomes confusing why individuals with autism do not present with impairment in facial expression recognition when the control group is matched on verbal mental age (Adolphs et al., 2001; Baron-Cohen et al., 1997); other’s emotional states are at least as much mental states as judgments of trustworthiness.

There does appear to be an emotional component to trustworthiness judgments. This is evidenced by the significant correlations of the participants’ trustworthiness ratings with the participants’ emotion ratings of the same stimuli. Interestingly, the multiple regression analysis indicated that the participants were using their judgment of the anger in the face stimuli presented to them to determine their judgment of that face’s trustworthiness. Of course, none of these faces were displaying angry affect. It is instead likely that the participants were estimating the face’s potential hostility (i.e., its potential predilection for the use of aggression) and using this estimate as a basis for their trustworthiness judgment. Previous work has indicated that individuals with psychopathy are impaired in processing the emotional expressions of others (Blair et al., 1997, 2001, 2004; Kosson, Suchy, Mayer, & Libby, 2002). However, their impairment is for the processing of fearful, sad and possibly disgusted expressions. No study has found any indication that the processing of the anger of others is impaired in individuals with psychopathy. In short, the results of the current study demonstrating intact trustworthiness judgments on the basis of face information in individuals with psychopathy is compatible with previous work reporting intact processing of angry expressions in individuals with psychopathy.

There has recently been a move to suggest an extended system for social cognition within which the amygdala plays a crucial role (Adolphs, 2003; Baron-Cohen et al., 2000; Brothers, 1997). This position has been criticized for being under-specified. In addition, it frequently relies on animal data indicating that amygdala lesions impair social behavior (Kling & Brothers, 1992; Kluver & Bucy, 1997). This is problematic as more recent data has shown that these early results most likely reflected damage to fibers that travel through and around the amygdala from other brain regions rather than damage to the amygdala per se (Amaral et al., 2003). However, there is data, in humans at least, to support the suggestion that the amygdala plays a more general role in social cognition. As stated above, imaging and neuropsychological work has indicated amygdala involvement during complex mental state processing in the context of the Eyes Task (Baron-Cohen...
et al., 1999) and during trustworthiness judgments (Adolphs et al., 1998; Winston et al., 2002). From the current study, and earlier work (Richell et al., 2003), we can see that individuals with psychopathy are not impaired on these tasks. However, on tasks such as aversive conditioning, passive avoidance and augmentation of the startle reflex following visual threat primes, for which there is a considerable animal as well as human literature indicating amygdala involvement (Ambrogi Lorenzini, Baldi, Bucherelli, Sacchetti, & Tassoni, 1999; Davis, 2000; LeDoux, 2000), individuals with psychopathy present with pronounced impairment (Lykken, 1957; Newman & Kosson, 1986; Patrick et al., 1993). These are tasks that have been used as animal models of anxiety. Whilst individuals with psychopathy present with reduced anxiety (Cleckley, 1976; Leventon, Patrick, Bradley, & Lang, 2000; Verona et al., 2001), this does not appear to be the case in individuals with autism. It is possible that in humans the amygdala has acquired a broader role in social cognition. However, this role of the amygdala is not impaired in individuals with psychopathy whereas the more basic, emotional learning functions of the amygdala are affected.

In summary, the results of the present study indicate that individuals with psychopathy do not present with impairment in the evaluation of trustworthiness from faces alone. This is in contrast to individuals with the social disorder autism, who have previously been shown to be impaired at trustworthiness judgments (Adolphs et al., 2001). Despite psychopathy and autism being linked to impairments in social cognition, this study provides further evidence of dissociable cognitive deficits between the two disorders.

Acknowledgments

This work was supported by the Medical Research Council [ref. G9716841]. The authors would like to thank the Psychology Department and staff at HMP Wormwood Scrubs.

References


