Abstract

Typical individuals make rapid and reliable evaluations of trustworthiness from facial appearances, which can powerfully influence behaviour. However, the same may not be true for children with autism spectrum disorder (ASD). Using an economic trust game, the current study revealed that like typical children, children with ASD rationally modulate their trust behaviour based on non-face cues to partner trustworthiness (e.g., reputation information). Critically however, they are no more likely to place their trust in partners with faces that look trustworthy to them, than those that look untrustworthy. These results cannot be accounted for by any group differences in children's conceptualization of trustworthiness, ability to read trustworthiness from faces, or understanding of the experimental paradigm. Instead, they seem to suggest that there may be a selective failure to spontaneously use facial cues to trustworthiness to guide behaviour in an ecologically valid context.
Though most would agree that one shouldn’t judge a book by its cover, typical individuals are powerfully drawn to read trait information from faces (Hassin & Trope, 2000). For example, strong and reliable attributions of trustworthiness are made after very brief exposure to faces (Willis & Todorov, 2006) and these impressions influence behaviour. During economic interactions (e.g., trust games, Berg, Dickhaut & McCabe, 1995) participants trust their partners with significantly larger investments if they are perceived to have trustworthy-looking faces, rather than untrustworthy-looking faces (Chang, Doll, van’t Wout, Frank & Sanfey, 2010; Rezlescu, Duchaine, Olivola & Chater, 2012; van’t Wout & Sanfey, 2008).

There is a growing interest in social cognition and trait inferences in individuals with autism spectrum disorder (ASD), who often experience difficulties reading social information from faces (Webb, Faja & Dawson, 2011). Their perceptions of trustworthiness have received particular attention, because these judgments have been linked with the functioning of the amygdala (Adolphs, Sears & Piven, 2001; Winston, Strange, O’Doherty & Dolan, 2004), which may be atypical in ASD (Ashwin, Baron-Cohen, Wheelright, O’Riordan & Bullmore, 2007; Critchley et al., 2000; Kliemann, Dziobek, Hatri, Baudewig & Heekeren, 2012).

Perhaps surprisingly, studies have often failed to find evidence of atypical facial trustworthiness judgments in individuals (both children and adults) with ASD (Caulfield, Ewing, Avard & Rhodes, 2014; Mathersul, McDonald & Rushby, 2012; Pinkham, Hopfinger, Pelphrey, Piven & Penn, 2008; White, Hill, Winston & Frith, 2006). Despite the links to atypical amygdala functioning in ASD, only a few studies have reported selectively atypical (inflated) trust ratings of
untrustworthy-looking faces (Adolphs et al., 2001; Couture et al., 2010; Losh et al., 2009). We note, however, that these explicitly prompted judgments of trustworthiness constitute a somewhat superficial measure of trust processing. Cued ratings could remain intact in the presence of profoundly atypical spontaneous responses to face stimuli, e.g., potentially reflecting differences in the emotional and motivational salience of faces for individuals with ASD (see Schultz, 2005). Indeed, there have been reports of atypical neural (Pinkham et al., 2008) and autonomic responses (Mathersul et al., 2012) in adults with ASD when viewing trustworthy and untrustworthy face stimuli.

A critical issue to resolve is whether impressions of trustworthiness modulate behaviour in ASD, as they do for typical individuals. To our knowledge, only one study has explored selective trust behavior in children with ASD, and this experiment did not consider the role of facial appearances. Rather, Yi and colleagues (2013) investigated whether young school aged children with ASD might be predisposed to somewhat ‘blindly’ trust information provided to them by others. In the context of a hide and seek game, the authors reported that children with ASD were more – but not indiscriminately – trusting than were typical children when provided with information about the location of a hidden reward by unfamiliar adults. Unfortunately any influence of the informants’ facial appearances was not explored.

One particularly elegant and efficient way to measure the influence of different cues upon trust behavior is through the use of economic trust games (Berg et al., 1995). In these games, Player A is given funds to invest with Player B, who receives triple whatever is given to her/him and then decides how much (if any) to return to Player A. Willingness to invest/reciprocate in each role
signals participants’ trust/trustworthiness, respectively. Several recent studies have demonstrated that trust behaviour in these games is powerfully influenced by facial appearances in typical adults (Chang et al., 2010; Rezlescu et al., 2012; van’t Wout & Sanfey, 2008) and children (Ewing, Caulfield, Read & Rhodes, 2014). Here, we used an economic trust game to investigate the influence of facial appearances on trust behaviour in children with ASD.

In Token Quest, a developmentally appropriate trust game (see Ewing et al., 2014) participants interacted with different [bogus] partners under three experimental conditions: 1) when participants had no information about their partners, 2) when they had access to photographs of the faces of their partners, who appeared either very trustworthy or untrustworthy and 3) when they had access to reputation information about their partners indicating a history of past behaviour that sounded very trustworthy or untrustworthy. This same paradigm has been previously used with adults and typically developing children as young as 4 years of age to investigate the cues that influence trust behavior across development (Ewing et al., 2014).

We confirmed participants’ perceptions of our face stimuli with an explicit trustworthiness-rating task and asked whether, like typical children, the children with ASD would selectively invest more tokens in those partners that looked trustworthy, than untrustworthy. We included the non-face (reputation) condition to determine whether any atypicalities in the use of trust information was specific to facial cues, or might reflect a broader trust processing atypicality. On special bonus trials, we also gave participants the opportunity to purchase access to these face and non-face cues (using their investment capital), reasoning that willingness to purchase might reveal whether they explicitly valued access
to the faces and reputation information when making trust decisions, like typical children (Ewing et al., 2014) and adults (Eckel & Petrie, 2011). Together, these measures were intended to reveal how children with ASD read and use face and non-face cues to partner trustworthiness spontaneously and when explicitly prompted to do so.

Method

Participants

Nine cognitively-able boys with ASD aged 6 years 9 months to 11 years 10 months were recruited from the West Australian Register for Autism Spectrum Disorders, local schools and community groups (see Table 1). Each had received an independent diagnosis of an ASD by a multidisciplinary team following DSM-IV criteria (American Psychiatric Association, 2013) and were rated at or above the cut-off for clinically significant levels of symptomatology (score of 12, Corsello et al., 2007) by their parents on a retrospective measure of autism symptoms: the Social Communication Questionnaire (SCQ, Rutter, Bailey, Lord & Berument, 2003). All children also scored above the autism spectrum cut-off (score of 7) on Module 3 of the Autism Diagnostic Observation Schedule-2 (ADOS-2; Lord et al., 2012), which is a measure of current ASD symptomatology.

In addition, 9 typically developing children (6 male) aged between 6 years 6 months and 9 years 11 months were recruited from local schools and community groups. These children were well matched to the ASD sample on non-verbal IQ and full-scale IQ and did not significantly differ in chronological age or verbal IQ (see Table 1). No typically developing participant displayed clinically significant levels of autistic symptomatology, as indexed by scores below the cut-off on the SCQ.
Procedure

All parents provided written consent prior to their children’s participation in the project. All children also gave verbal and written consent before taking part. Token Quest was part of two or three extended testing sessions that were administered in the family home, at their school or at the University of Western Australia. During these sessions, children completed the Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999), the ADOS-2 and five other experimental tasks unrelated to this study, programmed in Superlab 4 (Cedrus Corporation, 2008) and presented on a 15-inch MacBook Pro laptop computer.

Token Quest. This task measured children’s willingness to trust others in an economic investment ‘game’ paradigm (see Ewing et al., 2014 for more detailed description of the task and stimuli, reliability is yet to be formally assessed). In the game, participants played a series of rounds in which they were given the opportunity to invest 6 tokens (which looked like pirate treasure) with different (bogus) partners. These partners were described as staff and students from our University that had previously told us how they would play the game. We told them that interactions with some partners would result in better returns than others and to avoid risking losses they were free to not give away many, or any, of their tokens. However, we also told them that this ‘risk free’ approach would prevent them from collecting additional tokens for their treasure chest and that the aim of the game was to gather as many tokens as possible.

They each played 3 standard rounds (8 partners/trials) and 2 bonus rounds (2 partners/trials) of the game. On each trial, a representation of a
partner appeared alongside the text “How many tokens would you like to give this partner?” (Figure 1), which remained on screen until the participant gave 0 - 6 tokens to the experimenter. During the standard rounds, participants saw partners represented as either a blank identity (head silhouette) image (Round 1, Figure 1A), a photograph of a very trustworthy- or untrustworthy-looking face (as judged by adults) sourced from the Internet (Round 2, Figure 1B) or reputation information (presented as text and read aloud) about previous trustworthy or untrustworthy behaviour during the game (Round 3, Figure 1C). During the bonus rounds (which were announced after Rounds 2 and 3), participants were told that if they would like to see the faces of 2 additional partners (after Round 2) or access reputation information about the past behaviour of two additional partners (after Round 3), it would cost them 3 of their 6 tokens available for that round (Figure 1D and E). If they elected not to purchase these cues, they viewed blank identities.

Token Quest began with an extended explanation of the aims and rules of the game. Note that we did not explicitly mention trust or trustworthiness during these instructions or at any time during the main Token Quest task. The concept was not introduced until the end of the game, when children were required to rate the trustworthiness of the stimuli they had just viewed (see below). After this introductory phase of Token Quest, children were encouraged to ask any questions they had about how the game worked. They then completed two blank-identity practice trials. After each of these trials, a feedback screen revealed how many tokens the partner chose to return to them. This feedback was pre-determined to give participants one high token return

INSERT FIGURE 1 ABOUT HERE
experience (7 tokens) and one low token return experience (0 tokens). During the remainder of the game, feedback was provided at the end of each round to ensure that token returns were not associated with any specific partners. In all rounds, the number of tokens ‘returned’ to participants was contingent upon the number of tokens invested. All investments were rewarded, regardless of the identity of the ‘trustee’, such that participants that invested all or most of their tokens made better returns than those who invested fewer tokens. This feedback was presented on the computer screen and transferred to a personalized paper progress chart so that participants could keep track of the tokens in their treasure chest across rounds.

After completing Token Quest, we administered the trust rating task. Participants were informed that they would now need to rate the faces on “how trustworthy (they) think each looks”. We asked each child, “What makes someone trustworthy?” and encouraged them to share their ideas in a brief discussion with the experimenter before we highlighted three key elements of interpersonal trustworthiness for them: honesty, reliability and emotional trust (Rotenberg, 1994; Rotenberg et al., 2005). Participants then rated the trustworthiness of characters in six brief behavioural vignettes to confirm that they understood the concept (3 x trustworthy, e.g., “When someone tells him a secret, Joel never tells anyone else, even though sometimes he really wants to. How trustworthy is Joel?”; 3 x untrustworthy, e.g., “Dan said he would lend his copy of the last Harry Potter book to his sister, but he gave it to his friend instead. How trustworthy is Dan?”). They used a 7-point scale consisting of numbered cups (1 = not very trustworthy, 7 = very trustworthy) to make these ratings, via the keyboard (see Cooper, Geldart, Mondloch & Maurer, 2006). They then used the
same scale to rate the trustworthiness of the 5 trustworthy and 5 untrustworthy stimuli presented during the game. When asked, “How trustworthy is (each) face?” The faces each appeared on screen individually (order randomized) until participants responded. Together, Token Quest and the rating task took approximately 20 minutes to complete.

Results

Understanding Trustworthiness

There was no indication that the children with ASD had difficulties understanding the concept of trustworthiness. A 2 x 2 mixed ANOVA investigating the effects of group (ASD, typical) and stimulus trustworthiness (trustworthy, untrustworthy) on participants’ trust ratings of the characters in our behavioural vignettes revealed a significant main effect of trustworthiness, \( F(1, 16) = 280.26, p < .001, \text{partial } \eta^2 = .94 \), with no significant effect of participant group \( F(1, 16) = .14, p = .70, \text{partial } \eta^2 = .01 \) (see Table 2) and no interaction between group and trustworthiness, \( F(1, 16) = 3.51, p = .07, \text{partial } \eta^2 = .18 \). The absence of significant differences between the groups must be interpreted cautiously, given our small sample size. Nevertheless children with and without ASD significantly differentiated between the characters in the trustworthy and untrustworthy vignettes, \( ts > 8.48, ps < .001, ds > 2.83 \), which signals that both groups grasped the key concepts.

Trustworthiness Perception

There was also no evidence that children with ASD had any difficulties reading trustworthiness information from faces. Using the same ANOVA on

---

1 Some participants might have only viewed 8 faces of the 10 faces during Token Quest, if they chose not to reveal the faces of the identities in the face identity bonus round.
children’s explicit trust ratings of the faces used in Token Quest, we found that trustworthy faces were rated significantly more highly than the untrustworthy faces, $F(1, 16) = 23.45, p < .001$, partial $\eta^2 = .59$, with no effect of participant group $F(1, 16) = .26, p = .61$, partial $\eta^2 = .01$ or interaction involving group $F(1, 16) = 1.85, p = .19$, partial $\eta^2 = .10$ (Table 2). Moreover, the children with and without ASD significantly differentiated between the trustworthy and untrustworthy face stimuli with their ratings, $t_s > 3.02, ps < .01, d_s > 1.14$.

**INSERT TABLE 2 ABOUT HERE**

**Trust Behaviour**

Most children in both participant groups entered into the spirit of Token Quest, investing all 6 of their possible 6 tokens in each of the three trust conditions: blank identity, Mdn = 6.0, SD = 0.4; faces, Mdn = 6.0, SD = 0.5; reputation, Mdn = 6.0, SD = 0.8). This ceiling effect suggests that participants were motivated to try to maximize their outcomes via token investments.

**Selective trust behavior: The influence of facial trustworthiness cues.** To assess how the appearance of trustworthiness influenced trust behaviour in children with and without ASD, we examined the number of tokens that each group invested in trustworthy-looking partners (ASD: $M = 2.6, SD = 1.1$, Typical: $M = 4.2, SD = 1.0$) and untrustworthy-looking partners (ASD: $M = 3.2, SD = 1.2$, Typical: $M = 1.8, SD = 1.0$). The difference between these values (i.e., tokens to trustworthy minus tokens to untrustworthy) served as our index of the influence of face cues on trust behaviour. There were no outliers, the data were normally distributed and the skew and kurtosis of the distributions for each group (z score < +/- 1.96) indicated that these data were appropriate for parametric analysis (Field, 2009).
A one-way ANOVA revealed that facial trustworthiness cues modulated trust behaviour in typical children significantly more than in children with ASD, \( F(1, 17) = 9.92, p < .01, \text{ partial } \eta^2 = .38 \) (Figure 2). One-sample t-tests comparing means to zero revealed a significant influence of face cues on trust behaviour in the typical group, \( t(8) = 3.77, p < .01, d = 1.25 \), but not the children with ASD, \( t(8) = 0.89, p = .29 \). Thus unlike the typical children, children with ASD failed to selectively invest tokens with trustworthy-looking partners rather than untrustworthy-looking partners. In line with this result, examination of the association between investment behavior and explicit trustworthiness ratings for each face stimulus revealed non-significant but strong correlation in the typical group \( \tau(7) = .63, p = .08 \) but not the children with ASD \( \tau(7) = .25, p = .54 \).

**INSERT FIGURE 2 ABOUT HERE**

During the face bonus round, fewer children with ASD (44%) than typical children (77%) chose to purchase access to face cues before making their investment decisions, although this difference was not statistically significant, \( \chi^2(1) = 2.10, p < .14 \).

**Selective trust behavior: The influence of non-face trustworthiness cues.**

Inconsistent with any notions of broadly atypical trust processing in children with ASD, there was no evidence of atypicalities in the ASD group when participants had access to non-face cues to partner trustworthiness\(^2\). We calculated the mean difference between tokens invested with partners with trustworthy reputations (ASD: \( M = 4.9, SD = 1.4 \), Typical: \( M = 5.7, SD = 0.5 \)) and

\(^2\) Furthermore, an additional 2 x 2 mixed ANOVA also revealed an interaction between the effects of participant group (ASD, typical) and trustworthiness cue (faces, reputation) on trust behaviour, but unsurprisingly, given the small sample, this effect was not significant, \( F(1, 16) = 3.42, p = .08, \text{ partial } \eta^2 = .17 \).
untrustworthy reputations (ASD: M = 0.6, SD = 0.7, Typical: M = 0.2, SD = 0.4), to serve as our index of the influence of reputation cues on trust behaviour (see Table 2). Again there were no outliers and although the data were not all normally distributed, the low skew and kurtosis of the distributions for each group (z score < +/- 1.96) mean that parametric analyses were still acceptable (Field, 2009).

Trust behaviour in children with and without autism was powerfully influenced by the reputation cues. One-sample comparisons to zero indicated that both groups selectively invested tokens in partners with a trustworthy reputation rather than an untrustworthy reputation, both ts > 6.50, ps < .001, ds > 2.16. There was no significant group difference in this preference for trustworthy-sounding partners, $F(1, 17) = 2.32, p = .14$, partial $\eta^2 = .12$ (Figure 3). Nor was there much difference in willingness to purchase reputation cues during the reputation bonus round in the ASD group (44%) relative to the typical group (55%), $\chi^2 (1) = 0.22, p = 0.63$.

**INSERT FIGURE 3 ABOUT HERE**

Discussion

These results reveal an intriguing profile of trust perception and behaviour based on facial appearances in children with ASD. Despite no evidence of difficulties understanding trustworthiness as a concept, or perceiving trustworthiness from faces (as in Caulfield et al., 2013; Mathersul et al., 2012; Pinkham et al., 2008; White et al., 2006) we observed a striking difference, relative to typical children, in the way that facial trustworthiness cues influenced behaviour. Specifically, during an economic trust game, children with
ASD were no more likely to place their trust in partners that looked trustworthy than those that looked untrustworthy.

Results in the reputation condition suggest that this atypicality did not reflect group differences in how children understood the paradigm or engaged with the task. There, both groups selectively invested more tokens in partners with trustworthy than untrustworthy reputations. Thus children with ASD appeared to have the capacity to invest their tokens rationally, i.e., in line with partner trustworthiness, when they chose to do so. Along with other evidence of developmental increases in rational trust behavior with age during Token Quest (Ewing et al., 2014), these results support the validity of our developmentally appropriate trust game and our interpretation of participants’ investments as an index of perceived partner trustworthiness. Moreover they seem to challenge any notion of a broad trust processing atypicality in ASD.

Our limited sample size warrants cautious interpretation of these results. Still we highlight that where significant group differences in behavior did emerge, they did so despite the relatively small number of participants tested. Nevertheless, it will clearly be important for future research to investigate how these results replicate and generalize to other children and adults with ASD, including those with different cognitive and ASD symptom profiles. When conducting these future studies, it could be interesting to consider how trust behavior in this (still quite artificial) computer game context relates to behavior in more realistic interpersonal interactions. It seems possible that children with ASD could be less sensitive to manipulations of the experimental context, such as having more authentic interactions with partners, than would typical children.
What mechanisms/s might drive atypical trust behaviour based on face cues? One possibility is that children with ASD differ from typical children in the extent to which they spontaneously draw inferences about people from their facial appearances. In our study, children with ASD demonstrated intact perceptions of trustworthiness when they were explicitly prompted to rate faces for this characteristic, but they did not appear to draw upon this information during the trust game, when use of these cues was prompted more implicitly. Reduced social interest and atypical looking at faces is often reported in children with ASD (Dawson et al., 1998; Osterling & Dawson, 1994; Palomo et al., 2006) and a limited ‘default’ level of scrutiny might limit spontaneous social inferences from faces. When explicitly prompted to make these judgments, however, they seemed to have the capacity to do so.

Another possibility is that they do spontaneously infer trustworthiness from faces, but fail to modulate behaviour in light of this information. This might occur for any number of reasons. For example, rational trust behaviour based on face cues might have required complex mental state attribution and/or perspective taking in order to try to anticipate whether each partner would repay their trust. Indeed, the development of theory of mind has been associated with the emergence of selective trust behavior in typical children (Koenig, Clément, & Harris, 2004; Vanderbilt, Liu & Heyman, 2011). If this were the case, then difficulties putting themselves in the place of each partner (Baron-Cohen, Leslie & Frith, 1985; Baron-Cohen, Tager-Flusberg & Cohen, 2000; Leekam & Perner, 1991) might explain the relatively indiscriminate investments observed in children with ASD in the face condition, where this would have been more of an issue than in the reputation condition (with its more transparent
trustworthiness cues). It also seems plausible that failure to use face cues could also reflect differences in the emotional and motivational salience of faces, associated with atypical amygdala functioning in ASD (see Schultz, 2005).

It seems pertinent to highlight here, that this processing atypicality need not constitute a processing weakness relative to typical children. If faces constitute honest signals of trustworthiness, as argued by some (Porter, England, Juodis, ten Brinke & Wilson, 2008; Stirrat & Perrett, 2010) then of course, a failure to use these cues might also make it harder for individuals with ASD to place their trust in those who most deserve it. However, if faces are not always honest signals of trustworthiness (see Rule et al., 2013) then a reduced influence of these impressions on trust behaviour could potentially constitute a social strength in this group. Irrespective of the accuracy of trustworthiness impressions, however, there is growing evidence that the typically developing population reliably read facial trustworthiness from a young age (Caulfield et al., 2014; Cogsdill, Todorov, Spelke & Banaji, 2014; Ewing et al., 2014). Thus any behavioural differences relative to typical children could introduce a degree of unpredictability into their interactions with others and contribute to social processing difficulties.

We also took an exploratory look at how children with ASD explicitly valued access to face stimuli when making investment decisions. When alerted to the idea that they could purchase access to these cues during the bonus round, almost half of the children with ASD chose to buy them and there was no significant group difference relative to typical children. These results are interesting, but must be interpreted cautiously given the limited data (it was a ‘one-shot’ measure) and the extent to which participants’ intentions when
purchasing were somewhat ambiguous. Nevertheless they suggest, again, that group differences in trust behaviour might be less evident if participants were explicitly prompted to consider faces as cues to partner trustworthiness, than when these cues were to be considered more independently.

In summary, the current study revealed that when explicitly prompted, a group of children with ASD demonstrated an intact understanding of trust and an intact ability to evaluate trustworthiness from faces. However, in an economic trust game paradigm, these same children failed to spontaneously use this information to guide their trust behaviour. Importantly, rational trust behaviour (when provided with access to reputation cues) ruled out a broad trust processing atypicality in ASD. This selective failure to modulate behaviour in light of facial trustworthiness cues may contribute to social processing difficulties, relative to typical children. Encouragingly, however, their intact capacity to make these inferences when explicitly prompted suggests that interventions targeting this ‘default’ difference in spontaneous trust behavior, could be effective.
References


Figure Captions

*Figure 1.* Examples of stimuli presented during the different rounds of Token Quest: A) blank identity, B) face identity (note – this identity was not used in the task): C) reputation information. We also show the screens from the two bonus rounds that introduced participants to the notion of paying for access to D) faces and E) 'hints', i.e., reputation information.

*Figure 2.* Mean (SEM) difference between token investments with partners with trustworthy faces and untrustworthy faces (i.e., investments to trustworthy minus untrustworthy faces) in children with ASD and typical children.

*Figure 3.* Mean (SEM) difference between token investments in partners with trustworthy reputations and untrustworthy reputations (i.e., investments to trustworthy- minus untrustworthy-sounding partners) in children with ASD and typical children.
FIGURE TWO

[Graph showing the influence on trust behavior between children with autism and typical children.]
FIGURE THREE

[Graph showing the reputation influence on trust behavior for children with autism and typical children, with data points and error bars.]