Abstract

Facial impressions of trustworthiness guide social decisions in the general population, as shown by financial lending in economic Trust Games. As an exception, autistic boys fail to use facial impressions to guide trust decisions, despite forming typical facial trustworthiness impressions (Ewing et al., 2015). Here, we tested whether this dissociation between forming and using facial impressions of trustworthiness extends to neurotypical men with high levels of autistic traits. Forty-six Caucasian men completed a multi-turn Trust Game, a facial trustworthiness impressions task, the Autism-Spectrum Quotient, and two Theory of Mind tasks. As hypothesized, participants' levels of autistic traits had no observed effect on the impressions formed, but negatively predicted the use of those impressions in trust decisions. Thus, the dissociation between forming and using facial impressions of trustworthiness extends to the broader autism phenotype. More broadly, our results identify autistic traits as an important source of individual variation in the use of facial impressions to guide behaviour. Interestingly, failure to use these impressions could potentially represent rational behaviour, given their limited validity.

Should I Trust You? Autistic Traits Predict Reduced Appearance-Based Trust Decisions

Human faces convey rich social information about an individual's perceived personality (Zebrowitz, 2005). For example, a 100ms exposure to a novel face is enough for consistent first impressions of trustworthiness to form (Todorov, Pakrashi, & Oosterhof, 2009; Willis & Todorov, 2006). Rapid evaluation of trustworthiness from faces is important for survival as it can signal whether someone should be avoided or approached. These evaluations are well-understood (Todorov, Olivola, Dotsch, & Mende-Siedlecki, 2015), and based on underlying cues such as emotional expression, gender, and attractiveness (Oosterhof & Todorov, 2008; Sutherland et al., 2013; 2016). Although facial impressions of trustworthiness are formed automatically, they are not especially accurate predictors of trustworthy behaviour (Rule, Krendl, Ivcevic, & Ambady, 2013).

Regardless of their validity, facial impressions of trustworthiness can have a critical impact on social decisions in everyday life (Olivola, Funk, & Todorov, 2014). For example, more trustworthy-appearing host profile photos on Airbnb receive more bookings (Ert, Fleischer, & Magen, 2016). Surprisingly, host profile photos still influence booking decisions even when more valid information is available in the form of guest reviews. In a financial setting, borrowers on a peer-to-peer lending site are more likely to be funded if they had trustworthy-appearing photographs (Duarte, Siegel, & Young, 2012). In economic Trust Games run in controlled laboratory settings, facial impressions of partner trustworthiness also guide trust decisions, even when accurate information about trustworthy or untrustworthy behaviour is also available (e.g. Chang, Doll, van't Wout, Frank, & Sanfey, 2010). Together, this evidence suggests that facial impressions of trustworthiness are readily formed from faces and can have powerful effects on trust.

Importantly, although people show agreement in their facial impressions of trustworthiness (Rule et al., 2013), emerging research also indicates considerable individual variation in these impressions (Hehman, Sutherland, Flake, & Slepian, 2017). Hehman and colleagues (2017) found that variation in observers influenced facial impressions of trustworthiness as much as variation in the faces themselves. Understanding how individuals vary is one of the oldest psychological questions, yet critically, we have little understanding of the factors driving this observer variation in facial impressions of trustworthiness. Even more importantly, we also lack a clear understanding of how and why people differ in the extent to which they use these impressions to guide trust decisions, which is surprising given the critical social consequences of misplaced trust.

One key potential source of this observer variation is autistic traits. Specifically, autism (American Psychiatric Association, 2013) may represent one extreme end of a continuous distribution of differences in social and cognitive style which extends to levels of non-clinical significance in relatives and the typical population: the ‘broader autism phenotype’ (Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001; Landry & Chouinard, 2016; Piven, Palmer, Jacobi, Childress, & Arndt, 1997). Given the social communication and interaction differences associated with autism, variation in autistic traits is a good candidate for a source of individual variation in the formation and use of facial impressions of trustworthiness.

Indeed, there is strong behavioural and neuroimaging evidence for divergent face processing in autistic individuals as well as neurotypical individuals with high levels of autistic traits (Dawson, Webb, & McPartland, 2005; Deruelle, Rondan, Gepner, & Tardif, 2004; Halliday, MacDonald, Sherf, & Tanaka, 2014; Pierce, Müller, Ambrose, Allen, & Courchesne, 2001). For example, compared to their non-autistic peers, autistic individuals show reduced ability and/or atypical patterns of neural activation when judging facial emotional expressions (Critchley et al., 2000; Uljarevic & Hamilton, 2013) and when inferring social intentions from faces (Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001). Given that a key facial cue to trustworthiness impressions is emotional expression and that facial impressions involve judging intentions (Oosterhof & Todorov, 2008; Sutherland et al., 2016), autistic traits could affect facial impressions of trustworthiness. Consistent with this view, autistic people show atypical neural activation in brain regions associated with facial trustworthiness processing, such as the amygdala (Engell, Haxby, & Todorov, 2007).

Despite this evidence for atypical face processing, autistic people do not differ from their non-autistic peers when explicitly asked to form trustworthy impressions (Ewing, Caulfield, Read, & Rhodes, 2015; Mathersul, McDonald, & Rushby, 2013; S. White, Hill, Winston, & Frith, 2006) and their impressions show the usual influence of emotional expressions (Caulfield, Ewing, Burton, Avard, & Rhodes, 2014). Critically, however, they do not appear to use these impressions to guide social decisions. In an age-appropriate Trust Game, autistic boys did not use facial impressions of trustworthiness to guide their decisions, despite forming similar facial impressions of trustworthiness to neurotypical children when directly asked to do so (Ewing et al., 2015). Ewing and colleagues (2015) proposed that autistic children show a functional dissociation between forming facial impressions of trustworthiness when explicitly asked to do so and implicitly using these impressions to guide behaviour. Interestingly, if facial impressions are not especially accurate predictors of actual trustworthiness (Rule et al., 2013), failure to use them to guide trust decisions may actually represent highly rational social behaviour.

Current Study

Here, we ask whether neurotypical adults with high levels of behavioural traits associated with autism show the same functional dissociation between forming explicit facial impressions of trustworthiness and using these impressions when making trust decisions, as found in autistic children (Ewing et al., 2015). If autistic traits exist on a continuum (Baron-Cohen, Wheelwright, Hill, et al., 2001), then the functional dissociation observed in autism may also influence behaviour - to varying, milder degrees - in the general population. If so, this finding would pinpoint an important source of observer variation in the use of facial trustworthiness impressions.

Individual differences in autistic traits were measured using the Autism-Spectrum Quotient (AQ), a well-established self-report questionnaire (Baron-Cohen, Wheelwright, Hill, et al., 2001). Typicality of facial impressions of trustworthiness was measured by comparing participants' ratings of facial trustworthiness to those from an independent group of raters, so that each participant's score reflected their agreement with typical trustworthiness impressions (following Mathersul et al., 2013; S. White et al., 2006; Zebrowitz, Franklin Jr, Hillman, & Boc, 2013).

Appearance-based trust decisions were measured using a modified Trust Game based on previous research (Chang et al., 2010; Ewing et al., 2015; Rezlescu, Duchaine, Olivola, & Chater, 2012; van't Wout & Sanfey, 2008). In the Trust Game, participants had to decide how much money to transfer to virtual 'partners' that appeared trustworthy or untrustworthy. Appearance-based trust was defined as the money transferred to trustworthy-appearing minus untrustworthy-appearing partners (following Chang et al., 2010). Facial impressions were not accurate predictors of partner behaviour, so that the most rational strategy was to ignore them and rely only on one's personal experience with partner behaviour. We also examined the effect of social experience by manipulating trustworthy or untrustworthy partner behaviour during the Trust Game (i.e., return on investments), in addition to appearance (following Chang et al., 2010). Participants played multiple turns with the same partner, allowing them to experience how trustworthy their partner's behaviour was. Trustworthy partners transferred a fair share of money back and untrustworthy partners unfairly kept the money.

Participants' experience-based trust decisions were defined as the money transferred to the fair minus the unfair partners (following Chang et al., 2010).

We additionally included a brief IQ measure (the WASI-II) to control for the effects of general intelligence on responses during the facial trustworthiness impressions task and Trust Game (following Ewing et al., 2015). We also included a version of the Trust Game with slot machines rather than faces (Slot Machine Game) to control for individual differences in non-social task factors, such as reward sensitivity and general willingness to invest (following Chang et al., 2010). Finally, we included two tasks measuring theory of mind (ToM), to assess the ability to infer the beliefs, intentions and perspectives of others. Higher AQ has been linked with reduced perspective taking and ToM (Baron-Cohen, Wheelwright, Hill, et al., 2001), and reduced ToM has been linked with atypical Trust Game play (Chan & Chen, 2011), suggesting that ToM may mediate the relationship between autistic traits and trustworthiness processing. We used the Reading the Mind in the Eyes Test (RMET) (Baron-Cohen, Wheelwright, Hill, et al., 2001) because it measures ability to infer complex mental attributions from faces, similar to judging intentions. We also used the False-belief/Deception Comprehension task (Langdon, Connors, & Connaughton, 2017), which has been used previously to demonstrate poor first- and second-order ToM abilities in adults with schizophrenia compared to healthy controls. Unlike the RMET, this second ToM task provides a measure of ability in inferring mental-states independent of face perception. We focused on males, because the dissociation between formation and use of facial trustworthiness impressions was previously found in autistic boys (Ewing et al., 2015).

Method

Participants

Forty-six men (Age: M = 21.15, SD = 3.71, range = 13) participated, after giving informed consent to procedures approved by (removed for purposes of blind review) Human Research Ethics Committee. Sample size was based on an a-priori power analysis (G*power; Faul, Erdfelder, Buchner, & Lang, 2009), which confirmed that 23 males would give 0.90 power to detect the effect size found in Ewing et al.'s (2015) sample. Given our non-clinical sample and correlational design, we aimed to recruit approximately double this minimum sample size to allow for a more conservative effect size.

Only native English-speaking Caucasian participants were recruited, to ensure that participants fully understood the tasks and to avoid other-race effects that may impact face processing, as the face stimuli in the tasks were Caucasian (e.g., Hancock & Rhodes, 2008). We tested young adults, since age may affect facial impressions of trustworthiness (Castle et al., 2012). Participants were recruited via a first year Psychology undergraduate volunteer pool, personal invitation, and poster advertisement. They were given payment based on performance in the Trust Game and Slot Machine Game, to ensure these games were consequential and the results generalizable to real financial lending. Specifically, participants were told they would receive $5, $10, $15 or $20 based on broad quartiles of the total averaged money made in the Trust Game and Slot Machine Game. First year Psychology undergraduates were also given credit points.

Tasks

Autistic traits (AQ). The autism spectrum quotient (AQ) is a self-report measure of traits associated with autism in the general population, with good reliability and validity (Baron-Cohen, Wheelwright, Skinner, et al., 2001). The 50 items enquire about everyday behaviours and preferences (e.g., "I like to plan any activities I participate in carefully"). We used the full scale scoring (1 = definitely agree to 4 = definitely disagree) to maximise information about individual differences (Austin, 2005). The highest possible AQ score was 200 and the lowest possible was 50. Higher AQ scores indicate higher autistic trait levels (Baron-Cohen, Wheelwright, Skinner, et al., 2001).

Facial trustworthiness impressions task. After completing two practice trials, participants rated 40 female faces on their trustworthiness in one pseudo-randomised order, using a 9-point scale (1 = not at all trustworthy to 9 = extremely trustworthy). We used ambient face images (see Sutherland et al., 2013 for a discussion on the importance of using naturalistic stimuli) taken from the 10K US adult face database (Bainbridge, Isola, & Oliva, 2013). Faces were Caucasian females, emotionally neutral, front facing, and had direct gaze. We used a single face ethnicity and gender to avoid potential race or gender effects on facial impressions of trustworthiness (Buchan, Croson, & Solnick, 2008; Sutherland et al., 2018). This principle is used in many standardized tests (e.g. the Cambridge Face Memory Test: Duchaine & Nakayama, 2006).

Typicality of facial impressions was measured as the correlation between each individual participant' trustworthiness ratings and consensus mean trustworthiness ratings across the 40 faces (following Mathersul et al., 2013; S. White et al., 2006; Zebrowitz et al., 2013). Consensus ratings were taken from a large group of independent raters (N = 47 Caucasian adults recruited online) and had excellent reliability at the group level (Cronbach's $\alpha = .95$).

Trust Game. Participants played a multi-turn Trust Game with virtual partners, closely based on a previous study (Chang et al., 2010; see Figure 1 for the procedure). On each turn, participants received $10 of virtual money. They then viewed their partner's face for 3.5 s and were required to choose an amount between $0 and $10 to give to their partner. Any investment the participant gave to their partner was quadrupled. Their partner would then ‘decide’ to either return an equal split of the money (trustworthy behaviour) or keep the money and return nothing (untrustworthy behaviour). Thus, if the partner was trustworthy the participant gained double their investment, and if the partner was untrustworthy the participant lost their investment. Participants additionally kept any amount of the virtual $10 not invested. If participants took longer than eight seconds to make an offer, they forfeited their money for that turn (this event occurred only eight times during testing and these trials were treated as missing data in our analysis).

After completing two practice trials with cartoon partners, participants played eight turns with four partners (eight turns × four partners = 32 trials). Two of the four partners were trustworthy-looking and two were untrustworthy-looking and partners were either fair or unfair in behaviour. Partner behaviour was fully crossed with partner facial appearance (i.e. a trustworthy-appearing partner could be fair or unfair in behaviour) and assignment of partner behaviour to facial appearance was counterbalanced across participants. To increase the sample size of images, two sets of possible face photographs were used (participants viewed one set, following Chang et al., 2010). Each partner was not presented more than twice in a row and the same category of facial trustworthiness or behaviour was not presented more than three times in a row. Each participant received the same pseudo-random order of partners because the task was designed to measure individual differences.

Partner facial trustworthiness. Partner images were chosen to appear trustworthy or untrustworthy, based on pilot categorisation judgments (N = 39 Caucasian adults recruited online). At least 75% of pilot judges agreed that the ‘trustworthy’ face images looked trustworthy and the ‘untrustworthy’ faces looked untrustworthy. As above, face stimuli were taken from the 10K US adult face database (Bainbridge et al., 2013).

Partner trustworthy behaviour. Fair partners returned an equal split of money to participants on seven out of eight turns. Conversely, unfair partners unfairly kept all of the money on seven out of eight turns. Behaviour was probabilistic to prevent the game from being too easy and to better replicate real world social interaction (following Chang et al., 2010). Appearance and behaviour were unrelated.

Manipulation checks. Participants were asked some additional questions at the end of the Trust Game, to check they perceived the faces as intended and understood the game. They were asked to rate each partner's face on their trustworthiness levels between 1 and 10 (1 = not at all trustworthy and 10 = extremely trustworthy). They were also required to rate partners on their fairness levels, defined as the percentage of times their partner returned money (0 = 0% to 9 = 90% in 10% increments).

Control tasks. We included a non-social version of the Trust Game, the Slot Machine Game, to control for any non-social determinants of performance (e.g. reward sensitivity, investment generosity, and understanding of mathematical probabilities). The Slot Machine Game was identical to the Trust Game, except that participants played with four different coloured slot machines instead of 'human' partners (following Chang et al., 2010).

The Intelligence Quotient (IQ) was also used as a general control measure, in case intelligence impacted participants' concentration levels, ability to grasp gameplay, or understanding of trustworthiness concepts. IQ was measured using the two-subtest form of the Revised Wechsler Abbreviated Scale of Intelligence (WASI-II; Wechsler, 2011).

Theory of Mind (ToM) tasks. We included two tasks to measure ToM: the revised, "Reading the Mind in the Eyes" task (RMET; Baron-Cohen, Wheelwright, Hill, et al., 2001) and a revised, adult version of the false-belief Sally-Anne task (the False-belief/Deception Comprehension task: (Langdon et al., 2017). RMET scores were calculated by totalling the number of correct answers for each participant. False-belief scores were calculated by summing the first and second order ToM accuracy scores, representing accurate prediction and explanation of the story characters' behaviour by understanding mental states. Each story also included a comprehension question that was correctly answered by all participants, demonstrating that they understood each story.

General Procedure

All tasks were performed on an iMac computer (21.5 inch screen), except for the WASI-II, which was administered individually by the experimenter. Participants completed the Trust Game (15 min), Slot Machine Game (15 min), facial trustworthiness impressions task (five min), two Theory of Mind tasks (10 min), WASI-II (15 min), and the AQ (five min), in that order. Additional measures were completed before the WASI-II and after the AQ, as part of a different study. The facial trustworthiness impressions task was completed after the Trust Game and Slot Machine Game to avoid priming participants on the concept of 'trustworthiness'. The AQ was completed last so that any self-reflection produced by the questionnaire would not affect decision-making in the Trust Game and Slot Machine Game. No deception was used and participants were fully debriefed after the experiment. The testing session lasted approximately one and a half hours.

Results

Data Screening

One participant's data was excluded from the facial trustworthiness impressions task, due to a lack of variability in their responses (i.e. pressing the same button). Due to computer error, one participant's data was also missing from the manipulation checks at the end of the Trust Game. There were no univariate or multivariate outliers identified using the Tukey fences method and via checking the Mahalanobis distance scores. All variables were appropriately distributed for parametric analyses: skew range |0.07 – 1.51|, kurtosis range |0.27 – 3.50| (see Table 1 for descriptive statistics and the full table of correlations in Table

S1 in the supplementary materials). Data will be made available online with the published paper.

Manipulation Checks

To check that facial trustworthiness and partner behaviour in the Trust Game were perceived as expected, participants rated the facial trustworthiness and fairness of partners at the end of the Trust Game. Overall, as expected, participants gave significantly higher facial trustworthiness ratings to trustworthy-appearing partners (M = 5.27, SD = 1.44) than untrustworthy-appearing partners (M = 4.11, SD = 1.13): $t(44) = 3.94$, $p < .001$, Cohen's d = 0.90. Participants also gave significantly higher fairness percentages to fair partners (M = 72.0%, SD = 15.4%) than unfair partners (M = 20.8%, SD = 13.3%): $t(44) = 15.08$, $p < .001$, Cohen's d = 3.56. Therefore, the Trust Game manipulations worked as expected.

Autistic traits and Facial Impressions of Trustworthiness

Scores on the AQ showed good reliability (Cronbach's $\alpha = .82$), were comparable to previous research with adults (M = 112.8, SD = 13.9: Hoekstra, Bartels, Cath, & Boomsma, 2008), and had a broad range (see Figure 2). Facial trustworthiness impression scores were calculated as a Pearson product-moment correlation between participants' trustworthiness ratings of the faces and consensus ratings of the same faces, collected from an independent sample. Correlations were Fisher transformed before analysis.

As expected, participant levels of autistic traits did not significantly predict facial trustworthiness impression scores, $\beta = -.03$, $p = .859$, N = 45 (see Figure 3), even after controlling for general intelligence (IQ composite scores from the WASI-II), $\beta = .01$, p = .959 (see Table 2). Thus, autistic traits had no observed effect on the impressions formed.

Autistic Traits and Appearance-Based Trust Decisions

Appearance-based trust decisions were calculated as the difference between money transferred to trustworthy-appearing partners minus untrustworthy-appearing partners, on average on the initial four trials of the Trust Game (i.e. before feedback about partner behaviour became available). We first assessed the effect of appearance-based trust decisions across the whole group, to establish that our Trust Game was sensitive to appearance-based trust decisions and to replicate previous findings (e.g. Chang et al., 2010). Participants transferred significantly more money on average to trustworthy-appearing partners (*M* = 5.97, *SD* = 1.94) than untrustworthy-appearing partners (*M* = 3.33, *SD* = 2.37), on initial Trust Game trials: paired samples t-test $t(45) = 5.86$, $p < .001$, Cohen's $d = 1.22$. Thus overall, participants relied on the apparent facial trustworthiness of their partners to guide their trust decisions at the start of the Trust Game (replicating previous studies: Chang et al., 2010; Rezlescu et al., 2012; van't Wout & Sanfey, 2008).

As expected, autistic traits significantly negatively predicted appearance-based trust decisions, $\beta = -.32$, $p = .028$, N = 46 (see Figure 4a), which remained significant after controlling for general intelligence and initial money transferred to the first four slot machines in the Slot Machine Game, R^2 change = .11, $\beta = -.34$, $p = .030$ (see Table 3). Thus, autistic traits were significantly negatively related to appearance-based trust decisions, even when controlling for non-social aspects of gameplay. Importantly, autistic traits predicted appearance-based trust decisions significantly more than facial trustworthiness impressions: $t(87) = 2.32$, p = .023. This finding supports the predicted dissociation between forming facial trustworthiness impressions and using these impressions to make trust decisions.

To explore the types of autistic traits driving the significant negative association between autistic traits and appearance-based trust decisions, we examined the relationship between appearance-based trust decisions and the 'social interaction' and 'attention to detail' subscales of the AQ (as used in Rhodes, Jeffery, Taylor, & Ewing, 2013, based on a factor analysis of the AQ: Hoekstra et al., 2008). These subscales target social and non-social traits, respectively, and had reasonable internal consistency (Cronbach's $\alpha = .83$ and .69

respectively). The analyses were Bonferroni-corrected (adjusted alpha level of .025) for multiple comparisons because they were exploratory. The ‘attention to detail’ subscale significantly negatively predicted appearance-based trust decisions, β = -.42, *p* = .004, N = 46. That is, participants who scored highly on items like “I tend to notice details that others do not” were less likely to be influenced by their partners’ appearance during the trust game. By contrast, the ‘social interaction’ subscale did not significantly predict appearance-based behaviour, β = -.21, *p* = .165, N = 46. The effect of ‘attention to detail’ remained significant when controlling for general intelligence and initial money transferred to slot machines, R^2 change = .18, β = -.42, *p* = .004.

Autistic traits were not significantly correlated with either ToM measure: RMET scores, *r* = -.03, *p* = .831, N = 46, False-belief scores, *r* = -.10, *p* = .501, N = 46 (unlike in previous research: e.g. Baron-Cohen, Wheelwright, Hill, et al., 2001), providing no evidence that ToM mediated the observed relationship between autistic traits and appearance-based trust decisions.

Autistic Traits and Experience-based Trust Decisions

Experience-based trust decisions were calculated as the difference in money transferred to fair partners minus unfair partners, on average across all trials of the Trust Game where feedback on partner fair or unfair behaviour was available (i.e. all trials except the first four). We first assessed the effect of experience-based trust decisions across the whole group, to ascertain that our Trust Game was sensitive to experience-based trust decisions and to replicate previous findings (Chang et al., 2010). Participants transferred significantly more money on average to fair partners (*M* = 6.66, *SD* = 1.55) than unfair partners (*M* = 2.35, *SD* = 1.60): paired samples t-test *t*(45) = 13.73, *p* < .001, Cohen’s *d* = 2.74. Overall, participants relied on experience with partner behaviour to guide their trust decisions (consistent with previous research: Chang et al., 2010).

Importantly, autistic traits did not significantly predict experience-based trust decisions: β = -.19, p = .196, N = 46 (see Figure 4b; Table 4), nor did this result change when controlling for general intelligence and money transferred based on slot machine ‘behaviour’ (i.e. difference in the averaged money transferred to fair and unfair slot machines across all Slot Machine Game trials except the first four), β = -.07, p = .549. Therefore, there was no evidence that autistic traits promoted less rational social decisions based on experience in the Trust Game.

Discussion

Facial impressions of trustworthiness can have a powerful impact on social decisions to trust others (Duarte et al., 2012; Ert et al., 2016), but can have a reduced impact on the social decisions of autistic people (Ewing et al., 2015). Here for the first time, we found that neurotypical men with higher levels of autistic traits were less likely to use facial impressions to guide trust decisions, despite their ability to form typical facial impressions of trustworthiness when explicitly asked to do so. Thus the functional dissociation between forming and using facial impressions of trustworthiness found in autistic boys (Ewing et al., 2015) extends to the broader autism phenotype. More generally, our results identify an important source of variation in the use of facial impressions to guide trusting behaviour, a key social task.

Autistic Traits and the Functional Dissociation for Facial Trustworthiness

Autistic trait levels had no effect on explicit facial impressions of trustworthiness. This pattern is consistent with previous findings that autistic individuals show typical facial impressions of trustworthiness (Ewing et al., 2015; Mathersul et al., 2013; S. White et al., 2006). Importantly, however, autistic traits were negatively related to trust decisions based on partner appearance, showing that neurotypical individuals with higher levels of autistic traits were less influenced by partner appearance when making trust decisions. Thus, the functional dissociation between forming and using facial impressions of trustworthiness found in autistic boys (Ewing et al., 2015) extends to the broader autism phenotype (Piven et al., 1997). Not surprisingly, the effect size relating autistic traits to facial appearance biases was weaker than found when examining children with and without autism (Ewing et al, 2015). Nevertheless, it was reasonably strong, with autistic traits explaining 11% of the variance in those biases (after controlling for general intelligence and generosity).

What potential mechanisms might explain the functional dissociation between forming and using facial impressions of trustworthiness? From a perceptual perspective, this dissociation could be due to less spontaneous formation of facial impressions of trustworthiness in individuals with higher AQ scores, when not explicitly asked to judge facial trustworthiness. An alternative explanation for this dissociation is that men with autistic traits can spontaneously evaluate facial trustworthiness, but are less likely to functionally apply these evaluations to complex trust decisions.

Our exploratory findings suggest that a heightened attention to detail may be the core characteristic that drives this pattern of cognitive processing and behaviour. Potentially, greater attention to local details may make it more difficult for men with higher levels of autistic traits to spontaneously process a holistic facial impression, such as trustworthiness (Santos & Young, 2011; Todorov, Loehr, & Oosterhof, 2010). This pattern fits with other studies in which atypical face processing in autistic people can be explained by a bias towards processing local rather than global features of faces (Behrmann et al., 2006; Happé & Frith, 2006). Although Theory of Mind (ToM) requires the functional attribution of mental states to others, we found no evidence that it can explain the resistance to use facial impressions of trustworthiness to guide behaviour in men with high levels of autistic traits.

Is Reduced Use of Facial Trustworthiness a Social Disadvantage?

If facial appearance accurately reflects trustworthy behaviour, then relying on first impressions of trustworthiness to guide trust decisions may be a beneficial social strategy. However, facial cues may not be accurate indicators of general trustworthiness (Rule et al., 2013). When facial appearance is misleading, relying on appearance to guide trust decisions will lead to poor social judgment. In that case, the reduced reliance on facial trustworthiness found in men with higher levels of autistic traits may actually represent more rational social behaviour.

Moreover, even potentially rational trust behaviour could cause autistic individuals to be unpredictable in social interactions, due to its atypicality, contributing to interpersonal difficulties when interacting with the neurotypical majority. For example, choosing a trustworthy-looking profile picture for use online is arguably now a critical social skill (D. White, Sutherland, & Burton, 2017) with important real-life ramifications (Olivola et al., 2014). It is important to understand the social impact of a reduced reliance on facial trustworthiness, given that these judgements can have real consequences. Indeed, we also found that facial trustworthiness guided initial trust decisions at the group level even though people knew that the Trust Game had real financial penalties (consistent with previous work: Chang et al., 2010; Rezlescu et al., 2012; van't Wout & Sanfey, 2008).

In contrast to facial impressions, experience with an individual's fair or unfair behaviour is a valid predictor of their trustworthiness (King-Casas et al., 2005). Therefore, our Trust Game was designed so that experience with behaviour predicted a partner's trustworthiness whereas facial appearance did not, and, as a result, using experience with partner behaviour to guide trust decisions was the most rational game strategy. We found no evidence that autistic traits related to trust decisions based on experience with partner behaviour. In this sense, men with higher levels of autistic traits made trust decisions that were just as rational as those in men with lower levels; they used trust information when it was useful (i.e. experience with behaviour) but not when it was not useful (i.e. misleading facial appearance).

Limitations and Future Directions

Trust Games can model how facial impressions of trustworthiness guide social decisions in controlled laboratory settings (e.g. van't Wout & Sanfey, 2008) and our game employed pre-programmed 'partners', in order that we could carefully control facial appearance. It is possible that the resistance of individuals with higher autistic traits to using facial impressions reflects a lack of engagement in the game, at least when playing with 'partners' (and not slot machines). It will be interesting to see whether the results generalise to live social interactions.

Trust Games themselves involve investment decisions that are motivated by extrinsic monetary rewards (Johnson & Mislin, 2011). Consequently, our current findings relate especially to social settings where facial trustworthiness impacts on financial decisions, such as trading on peer-to-peer lending sites (Duarte et al., 2012) or choosing accommodation on Airbnb (Ert et al., 2016). Another interesting future direction would be to examine how autistic traits relate to trust behaviour or facial appearance biases in social settings that do not involve financial decisions, such as electoral voting or jury verdicts, given that these settings are also subject to facial trustworthiness biases (Todorov, Mandisodza, Goren, & Hall, 2005; Wilson & Rule, 2015). Interestingly, Bayliss and Tipper (2006) also previously found, using a gaze-cuing paradigm, that individuals with higher autistic traits were less likely to use facial cues in their trustworthiness decisions, where trustworthiness was (accurately) cued by predictive or deceptive gaze. Thus, our finding that individuals with higher autistic traits were also less guided by facial cues to trustworthiness in a trust game, may represent a more general pattern of behaviour that generalises across contexts.

Finally, autistic traits appear to represent an important individual difference underlying how facial impressions guide trust decisions, a relatively new area of study. In future, it will be important to chart the contributions of other individual differences, such as other key personality traits (Chan & Chen, 2011) or beliefs around the validity of facial appearance as a guide to trustworthy character (Suzuki, Tsukamoto, & Takahashi, 2017).

Conclusions

Variation in autistic traits in neurotypical individuals is associated with the extent to which facial impressions of trustworthiness influence social decision-making. When making trust decisions, neurotypical men with higher levels of autistic traits relied less on facial appearance information than those with lower levels, despite forming typical impressions when prompted to do so. Interestingly, failure to use these impressions could potentially represent rational behaviour, given their limited validity. Variation in autistic traits did not influence trust decisions based on actual partner behaviour, or in a non-social version of the game, and therefore did not represent a general disadvantage when making trust decisions. These results demonstrate that the dissociation between formation and use of facial impressions of trustworthiness found in autism (Ewing et al, 2015) extends to the broader phenotype. Neurotypical men with higher levels of autistic traits may therefore show more resistance to the influence of unreliable facial cues to trustworthiness in guiding everyday behaviour.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Austin, E. J. (2005). Personality correlates of the broader autism phenotype as assessed by the Autism Spectrum Quotient (AQ). *Personality and Individual Differences*, *38*(2), 451–460. http://dx.doi.org/10.1016/j.paid.2004.04.022

Bainbridge, W. A., Isola, P., & Oliva, A. (2013). The intrinsic memorability of face photographs. *Journal of Experimental Psychology: General*, *142*(4), 1323. http://dx.doi.org/10.1037/a0033872

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry*, *42*(2), 241–251. http://dx.doi.org/10.1111/1469-7610.00715

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): Evidence from asperger syndrome/high-functioning autism, malesand females, scientists and mathematicians. *Journal of Autism and Developmental Disorders, 31*(1), 5–17. http://dx.doi.org/10.1023/A:1005653411471

Bayliss, A. P., & Tipper, S. P. (2006). Predictive gaze cues and personality judgments: Should eye trust you? *Psychological Science*, *17*(6), 514–520. https://doi.org/10.1111/j.1467-9280.2006.01737.x

Behrmann, M., Avidan, G., Leonard, G. L., Kimchi, R., Luna, B., Humphreys, K., & Minshew, N. (2006). Configural processing in autism and its relationship to face processing. *Neuropsychologia, 44*(1), 110–129. http://dx.doi.org/10.1016/j.neuropsychologia.2005.04.002

Buchan, N. R., Croson, R. T. A., & Solnick, S. (2008). Trust and gender: An examination of behavior and beliefs in the Investment Game. *Journal of Economic Behavior & Organization*, *68*(3–4), 466–476. http://dx.doi.org/10.1016/j.jebo.2007.10.006

Castle, E., Eisenberger, N. I., Seeman, T. E., Moons, W. G., Boggero, I. A., Grinblatt, M. S., & Taylor, S. E. (2012). Neural and behavioral bases of age differences in perceptions of trust. *Proceedings of the National Academy of Sciences*, *109*(51), 20848–20852. http://dx.doi.org/10.1073/pnas.1218518109

Caulfield, F., Ewing, L., Burton, N., Avard, E., & Rhodes, G. (2014). Facial trustworthiness judgments in children with ASD are modulated by happy and angry emotional cues. *PLoS ONE*, *9*(5), e97644. https://doi.org/10.1371/journal.pone.0097644

Chan, K. K. S., & Chen, E. Y. H. (2011). Theory of mind and paranoia in schizophrenia: a game theoretical investigation framework. *Cognitive Neuropsychiatry*, *16*(6), 505–529. https://doi.org/10.1080/13546805.2011.561576

Chang, L. J., Doll, B. B., van't Wout, M., Frank, M. J., & Sanfey, A. G. (2010). Seeing is believing: Trustworthiness as a dynamic belief. *Cognitive Psychology*, *61*(2), 87–105. http://dx.doi.org/10.1016/j.cogpsych.2010.03.001

Critchley, H. D., Daly, E. M., Bullmore, E. T., Williams, S. C. R., Van Amelsvoort, T., Robertson, D. M., … Howlin, P. (2000). The functional neuroanatomy of social behaviour: changes in cerebral blood flow when people with autistic disorder process facial expressions. *Brain*, *123*(11), 2203–2212. http://dx.doi.org/10.1093/brain/123/11.2203

Dawson, G., Webb, S. J., & McPartland, J. (2005). Understanding the nature of face processing impairment in autism: insights from behavioral and electrophysiological studies. *Developmental Neuropsychology*, *27*(3), 403–424. http://dx.doi.org/10.1207/s15326942dn2703_6

Deruelle, C., Rondan, C., Gepner, B., & Tardif, C. (2004). Spatial frequency and face processing in children with autism and Asperger syndrome. *Journal of Autism and*

Developmental Disorders, *34*(2), 199–210. http://dx.doi.org/10.1023/B:JADD.0000022610.09668.4c

Duarte, J., Siegel, S., & Young, L. (2012). Trust and credit: The role of appearance in peer-to-peer lending. *Review of Financial Studies*, *25*(8), 2455–2484. https://doi.org/10.1093/rfs/hhs071

Duchaine, B., & Nakayama, K. (2006). The Cambridge Face Memory Test: Results for neurologically intact individuals and an investigation of its validity using inverted face stimuli and prosopagnosic participants. *Neuropsychologia*, *44*(4), 576–585. http://dx.doi.org/10.1016/j.neuropsychologia.2005.07.001

Engell, A. D., Haxby, J. V., & Todorov, A. (2007). Implicit trustworthiness decisions: automatic coding of face properties in the human amygdala. *Journal of Cognitive Neuroscience*, *19*(9), 1508–1519. http://dx.doi.org/10.1162/jocn.2007.19.9.1508

Ert, E., Fleischer, A., & Magen, N. (2016). Trust and reputation in the sharing economy: The role of personal photos on Airbnb. *Tourism Management*, *55*, 62–73. http://dx.doi.org/10.1016/j.tourman.2016.01.013

Ewing, L., Caulfield, F., Read, A., & Rhodes, G. (2015a). Appearance-based trust behaviour is reduced in children with autism spectrum disorder. *Autism*, *19*(8), 1002–1009. http://dx.doi.org/10.1177/1362361314559431

Ewing, L., Caulfield, F., Read, A., & Rhodes, G. (2015b). Perceived trustworthiness of faces drives trust behaviour in children. *Developmental Science*, *18*(2), 327–334. https://doi.org/10.1111/desc.12218

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. http://dx.doi.org/10.3758/BRM.41.4.1149

Halliday, D. W. R., MacDonald, S. W. S., Sherf, S. K., & Tanaka, J. W. (2014). A reciprocal model of face recognition and autistic traits: evidence from an individual differences perspective. *PloS ONE, 9*(5), e94013. https://doi.org/doi:10.1371/journal.pone.0094013

Hancock, K. J., & Rhodes, G. (2008). Contact, configural coding and the other-race effect in face recognition. *British Journal of Psychology*, *99*(1), 45–56. https://doi.org/10.1348/000712607X199981

Happé, F., & Frith, U. (2006). The weak coherence account: detail-focused cognitive style in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *36*(1), 5–25. http://dx.doi.org/10.1007/s10803-005-0039-0

Hehman, E. A., Sutherland, C. A. M., Flake, J. K., & Slepian, M. L. (2017). The unique contributions of perceiver and target characteristics in person perception. *JPSP*, *113*(4), 513–529. http://dx.doi.org/10.1037/pspa0000090

Hoekstra, R. A., Bartels, M., Cath, D. C., & Boomsma, D. I. (2008). Factor structure, reliability and criterion validity of the Autism-Spectrum Quotient (AQ): a study in Dutch population and patient groups. *Journal of Autism and Developmental Disorders*, *38*(8), 1555–1566. http://dx.doi.org/10.1007/s10803-008-0538-x

Johnson, N. D., & Mislin, A. A. (2011). Trust games: A meta-analysis. *Journal of Economic Psychology*, *32*(5), 865–889. https://doi.org/10.1016/j.joep.2011.05.007

King-Casas, B., Tomlin, D., Anen, C., Camerer, C. F., Quartz, S. R., & Montague, P. R. (2005). Getting to know you: reputation and trust in a two-person economic exchange. *Science*, *308*(5718), 78–83. http://dx.doi.org/10.1126/science.1108062

Landry, O., & Chouinard, P. A. (2016). Why we should study the broader autism phenotype in typically developing populations. *Journal of Cognition and Development*, *17*(4), 584–595. https://doi.org/10.1080/15248372.2016.1200046

Langdon, R., Connors, M., & Connaughton, E. (2017). Indirect task instructions better reveal theory-of-mind impairment, independent of executive dysfunction, in schizophrenia. *Psychiatry Research*, *256*, 342–344. https://doi.org/10.1016/j.psychres.2017.06.064

Mathersul, D., McDonald, S., & Rushby, J. A. (2013). Psychophysiological correlates of social judgement in high-functioning adults with autism spectrum disorder. *International Journal of Psychophysiology*, *87*(1), 88–94. http://dx.doi.org/10.1016/j.ijpsycho.2012.11.005

Olivola, C. Y., Funk, F., & Todorov, A. (2014). Social attributions from faces bias human choices. *Trends in Cognitive Sciences*, *18*(11), 566–570. https://doi.org/doi:10.1016/j.tics.2014.09.007

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *PNAS*, *105*(32), 11087–11092. https://doi.org/10.1073/pnas.0805664105

Pierce, K., Müller, R. A., Ambrose, J., Allen, G., & Courchesne, E. (2001). Face processing occurs outside the fusiform "face area" in autism: Evidence from functional MRI. *Brain*, *124*(10), 2059–2073. http://dx.doi.org/10.1093/brain/124.10.2059

Piven, J., Palmer, P., Jacobi, D., Childress, D., & Arndt, S. (1997). Broader autism phenotype: Evidence from a family history study of multiple-incidence autism families. *The American Journal of Psychiatry*, *154*(2), 185–190. http://dx.doi.org/10.1176/ajp.154.2.185

Rezlescu, C., Duchaine, B., Olivola, C. Y., & Chater, N. (2012). Unfakeable facial configurations affect strategic choices in trust games with or without information about past behavior. *PLoS ONE*, *7*(3), e34293. https://doi.org/10.1371/journal.pone.0034293

Rhodes, G., Jeffery, L., Taylor, L., & Ewing, L. (2013). Autistic traits are linked to reduced adaptive coding of face identity and selectively poorer face recognition in men but not women. *Neuropsychologia*, *51*(13), 2702–2708. https://doi.org/10.1016/j.neuropsychologia.2013.08.016

Rule, N. O., Krendl, A. C., Ivcevic, Z., & Ambady, N. (2013). Accuracy and consensus in judgments of trustworthiness from faces: Behavioral and neural correlates. *Journal of Personality and Social Psychology*, *104*(3), 409–426. https://doi.org/10.1037/a0031050

Santos, I. M., & Young, A. W. (2011). Inferring social attributes from different face regions: Evidence for holistic processing. *The Quarterly Journal of Experimental Psychology*, *64*(4), 751–766. https://doi.org/doi/full/10.1080/17470218.2010.519779

Sutherland, C. A. M., Liu, X., Zhang, L., Chu, Y., Oldmeadow, J. A., & Young, A. W. (2018). Facial first impressions across culture: Data-driven modeling of Chinese and British perceivers' unconstrained facial impressions. *Personality and Social Psychology Bulletin*, *44*(4). https://doi.org/10.1177/0146167217744194

Sutherland, C. A. M., Oldmeadow, J. A., Santos, I. M., Towler, J., Burt, D. M., & Young, A. W. (2013). Social inferences from faces: Ambient images generate a three-dimensional model. *Cognition*, *127*(1), 105–118. https://doi.org/10.1016/j.cognition.2012.12.001

Sutherland, C. A. M., Young, A. W., & Rhodes, G. (2016). Facial first impressions from another angle: Social judgements are influenced by both emotional expression and viewpoint. *British Journal of Psychology, 108*(2), 397–415. https://doi.org/10.1111/bjop.12206

Suzuki, A., Tsukamoto, S., & Takahashi, Y. (2017). Faces tell everything in a just and biologically determined World: Lay theories behind face reading. *Social Psychological and Personality Science*, 1948550617734616. https://doi.org/10.1177/1948550617734616

Todorov, A., Loehr, V., & Oosterhof, N. N. (2010). The obligatory nature of holistic processing of faces in social judgments. *Perception*, *39*(4), 514–532. https://doi.org/DOI: 10.1068/p6501

Todorov, A., Mandisodza, A. N., Goren, A., & Hall, C. C. (2005). Inferences of competence from faces predict election outcomes. *Science*, *308*(5728), 1623–1626. https://doi.org/10.1126/science.1110589

Todorov, A., Olivola, C. Y., Dotsch, R., & Mende-Siedlecki, P. (2015). Social attributions from faces: Determinants, consequences, accuracy, and functional significance. *Annual Review of Psychology*, *66*(1), 519–545. https://doi.org/10.1146/annurev-psych-113011-143831

Todorov, A., Pakrashi, M., & Oosterhof, N. N. (2009). Evaluating faces on trustworthiness after minimal time exposure. *Social Cognition, 27*(6), 813–833. https://doi.org/10.1521/soco.2009.27.6.813

Uljarevic, M., & Hamilton, A. (2013). Recognition of emotions in autism: a formal meta-analysis. *Journal of Autism and Developmental Disorders*, *43*(7), 1517–1526. http://dx.doi.org/10.1007/s10803-012-1695-5

van’t Wout, M., & Sanfey, A. G. (2008). Friend or foe: The effect of implicit trustworthiness judgments in social decision-making. *Cognition*, *108*(3), 796–803. http://dx.doi.org/10.1016/j.cognition.2008.07.002

Wechsler, D. (2011). *WASI II: Wechsler Abbreviated Scale of Intelligence* (2nd ed.). San Antonio: Psychological Corporation.

White, D., Sutherland, C. A., & Burton, A. L. (2017). Choosing face: The curse of self in profile image selection. *Cognitive Research: Principles and Implications*, *2*(1), 23. https://doi.org/10.1186/s41235-017-0058-3

White, S., Hill, E., Winston, J., & Frith, U. (2006). An islet of social ability in Asperger Syndrome: judging social attributes from faces. *Brain and Cognition*, *61*(1), 69–77. http://dx.doi.org/10.1016/j.bande.2005.12.007

Willis, J., & Todorov, A. (2006). First impressions: making up your mind after a 100-ms exposure to a face. *Psychological Science*, *17*(7), 592–598. https://doi.org/10.1111/j.1467-9280.2006.01750.x

Wilson, J. P., & Rule, N. O. (2015). Facial trustworthiness predicts extreme criminal-sentencing outcomes. *Psychological Science*, *26*(8), 1325–1331. https://doi.org/10.1177/0956797615590992

Zebrowitz, L. A. (2005). Appearance DOES Matter. *Science, 308*, 1565–1566. https://doi.org/10.1126/science.1114170

Zebrowitz, L. A., Franklin Jr, R. G., Hillman, S., & Boc, H. (2013). Older and younger adults' first impressions from faces: Similar in agreement but different in positivity. *Psychology and Aging, 28*(1), 202–212. http://dx.doi.org/10.1037/a0030927

Tables

Table 1

Descriptive statistics for task variables. All N = 46 apart from the facial trustworthiness impressions (N = 45).

Variables	Mean	SD	Min	Max	Skew	Kurtosis
AQ total	112.78	13.87	88.00	155.00	0.73	1.14
AQ Attention to detail	24.78	4.54	17.00	37.00	0.17	-.38
AQ Interpersonal	88.00	12.43	69.00	127.00	1.11	1.73
RMET total score	26.24	3.35	17.00	33.00	-0.67	0.35
False-belief Stories (1st and 2nd order ToM)	8.89	2.20	4.00	12.00	-0.14	-0.98
IQ (WASI composite score)	110.74	9.27	92.00	138.00	0.60	0.80
Facial trustworthiness impressions	0.33	0.20	-0.02	0.80	0.28	-0.67
Money ($) to trustworthy-looking partners (initial trials)	5.97	1.94	2.00	10.00	0.10	-0.67
Money ($) to untrustworthy-looking partners (initial trials)	3.33	2.37	0.00	9.00	0.42	-0.27
Money ($) to trust minus untrust-looking partners (initial trials)	2.64	3.06	-4.00	10.00	0.07	-0.44
Money ($) to fair partners (all trials except initial)	6.66	1.55	2.36	10.00	-0.57	0.67
Money ($) to unfair partners (all trials except initial)	2.35	1.60	0.21	8.14	1.51	3.50
Money ($) to fair minus unfair partners (all trials except initial)	4.32	2.13	-1.14	8.21	-0.50	0.49
Slot Machine control (money transferred, initial trials)	5.24	1.47	2.50	9.50	0.40	0.77
Slot Machine control (money to fair versus unfair slot machines)	5.15	2.63	0.00	9.64	-0.11	-0.89

Table 2

Hierarchical multiple regression predicting facial trustworthiness impression scores from autistic traits (AQ scores; N = 45).

Predictor	B ± SE	β	*t*	*p*	Partial *r*	R^2
Step 1						.06
IQ	.01 ± .00	.24	1.62	.112	.24	
Step 2						.06
IQ	.01 ± .00	.24	1.59	.118	.24	
AQ	.00 ± .00	.01	.05	.959	.01	

Table 3

Hierarchical multiple regression predicting money transferred based on partner facial trustworthiness (initial four trials of the trust game) from Autistic Traits (AQ scores; N = 46).

Predictor	B ± SE	β	*t*	*p*	Partial *r*	R^2
Step 1						.00
IQ	.02 ± .05	.05	.32	.750	.05	
Slot Machine Game control	-.05 ± .34	-.02	-.15	.882	-.02	
Step 2						.11
IQ	-.01 ± .05	-.02	-.14	.892	-.02	
Slot Machine Game control	-.17 ± .33	-.08	-.53	.601	-.08	
AQ	-.07 ± .03	-.34	-2.25	.030*	-.33	

Note. Slot Machine Game control is the average money transferred to the first four slot machines in the Slot Machine Game. **p* < .05.

Table 4

Hierarchical multiple regression predicting money transferred based on partner behaviour (throughout the Trust Game, except the initial four trials) from Autistic Traits (AQ scores; N = 46).

Predictor	B ± SE	β	*t*	*p*	Partial *r*	R^2
Step 1						.42
IQ	.08 ± .03	.35	3.05	.004**	.42	
Slot Machine Game control	.47 ± .09	.58	4.98	<.001***	.60	
Step 2						.43
IQ	.08 ± .03	.34	2.88	.006**	.41	
Slot Machine Game control	.46 ± .10	.57	4.81	<.001***	.60	
AQ	-.01 ± .02	-.07	-.60	.549	-.09	

Note. Slot Machine Game control is the difference in money transferred to fair versus unfair slot machines, averaged throughout the Slot Machine Game except for the first four trials.

$^{**}p < .01$. $^{***}p < .001$.

Figures

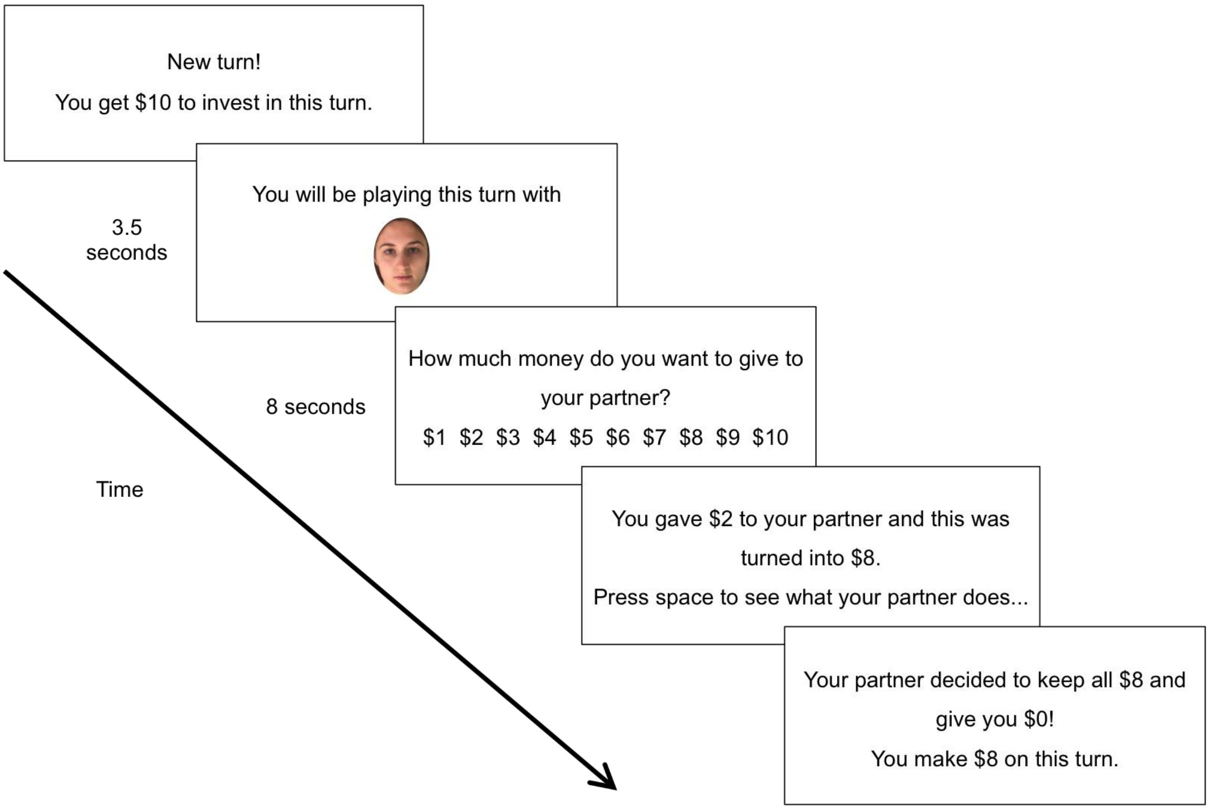

Figure 1. Example of a single turn from the Trust Game. *Note*. Instead of transferring money, participants could also choose to transfer no money by pressing a key labelled as 'no money'. The face above is an example due to copyright reasons and was not one of the faces used in the Trust Game.

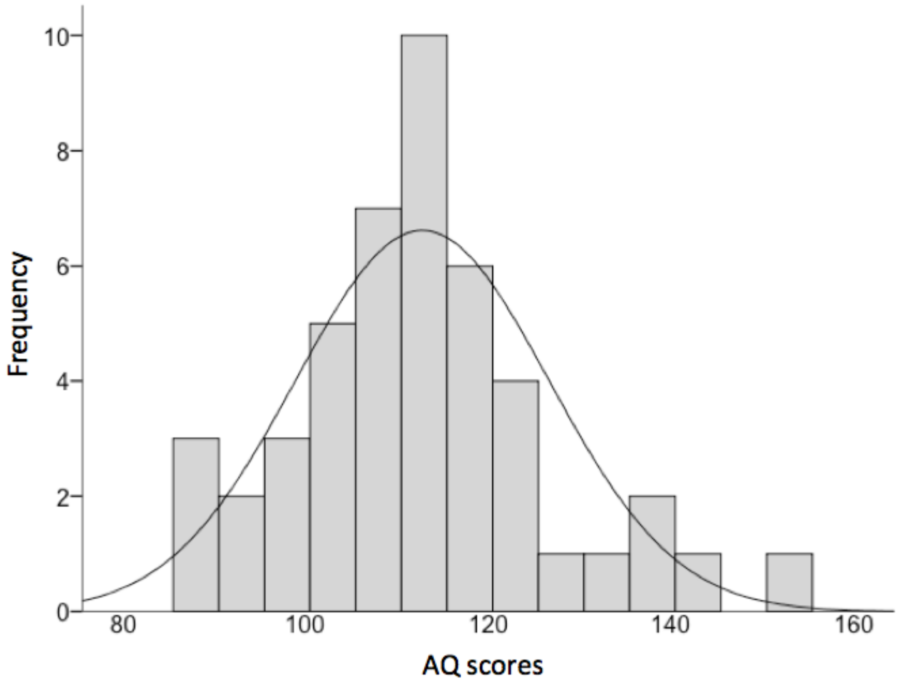

Figure 2. Histogram of autistic traits (AQ scores; N = 46). The normal distribution curve is shown.

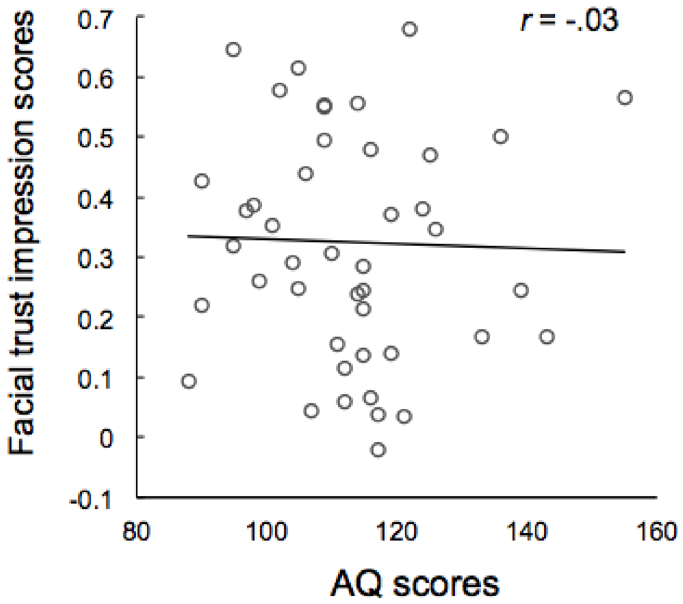

Figure 3. Scatter plot of autistic traits (AQ scores) and facial trustworthiness impression scores, representing the extent to which each participant agreed with independent consensus ratings, as measured by individual Pearson's correlations (N = 45). The best-fitting regression line is shown and the overall Person's correlation.

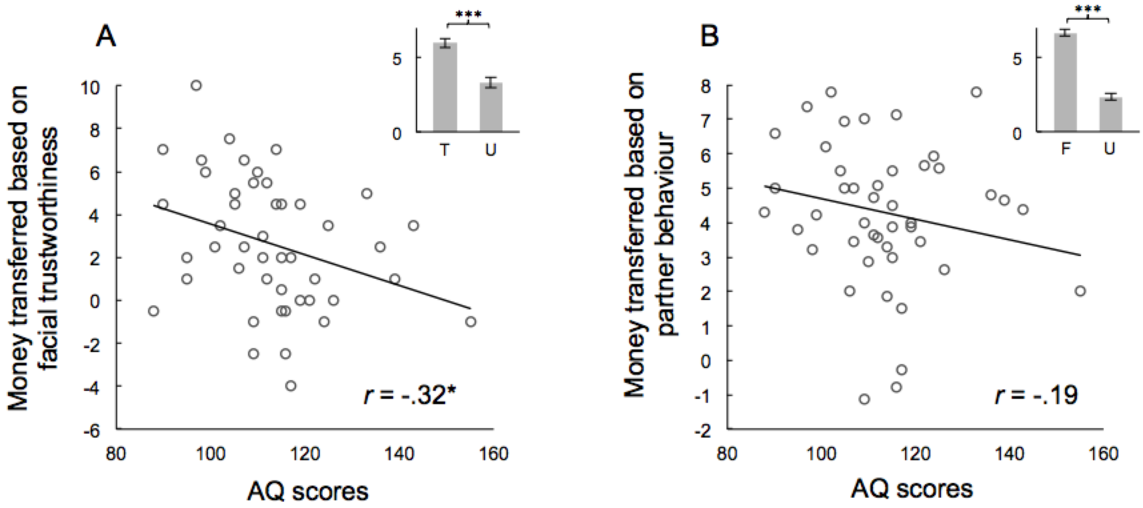

Figure 4. A) Scatter plot of autistic traits (AQ scores) and appearance-based trust decisions (on average across the initial four Trust Game trials, before feedback about behaviour became available). The best-fitting regression line is shown (N = 46). *B)* Scatterplot of autistic traits (AQ scores) and experience-based trust decisions (on average across all Trust Game trials where feedback about fair or unfair behaviour was available). The best-fitting regression line is shown and the overall Person's correlations (N = 46). *Note.* The top right of each figure displays the money transferred to trustworthy-appearing (T) and untrustworthy-appearing (U) partners, or fair-behaving (F) and unfair-behaving (U) partners, collapsed across all participants. SE bars are shown. * p <.05, *** p < .001.

Supplementary materials

Table S1

Pearson's correlations and associated p-values between main task variables. All N = 46 apart from the facial trustworthiness impressions (N = 45).

	AQ	IQ	RMET	False-belief Stories	Facial trust. impress.	Money (looks)	Money (fairness)	Slot Game control (total initial)
IQ (WASI composite)	-.156							
	.299							
RMET (total score)	-.032	.271						
	.831	*.068*						
False-belief Stories (1st and 2nd order ToM)	-.102	.26	.406**					
	.501	*.081*	*.005*					
Facial trustworthiness impressions	-.027	.24	.344*	.461**				
	.859	*.112*	*.021*	*.001*				
Money ($) to trust minus untrust looking partners (initial trials)	-.324*	.061	.094	.060	.039			
	.028	*.686*	*.534*	*.692*	*.798*			
Money ($) to fair minus unfair partners (all trials except initial)	-.194	.306*	.253	.240	.318*	.282		
	.196	*.039*	*.09*	*.109*	*.033*	*.058*		
Slot Game control initial money ($) transferred	-.095	-.366*	-.246	-.297*	-.407**	-.044	-.102	
	.528	*.012*	*.10*	*.045*	*.006*	*.774*	*.499*	
Slot Game control money ($) transferred (fair v unfair slot)	-.121	-.084	.329*	.300*	.136	.191	.548**	.024
	.423	*.58*	*.026*	*.043*	*.371*	*.203*	*<.001*	*.874*

*** $p < .001$, *$p < .05$