Appearance-based trust processing in schizophrenia

Sutherland, C. A. M.,[1,2*] Rhodes, G.,[1] Williams, N.,[3] Connaughton, E.,[3] Ewing, L.,[4] Caruana, N.,[3] Langdon, R.[3]

[1] School of Psychological Science, University of Western Australia, 35 Stirling Hwy, Crawley, 6009, WA, Australia.

[2] School of Psychology, University of Aberdeen, Aberdeen, AB24 3FX, UK.

[3] Department of Cognitive Science, 16 University Avenue, Macquarie University, NSW 2109, Australia.

[4] School of Psychology, University of East Anglia, Norwich Research Park, NR4 7TJ, UK.

*Requests for reprints should be addressed to Clare Sutherland (clare.sutherland@abdn.ac.uk)

Abstract

Objectives. Schizophrenia is characterised by impaired social interactions and altered trust. In the general population, trust is often based on facial appearance, with limited validity but enormous social consequences. The aim was to examine trust processing in schizophrenia, and specifically to examine how people with schizophrenia use facial appearance as well as actual partner fairness to guide trusting decisions.

Design. An experimental economic game study.

Methods. Here we tested how schizophrenia patients and control participants (each N = 24) use facial trustworthiness appearance and partner fairness behaviour to guide decisions in a multi-round Trust Game. In the Trust Game, participants lent money to 'partners' whose facial appearance was either untrustworthy or trustworthy, and who either played fairly or unfairly. Clinical symptoms were measured as well as explicit trustworthiness impressions.

Results. Overall, the schizophrenia patients showed unimpaired explicit facial trustworthiness impressions and unimpaired facial appearance biases in the Trust Game. Crucially, patients and controls significantly differed so that the schizophrenia patients did not learn to discriminate in the Trust Game based on actual partner fairness, unlike control participants.

Conclusion. A failure to discriminate trust has important implications for everyday functioning in schizophrenia, as forming accurate trustworthiness beliefs is an essential social skill. Critically, without relying on more valid trust cues, people with schizophrenia may be especially susceptible to the misleading effect of appearance when making trusting decisions.

Keywords: schizophrenia, psychosis, facial impressions, trust behaviour, facial trustworthiness

Practitioner Points

- People with schizophrenia made very similar facial trustworthiness impressions to healthy controls and also used facial appearance to guide trust decisions similarly to controls.

- However, the patient group were less able to explicitly distinguish between fair and unfair partners based on their behaviour compared to the control group.

- Moreover, people with schizophrenia failed to use actual partner fairness to guide their financial decisions in the Trust Game, unlike controls, and this impairment was specific to a social task.

- People with schizophrenia may be particularly reliant on facial appearance when trusting others, as they may struggle to incorporate more valid trustworthiness information in their decision making, such as actual partner fairness.

Appearance-based trust processing in schizophrenia

Schizophrenia is characterised by impaired social interactions and altered trust (Brüne & Wischniewski, 2011; Freeman, 2007), particularly for patients with paranoia (Gromann et al., 2013). These differences manifest early, are particularly resistant to intervention and are especially predictive of everyday functioning (Chan & Chen, 2011; Green, Horan, & Lee, 2015). Moreover, the association between trust and patient symptomology is not always straightforward, requiring further investigation (McIntosh & Park, 2014; Prevost, Brodeur, Onishi, Lepage, & Gold, 2015). Investigating trust impairment in schizophrenia is therefore vital to both understanding and treatment of the disorder.

A new approach to investigating trust in schizophrenia is to employ economic games, which operationalise trust through interactive financial lending (Brüne & Wischniewski, 2011; Chan & Chen, 2011). Although underutilised in psychiatry research, economic games have the benefit of measuring trust unobtrusively, are engaging, and have real-life applicability. Moreover, these games may be especially sensitive to social impairment as they are interactive and occur in real time (Chan & Chen, 2011; Fett et al., 2012).

In economic game studies, individuals with schizophrenia or high schizotypy often show increased mistrust, indicated by lower financial lending (Fett et al., 2012, 2016; Gromann et al., 2013), although not always (van't Wout & Sanfey, 2011), and sometimes patients are hyper-fair (Agay, Kron, Carmel, Mendlovic, & Levkowitz, 2008; Wischniewski & Brüne, 2011). More recently, two studies have examined trust over repeated interactions, finding that individuals with schizophrenia fail to learn partner trustworthiness (Fett et al., 2012, 2016).

However, these economic game studies have largely focused on explicit, top-down cues to trustworthiness, such as partner behaviour or reputational information (e.g. Fett et al., 2012). In doing so, studies have neglected bottom-up, implicit trust cues, like partner facial

appearance. Trust behaviour based on facial appearance is widespread in the general population (Olivola, Funk, & Todorov, 2014), although facial appearance is not a highly accurate trustworthiness cue (Rule, Krendl, Ivcevic, & Ambady, 2013). Trustworthy-looking partners are given more money than untrustworthy-looking partners on real financial websites (Duarte, Siegel, & Young, 2012) and in lab-based trust games (Chang, Doll, van't Wout, Frank, & Sanfey, 2010; Ewing, Caulfield, Read, & Rhodes, 2015). It is unclear if people with schizophrenia similarly use facial appearance to guide their trust decisions.

Current study

The overall aim was to examine trust processing in schizophrenia, and specifically to examine how people with schizophrenia use facial appearance and actual fairness to guide trusting decisions. We measured trusting behaviour using an economic Trust Game where decisions could be based on partner facial appearance and/or behaviour. In the Trust Game (based on Chang et al., 2010; Ewing et al., 2015), trusting decisions were indexed by money transferred to four 'partners', who looked untrustworthy or trustworthy, and played fairly (returned an equal share of invested money) or unfairly (kept all the money). Participants played multiple rounds of the Trust Game with the same partners. This design allowed us to investigate whether people with schizophrenia differed from controls in their reliance on facial appearance as well as actual partner fairness in guiding trust.

Regarding appearance-based trust, social behaviour in schizophrenia and autism is often similar (Frith & Johnstone, 2003), and autistic boys are less guided by facial appearance in trust games than non-autistic boys (Ewing et al., 2015). Following the pattern in autism, we would predict that people with schizophrenia are less guided by appearance than controls. Yet, to the extent that people with schizophrenia may form more extreme impressions from facial appearance than controls (Trémeau et al., 2016), this account would

instead predict that people with schizophrenia are relatively *more* guided by appearance. We tested these opposing hypotheses here. We also tested explicit trustworthiness impressions based on facial appearance, to inform the Trust Game results. We did not have a strong prediction for explicit facial impressions as evidence is mixed (Hall et al., 2004, 2010; Haut & MacDonald III, 2010; Marwick & Hall, 2008; McIntosh & Park, 2014; van't Wout et al., 2007).

Regarding experience-based trust, we predicted that the participants with schizophrenia would be less influenced than controls by partner actual fairness when making trusting decisions (following Fett et al., 2012, 2016). Capacity to learn partner fairness is likely underpinned by ability to utilise experience when reasoning, and theory of mind (ToM), skills that are impaired in schizophrenia (Langdon, 2005; Woodward, Moritz, Cuttler, & Whitman, 2006). As a supplementary analysis (see supporting information), we therefore measured cognitive reasoning (Moritz & Woodward, 2005; Woodward et al., 2006) and ToM (Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001; Langdon, 2005) to ascertain any links between these capacities and trust behaviour. Finally, we included a version of the Trust Game with slot machines rather than partners, to control for non-social aspects of gameplay (Chang et al., 2010) in order to understand if any impairment was specific to a social interaction context.

Methods

Participants

We recruited 24 individuals with schizophrenia or schizoaffective disorder from volunteer registers in the [blinded for review]. We recruited 24 healthy controls from the general community. Exclusion criteria for all participants included history of brain injury or other neurological disorder, or current or previous persistent substance abuse. Controls with a

personal or familial history of psychotic disorder were also excluded. Clinical participants had a DSM5 diagnosis of schizophrenia (N = 17) or schizoaffective disorder (N = 7) (American Psychiatric Association, 2013). Diagnosis was confirmed using the Diagnostic Interview for Psychosis (Castle, Jablensky, & McGrath, 2006) and clinical history. Current symptoms were rated using the Scales for Assessing Positive and Negative Symptoms of Schizophrenia (SAPS/SANS: Andreasen, 1983, 1984). Thirteen clinical participants had current delusions and eleven did not. We also measured general proneness to delusional ideation using the Peters et al. Delusions Inventory (PDI, Peters, Joseph, Day, & Garety, 2004), and recorded the persecution subscore (Verdoux et al., 1998) given the relevance of paranoia for trust (Freeman, 2007). All clinical participants were on medication at the time of testing. Of the 24 clinical participants taking antipsychotic medication, two were taking typical antipsychotic medication only, 20 were taking atypical antipsychotic medication only, and two were taking a combination of typical and atypical medications. Additionally, 14 clinical participants were also taking an antidepressant.

Clinical and control groups did not differ in premorbid intelligence (on the National Adult Reading Test: NART), current intelligence (on the Matrix Reasoning component of the Weschler Adult Intelligence Scale) or age at testing (Table 1). Participants gave written informed consent to procedures approved by the Human Research Ethics Committee of [blinded for review].

Table 1 about here

Tasks

Facial impressions: Participants rated the trustworthiness of 40 faces in a pseudo-

randomised order[1] (after two practice images) using a 9-point scale (1 = not at all trustworthy, 9 = very trustworthy). We measured mean trustworthiness impressions and agreement with consensus trustworthiness impressions (from Hooper et al., 2018). Face images were taken from the US10K (Bainbridge, Isola, & Oliva, 2013) and restricted to Caucasian females to avoid stereotyping (Hooper et al., 2018).

Trust Game: Participants played a multi-turn Trust Game previously used with students (Hooper et al., 2018; see **Figure 1**). After two practice trials with cartoon partners, participants played eight turns with four virtual 'partners' (32 trials). On each turn, participants received $10 of virtual money, then viewed their partner's face for 3.5s and were required to invest between $0 and $10 with their partner. Any investment was quadrupled. Partners returned an equal split of the money (fair behaviour) or kept the money (unfair behaviour). Thus, if the partner was fair, the participant doubled their investment, and if unfair, the participant lost their investment. Participants additionally kept any amount not invested. If participants took longer than 8s to make an offer, they forfeited all their money on that turn (seven and eight turns across both games for clinical and control participants respectively). These trials were treated as missing data to ensure money transferred was not confounded by impulsivity.

Participants viewed one of two possible sets of partners, which were counterbalanced across participants. Partners were Caucasian females only, to avoid stereotyping, and were taken from the US10K (Bainbridge et al., 2013). Two partners were trustworthy-looking and two were untrustworthy-looking (Hooper et al., 2018).

Assignment of partner behaviour was crossed with facial appearance and counterbalanced across participants. Fair partners returned money to participants on 7/8

---

[1] Due to experimenter error, two patients and two controls were shown the images in a different order.

turns; unfair partners returned money on 1/8 turns. Behaviour was probabilistic to prevent the game from being too easy and to better replicate real-world interaction (Chang et al., 2010).

Figure 1 about here

Slot Machine game: We included a non-social version of the Trust Game, which was identical except that participants played with four coloured slot machines. Including this game allowed us to control for task-specific aspects of the Trust Game (e.g. reward sensitivity).

Explicit beliefs: At the end of the Trust Game and Slot Machine tasks, we examined participants' beliefs about how trustworthy their partners looked, as well as how fairly the partners or slot machines behaved. Participants 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 or slot machines on their fairness levels, defined as the percentage of times their partner or the slot machine returned money (0 = 0% to 9 = 90% in 10% increments).

Additional measures: we also measured cognitive biases and theory of mind (see **supporting information** for more details).

General Procedure

Computer-based tasks were performed on a Dell PC running Windows 7. The facial trustworthiness task was completed after the trust games to avoid priming participants. No deception was used and participants were fully debriefed afterwards. Testing sessions lasted 2-3 hours.

Analysis

9

We first tested for a clinical difference in explicit trustworthiness facial impressions by correlating each of the clinical and control groups' impressions with independent consensus impressions. We statistically compared the clinical and control groups' correlations using a Fisher z test.

In order to understand if there was a clinical difference in trusting decisions based on partners' appearance and/or actual fairness over time, we then analysed trusting decisions in the trust game (money transferred) with a four-way mixed ANOVA. The ANOVA had Group (clinical or control) as a between-subjects factor, and Partner Appearance (trustworthy or untrustworthy), Partner Behaviour (fair or unfair) and Time (first or last block) as within-subjects factors. The four-way interaction was not significant, thus we ran a three-way ANOVA for Group, Appearance, and Time after collapsing over Behaviour to examine the effect of Appearance. We then ran a three-way ANOVA for Group, Behaviour, and Time after collapsing over Appearance to examine the effect of Behaviour. We ran an identical ANOVA for the control Slot Machine task, but without the Appearance factor.

Finally, in order to understand if there was a clinical difference in explicit appearance and fairness beliefs, we ran two-way mixed ANOVAs on the explicit appearance and fairness ratings of the partners in the Trust and Slot Machine Games. These ANOVAs had Group (clinical or control) as a between-subjects factor and either Partner Appearance (trustworthy or untrustworthy) or Partner or Slot Machine Behaviour (fair or unfair) as within-subjects factors.

Across all analyses, we followed up any significant effects for Appearance or Behaviour using paired t-tests to compare trustworthy-looking versus untrustworthy-looking partners or fair versus unfair partners respectively. Where the clinical Group factor interacted with Appearance or Behaviour conditions, we ran separate paired t-tests for clinical and control groups separately in order to understand the pattern for each group. Where the Time

factor interacted, we ran separate paired t-tests for first and last blocks separately in order to understand the pattern at each time point.

<div align="center">

Results

</div>

Explicit Facial Trustworthiness Impressions

We first examined how people with schizophrenia made explicit trust impressions from faces, in order to contextualise the Trust Game results.

Main facial impressions task: One participant in each group was excluded from the facial impressions agreement analysis because they failed to discriminate between the faces. Both groups showed significant agreement with consensus facial trustworthiness judgements: clinical mean $r = .31$, SD $r = .20$, $t(22) = 7.06$, $p < .001$, control mean $r = .27$, SD $r = .16$, $t(22) = 7.72$, $p < .001$, with no significant group difference; $t(44) = 0.96$, $p = .343$, $d = 0.28$ (all correlations reflect Pearson's r; t-test after Fisher-transformation). We also did not find a difference in mean impressions: clinical mean = 5.31, SD = 1.12; control mean = 5.03, SD = 0.90; $t(46) = -0.97$, $p = .336$, $d = 0.28$ or in impression variance, $t(46) = 0.48$, $p = 0.492$. All d values (here and elsewhere) reflect Cohen's d for the group difference.

Facial impressions of Trust Game partners: As a manipulation check, participants rated the facial trustworthiness of their partners after the Trust Game. We analysed these explicit impressions in a two-way mixed ANOVA with Group (clinical versus control) as a between-subjects factor and Partner Appearance (Trustworthy versus Untrustworthy) as a within-subjects factor. The main effect of Appearance was significant, indicating that participants perceived the trustworthy-looking partners as more trustworthy (M = 4.8, SD = 2.1) than the untrustworthy-looking partners (M = 3.5, SD = 1.6), as expected: $F(1,46) = 20.30$, $p < .001$, $\eta_p^2 = 0.31$). There was no main effect or interaction with Group: both

$F(1,46) < 0.27$, $p > .60$, $\eta_p{}^2 < 0.01$), thus the clinical group did not show lower overall trust, agreeing with the main trustworthiness impressions task.

Trust Game Decisions

Our main aim was to examine how people with schizophrenia use facial appearance and/or experience with actual fairness to guide trusting decisions. We conducted a four-way mixed ANOVA with Group (clinical versus control) as a between-subjects factor, and Partner Appearance (trustworthy versus untrustworthy), Partner Behaviour (fair or unfair) and Time (first versus last block) as within-subjects factors. We could compare the first and last block directly because these trials were identical. The four-way interaction was not significant: $F(1,42) = 0.18$, p = .37, $\eta_p{}^2 = .02$ (note that four participants had to be excluded from this overall analysis due to missing trials). As the overall four-way interaction was not significant, for simplicity, we examined trust based on appearance and behaviour separately (i.e. after collapsing across behaviour/appearance respectively). These analyses allowed us to use the whole dataset and paralleled the slot machine game (see the supplementary materials for the four-way ANOVA results, which were identical).

Appearance-based Trust Decisions: To test for a group difference in appearance-based trust, we conducted a three-way mixed ANOVA on money transferred in the Trust Game, with Group (clinical versus control) as a between-subjects factor and Partner Appearance (trustworthy versus untrustworthy) and Time (first versus last block) as within-subjects factors (collapsed across Partner Behaviour). Critically, by the last block in the Trust Game, participants should have stopped using facial appearance to guide decision-making, given that it did not accurately cue partner fairness.

Interestingly, the only significant effect was a main effect for Partner Appearance: $F(1,46) = 7.12$, $p = .011$, $\eta_p{}^2 = 0.13$. On average, participants transferred more money to

trustworthy-looking (mean = \$5.34, SD = \$2.17) than untrustworthy-looking partners (mean = \$4.64, SD = \$2.55) in the Trust Game. No other effects were significant: $F(1,46) < 1.36$, $p > .25$, $\eta_p^2 < 0.03$ (Figure 2). In other words, misleading facial appearance was used to guide decision-making, even when other, more objective information (partner fairness) was available.

Figure 2 about here

Experience-based Trust Decisions: To test whether the groups learned to discriminate partner fairness, we conducted a three-way mixed ANOVA on money transferred in the Trust Game, with Group (clinical versus control) as a between-subjects factor and Partner Behaviour (fair versus unfair) and Time (first versus last block) as within-subjects factors (collapsed across Partner Appearance). There was a significant main effect of Partner Behaviour: $F(1,46) = 19.89$, $p < .001$, $\eta_p^2 = 0.30$, and two-way interactions between Group and Partner Behaviour, as well as Time and Partner Behaviour: both $F(1,46) > 8.05$, $p < .007$, $\eta_p^2 > 0.15$ (Figure 3).

Critically, there was also a significant three-way interaction, suggesting that the groups had different responses to partner behaviour over time: $F(1,46) = 6.13$, $p = .017$, $\eta_p^2 = 0.12$ (Figure 3). We ran paired t-tests to follow up on this interaction. As expected, neither the control nor clinical group discriminated based on fair versus unfair partner behaviour in the first block, before partner behaviour was known: control $t(23) = 1.18$, $p = .251$, $d = 0.22$, clinical: $t(23) = 0.51$, $p = .613$, $d = 0.09$. Strikingly, whereas controls did discriminate based on fair versus unfair partner behaviour by the last block: $t(23) = 5.81$, $p < .001$, $d = 1.47$, the clinical group did not: $t(23) = 1.05$, $p = .303$, $d = 0.23$. Thus, unlike controls, the patients with schizophrenia failed to learn to use partner behaviour to guide their trust decisions.

13

Figure 3 about here

Non-social control task

The clinical group struggled to discriminate partner actual fairness in the Trust Game, raising the question of whether this effect is specific to a social task involving human partners, or whether it reflects a more general problem. To answer this question, we examined whether the groups learned to discriminate between rewarding and non-rewarding slot machines over the course of the Slot Game. Thus, we conducted a three-way mixed ANOVA on money transferred, with Group (clinical versus control) as a between-subjects factor and Slot Machine (rewarding versus unrewarding) and Time (first versus last block) as within-subjects factors (Figure 4). There was a significant main effect of Slot Machine: $F(1,46) = 24.55$, $p < .001$, $\eta_p^2 = 0.35$, reflecting more money transferred to the rewarding slot machine (M = \$6.44, SD = \$2.54) versus the unrewarding slot machine (M = \$4.88, SD = \$3.09). There was also a main effect of Group: $F(1,46) = 6.06$, $p = .018$, $\eta_p^2 = 0.12$, reflecting a tendency for the clinical participants to transfer more money on average (M = \$6.28, SD = \$2.95) than controls (M = \$5.04, SD = \$2.79). There was a significant two-way interaction between Slot Machine and Time, reflecting learning which slots were rewarding: $F(1,46) = 20.24$, $p < .001$, $\eta_p^2 = 0.31$. Importantly, the three-way interaction was not significant, $F(1,46) = 0.09$, $p = .771$, $\eta_p^2 = 0.002$. At the start, neither group discriminated between rewarding and unrewarding slot machines: both $t(23) < 1.03$, $p > .314$, $d < 0.12$. By the end, both groups had learnt to discriminate between rewarding and unrewarding slot machines: both $t(23) > 3.54$, $p < .002$, $d > 0.94$ (Figure 4).

Figure 4 about here

Explicit fairness beliefs

Finally, given the suggestion of possible dissociations between explicit and implicit social-cognitive processes in schizophrenia (e.g. Frith, 2004), we investigated whether the clinical group's inability to distinguish partner fairness in the Trust Game reflected an explicit or implicit problem in distinguishing fairness. To measure their explicit fairness beliefs, participants rated the perceived fairness of their partners and slot machines at the end of the games. We analysed the Trust Game fairness ratings in a two-way mixed ANOVA with Group (clinical versus control) as a between-subjects factor and Partner Behaviour (fair versus unfair) as a within-subjects factor. There was a significant main effect of Behaviour, thus participants correctly distinguished fair and unfair partners: $F(1,46) = 38.10$, $p < .001$, $\eta_p^2 = 0.45$; however, this pattern was qualified by a significant interaction with Group: $F(1,46) = 10.07$, $p = .003$, $\eta_p^2 = 0.18$. Both the clinical and control participants correctly judged the fairness of their partners in the Trust Game, although fairness discrimination was weaker in the clinical group (fair M = 5.2, SD = 1.7, unfair M = 4.1, SD = 1.7, $t(23) = 2.45$, $p < .023$, $d = 0.64$), compared to the control group (fair M = 6.2, SD = 1.5, unfair M = 2.9, SD = 1.7, $t(23) = 5.91$, $p < .001$, $d = 2.09$). Thus, the clinical participants were less able to explicitly distinguish fair and unfair partners than controls, mirroring the pattern for money transferred. The main effect of Group was not significant, thus the clinical and control groups did not differ in absolute fairness ratings: $F(1,46) = 0.04$, $p = .842$, $\eta_p^2 < 0.00$.

In the Slot Game, there was only a significant main effect of Behaviour, $F(1,46) = 46.25$, $p < .001$, $\eta_p^2 = 0.50$, with no other significant effects: both $F(1,46) < 0.47$, $p > .50$, $\eta_p^2 < 0.01$. Overall, all participants correctly distinguished rewarding (M = 6.0, SD = 1.7) and unrewarding slot machines (M = 3.3, SD = 1.7) in their fairness ratings. Thus, as for money transferred, the group difference in correctly judging fairness was specific to the social Trust Game.

Discussion

Overall, individuals with schizophrenia and healthy controls formed very similar explicit facial trustworthy impressions. Both groups also relied on facial trustworthy appearance to guide financial decisions in the Trust Game. Crucially, however, the participants with schizophrenia failed to use actual partner fairness to guide their financial decisions in the Trust Game, unlike controls, and this impairment was specific to a social task. Moreover, the clinical group were also less able to explicitly distinguish between fair and unfair partners based on their behaviour compared to the control group. Our findings agree with two recent studies on trust reciprocation, which have also shown impairments in trust updating in schizophrenia and psychosis (Fett et al., 2012, 2016). Critically, here we also show that people with schizophrenia appear to anchor on an initial, potentially misleading trust decision based on appearance, and then, unlike control participants, fail to update their strategy to incorporate more valid information about actual partner trustworthiness.

A failure to learn trustworthiness in schizophrenia

Our Trust Game results are also strikingly similar to cognitive reasoning biases found previously in schizophrenia (Moritz & Woodward, 2005; Woodward et al., 2006), suggesting that reasoning biases may be contributing to social impairment in this disorder. Our distinction between appearance-based and experience-based trust is also consistent with recent theories of low- and high-level processes of predictive coding and their potential disturbance in psychosis (see Sterzer et al., 2018; Sterzer, Voss, Schlagenhauf, & Heinz, 2019 for reviews). In predictive coding accounts of schizophrenia, psychosis may reflect a problem with weaker prior expectations for lower-level visual input (here, appearance), which in turn may cause stronger prior expectations for higher-level beliefs, or conceptual

input (here, learning from experience). This account would likely predict less strong visual appearance biases, which we did not find (nor did other recent predictive coding studies: Kaliuzhna et al., 2018; Palmer, Caruana, Clifford, & Seymour, 2018); nevertheless, it would also predict less updating of conceptual beliefs based on actual experience, which is consistent with our results.

The clinical group did not show a failure in experience-based trust for the control (non-social) slot machine game. Why not? Certainly, there is a wealth of literature showing aberrant non-social reward processing in schizophrenia (see Strauss, Waltz, & Gold, 2014 for a review). Some predictive coding accounts argue that social processing may be particularly disrupted in psychosis as aberrant predictive coding would make other people's intentions seem unreliable (Sterzer et al., 2018). Given that we can never truly verify another's intentions, social cues may be relatively more uncertain than non-social ones (Sterzer et al., 2018). Potentially, forming (and overcoming) an initial impression is also critical: in the social Trust Game, it is possible to base a decision on the initial and misleading evidence of partner facial appearance, in contrast to the slot machine game. Similarly, non-social learning studies have found that habitual decision-making is less disrupted in schizophrenia than flexible decision-making (Culbreth, Westbrook, Daw, Botvinick, & Barch, 2016; Strauss et al., 2014). Finally, it is also possible that the slot machine game may simply have been easier or influenced by practice, given that it was presented after the Trust Game to avoid influencing performance on our measure of key interest. Future research should tease apart these explanations, because successful learning in the slot machine game could provide a useful therapeutic tool if it is indeed easier and training generalises.

Interestingly, there was no evidence of overall higher suspiciousness for the clinical group in the Trust Game. The clinical group transferred as much money as did controls and rated their partners equally in fairness. Previous studies measuring gameplay in schizophrenia

and psychosis have been mixed (Agay et al., 2008; Fett et al., 2012; Gromann et al., 2013; van't Wout & Sanfey, 2011; Wischniewski & Brüne, 2011). Overall, previous studies are heterogeneous in both patient symptoms and gameplay, including incentive structure and live versus pre-programmed gameplay. For this reason, we deliberately chose a game design which has been used in previous work with neurotypical adults (Chang et al., 2010; Hooper et al., 2018). However, future research should systematically test these factors and could also use more complex and/or naturalistic trust game designs.

Unimpaired trustworthiness appearance processing in schizophrenia

The group with schizophrenia made very similar explicit facial impressions to healthy participants and also used facial appearance to guide trust decisions. The finding of unimpaired facial trustworthiness impressions contradicts some previous studies (Hall et al., 2004; Haut & MacDonald III, 2010; Pinkham, Hopfinger, Pelphrey, Piven, & Penn, 2008; Trémeau et al., 2016) but agrees with others (Marwick & Hall, 2008; McIntosh & Park, 2014; Mukherjee et al., 2014). Our study used naturalistic face images, which have greater ecological validity and likely preserve more trustworthiness cues relative to lab-based images used in some studies (e.g. Pinkham et al., 2008). Moreover, even studies that find clinical impairments observe stronger differences for judgements other than trust (e.g., intelligence: Hall et al., 2004), suggesting that any such differences are not specific to trust per se. Interestingly, the pattern here also contrasts with the pattern found in autism (Ewing et al., 2015) and for men with high autistic traits (Hooper et al., 2018), who showed less appearance-based trust decisions but similar experience-based trust compared to controls, suggesting important social processing differences between schizophrenia and autism.

Future research

Future research should test whether the interpersonal trust issues found here are also present at other stages of illness (e.g., early psychosis) and in other clinical samples (e.g.,

more functionally impaired groups). Effects on everyday social functioning and implications for treatment likewise warrant further research. A promising future direction will be to establish whether training on economic games can improve everyday social interaction problems in schizophrenia (Chan & Chen, 2011). Anecdotally, these games show promise as a therapeutic tool, as the participants with schizophrenia reported enjoying the tasks, which did not take long (15 mins).

Conclusions

Facial appearance guides trust across a wide variety of social contexts, although it is often misleading (Olivola et al., 2014). Critically, people with schizophrenia may be particularly subject to the misleading effect of appearance when making everyday trusting decisions, as they may struggle to incorporate more valid information in their decision-making, such as the actual fairness of their social partners.

References

Agay, N., Kron, S., Carmel, Z., Mendlovic, S., & Levkowitz, Y. (2008). Ultimatum bargaining behavior of people affected by schizophrenia. *Psychiatry Research*, *157*(1–3), 39–46. https://doi.org/10.1016/j.psychres.2006.03.026

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

Andreasen, N. C. (1983). *The Scale for the Assessment of Negative Symptoms (SANS).* Iowa City, IA: The University of Iowa.

Andreasen, N. C. (1984). *The Scale for the Assessment of Positive Symptoms (SAPS).* Iowa City, IA: The University of Iowa.

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

Brüne, M., & Wischniewski, J. (2011). Complex social cognition and the appreciation of

    social norms in psychiatric disorders: Insights from evolutionary game theory. In R.

    Ebstein, S. Shamay-Tsoory, & S.-H. Chew (Eds.), *From DNA to Social Cognition*

    (pp. 215–231). https://doi.org/10.1002/9781118101803.ch13

Castle, D. J., Jablensky, A., & McGrath, J. J. (2006). The diagnostic interview for psychoses

    (DIP): development, reliability and applications. *Psychological Medicine*, *36*, 69–80.

    https://doi.org/10.1017/S0033291705005969

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

Culbreth, A. J., Westbrook, A., Daw, N. D., Botvinick, M., & Barch, D. M. (2016). Reduced

    model-based decision making in schizophrenia. *Journal of Abnormal Psychology*,

    125(6), 777–787. https://doi.org/doi:10.1037/abn0000164

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

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

Fett, A.-K. J., Shergill, S. S., Joyce, D. W., Riedl, A., Strobel, M., Gromann, P. M., & Krabbendam, L. (2012). To trust or not to trust: the dynamics of social interaction in psychosis. *Brain*, *135*(3), 976–984. https://doi.org/10.1093/brain/awr359

Fett, A.-K. J., Shergill, S. S., Korver-Nieberg, N., Yakub, F., Gromann, P. M., & Krabbendam, L. (2016). Learning to trust: trust and attachment in early psychosis. *Psychological Medicine*, *46*(7), 1437–1447. https://doi.org/10.1017/S0033291716000015

Freeman, D. (2007). Suspicious minds: the psychology of persecutory delusions. *Clinical Psychology Review*, *27*(4), 425–457. https://doi.org/10.1016/j.cpr.2006.10.004

Frith, C. (2004). Schizophrenia and theory of mind. *Psychological Medicine*, *34*(3), 385–389. https://doi.org/doi:10.1017/S0033291703001326

Frith, C., & Johnstone, E. C. (2003). *Schizophrenia: A very short introduction*. OUP Oxford.

Green, M. F., Horan, W. P., & Lee, J. (2015). Social cognition in schizophrenia. *Nature Reviews Neuroscience*, *16*(10), 620–631. https://doi.org/10.1038/nrn4005

Gromann, P. M., Heslenfeld, D. J., Fett, A.-K., Joyce, D. W., Shergill, S. S., & Krabbendam, L. (2013). Trust versus paranoia: abnormal response to social reward in psychotic illness. *Brain: A Journal of Neurology*, *136*(Pt 6), 1968–1975. https://doi.org/10.1093/brain/awt076

Hall, J., Harris, J. M., Sprengelmeyer, R., Sprengelmeyer, A., Young, A. W., Santos, I. M.,

    … Lawrie, S. M. (2004). Social cognition and face processing in schizophrenia. *The*

    British Journal of Psychiatry, *185*(2), 169–170. https://doi.org/10.1192/bjp.185.2.169

Hall, J., Whalley, H. C., McKirdy, J. W., Sprengelmeyer, R., Santos, I. M., Donaldson, D. I.,

    … others. (2010). A common neural system mediating two different forms of social

    judgement. *Psychological Medicine*, *40*(07), 1183–1192.

    https://doi.org/10.1017/S0033291709991395

Haut, K. M., & MacDonald III, A. W. (2010). Persecutory delusions and the perception of

    trustworthiness in unfamiliar faces in schizophrenia. *Psychiatry Research*, *178*(3),

    456–460. https://doi.org/10.1016/j.psychres.2010.04.015

Hooper, J. J., Sutherland, C. A. M., Ewing, L., Langdon, R., Caruana, N., Connaughton, E.,

    … Rhodes, G. (2018). Should I trust you? Autistic traits predict reduced appearance-

    based trust decisions. *British Journal of Psychology*.

    https://doi.org/10.1111/bjop.12357

Kaliuzhna, M., Stein, T., Rusch, T., Sekutowicz, M., Sterzer, P., & Seymour, K. J. (2018).

    No evidence for abnormal priors in early vision in schizophrenia. *Schizophrenia*

    Research. https://doi.org/10.1016/j.schres.2018.12.027

Langdon, R. (2005). Theory of mind in schizophrenia. In B. F. Malle & S. D. Hodges (Eds.),

    Other minds: How humans bridge the divide between self and others (pp. 323–342).

    New York: The Guilford Press.

Marwick, K., & Hall, J. (2008). Social cognition in schizophrenia: a review of face

    processing. *British Medical Bulletin*, *88*(1), 43–58.

    https://doi.org/10.1093/bmb/ldn035

McIntosh, L. G., & Park, S. (2014). Social trait judgement and affect recognition from static facea and video vignettes in schizophrenia. *Schizophrenia Research*, *158*, 170–175. https://doi.org/10.1016/j.schres.2014.06.026

Moritz, S., & Woodward, T. S. (2005). Jumping to conclusions in delusional and non-delusional schizophrenic patients. *British Journal of Clinical Psychology*, *44*(2), 193–207. https://doi.org/10.1348/014466505X35678

Mukherjee, P., Whalley, H. C., McKirdy, J. W., Sprengelmeyer, R., Young, A. W., McIntosh, A. M., … Hall, J. (2014). Altered amygdala connectivity within the social brain in schizophrenia. *Schizophrenia Bulletin*, *40*(1), 152–160. https://doi.org/10.1093/schbul/sbt086

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

Palmer, C. J., Caruana, N., Clifford, C. W. G., & Seymour, K. J. (2018). Adaptive sensory coding of gaze direction in schizophrenia. *Royal Society Open Science*, *5*(12), 180886. https://doi.org/10.1098/rsos.180886

Peters, E. R., Joseph, S. A., Day, S., & Garety, P. A. (2004). Measuring delusional ideation: the 21-item Peters et al. Delusions Inventory (PDI). *Schizophrenia Bulletin*, *30*, 1005–1022. http://dx.doi.org/10.1093/oxfordjournals.schbul.a007116

Pinkham, A. E., Hopfinger, J. B., Pelphrey, K. A., Piven, J., & Penn, D. L. (2008). Neural bases for impaired social cognition in schizophrenia and autism spectrum disorders. *Schizophrenia Research*, *99*(1–3), 164–175. https://doi.org/10.1016/j.schres.2007.10.024

Prevost, M., Brodeur, M., Onishi, K. H., Lepage, M., & Gold, I. (2015). Judging strangers'

    rrustworthiness is associated with theory of mind skills. *Frontiers in Psychiatry*,

    6(52), 1–6. https://doi.org/10.3389/fpsyt.2015.00052

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

Sterzer, P., Adams, R. A., Fletcher, P., Frith, C., Lawrie, S. M., Muckli, L., … Corlett, P. R.

    (2018). The predictive coding account of psychosis. *Biological Psychiatry*, *84*(9),

    634–643. https://doi.org/10.1016/j.biopsych.2018.05.015

Sterzer, P., Voss, M., Schlagenhauf, F., & Heinz, A. (2019). Decision-making in

    schizophrenia: A predictive-coding perspective. *NeuroImage*, *190*, 133–143.

    https://doi.org/10.1016/j.neuroimage.2018.05.074

Strauss, G. P., Waltz, J. A., & Gold, J. M. (2014). A review of reward processing and

    motivational impairment in schizophrenia. *Schizophrenia Bulletin*, *40*(2), S107–S116.

    https://doi.org/doi:10.1093/schbul/sbt197

Trémeau, F., Antonius, D., Todorov, A., Rebani, Y., Ferrari, K., Lee, S. H., … others. (2016).

    What can the study of first impressions tell us about attitudinal ambivalence and

    paranoia in schizophrenia? *Psychiatry Research*, *238*, 86–92. https://doi.org/doi:

    10.1016/j.psychres.2016.02.014

van't Wout, M., Aleman, A., Kessels, R. P. C., Cahn, W., de Haan, E. H. F., & Kahn, R. S.

    (2007). Exploring the nature of facial affect processing deficits in schizophrenia.

    Psychiatry Research, *150*(3), 227–235.

    https://doi.org/10.1016/j.psychres.2006.03.010

van't Wout, M., & Sanfey, A. G. (2011). Interactive decision-making in people with

      schizotypal traits: A game theory approach. *Psychiatry Research*, *185*(1), 92–96.

      https://doi.org/10.1016/j.psychres.2010.05.013

Verdoux, H., van Os, J., Maurice-Tison, S., Gay, B., Salamon, R., & Bourgeois, M. (1998). Is

      early adulthood a critical developmental stage for psychosis proneness? A survey of

      delusional ideation in normal subjects. *Schizophrenia Research*, *29*, 247–254.

      https://doi.org/10.1016/S0920-9964(97)00095-9

Wischniewski, J., & Brüne, M. (2011). Moral reasoning in schizophrenia: an explorative

      study into economic decision making. *Cognitive Neuropsychiatry*, *16*(4), 348–363.

      https://doi.org/10.1080/13546805.2010.539919

Woodward, T. S., Moritz, S., Cuttler, C., & Whitman, J. C. (2006). The contribution of a

      cognitive bias against disconfirmatory evidence (BADE) to delusions in

      schizophrenia. *Journal of Clinical and Experimental Neuropsychology*, *28*(4), 605–

   617. https://doi.org/10.1080/13803390590949511
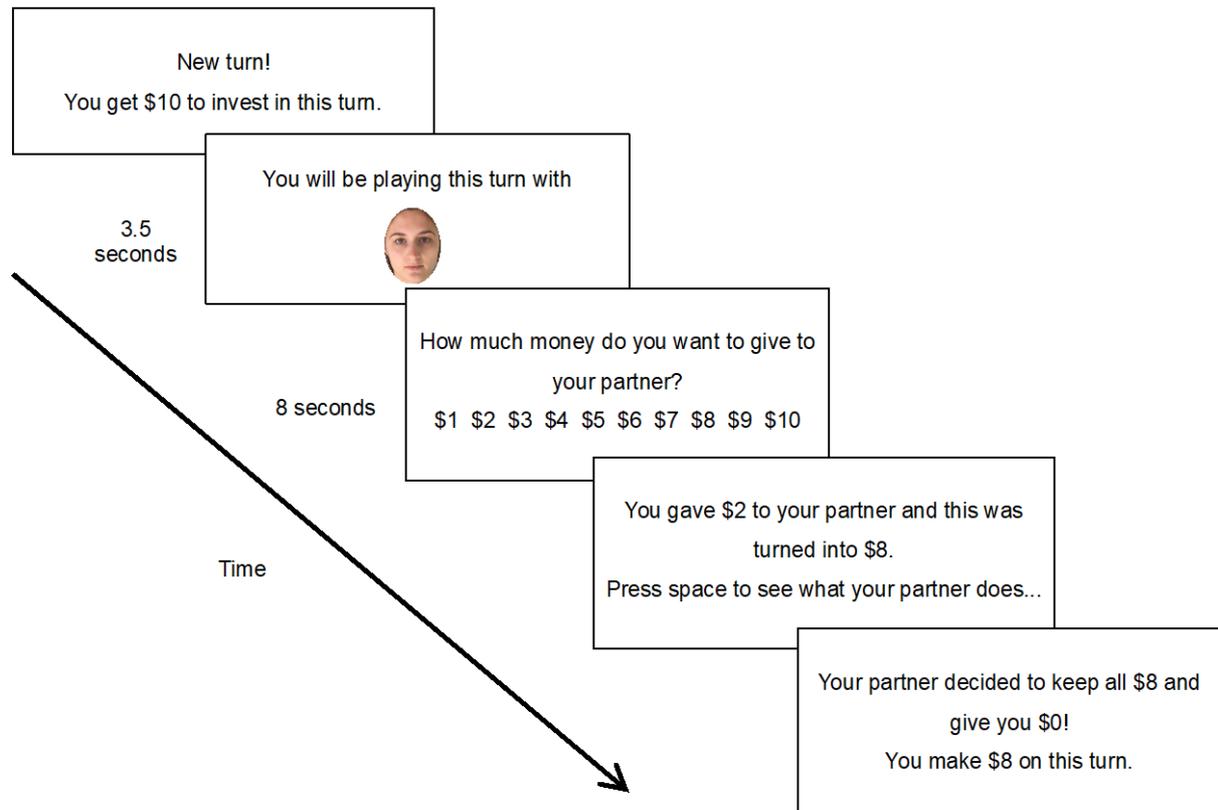
Figures

Figure 1. Example of a single Trust Game turn. Note. Instead of transferring money, participants could also choose to transfer no money by pressing a key labelled "no money". The face shown in the game here is an example due to copyright reasons and was not one of the faces used in the Trust Game, although it is representative. Figure previously published in Hooper et al., British Journal of Psychology, 2018.
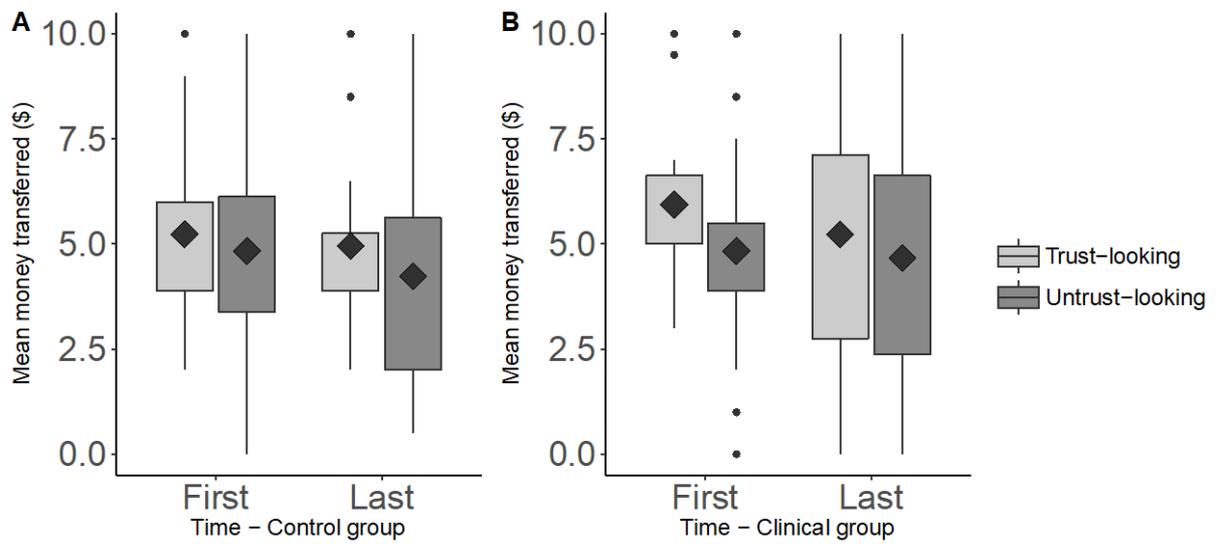
Figure 2. Trust Game and partner appearance. Money transferred in the Trust Game on average for A) control (left hand panel) and B) clinical participants (right hand panel), based on partner trustworthy (light grey) or untrustworthy (dark grey) appearance in the first and last blocks of the Trust Game. Figures depict boxplots with the mean superimposed. The only significant effect was of partner appearance, so that more money was transferred to trustworthy-looking than untrustworthy-looking partners.
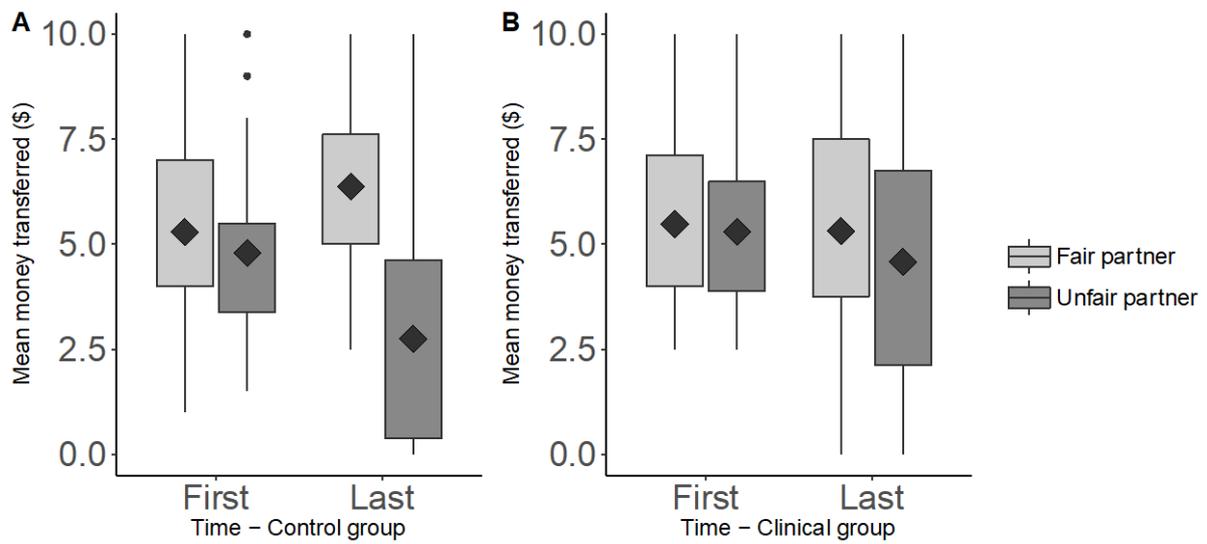
Figure 3. Trust Game and partner fairness. Money transferred in the Trust Game on average for A) control (left hand panel) and B) clinical participants (right hand panel), based on fair (light grey) or unfair (dark grey) partner behaviour in the first and last blocks of the Trust Game. Figures are boxplots with the mean superimposed. In the first block, participants have not experienced the unfair or fair behaviour of their partners, thus neither control nor clinical participants discriminate based on partner behaviour. By the last block, control participants have successfully learnt which partners are trustworthy, as they discriminate based on fair or unfair partner behaviour. Crucially, participants with schizophrenia fail to learn which partners are trustworthy.
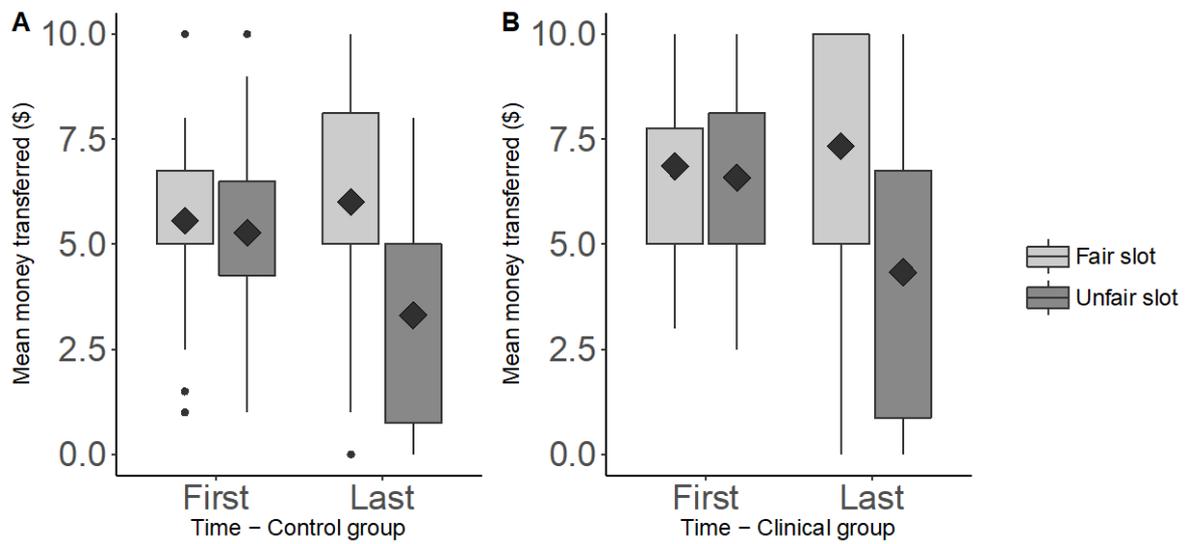
Figure 4. Slot machine game. Money transferred in the Slot Machine Game on average for A) control (left hand panel) and B) clinical participants (right hand panel), based on fair (light grey) or unfair (dark grey) partner behaviour in the first and last blocks of the Slot Machine Game. Figures are boxplots with the mean superimposed. In the first block, participants have not experienced the unfair or fair behaviour of their partners, thus neither control nor clinical participants discriminate based on partner behaviour. By the last block, clinical and control groups have successfully learnt which slot machines are fair or unfair.

Table 1

Mean and SD for demographic and IQ variables for clinical and control groups, and clinical

symptomology for the clinical group. Effect sizes are Cohen's *d* values for the group

difference.

	Clinical group	Healthy control group	Significance Test
N	24	24	-
Males:females	14:10	14:10	-
Age at testing (years)	51.9 (8.9)	45.6 (13.8)	$t(39.4) = 1.88, p = .068, d = 0.5$
NART full score	108.0 (9.7)	106.5 (10.1)	$t(46) = 0.51, p = .612, d = 0.2$
WASI matrix reasoning (raw score)	17.6 (4.3)	19.7 (4.8)	$t(46) = 1.60, p = .116, d = 0.5$
WASI matrix reasoning (scaled score)	9.3 (3.3)	10.6 (3.2)	$t(46) = 1.36, p = .179, d = 0.4$
Age at diagnosis (years)	25.7 (9.1)	-	-
SAPS global (mean)	1.2 (0.7)	-	-
SANS global (mean)	2.4 (0.6)	-	-
PDI score (sum of paranoia items: Q1, Q3-5)	2.5 (1.3)	-	-