Social Cognitive and Affective Neuroscience, 2019, 861–870

doi: 10.1093/scan/nsz058
Advance Access Publication Date: 11 September 2019
Original article

The neural mechanisms of social reward in early psychosis

Anne-Kathrin J. Fett,[1,2,3,] Elias Mouchlianitis,[2] Paula M. Gromann,[1] Lucy Vanes,[2,3] Sukhi S. Shergill,[2,†] and Lydia Krabbendam[4,†]

[1]Department of Psychology, City University of London, London ECIV 0HB, United Kingdom, [2]King's College London, Institute of Psychiatry, Psychology and Neuroscience, Department of Psychosis Studies, De Crespigny Park, SE5 8AF, London, United Kingdom, [3]Wellcome Centre for Human Neuroimaging, Institute of Neurology, University College London, 12 Queen Square, London, United Kingdom and [4]Department of Developmental and Clinical Psychology, Faculty of Behavioural and Movement Sciences, VU Amsterdam, Van der Boechorststraat 1s, 1081 BT Amsterdam, the Netherlands

Correspondence should be addressed to A.-K.J. Fett, Department of Psychology, City University of London, London ECIV 0HB, United Kingdom. E-mail address: anne-kathrin.fett@city.ac.uk

Abstract

In chronic psychosis, reduced trust is associated with a neural insensitivity to social reward and reduced theory of mind (ToM). Here we investigate whether these mechanisms could underlie emerging social impairments in early psychosis. Twenty-two participants with early psychosis and 25 controls (male, 13–19 years) participated in two interactive trust games against a cooperative and unfair partner. Region of interest neuroimaging analyses included right caudate, medial prefrontal cortex (mPFC) and right temporoparietal junction (rTPJ), involved in reward and ToM processing. Both groups showed similar levels of trust (i.e. investments). However, individuals with psychosis failed to activate the caudate differentially in response to cooperation and unfairness while making decisions to trust. During cooperative returns, patients showed reduced and controls increased caudate activation. Patients demonstrated greater rTPJ activation than controls, possibly pointing towards compensatory mechanisms. Effects were associated with Wechsler Abbreviated Scale of Intelligence vocabulary scores. No group differences emerged in mPFC activation. Early psychosis is associated with an aberrant neural sensitivity to social reward. This could foster reduced social motivation and social isolation. Absent behavioural differences in early, relative to chronic psychosis could indicate that trust is achieved through increased compensatory demand on ToM.

Key words: adolescence; early psychosis; fMRI; neuroeconomics; social cognition; trust

The rewarding nature of social contact drives human social behaviour (Krach *et al.*, 2010). The sensitivity to others' social signals is fundamental in understanding their behaviour in social interactions. Theory of mind (ToM), i.e. the ability to take another person's perspective into account, and social reward sensitivity, i.e. the ability to process positive or negative behavioural cues from others, enable us to build a mental model of them during our social encounters. Both social cognitive processes are therefore fundamental in social relationships, and impairment can lead to problematic social interactions.

Deficits in ToM and aberrant dopamine function (Kapur *et al.*, 2005), which is closely linked to reward processing, have been

Received: 25 June 2018; **Revised:** 8 May 2019; **Accepted:** 23 July 2019

suggested to underlie paranoia and social disconnection in psychosis (Kapur *et al.*, 2005; Eisenberger & Cole, 2012). Supporting evidence for a direct association between aberrant social decision making and reward processing mechanisms comes from research with the interactive trust game (Berg *et al.*, 1995). In this paradigm, the first player (investor) receives an endowment that he can share with the second player (trustee). The amount is tripled, and the trustee then decides whether to return a share of this overall amount or not. Cooperation yields the best payoff for both players, but initially the best payoff for the trustee occurs through not cooperating. For successful social interactions, it is crucial that intentions and goals of others are represented to optimise the mutual interaction (Yoshida *et al.*, 2008). Theory of mind is necessary to anticipate the effects of one's own trusting decisions and to decipher the trustee's response, and better ToM abilities have been associated with a superior ability to adapt one's own decisions to the decision-making strategy of the trustee (Fett *et al.*, 2014b). The ability to learn from the trustee's response through social reward or punishment is vital to establish whether trust pays off.

In the trust game, individuals with chronic nonaffective psychosis exhibit lower trust towards others than controls. Lower trust is associated with paranoid delusions, and the evidence suggests that low trust is maintained by reduced sensitivity to positive social information and reduced sensitivity to the game partner's actual trustworthy behaviour. In support of dysfunctional social reward and ToM processing as underlying mechanisms of reduced trust, neuroimaging research has associated the loss of trust with reduced activation of the right caudate nucleus, a key area of reward processing (King-Casas *et al.*, 2005; Phan *et al.*, 2010; Fett *et al.*, 2012; Bhanji & Delgado, 2014) and the right temporoparietal junction (rTPJ), which underlies ToM (Gromann *et al.*, 2013). Yet, the authors also reported normal activation patterns in the medial prefrontal cortex (mPFC), another area that has been related to ToM (Sanfey, 2007; Fett *et al.*, 2015; Krueger & Meyer-Lindenberg, 2018; Porcelli *et al.*, 2018). The aberrant neural mechanisms might lead to social impairment; however, they could also be secondary to exposure to other social or environmental factors that are associated with chronicity of illness. Typically, social impairments seem to emerge early in psychotic disorders, and adolescence is a crucial period for these changes (Velthorst *et al.*, 2016). Social cooperation and ToM continue to develop during this stage (Blakemore, 2008; Fett *et al.*, 2014a), emphasising its importance as a sensitive period for interventions that aim to tackle social difficulties at their roots.

Previously, early psychosis has been associated with reduced trust towards others. However, in an important difference from patients with a long-standing illness, patients in the early phase of the illness have been found able to overcome initial distrust during repeated social interactions with trustworthy, cooperative others. This finding shows that the capacity to build trust through positive experiences with others is retained (Fett *et al.*, 2016; Lemmers-Jansen *et al.*, 2018) and could suggest an intact sensitivity to social reward. Alternatively, other compensatory cognitive mechanisms may be operating to counteract deficits in trust that are due to impaired social reward sensitivity (Brüne *et al.*, 2011). One possible cognitive mechanism is ToM. Some studies show that ToM is still relatively preserved in adolescents in the early stages of the psychotic illness (Achim *et al.*, 2012; Korver-Nieberg *et al.*, 2013; Ho *et al.*, 2015; Canty *et al.*, 2017; Bartholomeusz *et al.*, 2018). Despite this, the findings of Bartholomeusz *et al.* (2018) on neural processing differences in regions related to ToM suggest that subtle changes in neural mechanisms may precede overt behavioural change in ToM.

We used an interactive trust game with a trustworthy, cooperative and not trustworthy, unfair partner during functional magnetic resonance imaging (fMRI) in a sample of 22 patients with early psychosis and 25 controls. We hypothesised that we would see lower basic trust (first investment) in patients compared with controls, but similar levels of average trust towards the game partners. Based on the hypothesis that aberrant dopaminergic signalling in response to social reward underlies psychosis and emerging social impairment, we expected to see reduced activation in the right caudate. Given the unimpaired response to cooperative behaviour, we hypothesised that compensatory processing would occur through ToM and be reflected in increased activation in associated brain regions (rTPJ and mPFC). In line with the hypothesised link between social reward sensitivity, paranoia, and social motivation, we expected associations between higher positive symptoms (particularly paranoid ideation) and negative symptoms with lower trust towards others and with reduced caudate activation.

Methods

Subjects

The sample included right-handed male adolescents, 25 healthy and 22 with early psychosis, which was defined as an illness duration of less than 3 years. We included only males to have a homogeneous sample that was comparable to our previous study in males with chronic nonaffective psychosis (Gromann *et al.*, 2013). Behavioural trust game data of a larger adolescent sample, which included behavioural data of participants of this fMRI study, have previously been reported (Fett *et al.*, 2014a; Fett *et al.*, 2014b; Fett *et al.*, 2016). Participants took part in the behavioural arm of the study if they were ineligible for fMRI scanning (e.g. due to braces, metal implants, etc.), or if they did not want to undergo fMRI scanning, and after the intended fMRI sample size was reached. Informed consent was obtained from all adolescents and their parents/guardians if participants were younger than 16 years. Inclusion criteria were age between 13 and 19 years, fluency in English (i.e. able to read and understand the testing material and interview questions), and being able and willing to give written informed consent. Patients had experienced at least one psychotic episode according to *International Classification of Diseases, 10th Revision* criteria (WHO, 1992) and an *International Classification of Diseases, 10th Revision*, diagnosis of schizophrenia, schizotypal and delusional disorders, or mood disorders with psychosis, as diagnosed by their treating psychiatrist [primary diagnoses: 18 × nonaffective (of which five schizophrenia) and 4 × affective psychosis]. Patients had 1.1 hospital admissions on average (range, 0–4). All but one took antipsychotic medication. Those who took medication were on atypical antipsychotics. One patient took an additional typical antipsychotic. Exclusion criteria were diagnosed substance use/abuse or neurological conditions. Additional exclusion criteria for controls were having a history of a psychiatric diagnosis and a family history of psychosis. Patients were recruited through consultant psychiatrists and the Mental Health Research Network in the SLAM, Oxleas, NELFT, and SEPT NHS Foundation Trusts. Control participants were recruited from local schools, the Institute of Psychiatry volunteer database 'Mindsearch,' via colleagues and previous participants. The study was approved by the South West London REC (10/H0806/38).

Assessments

Symptoms were assessed with the Positive and Negative Syndrome Scale (PANSS; Kay *et al.*, 1987), which consists of a positive, negative and general symptoms scale ranging from 1 (absent) to 7 (extreme). The Green Paranoid thought scale (Green *et al.*, 2008) was used to measure ideas of social reference and persecution. Each subscale included 16 items scored from 1 ('not at all') to 5 ('totally'), with higher scores reflecting higher delusions. The vocabulary subtest of the Wechsler Abbreviated Scale of Intelligence (WASI) was used as an indicator of estimated cognitive ability (WASI; Lezak, 2004).

Experimental design

Participants were investors in a multiround trust game (Gromann *et al.*, 2013) and interacted with computers, but were told that they would play with two human partners. These were preprogrammed to behave cooperatively and unfair, either reinforcing investor trust (i.e. increasing investments) with higher repayments or responding to investor trust with lower repayments (see Supplementary Material 1 for algorithms). At the beginning of each experimental round, participants saw the investment cue of £10 (2 s). They then decided how much between £0 and £10 they wanted to share with the other player (investment phase; maximum 4 s). The invested amount was shown (2 s), followed by a waiting period with a bar slowly filling itself with dots (2–4 s) and a fixation cross (500 ms). The shared amount was tripled, and the second player made a repayment. The partner's response (repayment phase) was displayed (3 s), followed by the totals (2.5–4.5 s). The trial ended with a fixation cross (500 ms) and lasted 18.5 s in total. Each condition (cooperative/unfair) consisted of 20 randomly interspersed experimental and 20 control trials (Figure 1). Control trials were included as baseline condition in the fMRI analysis. Here, participants were asked to select the number that was indicated by an arrow (control investment phase; maximum 4 s). All other phases of the control condition included bars of the same height to keep the visual stimuli comparable to the experimental trials. After the decision-making phase, a bar was shown with the text 'rest,' indicating that the participant did not have to do anything (2 s); this was followed by a short waiting period (2–4 s) and a fixation cross (500 ms). Next, two bars were shown with the text 'rest' (3 s control repayment phase and 2.5–4.5 s total control phase). The control trial ended with a fixation cross (500 ms) and lasted 18.5 s in total. The order of the cooperative and unfair condition was counterbalanced across subjects. Order of presentation had no effect on first or average investments during both conditions (all $P > 0.11$). After the session, participants completed a questionnaire (free text entry) asking them whether they had any doubts that their counterpart was real.

Scanning parameters

Imaging data were acquired using a 3 T GE Signa Neuro-optimised MR System at the Centre of Neuroimaging Sciences of the Institute of Psychiatry, Psychology and Neuroscience, King's College London. A quadrature birdcage head coil was used for radiofrequency transmission and reception. Foam padding was placed around the participant's head in the coil to minimise head movement. Participants made their responses using two buttons on a two-button box with their index and middle fingers of their right hand. Three hundred seventy T_2*-weighted whole-brain echo-planar images sensitive to the blood oxygen level–dependent contrast were acquired with the following parameters: slice thickness = 2.4 mm, gap = 1 mm, repetition time = 2 s, echo time = 25 ms, flip angle = 75°, in-plane resolution = 3.4 mm, number of slices = 38, number of slices/DDAs = 4, matrix = 64 × 64. For anatomical reference, a coronal fast spoiled gradient echo image of the whole brain was obtained for each subject, which consisted of 196 slices acquired with the following parameters: thickness = 1.1 mm, gap = 0, repetition time = 7 s, echo time = 2.8 ms, flip angle = 20°, matrix = 256 × 256.

Data analysis

The behavioural data were analysed in Stata 14 (StataCorp, 2015). Group differences in demographics and estimated cognitive ability were analysed with *t*-tests. Group differences in first investments across the two trust game conditions (basic trust) and in average trust towards each game partner were analysed with multilevel random regression analyses (XTREG), to account for multiple observations [investments (level 1); within participants (level 2)]. We report all analyses excluding and including age and WASI vocabulary score. Within patients, we investigated the relationships between trust (height of investments) and positive and negative symptoms as well as ideas of social reference and persecution with multilevel random regression analyses.

The neuroimaging analysis was completed using FSL FEAT (FMRI Expert Analysis Tool) version 6.00 (FMRIB's Software Library, www.fmrib.ox.ac.uk/fsl). Registration to high-resolution structural and/or standard space images was carried out using FLIRT (Jenkinson & Smith, 2001; Jenkinson *et al.*, 2002). The following prestatistics processing was applied, motion correction using MCFLIRT (Jenkinson *et al.*, 2002), nonbrain removal using BET (Smith, 2002), spatial smoothing using a Gaussian kernel of FWHM 6 mm, grand-mean intensity normalisation of the entire 4D dataset by a single multiplicative factor, and high-pass temporal filtering (Gaussian-weighted least-squares straight line fitting, with sigma = 45.0 s). Time-series statistical analysis was carried out using FILM with local autocorrelation correction (Woolrich *et al.*, 2001). For the first-level analysis, the investment and repayment phases of experimental and control trials were modelled separately. Each phase was modelled as an epoch: the investment phase from onset until button press and the repayment phase with a duration of 3 s. Both were convolved with a canonical haemodynamic response function. Six standard motion parameters were added as regressors of no interest, as well as a motion artefact confound matrix, which identified motion-corrupted volumes. Volumes detected as corrupted were calculated by DVARS metric as implemented by FSL Motion Outliers in FSL (https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FSLMotionOutliers).

Contrasts of interest for each subject were created by comparing mean BOLD signal of the investment (real investment vs. control investment) and repayment phase (real repayment vs. control repayment) in the experimental trials to their respective phase in the control trials [see Figure 1 for an illustration of the experimental (real) and control trials]. For the group-level analysis, we used a priori region of interest (ROI) analyses as specified in Gromann *et al.* (2013). Regions of interest were based on research identifying robust reward- and ToM-related activation in independent samples for the right caudate (Talairach coordinates 10, 9, 4; Knutson *et al.*, 2003), the right rTPJ (Talairach coordinates 51, −54, 27; Saxe & Kanwisher, 2003), and the mPFC (Talairach coordinates −3, 64, 20; Hampton *et al.*, 2008). Regions of interest were created with a 5-mm sphere centred on the

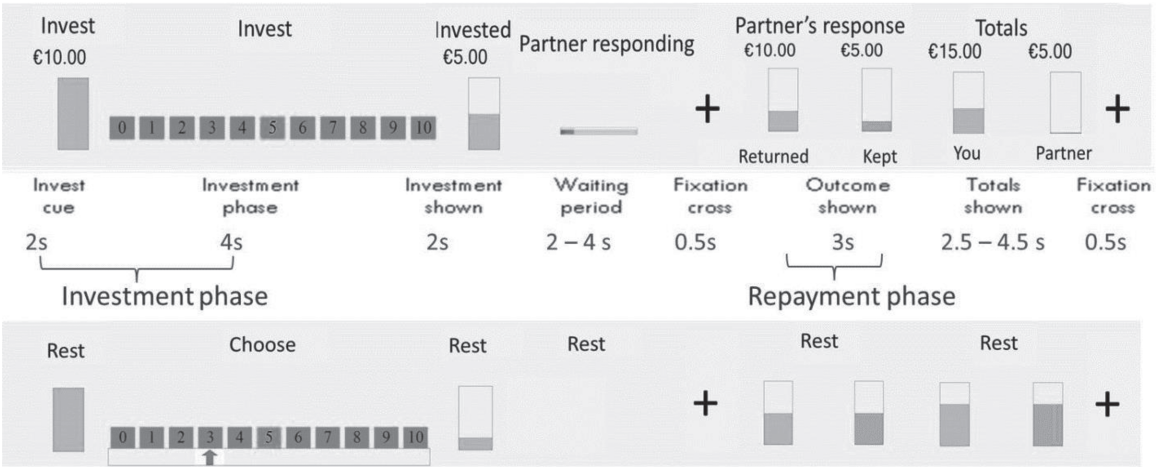

Figure 1. Experimental set-up of the trust game (top panel experimental trial, bottom panel control trial)

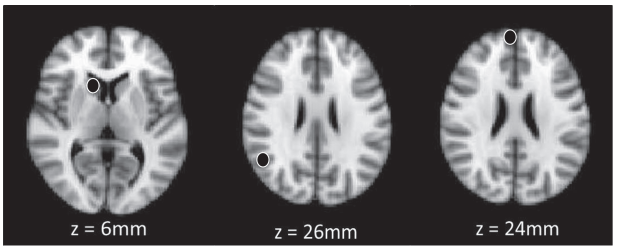

Figure 2. ROIs of the right caudate (left), right TPJ (middle), and mPFC (right).

coordinates (Figure 2). Contrast estimates were extracted from these ROIs for each participant. To analyse group differences, we used a mixed-factorial design, with between-subjects factor group (patients vs. controls) and within-subjects factor condition (cooperative vs. unfair). Analyses are reported excluding and including age and WASI vocabulary score. Cohen's *d* effect sizes were computed for ease of interpretation. Associations between ideas of social reference and persecution, positive and negative symptoms, investments, and brain activation were analysed with parametric regression analyses. Exploratory whole-brain analysis was performed to investigate group differences outside the a priori–defined ROIs. *Z* (Gaussianised T/F) statistic images were thresholded using clusters determined by $Z > 2.3$ and a family-wise-error whole-brain corrected cluster significance threshold of $P = 0.05$ (Supplementary Material 2; Worsley, 2001).

Results

Behavioural results

Patients had a lower estimated cognitive ability than controls but did not differ significantly in age. The groups did not differ significantly in first investments (basic trust) or average investments during cooperative or unfair interactions (Table 1).

Within patients, higher PANSS positive ($b = -1.28$, $P = 0.02$) and negative symptoms ($b = -1.35$, $P = 0.005$) were associated with lower basic trust. Higher persecutory delusions ($b = -1.24$, $P = 0.01$) and PANSS positive ($b = -0.9$, $P = 0.03$) and negative symptoms ($b = -1.08$, $P = 0.005$) were associated with lower investments during cooperation. Negative symptoms were associated with lower investments during unfair interactions ($b = -0.77$, $P = 0.03$). *P* values <0.0125 survive Bonferroni correction ($\alpha = 0.05/4$).

Neuroimaging Results

Caudate. During the investment phase, there was a significant group-by-condition interaction [$F(1,44) = 5.32$, $P = 0.027$, $d = 0.67$] and a significant main effect of condition [$F(1,44) = 7.4$, $P = 0.009$, $d = 0.79$]. The main effect of group was nonsignificant [$F(1,44) = 0.09$, $P = 0.76$, $d = 0.08$]. Analyses by group showed a significant condition effect in controls [$t(24) = -3.71$, $P = 0.0006$, $d = 1.09$], but not patients [$t(21) = -0.30$, $P = 0.76$, $d = 0.09$]. Specifically, controls showed greater caudate activation during cooperative and lower activation during unfair interactions, compared to the control condition. Analyses by condition showed a group effect on caudate activation that only trended towards significance during cooperative interactions [$t(44) = -1.87$, $P = 0.06$, $d = 0.40$], and that was nonsignificant during unfair interactions [$t(45) = 1.62$, $P = 0.11$, $d = 0.37$], see Figure 3A, left panel. The group-by-condition interaction did not change when age and vocabulary score were included in the model, and effects of age and vocabulary scores were nonsignificant (both $P > .52$). The trend effect of group during cooperation remained ($P = 0.06$, $d = 0.39$). Neither age nor vocabulary score was associated with ROI activation ($P > 0.51$).

During the repayment phase, there was a group-by-condition interaction and a group main effect that trended towards significance [$F(1,44) = 3.14$, $P = 0.08$, $d = 0.53$ and $F(1,44) = 3.27$, $P = 0.07$, $d = 0.52$]. The main effect of condition was nonsignificant [$F(1,44) = 0.32$, $P = 0.57$, $d = 0.17$]. Analyses by group showed a trend effect of condition in controls [$t(24) = -1.73$, $P = 0.09$, $d = 0.35$], but not patients [$t(21) = 0.82$, $P = 0.42$, $d = 0.17$]. Patients showed lower caudate activation in interactions with both game partners. Controls showed higher activation in cooperative and lower activation in unfair interactions. Analyses by condition showed that patients' ROI activation differed significantly from controls in the cooperative [$t(44) = -2.37$, $P = 0.02$, $d = 0.7$], but not the unfair condition [$t(45) = -0.44$, $P = 0.66$, $d = 0.13$], see Figure 3A, right panel. Including age and vocabulary score the interaction changed to [$P = 0.07$, $d = 0.57$]. Only vocabulary scores were significantly associated with ROI activation ($P < 0.01$). Group differences during cooperation became nonsignificant [$P = 0.25$, $d = 0.15$]. Effects in the caudate with $P > 0.017$ did not survive a more stringent Bonferroni-corrected threshold of $\alpha = 0.05/3$.

Table 1. Sample characteristics and behavioural trust game results

	Patients n = 22 Mean	(SD)	Controls n = 25 Mean	(SD)	Test statistic
Age	17.57	(1.27)	16.80	(1.59)	$t(45) = -1.82$, $P = 0.07$
WASI vocabulary	45.78	(10.32)	56.72	(9.91)	$t(45) = 3.55$, $P < 0.001$
PANSS positive average Range (1–4.7)	1.91	(0.91)			
PANSS negative average Range (1–4.2)	2.01	(1.04)			
GPTS social reference	2.14	(0.97)			
GPTS persecution	1.80	(0.84)			
Basic trust	5.14	(2.68)	5.68	(2.47)	$t(45) = -1.01$, $P = 0.32$
Trust during cooperation	6.57	(2.14)	6.84	(1.82)	$t(45) = -0.66$, $P = 0.50$
Trust during unfairness	4.51	(1.78)	4.13	(2.11)	$t(44) = 0.96$, $P = 0.34$

Temporoparietal junction. During the investment phase, there was no significant group-by-condition interaction [$F(1,44) = 0.37$, $P = 0.54$, $d = 0.18$]. The group main effect trended towards significance [$F(1,44) = 3.08$, $P = 0.08$, $d = 0.52$], with lower temporoparietal junction (TPJ) activation in patients than controls. The condition main effect was nonsignificant [$F(1,44) = 0.51$, $P = 0.48$, $d = 0.21$], see Figure 3B, left panel. Including age and vocabulary score, the trend effect of group became nonsignificant [$P = 0.21$, $d = 0.39$]. Neither age nor vocabulary score was significantly associated with TPJ activation (both $P > 0.60$).

During the repayment phase, there was a significant group-by-condition interaction [$F(1,44) = 3.91$, $P = 0.05$, $d = 0.60$]. Main effects of group [$F(1,44) = 3.07$, $P = 0.09$, $d = 0.51$] and condition [$F(1,44) = 1.26$, $P = 0.27$, $d = 0.32$] were nonsignificant. Analyses by group showed no condition effect in controls [$t(24) = 0.63$, $P = 0.53$, $d = 0.11$]. Patients, however, demonstrated greater TPJ activation in the cooperative than the unfair condition [$t(21) = -2.10$, $P = 0.04$, $d = 0.57$]. Group differences in TPJ activation were significant in the cooperative [$t(44) = 2.60$, $P = 0.01$, $d = 0.77$], but not the unfair condition [$t(45) = -0.11$, $P = 0.91$, $d = 0.03$], see Figure 3B, right panel. Including age and vocabulary score, the interaction reduced to $P = 0.08$, $d = 0.54$. Neither age nor vocabulary score was significantly associated with TPJ activation (both $P > 0.95$). Group differences during cooperation remained at trend level ($P = 0.09$, $d = 0.49$). Effects in the TPJ with $P > 0.017$ did not survive a more stringent Bonferroni-corrected threshold of $\alpha = 0.05/3$.

Medial prefrontal cortex. There were no significant effects of group {investment: [$F(1,44) = 0.06$, $P = 0.81$, $d = 0.07$]; repayment [$F(1,44) = 0.2$, $P = 0.65$, $d = 0.13$]} and condition {investment: [$F(1,44) = 0.39$, $P = 0.53$, $d = 0.19$]; repayment [$F(1,44) = 0.09$, $P = 0.76$, $d = 0.08$]} and no significant interactions {investment: [$F(1,44) = 0.10$, $P = 0.75$, $d = 0.09$]; repayment [$F(1,44) = 0.05$, $P = 0.83$, $d = 0.06$]} for the mPFC, see Figure 3C.

Additional analyses

Sixteen percent of participants indicated that they had doubts that the other player was real. The groups did not differ significantly [$\chi^2(1) = 2.51$, $P = 0.11$]. Having doubts was unrelated to ROI activation (all $P > 0.13$), except for the mPFC during the unfair repayment phase [$t(35) = -2.34$, $P = 0.04$, $d = 0.77$].

ROI activation and investments. First investments (basic trust) and average investments were unrelated to ROI activation during the cooperative and unfair investment phase (all $P > 0.05$).

ROI activation and symptoms. There were no significant associations between the ROI contrast estimates and PANSS positive or negative symptoms (all $P > 0.05$). During cooperative and unfair repayments, higher persecutory delusions were associated with lower rTPJ activation ($\beta = -0.54$, $P = 0.02$, and $\beta = -0.55$, $P = 0.03$). No other associations between GPTS subscales and brain activation were significant.

Exploratory whole-brain analysis. Group differences were absent at the whole-brain level. Main clusters of task activation are shown in Table 2. For task activation maps across groups and detailed local maxima of the clusters in Table 2, see Supplementary Material 2.

Discussion

We used an interactive trust game during fMRI to examine the neural mechanisms of social interactions in early psychosis. The groups did not differ in trusting behaviour but showed differential neural activation patterns. During cooperative interactions, psychosis was associated with reduced activation in the right caudate, a key area of reward processing, suggesting that abnormal sensitivity to social reward is present early in the disorder. While making decisions to trust, patients showed lower rTPJ activation, perhaps indicating reduced mentalising. When cooperative partner feedback was revealed, this pattern was reversed. This could reflect a compensatory mechanism as part of an increased effort to interpret others' social signals.

Behavioural findings

Reward processing in response to social information guides (social) decision making in human interactions. In the brain, this process is mediated by dopamine, which is implicated in the pathophysiology of psychosis and its core symptoms, such as paranoid delusions and reduced social motivation (Kapur *et al.*, 2005; Kirschner *et al.*, 2016). Chronic psychosis is associated with reduced trust and insensitivity to the rewarding aspects of social cooperation (Fett *et al.*, 2012; Gromann *et al.*, 2013). While impairments in reward processing might underlie social problems, they could also be a consequence of other disorder-related factors, such as negative social experiences or long-standing medication use. Absent differences in trust between patients with early psychosis and controls in this study compared to previous research support the interpretation

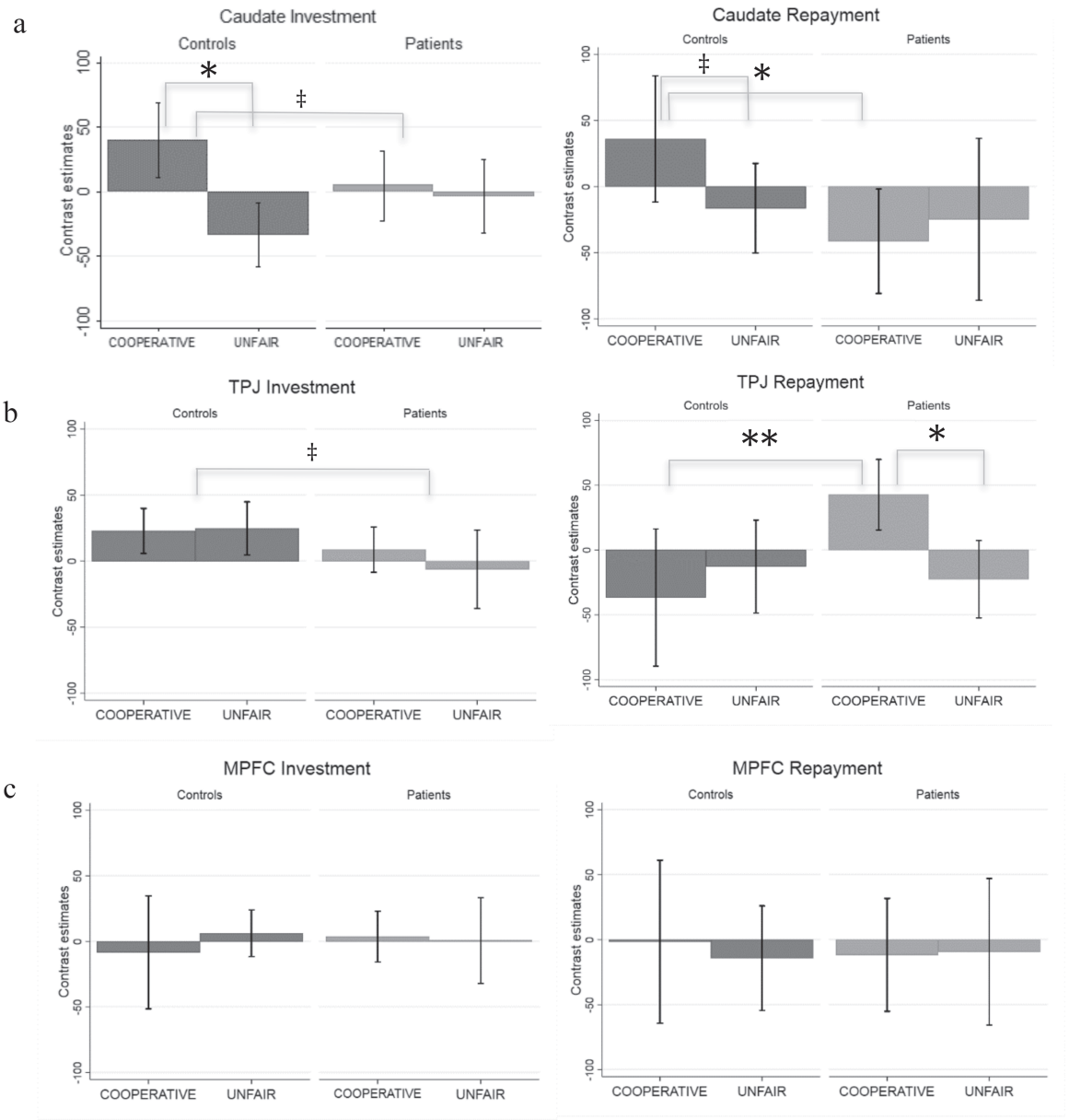

Figure 3. (a) Right caudate, (b) rTPJ and (c) mPFC activation by game phase, condition and group Note. Left panel shows activation during investments, right panel shows activation during repayments (error bars show standard deviations). $^{**}p < 0.01$, $^{*}p < 0.05$, $\ddagger p < 0.09$.

that later changes in trust may be secondary to the disorder. However, it is important to note that preexisting processing impairments and disorder-related mechanisms are not mutually exclusive as explanatory factors. Rather, they could interact or have additive effects. For instance, in the early phase of psychosis, impaired trust could still be compensated through other cognitive mechanisms; over time, typical age-related increases in trust observed in healthy individuals (Fett *et al.*, 2014a) could be compounded by negative social experiences or long-standing medication use, which could lead to reduced functional plasticity. Thus, disorder-related factors might limit the ability or willingness to exert effort to overcome distrust and suspiciousness against others in addition to preexisting vulnerabilities.

The current study did not show reduced basic trust in patients in early psychosis; however, altered trust in the early illness stage has previously been reported by others (Campellone *et al.*, 2016; Fett *et al.*, 2016; Lemmers-Jansen *et al.*, 2018). One possible explanation for the divergent finding could be the low illness severity in the current sample. Our finding that lower basic trust was associated with higher positive and negative symptoms supports this interpretation and suggests that group differences could emerge between controls and more acute patients. During repeated interactions, lower trust towards the cooperative game partner was associated with persecutory delusions and positive and negative symptoms. Negative symptoms were also associated with lower trust towards the unfair game partner. Associations between negative symptoms and persecutory delusions and lower trust during cooperative interactions were most robust. These findings confirm the important link between exaggerated suspiciousness, social withdrawal, and problematic social functioning early in the disorder.

Table 2. Whole-brain task activation in the cooperative and unfair condition by game phase

Condition Contrast	Brain region	Hemisphere	Talairach Coordinates X	Y	Z	Cluster size	Z
Cooperative							
Investment vs. control							
	Cingulate gyrus	L	−1	22	35	8767	6.22
	Inferior parietal lobule	R	38	−55	36	688	3.93
Repayment vs. control							
	Fusiform gyrus	L	−24	−82	−13	16 203	6.85
	Middle frontal gyrus	R	46	24	27	5836	5.57
	Precentral gyrus	R	−39	−1	30	2425	5.87
	Medial frontal gyrus	R	1	17	42	672	4.37
Unfair							
Investment vs. control	Cingulate gyrus	R	5	20	34	2784	5.53
Repayment vs. control	Lingual gyrus	L	−15	−84	−13	9990	6.08
	Inferior frontal gyrus	R	42	5	27	2510	5.8
	Inferior parietal lobule	L	−34	−60	41	1348	5.28
	Precentral gyrus	L	−37	1	30	612	4.22

Note. Clusters were significant with a (corrected) cluster significance threshold of $P = 0.05$. The local maxima and corresponding brain areas of the broader clusters are reported in the Supplementary Material. FSL MNI coordinates were transformed to Talairach coordinates with the GingerAle 2.3.6 (http://www.brainmap.org/) convert foci option using FSL to Talairach. Brain regions were then identified with the Talairach Client (Lancaster *et al.*, 2000).

Neuroimaging findings

At the neural level, individuals with early psychosis showed right caudate activation patterns similar to those previously observed in chronic patients (Gromann *et al.*, 2013). The caudate responds to incentive salience (Berridge & Robinson, 1998) with greater activation in response to higher social rewards (Bhanji & Delgado, 2014) and increased activation occurs during trusting decisions in fair social interactions, when cooperation is anticipated or experienced (Rilling, 2002; King-Casas *et al.*, 2005; Phan *et al.*, 2010). In line with this, controls showed greater caudate activation during decisions to trust during cooperative than unfair interactions and a similar, albeit less pronounced pattern during the partner's response. Individuals with early psychosis did not activate the caudate differentially when trusting the cooperative or unfair partner. When partner cooperation was revealed, they showed less caudate activation during interactions with both partners than control trials. Our findings show that group differences in caudate activation are specific to positive social interactions, as previously found in adults in the chronic illness stages (Gromann *et al.*, 2013). Blunted caudate activation during trust game interactions with cooperative others has also been found in first-degree relatives of patients with psychosis (Gromann *et al.*, 2014), suggesting that the reduced neural sensitivity to social reward could be an endophenotype of psychosis. During unfair interactions, controls and patients showed reduced caudate activation. The intact neural sensitivity to negative social cues in patients is in line with previous research showing impaired reward processing and intact loss processing in psychosis (Waltz *et al.*, 2007) and contradicts the notion of general reward learning impairments in psychosis. The finding of specific insensitivity to positive social reward is important, because socially rewarding experiences motivate social behaviour (Phan *et al.*, 2010; Radke *et al.*, 2016). The absence thereof could be the root of impaired social functioning in psychosis. Alternative explanations of our findings in terms of prediction error signalling, where caudate activation should be reduced during expected as opposed to surprising partner responses, appear less plausible (Schultz *et al.*, 1997), given that participants were more likely to invest higher in the cooperative than the unfair partner. The adjustment of trust in line with the response style suggests that participants predicted the reciprocation patterns successfully. The caudate response was unrelated to symptoms and the height of the investments, indicating that the group differences were due to neither illness severity, nor the differential evaluation of monetary reward.

Successful social interactions depend not only on reward learning, but also on ToM, which is vital to infer others' intentions and to anticipate their (re)actions (Hampton *et al.*, 2008; Fett *et al.*, 2014b). The mPFC and rTPJ subserve this function (Saxe & Wexler, 2005; Frith & Frith, 2006). In line with previous neuroimaging studies (Sugranyes *et al.*, 2011), patients activated the rTPJ less than controls during trusting decisions. This could suggest reduced mentalising about the own investment decisions. Interestingly, patients showed greater TPJ activation than controls during cooperative repayments. Previous research reported complex patterns of hypomentalising and hypermentalising in early psychosis (Bliksted *et al.*, 2018). While speculative, it is possible that engaging the rTPJ allows for adjustment of trust based on more elaborate compensatory mentalising computation (e.g. thoughts such as 'Why did the other person make this decision?'). The idea of a compensatory mechanism is supported by observations of lower trust in relation to reduced caudate and rTPJ activation in chronic patients (Gromann *et al.*, 2013). During repayments, higher persecutory delusions were associated with lower rTPJ activation, suggesting hypomentalising about the others' behavioural signals in those with stronger paranoid delusions. The mPFC has also been linked to impairments of ToM

(Frith & Frith, 2006; Kronbichler *et al.*, 2017). However, similarly to patients in the chronic illness stages, individuals with early psychosis did not differ from controls in terms of mPFC activation. Gromann and colleagues suggested that functioning in this specific area of the social brain network might be relatively intact in nonaffective psychosis (Gromann *et al.*, 2013). This is contradicted by the fact that a variety of ToM-related tasks have been associated with lower activation in the mPFC in individuals with psychosis compared to controls (Sugranyes *et al.*, 2011; Fett *et al.*, 2015; Kronbichler *et al.*, 2017). An alternative explanation of our findings might be that task demands of the trust game on the mPFC are lower compared to more standard ToM tasks. This interpretation is supported by recent work showing no effect of transcranial direct current stimulation to the mPFC on trust game behaviour (Colzato *et al.*, 2015).

Limitations and methodological issues

Our findings need to be viewed in light of several important methodological issues. First, it is important consider that our sample included participants in early to later stages of adolescent development. Thus, any differences in brain activation in the current sample compared to longstanding psychosis may not only be due to differences in chonicity, but may also be accounted for by the developmental trajectory. To disentangle these effects, future studies should include a wide range of participants in different stages of their development from adolescence to adulthood, as well as in different stages of the psychotic disorder. Furthermore, the fact that the current sample was relatively young means that diagnoses may change. Longitudinal research will be necessary to provide evidence for trajectories of social cognition and reward processing and associations with social behaviour in psychotic disorders over time. Second, group differences reduced to nonsignificance when the WASI vocabulary scores were added to the statistical models, as index of general cognitive ability. Importantly, impairments in cognitive ability predate the onset of psychotic disorders (Meier *et al.*, 2014; Mollon & Reichenberg, 2018). They are by definition inherent to the neurodevelopmental disorder, and therefore, it can be questioned whether a measure of IQ is a suitable covariate (Dennis *et al.*, 2009). However, these results do give important insights into the close association between general cognitive impairment, social cognition, and reward processing in psychosis. Third, this study included only males, and the results are therefore not generalizable to the entire patient population. Sex differences in the social behaviour in the trust game have previously been reported in the adolescent population (Lemmers-Jansen *et al.*, 2017), and future fMRI studies on mechanisms of social interaction in psychosis should investigate possible gender effects. Fourth, all but one patient took atypical antipsychotic medication (one patient received an additional typical antipsychotic). This clinical reality is a methodological issue that affects most studies that investigate psychosis and could have influenced reward processing in patients. Importantly, two sources of evidence suggest that this is not the case. First, atypical antipsychotics actually seem to normalise the brain reward response (Juckel *et al.*, 2006; Schlagenhauf *et al.*, 2008). In addition, healthy first-degree relatives of patients with psychosis show differential processing of social reward, without any medication confounds (Gromann *et al.*, 2014). As such, the current findings might reflect an underestimation of the effects that would be present in unmedicated individuals. While studies in unmedicated individuals are extremely difficult to conduct, it would be valuable if future research attempted to study trust in social interaction in unmedicated patient cohorts to shed more light on this issue. Fifth, 16% of the participants indicated doubts that the other player was real. This could have influenced decisions to trust and mentalising processes. However, there were no significant associations between having doubts and brain activation in any of the ROIs except the mPFC, for which we did not observe group differences in activation. Sixth, many findings in the caudate and TPJ did not survive a more stringent Bonferroni correction of α (0.05/3) = 0.017. Given the lack of significant effects in the presence of a pattern of moderate to large effect sizes, the results need to be interpreted with caution and warrant replication in a larger sample before any more firm conclusions can be made. Finally, fMRI allows for the investigation of the role of particular brain regions in certain cognitive functions. However, caution is required when cognitive processes are inferred from activation. Our data add to the growing literature on social brain systems of reward processing and ToM in psychosis and can be regarded as guide for necessary future inquiries.

Conclusion

Characterising the social brain and behaviour link can aid the identification of the underlying factors of social impairment. This study suggests that abnormal operation of reward-based mechanisms during cooperative social interactions in early psychosis may underlie social impairments. Importantly and in contrast to patients who have been affected by long-standing psychosis (Gromann *et al.*, 2013), young patients did not show overt behavioural differences in trust, suggesting that they compensate reduced activation in the reward system via other mechanisms. This highlights the potential of this sensitive time for interventions aimed at preventing decline in social functioning.

Supplementary Data

Supplementary data are available at *SCAN* online.

Funding

This work was supported by a Netherland Organization for Scientific Research (NWO) [#451–13-035 to A.-K.J.F. and #453-11-005 to LK]; a NARSAD Young Investigator Award from the Brain and Behaviour Foundation [#24138 to A.-K.J.F.]; ERC Consolidator Grants #648082 to LK and #311686 to S.S.S.]. S.S.S. is also supported by the NIHR Mental Health Biomedical Research Center at the SLaM NHS Trust and King's College London.

Acknowledgements

We thank the participants, the MHR support team, clinical studies officers, and consultants at NELFT, SEPT, SLAM, and OXLEAS; Partha Banerjea, Uttom Chowdhury, Philip Collins, Sophia Frangou, Giovanni Giaroli, Daniel Hayes, Kah Mirza, Viviana Porcari, Kurinji Semmanpalayam, Derek Tracy, Tracy Collier, Farah Yakub, and the local schools.

References

Achim, A.M., Ouellet, R., Roy, M.-A., Jackson, P.L. (2012). Mentalizing in first-episode psychosis. *Psychiatry Research*, **196**(2–3), 207–13.

Bartholomeusz, C.F., Ganella, E.P., Whittle, S., *et al.* (2018). An fMRI study of theory of mind in individuals with first episode psychosis. *Psychiatry Research: Neuroimaging*, **281**, 1–11.

Berg, J., Dickhaut, J., McCabe, K. (1995). Trust, reciprocity and social history. *Games and Economic Behavior*, **10**, 122–42.

Berridge, K.C., Robinson, T.E. (1998). What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Research Reviews*, **28**(3), 309–69.

Bhanji, J.P., Delgado, M.R. (2014). The social brain and reward: social information processing in the human striatum. *Wiley Interdisciplinary Reviews: Cognitive Science*, **5**(1), 61–73.

Blakemore, S.-J. (2008). The social brain in adolescence. *Nature Reviews Neuroscience*, **9**(4), 267–77.

Bliksted, V., Frith, C., Videbech, P., *et al.* (2018). Hyper- and hypomentalizing in patients with first-episode schizophrenia: fMRI and behavioral studies. *Schizophr Bull*. 2019 Mar 7;**45**(2):377–385.

Brüne, M., Özgürdal, S., Ansorge, N., *et al.* (2011). An fMRI study of 'theory of mind' in at-risk states of psychosis: comparison with manifest schizophrenia and healthy controls. *NeuroImage*, **55**(1), 329–37.

Campellone, T.R., Fisher, A.J., Kring, A.M. (2016). Using social outcomes to inform decision-making in schizophrenia: relationships with symptoms and functioning. *Journal of Abnormal Psychology*, **125**(2), 310–21.

Canty, A.L., Neumann, D.L., Shum, D.H.K. (2017). Using virtual reality to assess theory of mind subprocesses and error types in early and chronic schizophrenia. *Schizophrenia Research: Cognition*, **10**, 15–9.

Colzato, L.S., Sellaro, R., van den, W.P.M., Hommel, B. (2015). tDCS of medial prefrontal cortex does not enhance interpersonal trust. *Journal of Psychophysiology*, **29**(4), 131–4.

Dennis, M., Francis, D.J., Cirino, P.T., Schachar, R., Barnes, M.A., Fletcher, J.M. (2009). Why IQ is not a covariate in cognitive studies of neurodevelopmental disorders. *Journal of the International Neuropsychological Society: JINS*, **15**(3), 331–43.

Eisenberger, N.I., Cole, S.W. (2012). Social neuroscience and health: neurophysiological mechanisms linking social ties with physical health. *Nature Neuroscience*, **15**(5), 669–74.

Fett, A.J., Shergill, S., Korver-Nieberg, N., Yakub, F., Gromann, P., Krabbendam, L. (2016). Learning to trust: trust and attachment in early psychosis. *Psychological Medicine*, **46**(7), 1437.

Fett, A.-K., Shergill, S., Krabbendam, L. (2015). Social neuroscience in psychiatry: unravelling the neural mechanisms of social dysfunction. *Psychological Medicine*, **45**(6), 1145–65.

Fett, A.-K.J., Gromann, P.M., Giampietro, V., Shergill, S.S., Krabbendam, L. (2014a). Default distrust? An fMRI investigation of the neural development of trust and cooperation. *Social Cognitive and Affective Neuroscience*, **9**(4), 395–402.

Fett, A.-K.J., Shergill, S.S., Gromann, P.M., Dumontheil, I., Blakemore, S.-J., Yakub, F., Krabbendam, L. (2014b). Trust and social reciprocity in adolescence—a matter of perspective-taking. *Journal of Adolescence*, **37**(2), 175–84.

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** (Pt 3), 976–84.

Frith, C.D., Frith, U. (2006). How we predict what other people are going to do. *Brain Research*, **1079**(1), 36–46.

Green, C., Freeman, D., Kuipers, E., Bebbington, P., Fowler, D., Dunn, G., Garety, P. (2008). Measuring ideas of persecution and social reference: the Green et al. Paranoid Thought Scales (GPTS). *Psychological Medicine*, **38**(1), 101–11.

Gromann, P.M., Heslenfeld, D.J., Fett, A.-K.J., Joyce, D.W., Shergill, S.S., Krabbendam, L. (2013). Trust versus paranoia: abnormal response to social reward in psychotic illness. *Brain*, **136**, 1968–75.

Gromann, P.M., Shergill, S.S., de Haan, L., Meewis, D.G.J., Fett, A.K.J., Korver-Nieberg, N., Krabbendam, L. (2014). Reduced brain reward response during cooperation in first-degree relatives of patients with psychosis: an fMRI study. *Psychological Medicine*, **44**(16), 3445–54.

Hampton, A.N., Bossaerts, P., O'Doherty, J.P. (2008). Neural correlates of mentalizing-related computations during strategic interactions in humans. *Proceedings of the National Academy of Sciences of the United States of America*, **105**(18), 6741–6.

Ho, K.K., Lui, S.S., Hung, K.S., Wang, Y., Li, Z., Cheung, E.F., Chan, R.C. (2015). Theory of mind impairments in patients with first-episode schizophrenia and their unaffected siblings. *Schizophrenia Research*, **166**(1–3), 1–8.

Jenkinson, M., Bannister, P., Brady, M., Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage*, **17**(2), 825–41.

Jenkinson, M., Smith, S. (2001). A global optimisation method for robust affine registration of brain images. *Medical Image Analysis*, **5**(2), 143–56.

Juckel, G., Schlagenhauf, F., Koslowski, M., *et al.* (2006). Dysfunction of ventral striatal reward prediction in schizophrenic patients treated with typical, not atypical, neuroleptics. *Psychopharmacology*, **187**(2), 222–8.

Kapur, S., Mizrahi, R., Li, M. (2005). From dopamine to salience to psychosis—linking biology, pharmacology and phenomenology of psychosis. *Schizophrenia Research*, **79**(1), 59–68.

Kay, S.R., Fiszbein, A., Opler, L.A. (1987). The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophrenia Bulletin*, **13**(2), 261–76.

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.

Kirschner, M., Hager, O.M., Muff, L., *et al.* (2016). Ventral striatal dysfunction and symptom expression in individuals with schizotypal personality traits and early psychosis. *Schizophr Bull*. 2018 Jan 13;**44**(1):147–157.

Knutson, B., Fong, G.W., Bennett, S.M., Adams, C.M., Hommer, D. (2003). A region of mesial prefrontal cortex tracks monetarily rewarding outcomes: characterization with rapid event-related fMRI. *NeuroImage*, **18**(2), 263–72.

Korver-Nieberg, N., Fett, A.-K.J., Meijer, C.J., Koeter, M.W., Shergill, S.S., de Haan, L., Krabbendam, L. (2013). Theory of mind, insecure attachment and paranoia in adolescents with early psychosis and healthy controls. *Australian and New Zealand Journal of Psychiatry*, **47**(8), 737–45.

Krach, S., Paulus, F.M., Bodden, M., Kircher, T. (2010). The rewarding nature of social interactions. *Frontiers in Behavioral Neuroscience*, **4**, 22.

Kronbichler, L., Tschernegg, M., Martin, A.I., Schurz, M., Kronbichler, M. (2017). Abnormal brain activation during theory of mind tasks in schizophrenia: a meta-analysis. *Schizophrenia Bulletin*, **43**(6), 1240–50.

Krueger, F., Meyer-Lindenberg, A. (2018). Toward a model of interpersonal trust drawn from neuroscience, psychology, and economics. *Trends Neurosci*. 2019 Feb;**42**(2):92–101.

Lancaster JL, Woldorff MG, Parsons LM, Liotti M, Freitas CS, Rainey L, Kochunov PV, Nickerson D, Mikiten SA, Fox PT, Auto-

mated Talairach Atlas labels for functional brain mapping. *Human Brain Mapping* **10**:120–131, 2000.

Lemmers-Jansen, I.L.J., Fett, A.-K.J., Hanssen, E., Veltman, D.J., Krabbendam, L. (2018). Learning to trust: social feedback normalizes trust behavior in first-episode psychosis and clinical high risk. *Psychol Med*. 2019 Apr;**49**(5):780–790. doi: 10.1017/S003329171800140X. Epub 2018 Jun 13.

Lemmers-Jansen, I.L.J., Krabbendam, L., Veltman, D.J., Fett, A.-K.J. (2017). Boys vs. girls: gender differences in the neural development of trust and reciprocity depend on social context. *Developmental Cognitive Neuroscience*, **25**, 235–45.

Lezak, M.D. (2004). *Neuropsychological Assessment*, USA: Oxford University Press.

Meier, M.H., Caspi, A., Reichenberg, A., *et al.* (2014). Neuropsychological decline in schizophrenia from the premorbid to the postonset period: evidence from a population-representative longitudinal study. *American Journal of Psychiatry*, **171**(1), 91–101.

Mollon, J., Reichenberg, A. (2018). Cognitive development prior to onset of psychosis. *Psychological Medicine*, **48**(3), 392–403.

Phan, K.L., Sripada, C.S., Angstadt, M., McCabe, K. (2010). Reputation for reciprocity engages the brain reward center. *Proceedings of the National Academy of Sciences of the United States of America*, **107**(29), 13099–104.

Porcelli, S., Van Der, N., van der Werff, S., *et al*. (2018). Social brain, social dysfunction and social withdrawal. *Neurosci Biobehav Rev*. 2019 Feb;**97**:10–33.

Radke, S., Seidel, E.-M., Eickhoff, S.B., Gur, R.C., Schneider, F., Habel, U., Derntl, B. (2016). When opportunity meets motivation: neural engagement during social approach is linked to high approach motivation. *NeuroImage*, **127**(Supplement C), 267–76.

Rilling, J.K. (2002). A neural basis for social cooperation. *Neuron*, **35**, 395–405.

Sanfey, A.G. (2007). Social decision-making: insights from game theory and neuroscience. *Science*, **318**(5850), 598–602.

Saxe, R., Kanwisher, N. (2003). People thinking about thinking people. The role of the temporo-parietal junction in 'theory of mind'. *NeuroImage*, **19**(4), 1835–42.

Saxe, R., Wexler, A. (2005). Making sense of another mind: the role of the right temporo-parietal junction. *Neuropsychologia*, **43**(10), 1391–9.

Schlagenhauf, F., Juckel, G., Koslowski, M., *et al*. (2008). Reward system activation in schizophrenic patients switched from typical neuroleptics to olanzapine. *Psychopharmacology*, **196**(4), 673–84.

Schultz, W., Dayan, P., Montague, P.R. (1997). A neural substrate of prediction and reward. *Science*, **275**(5306), 1593–9.

Smith, S.M. (2002). Fast robust automated brain extraction. *Human Brain Mapping*, **17**(3), 143–55.

StataCorp (2015) Stata statistical software. In: StataCorp editor. *Stata*, College Station, TX: StataCorp LP.

Sugranyes, G., Kyriakopoulos, M., Corrigall, R., Taylor, E., Frangou, S. (2011). Autism spectrum disorders and schizophrenia: meta-analysis of the neural correlates of social cognition. *PLoS One*, **6**(10), e25322.

Velthorst, E., Fett, A.K.J., Reichenberg, A., Perlman, G., Van Os, J., Bromet, E.J., Kotov, R. (2016). The 20-year longitudinal trajectories of social functioning in individuals with psychotic disorders. *Am J Psychiatry*. 2017 Nov 1; **174**(11): 1075–1085.

Waltz, J.A., Frank, M.J., Robinson, B.M., Gold, J.M. (2007). Selective reinforcement learning deficits in schizophrenia support predictions from computational models of striatal–cortical dysfunction. *Biological Psychiatry*, **62**(7), 756–64.

Wechsler, D. (1999). Wechsler Abbreviated Scale of Intelligence. San Antonio, TX: The Psychological Corporation.

WHO (1992). *The ICD-10 Classification of Mental and Behavioural Disorders: Clinical Descriptions and Diagnostic Guidelines*, Geneva: World Health Organization.

Woolrich, M.W., Ripley, B.D., Brady, M., Smith, S.M. (2001). Temporal autocorrelation in univariate linear modeling of FMRI data. *NeuroImage*, **14**(6), 1370–86.

Worsley, K. (2001). Statistical analysis of activation images. In: *Functional MRI: An Introduction to Methods*, Vol. **14**, pp. 251–70, Oxford University Press: Oxford.

Yoshida, W., Dolan, R.J., Friston, K.J. (2008). Game theory of mind. *PLoS Computational Biology*, **4**(12), e1000254.