



# City Research Online

## City, University of London Institutional Repository

---

**Citation:** Fett, A-K. ORCID: 0000-0003-0282-273X (2021). Neural, behavioural and real-life correlates of social context sensitivity and social reward learning during interpersonal interactions in the schizophrenia spectrum. *Australian and New Zealand Journal of Psychiatry*,

This is the accepted version of the paper.

This version of the publication may differ from the final published version.

---

**Permanent repository link:** <https://openaccess.city.ac.uk/id/eprint/25836/>

**Link to published version:**

**Copyright and reuse:** City Research Online aims to make research outputs of City, University of London available to a wider audience. Copyright and Moral Rights remain with the author(s) and/or copyright holders. URLs from City Research Online may be freely distributed and linked to.

---

City Research Online:

<http://openaccess.city.ac.uk/>

[publications@city.ac.uk](mailto:publications@city.ac.uk)

---

**Neural, behavioural and real-life correlates of social context sensitivity and social reward learning during interpersonal interactions in the schizophrenia spectrum**

Hanssen, E. <sup>1,2,3</sup>, van Buuren, M. <sup>1</sup>, Van Atteveldt, N. <sup>1</sup>, Lemmers-Jansen, I.L.J. <sup>1,2</sup>, Fett, A.-K.J. <sup>1,2,4</sup>

<sup>1</sup>Department of Clinical, Neuro and Developmental Psychology, Faculty of Behavioural and Movement Sciences, and Institute for Brain and Behaviour (IBBA) Amsterdam, Vrije Universiteit Amsterdam, the Netherlands

<sup>2</sup>CSI Lab, Institute of Psychiatry, Psychology and Neuroscience, Department of Psychosis Studies, King's College London, London, United Kingdom

<sup>3</sup>Hersencentrum Mental Health Institute, Amsterdam, the Netherlands

<sup>4</sup>Department of Psychology, City, University of London, London, United Kingdom

\* Corresponding author: Esther Hanssen

Postal address: Vrije Universiteit Amsterdam, Faculty of Behavioural and Movement Sciences, Van der Boechorststraat 7, 1081 BT Amsterdam

Email address: emehanssen@gmail.com

Word count main body of text: 5423 | Word count abstract: 335

2 tables, 3 figures, supplementary material A – E

## **Abstract**

**Objective** Recent findings suggest that diminished processing of positive contextual information about others during interactions may contribute to social impairment in the schizophrenia spectrum (SZ). This could be due to general social context processing deficits or specific biases against positive information. We studied the impact of positive and negative social contextual information during social interactions using functional neuroimaging and probed whether these neural mechanisms were associated with real-life social functioning in SZ.

**Methods** Patients with SZ (N=23) and controls (N=25) played three multi-round trust games during fMRI scanning, with no, positive and negative information about the counterpart's trustworthiness, while all counterparts were programmed to behave trustworthy. The main outcome variable was the height of the shared amount in the trust game, i.e. investment, representing an indication of trust. The first investment in the game was considered to be basic trust, since no behavioural feedback was given yet. We performed region-of-interest analyses and examined the association with real-life social functioning using the Experience Sampling Method.

**Results** Social contextual information had no effect on patients' first investments, whereas controls made the lowest investment after negative and the highest investments after positive contextual information was provided. Over trials, patients decreased investments, suggesting reduced social reward learning, whereas controls increased

investments in response to behavioural feedback in the negative context. Patients engaged the dorsolateral prefrontal cortex (dlPFC) less than controls during context presentation and showed reduced activity within the caudate during repayments. In patients, lower investments were associated with more time spent alone and social exclusion and lower caudate activation was marginally significantly associated with higher perceived social exclusion.

**Conclusion** The failure to adapt trust to positive and negative social contexts suggests that patients have a general insensitivity to prior social information, indicating top-down processing impairments. In addition, patients show reduced sensitivity to social reward, i.e. bottom-up processing deficits. Moreover, lower trust and lower neural activation were related to lower real-life social functioning. Together, these findings indicate that improving trust and social interactions in SZ needs a multi-faceted approach that targets both mechanisms.

## **1.1 Introduction**

The ability to integrate social contextual information and behavioural feedback from others is necessary for successful social interactions (Ruz et al., 2011) and is an important foundation of trust in social relationships (Lewicki and Wiethoff, 2006). Individuals with a schizophrenia spectrum (SZ) diagnosis, hereafter referred to as patients, show a reduced ability in judging social signals (Penn et al., 2008). Previous research suggests two underlying mechanisms. First, there are deficits in learning from others' behavioural feedback during social interactions, suggesting problems in bottom-up mechanisms. Second, there are problems with the integration of a-priori contextual information in a top-down way (Chung et al., 2010; Hooker et al., 2011). Here we investigate how these two mechanisms of social information processing impact on social behaviour in real-time interactions, using a modified neuroeconomic trust game (Fett et al., 2015).

Studies employing the trust game (Berg et al., 1995) in SZ have demonstrated that patients invested lower initial amounts, indicating lower trust towards others (Fett et al., 2012; Gromann et al., 2013; Lemmers-Jansen et al., 2018). In addition, SZ has been associated with a reduced ability to use others' cooperative behavioural feedback to adjust trusting behaviour, i.e. bottom-up processing (Fett et al., 2012; Gromann et al., 2013). Research on the effects of social contextual information on trust in the general population demonstrated that trust increases in response to a trustworthy interaction partner, showing a strong impact of a-priori information on trust in a top-down manner (Delgado et al.,

2005). Patients in contrast show a diminished sensitivity to such prior positive information (Fett et al., 2012). The reduced ability to use bottom-up information and not being able to overcome distrust during positive interactions in response to such information, may explain real-life social impairment seen in SZ (Velthorst et al., 2016) and could be tackled with cognitive bias modification or other ways of cognitive remediation. However, it has not been investigated whether the insensitivity to social context reflects a more general processing deficit or a specific bias against positive social information.

Three core cognitive mechanisms have been suggested to underlie trust and decision making in social contexts (Declerck et al., 2013). First, context processing and cognitive control, which are subserved by the dorsolateral prefrontal cortex (dlPFC) (MacDonald et al., 2000). Second, theory of mind, i.e. the ability to infer the mental states of others, in which medial prefrontal cortex (mPFC) and temporo-parietal junction (TPJ) are implicated as key regions (Carrington and Bailey, 2009; Schurz et al., 2014). Third, reward processing, which strongly involves the caudate nucleus (Sanfey, 2007; Krach et al., 2010).

Deficits in mentalizing, social reward processing (bottom-up) and social context processing (top-down) have been suggested to underlie lower trust, paranoia and social disconnection in SZ (Couture et al., 2006; Velthorst et al., 2016; Kapur et al., 2005). In SZ, prior studies have reported reduced activation within the dlPFC during context

processing (Niendam et al., 2014; Barch and Ceaser, 2012), within the mPFC and TPJ during mentalising (Pinkham et al., 2008; Green et al., 2015; Lee et al., 2004), and within the caudate during both non-social (Juckel et al., 2006b; Murray et al., 2008) and, more importantly, social reward processing (Gromann et al., 2013; Fett et al., 2019). This earlier work leads to the hypotheses that the dlPFC, mPFC, TPJ and caudate play an important role in disturbed social decision making and context processing in SZ.

In this first-time investigation of the underlying mechanisms of disturbed trust and social interaction in SZ, we therefore probed the impact of different social contexts and investigated the underlying neural correlates in patients with SZ and healthy controls , using a modified version of an interactive trust game while measuring brain activity with fMRI. Since social interactions are embedded in peoples' daily lives in a complex way, it is important to elucidate the association between the neural processes underlying social interactions and daily-life social engagement in SZ. To investigate this, we combined fMRI and the Experience Sampling Method (ESM), a diary method (Delespaul, 1995). Initial ESM studies (Moran et al., 2019; Kluge et al., 2018) started to investigate how brain activation during task-based fMRI translates to real-world functioning. This method ensures high ecological validity because it allows for real-time monitoring of behaviour in daily-life contexts. In patients, Moran et al. (2019) found that greater hemodynamic signal change during (non-social) reward anticipation in caudate, insula and anterior

cingulate was associated with greater anticipated pleasure and motivation for daily-life activities.

We hypothesized that: a) patients would show a general reduced sensitivity to prior information about the counterpart reflected in no differences in baseline trust between conditions, i.e. first investments. Controls would increase investments from the negative to the positive context; b) patients would not increase trust over trials in response to benevolent behavioural feedback, whereas controls would do so; c) patients would engage the regions-of-interest (ROIs) to a lesser extent than controls: i) in the left dlPFC during context presentation and investment (i.e. trusting) decisions, because of its specific role in context processing; ii) in the mPFC and right TPJ during context presentation and investment decisions, given their role in mentalising mechanisms, which we expected to a greater extent in controls while processing prior social information compared to no information; iii) in the right caudate nucleus during the partner's repayments (i.e. receipt of social reward); d) for patients, reduced trust and reduced activation in the ROIs during the trust game is associated with lower daily life social functioning, i.e. more time spent alone, higher perceived social exclusion and lower perceived relationship quality, measured by ESM.

## **2.1 Methods**

### *2.1 Subjects*



Twenty-five patients with a SZ diagnosis and 26 controls without a personal or family history of SZ were included (for recruitment see **supplement - A**). Inclusion criteria were: 1) age 18 to 65, 2) good understanding of the English language, 3)  $IQ > 70$ . An additional criterion for patients was a SZ diagnosis according to ICD-10 (WHO, 1992), which was confirmed with the treating NHS clinician. Exclusion criteria were: 1) a history of any neurological conditions, 2) a diagnosis of alcohol/drug dependence within six months. One control subject did not complete MRI scanning due to anxiety. Two patients were excluded from analyses due to excessive movement (framewise displacement  $\geq 1.5$  mm in  $\geq 20\%$  of the volumes per run). Therefore, analyses were performed on 23 patients and 25 controls. Forty-four participants completed the ESM measurements (20 patients, 24 controls). The London-Harrow Research Ethics Committee [14/LO/0071] approved this study.

## *2.2 Measures*

### *2.2.1 Estimated cognitive ability*

To assess an estimated cognitive ability, an abbreviated two-test version of the Wechsler Abbreviated Scale of Intelligence (WASI) was used (Wechsler, 1999), which consisted of the vocabulary subtest and the matrix reasoning subtest. WASI scores are reported in **Table 1**.

### *2.2.2 Symptoms*

The Positive and Negative Syndrome Scale (PANSS) semi-structured interview was used to measure symptom severity in the two weeks prior to testing in patients (Kay et al., 1987). Fourteen items evaluate the severity of positive and negative symptoms (1 = absent to 7 = extreme). PANSS scores are reported in **Table 1**.

### *2.2.3 Trust game*

To measure the impact of social context processing in social interactions, we employed a modified multi-round trust game (Berg et al., 1995; King-Casas et al., 2005; Gromann et al., 2013). In a multi-round classic trust game, the first player, i.e. the investor, is given an initial endowment of £10 and has to invest a chosen amount between zero and ten pounds. This amount is tripled and given to the second player, i.e. the trustee. The trustee then decides whether and how much of the tripled amount he or she wants to give back to the investor. The chosen amount to invest by the investor reflects trust (given that the trustee can choose not to return any money). In our study, participants played the role of the investor and played the trust game three times, with three different hypothetical counterparts. In one game (condition), the trust game was presented without prior social contextual information, i.e. a classic multi-round trust game, while the other two conditions were modified to examine social context processing. In the negative and positive context condition, participants first played three ‘blind’ rounds, without seeing

the repayments of the interaction partner. These rounds were implemented to establish the cooperativeness of the trustee (Fett et al., 2012). When making their blind investments participants saw the following the message: ‘Determining E.H.’s initial average returns’. After these three ‘blind’ investments, they viewed the following message on the screen pointing out the average returns (more or less than invested) of the trustee during the blind investments (i.e. the social context): ‘On average your partner E.H. returned more/less than you invested’. This message was shown before each trial in the trust game. Participants completed a total of 120 trials (60 experimental, 60 control), equally divided over the three conditions (no context, positive context and negative context). Control trials were included to control for general effects of motor and visual activation elicited by the task. An experimental and control trial with the respective phases and timings are displayed in Figure 1.

Participants were instructed that they played with a real human player via the Internet but were actually playing with a pre-programmed computer that behaved in the same benevolent probabilistic manner in all three contexts (for algorithm see **supplement - B**). After completion of the trust game, participants were asked whether they thought the other players were real and trustworthy, on a 7-point Likert scale. Ratings on whether the other players were unreal did not differ between groups (16 % controls and 13 % patients,

$p > .83$ ), and were unrelated to investments in the trust game ( $p > .82$ ). Also, the ratings of trustworthiness of the interacting partners did not differ between groups ( $p > .59$ ).

#### *2.2.4 Experience Sampling Method (ESM) – Measurement of social engagement*

ESM (Palmier-Claus et al., 2011), a structured diary technique, was used to measure social engagement in daily life. The ESM device (iPod) gave a signal to fill in the questionnaire 10 times a day, by means of a pseudo-random ‘beep’ on 7 consecutive days. We included several questions to probe real-life social functioning in terms of social engagement, social exclusion and quality of social relationships. The question that was used assessing social engagement was ‘*Are you alone?*’ (yes/no). Perceived social exclusion was assessed when individuals were alone using an average of the two items: ‘*I feel lonely*’ and ‘*I feel excluded*’ (Cronbach’s alpha 0.82). The perceived quality of social relationships was assessed when individuals were in social company with an average of the four items: ‘*I like the person(s) I am with*’, ‘*I feel close to them*’, ‘*They are dependable*’, ‘*I trust them*’ (Cronbach’s alpha 0.90). ESM items were rated on a 7-point Likert scale.

#### *2.3 Procedure*

The two sessions took place at the Institute of Psychiatry, Psychology, Neuroscience (IoPPN), King's College London. Participants gave written informed consent before the study.

Participants first completed a demographic questionnaire and several practice trials before playing the trust game in the MRI scanner. They were told that they would receive the earnings from one randomly selected round to keep them motivated (between £0 and £30). For fairness reasons all participants received a payment of £5. After the trust game, they completed the questionnaire on their perception of the game partners. Next, the PANSS interview was administered. Last, an explanation of the iPod was given to participants and they completed one practice questionnaire. The morning after the first session the 7-days ESM data collection started. All participants were contacted by phone on day 2 for guidance in case of any problems or difficulties with the iPod. The second session consisted of assessment of the WASI subtests. Participants handed in the iPod and experiences were discussed. At the end of the study each participant received payment (£60 + £5) for participation.

## *2.4 Statistical analyses*

### *2.4.1 Behavioural analyses*

Statistical analyses were performed using STATA version 14 (StataCorp, 2015). We examined group differences in demographics using chi-square tests and regression analyses. Investments were analysed using mixed effects multilevel random regression analyses (MIXED) to account for repeated measurements within persons, with (first) investment as dependent variables and with group (control, patient) and context (negative, no and positive) and their respective interactions as independent variables. For investments over trials, trial number (1-20) was added to the model. Mixed effects multilevel random regression analyses (MIXED) were used to examine associations between investments and ESM indices of real-life social functioning (%alone, social exclusion and quality of social relationships) across contexts. Interactions were probed with the CONTRAST command. Analyses with a significant group effect are additionally reported with estimated IQ as a covariate. Additional analyses on association with symptoms are reported in **supplement - E**.

#### *2.4.2 fMRI data acquisition & scanning parameters*

Imaging data were acquired using a 3 T GE Signa Neuro-optimized MR System at the Centre of Neuroimaging Science of the IoPPN. Functional images were acquired by a T2\*-weighted echo-planar imaging sequence scanning 39 axial slices of 3.0 mm thick with 0.3 mm gap. The in-plane resolution was  $3.3 \times 3.3$  mm (FOV  $211 \times 211$ ), flip angle:  $75^\circ$ , TR = 2.00 s; TE = 30 ms. There were 413 volumes per run. For anatomical reference,

a T1-weighted image (196 slices; isotropic voxels of 1.2 mm; TR 7.312 ms; TE 3.016 ms; flip angle: 11°; FOV 270 mm) was acquired.

Imaging data were analysed using Statistical Parametric Mapping 12 (SPM, 2014). Pre-processing of the functional images consisted of realign and unwarp, and coregistration to individual anatomical images. Next, using unified, segmentation, anatomical images were segmented and normalization parameters estimated. These parameters were used to transform functional and anatomical images to a Montreal Neurological Institute (MNI) template. Subsequently, smoothing was applied (Gaussian kernel 6 mm full-width at half-maximum), and the last three volumes were removed, at the ending of the task.

#### *2.4.3 ROI analyses*

A general linear model was used per run in which the three phases in the game were modelled as regressors of interest for the experimental and control condition separately (see Figure 1). The cue phase was time-locked to the start of each trial (duration 2 s), the investment phase started after the investment cue (duration 4 s) and the repayment phase started at the beginning of the repayments shown and was modelled until the end of the displayed totals (duration 5 - 8 s). All other game phases were combined into one regressor of no interest (the investment cue, the invested amounts, waiting for the

partners response and the two fixation crosses). All phases were modelled using a box-car function convolved with the hemodynamic response function (Friston et al., 1995). To correct for motion, the six realignment parameters and regressors for volumes detected as motion-corrupted, calculated by DVARS metric as implemented in FSL, version 6.00 (FMRIB's Software Library, [www.fmrib.ox.ac.uk/fsl](http://www.fmrib.ox.ac.uk/fsl)) by FSL Motion Outliers (<https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FSLMotionOutliers>) were included in the design matrix, making the total number of regressors in the model variable for each individual, with a minimum of ten regressors (cue phase, investment phase, repayment phase, regressor of no interest and six motion parameters). A high pass filter of 128 seconds was used. Subsequently, for each phase of interest, contrast images were created by contrasting a specific phase of the experimental condition with the corresponding phase in the control condition.

A-priori ROI analyses were performed. Talairach coordinates were converted to MNI space (<https://bioimagesuiteweb.github.io/webapp/mni2tal.html>), resulting in the following ROI MNI coordinates: right caudate (17, 20, 3), right TPJ (50, -56, 27), mPFC (-3, 64, 24) (Gromann et al., 2014) and the left dlPFC (-43, 18, 29) (MacDonald and Carter, 2003). ROIs were created in MarsBar with an 8 mm sphere (version 0.44; <http://marsbar.sourceforge.net>). For each ROI and each subject, average signal change (beta estimate) was extracted to investigate group and context effects, and to test association between the fMRI and the ESM data. Additional analyses on association with



symptoms (PANSS positive scale , PANSS negative scale, PANSS suspiciousness, PANSS amotivation factor and PANSS diminished expression factor; **supplement - E**) and IQ are reported. Mixed effects multilevel random regression analyses (MIXED) were used to examine associations between ROI beta estimates and ESM indices of real-life social functioning (%alone, social exclusion and quality of social relationships) across contexts. The results of the ROI-based analyses were Bonferroni corrected at alpha levels of 0.0125 per test (0.05/4, as tests were performed with data from four ROIs).

#### *2.4.4 Exploratory analyses*

We also performed exploratory whole-brain analyses, investigating neural activation beyond the predefined ROIs for all three game phases. Analyses were corrected at family-wise-error (FWE) whole-brain cluster significance threshold of  $p = 0.05$  (see **supplement - C**).

### **3. Results**

#### *3.1 Behavioural analysis*

Groups did not differ in age and gender. Patients had a lower estimated cognitive ability than controls. Percentage of time spent alone was higher in patients than controls. Patients

felt lonelier and more excluded, but reported a similar quality of their social relationships compared to controls (see **Table 1**).

### *3.1.1 Baseline trust: context effect and group differences in first investments*

First investments were examined to establish context effects on baseline trust (Table 2). There was a significant group-by-context interaction ( $\chi^2(2) = 7.34, p = 0.02$ ), which remained significant when estimated IQ was added to the model ( $\chi^2(2) = 7.34, p = 0.02$ ). Estimated IQ was not significantly associated with baseline trust ( $p = .32$ ). The context effect was only significant in controls, who made lower investments in the negative than the no-context condition ( $b = -1.28, 95\% \text{ CI } [-2.23, -.33], p = .009$ ) and higher investments in the positive than the no-context condition ( $b = 1.32, 95\% \text{ CI } [.35, 2.28], p = 0.007$ ). Patients' first investments did not differ by context (both  $p > .75$ ). In all three contexts, first investments did not differ significantly between groups (all  $p > .07$ ).

### *3.1.2 Changes in trust over trials: context effect and group differences*

We examined interference of context information on changes in trust (i.e. investments) over time (Figures 2 a and b). The 3-way interaction between group, context and trial was marginally significant ( $\chi^2(2) = 5.07, p = .07$ ) and analyses by group showed a marginally significant interaction of context-by-trial number in controls ( $\chi^2(2) = 5.26, p = .07$ ), but not patients ( $p = .21$ ). Across trials, both groups showed a context effect with significant

differences between negative and no context, which was larger in controls than patients (controls:  $b = -.64$ , 95% CI [-.88, -.39],  $p < .0001$ , patients:  $b = -.39$ , 95% CI [-.67, -.11],  $p = .006$ ), but no differences between the positive and no context (both  $p > .41$ ). Both groups showed a significant main effect of trial number; controls increased investments over time, while patients decreased their investments (controls:  $b = .03$ , 95% CI [.01, .04],  $p = .002$ ; patients:  $b = -.04$ , 95% CI [-.06, -.02],  $p < .0001$ ).

Additional, analyses by context showed that the effect of trial number was most pronounced and in opposite directions in patients and controls in the negative context (group-by-trial number interaction ( $\chi^2(2) = 21.21$ ,  $p < .0001$ ), controls:  $b = .06$ , 95% CI [.03, .08],  $p < .001$ ; patients:  $b = -.05$ , 95% CI [-.08, -.02],  $p < .001$ ). In the no context condition, there was a group-by-trial number interaction ( $\chi^2(2) = 8.37$ ,  $p = .004$ ); no significant changes in investments were found controls ( $p = .51$ ), but patients invested significantly less over time ( $b = -.06$ , 95% CI [-.10, -.01],  $p = .02$ ). Groups did not differ significantly in the positive context condition and did not show investment changes over trials (all  $p > .30$ ). Additional analyses between symptoms and behavioural results are reported in **supplement - E**.

## 3.2 *fMRI analysis*

### 3.2.1 *ROI analyses by trust game phase*

*Cue phase.* There were no significant group-by-context interaction in any ROIs (all  $p > 0.12$ ). A significant group effect in the left dlPFC showed lower activation in patients than controls ( $b = -0.59$ , 95% CI [-1.00, -.18],  $p = .004$ , Figure 3a). The effect remained significant when the IQ estimate was added to the model as covariate ( $b = -0.71$ , 95% CI [-1.24, -.17],  $p = .009$ ). The IQ estimate was not significantly related to left dlPFC activation ( $p = .51$ ).

*Investment phase.* There were no significant group-by-context interactions (all  $p > .08$ ), nor main effects of group (all  $p > .50$ ) or context in any of the ROIs (all  $p > .11$ ).

*Repayment phase.* There were no significant group-by-context interactions or context effects for the right caudate or mPFC. A significant group effect showed lower right caudate activation in patients compared to controls ( $b = -.14$ , , 95% CI [-.25, -.03],  $p = .01$ , Figure 3b). This effect remained significant when the IQ estimate was added to the model as covariate ( $b = -0.16$ , 95% CI [-.30, .004],  $p = .028$ ). The IQ estimate was not significantly related to right caudate activation ( $p = .66$ ). Higher mPFC activation was found in patients than controls ( $b = .32$ , 95% CI [.003, .62],  $p = .047$ ). However, this result did not survive Bonferroni correction. There were no interaction, group or context effects in any other ROI (all  $p > .16$ ).

Given that we observed a behavioural group difference in investments over trials, we performed exploratory analyses probing the change in ROI activation over investment

trials. The results are reported in **supplement - D**. Additional analyses between symptoms and ROI results are reported in **supplement - E**.

### *3.2.2 Associations between investment, ROI activation and real-life social functioning in patients*

We were specifically interested to examine whether lower neural activation within the left dlPFC and the right caudate found in patients was related to the level of real-world social functioning. In addition, we explored this association in the less robustly increased mPFC signal in patients.

Across contexts, higher investments were significantly associated with less time spent being alone ( $b = -.02$ , 95% CI [-.05, -.001],  $p = .04$ ). A significant interaction between context and perceived relationship quality on investments ( $X(2) = 23.01$ ,  $p < .0001$ ), showed trend-level associations between higher relationship quality and higher investments in the positive ( $b = .85$ , 95% CI [-.06, 1.78],  $p = 0.07$ ) and negative ( $b = .94$ , 95% CI [-.06, 1.78],  $p = 0.06$ ), but not the no context condition ( $p = .79$ ). There was a significant interaction between context and social exclusion on investments ( $X(2) = 13.19$ ,  $p < .0001$ ), indicating an association between lower perceived social exclusion and higher investments the no context condition only ( $b = -1.17$ , 95% CI [-1.97, -.35],  $p = 0.005$ , other  $p > .23$ ).

There were no significant associations between any measure of real-life social functioning and dlPFC activation (all  $p > .21$ ). A significant interaction between social exclusion and context emerged for the right caudate ( $X(2) = 8.51, p = .01$ ); lower perceived social exclusion was marginally significantly associated with higher caudate activation in the positive context only ( $b = -.17, 95\% \text{ CI } [-.36, -.018], p = .07$ , all other  $p > .18$ ). No significant associations were present with social relationship quality or time spent alone (both  $p > .17$ ). Higher perceived social relationship quality was associated with higher mPFC activation ( $b = .27, 95\% \text{ CI } [.002, .54], p = .048$ ). There were no significant associations between mPFC activation and social exclusion or time spent alone (both  $p > .22$ ).

#### **4. Discussion**

Using a novel, modified trust game, we examined the impact of social context on trust and social reward and the underlying neural activation patterns during real-time social interactions. In addition, we probed the associations with real-life social functioning in SZ. Patients showed no differential effect of social context on first investments, regardless of the valence or the absence of a context, whereas controls showed the expected distinct context effect with highest levels of trust in response to positive and lowest levels of trust in response to negative social information. This suggests a general insensitivity to social

context instead of a bias against positive social contextual information in patients. Patients also did not increase trust in response to benevolent behavioural feedback whereas controls did. Within the patient group, the findings indicate an association between lower trust and lower real-world social functioning.

We found overall lower activation in the left dlPFC during context presentation and less engagement of the right caudate nucleus during repayments in patients compared to controls. These results suggest that SZ is associated with a general insensitivity to social contexts and with a reduced sensitivity to social reward. On the neural level, we also found an association between caudate activation and lower real-life social functioning in patients.

#### *4.1 Social context effect on baseline trust*

For baseline trust (i.e. the first investment where partner feedback has not yet been received), we found reduced sensitivity to positive and negative social contextual information in patients compared to controls. This seems to reflect a general insensitivity to social contextual information. These results strengthen and extend previous evidence of a social context top-down processing deficit in SZ (Fett et al., 2012; Niendam et al., 2014; Baez et al., 2013), however, this may not generalize to tasks concerning different types of information processing. This insensitivity points to persistent a-priori beliefs

about other people in SZ. It is important to unravel whether this context sensitivity is a risk factor for developing a SZ or secondary to the disorder or related factors, by examining whether this insensitivity is also found in first-episode patients and individuals at high risk for SZ. In the current study, patients tended to approach social interactions with similar trusting behaviour as controls, in line with findings in first-episode patients (Fett et al., 2019), but contrasting results have also been found in chronic and first-episode patients and individuals at clinical high risk (Fett et al., 2012; Gromann et al., 2013; Lemmers-Jansen et al., 2018).

#### *4.2 Trust over time: the effects of benevolent feedback*

As hypothesized, patients did not increase trust in response to benevolent partner feedback, in line with previous literature in patients with chronic psychosis (Fett et al., 2012; Gromann et al., 2013). Controls, however, increased trust in the negative context and showed stable levels of trust after no and positive information, suggesting that prior positive beliefs about others were matched by the benevolent partner feedback. Patients seem to have difficulties to overcome the given prior negative information, i.e. persistent a-priori negative beliefs about others. Patients showed a tendency to reduce trust, even though subjective ratings of trustworthiness of the interacting partner were similar for both groups. Other studies have shown that patients show reduced sensitivity to social



rewards, such as smiles (Catalano et al., 2018). Our results strengthen the evidence that patients have deficits in bottom-up processing of partner feedback, which might be due to an insensitivity to social reward (Fett et al., 2012; Gromann et al., 2013). The reduced ability to increase trust could explain patients' reduced motivation to engage in social behaviour (Krach et al., 2010).

#### *4.3 Neural findings during social interactions with contextual information*

In support of our hypotheses, we found reduced left dlPFC activation in patients compared to controls during social context presentation. The dlPFC has been implicated in deficits in non-social context processing in schizophrenia (Barch and Ceaser, 2012; Niendam et al., 2014) and is viewed as key region in top-down cognitive control (MacDonald et al., 2000). The impairment in top-down modulation of trust in response to contextual information may be related to reduced engagement of the dlPFC in SZ. This result that dlPFC activation was lower in all contexts, supports a general insensitivity rather than a specific bias. We did not see any group differences in ROI activation during investments. However, during the repayment phase, where reward processing takes place, we found context independent blunted activation in the right caudate nucleus in patients compared to controls. The caudate is highly innervated by dopaminergic neurons (Björklund and Dunnett, 2007). Aberrant regulation of dopamine is thought to play a key

role in SZ and reward processing (Howes and Kapur, 2009), and may account for the insensitivity to social feedback, i.e. reward. This finding adds to evidence showing deficits in social reward processing in SZ (Gromann et al., 2013; Lee et al., 2018) and may explain why patients fail to increase trust over time. A less robust effect was found in the mPFC; activation was higher in patients compared to controls. This was unexpected, but may be related to the role of the mPFC in reward-based action selection (Euston et al., 2012). We speculate that this may point to a compensation mechanism for reduced engagement of the caudate, which is tentatively supported by the correlations with real-world social outcomes, which are discussed in the following section.

#### *4.4 Associations with real-life social engagement*

Previous social neuroscience studies have started to investigate the link between the reward related processing in the brain and daily life (Bakker et al., 2019; Moran et al., 2019), yet associations with real-life social functioning have not been considered in schizophrenia, despite yielding meaningful insights in healthy controls (Bickart et al., 2011; Kanai et al., 2012; Lewis et al., 2011). We found that in real life, patients spent more time alone than healthy controls, in line with previous work (Pinkham and Penn, 2006; Velthorst et al., 2016). They reported higher feelings of loneliness and social exclusion, but a similarly good relationship quality as controls. On the behavioural level, higher investments were associated with higher social functioning in real life. A higher

level of trust may create an advantageous basis for engaging in meaningful social relationships (Campellone et al., 2016). This could positively impact on social support networks, which in turn could aid in recovery (Corrigan and Phelan, 2004). At the neural level, we found that in patients, lower activation in the caudate in the positive context was marginally significantly associated with higher perceived social exclusion. More engagement of this key reward area may be related to heightened experience of positive social interaction, i.e., social reward (Gromann et al., 2013). This could lead to a higher sense of belonging or inclusion in social relationships. In addition, higher activation in the mPFC was associated with higher reported relationship quality in daily life, which would support the role of mentalizing abilities (Schurz et al., 2014). However, it is possible that this finding also relates to higher feelings of (social) reward (Euston et al., 2012). Future evidence from larger replication studies will be needed to strengthen and support these initial results.

#### *4.5 Limitations*

The current results should be interpreted in light of the following limitations. First, our study is an initial investigation of social context processing in relation to neural activation and in relation to daily life social functioning in a relatively small sample of relatively stable patients with a SZ diagnosis. Specifically, our patient sample had relatively low symptom levels: negative and positive scores of respectively 14.74 (SD = 6.15) and 12.26

(SD = 3.25). We defined these scores as relatively low, since, with respect to the 7-point Likert scale of the PANSS items and the number of items in the positive and negative subscale, the item scores of 1 and 2 are respectively absent and minimal (Kay et al., 1987), (see also **supplement - E**). Consequently, our results need to be interpreted with caution when it comes to generalizability and larger replication studies are warranted before any firm conclusions can be made. Second, medication type has been found to have an effect on reward processing in SZ (Juckel et al., 2006a), however, the majority of participants in our sample was on atypical antipsychotics (see Table 1), which are thought to normalize reward processing (Nielsen et al., 2012; Schlagenhauf et al., 2008). In addition, healthy first-degree relatives of patients with SZ show reduced social reward processing, without any medication confounds (Gromann et al., 2014; Hanssen et al., 2020). Therefore, the current findings are not likely to represent an enhancement of the effects, they may even reflect an underestimation, compared to expected effects in unmedicated patients. Studying unmedicated patients is valuable, but poses a great challenge due to the clinical reality. Third, subjective self-report measures, like ESM, may raise the question of response accuracy or social desirable, which may be related to suspiciousness. However, ESM is a widely used and well-validated method in psychiatric and schizophrenia samples (Myin-Germeys et al., 2018; Delespaul, 1995) and has the advantage that it does not rely on retrospective recall (Potheegadoo et al., 2012). Moreover, our sample showed a good compliance and did not report any issues related to

the use of the ESM app during the debriefing. We therefore have no grounds to assume that responses were inaccurate. Last, not all participants believed they were playing with a real human player, however, this was not associated with investments and adequate changes in investments in controls after context information and behavioural feedback indicated that the experimental manipulation was effective.

#### *4.6 Concluding remarks*

This is the first study to investigate the neural correlates of social context processing and the link with real-life social functioning in SZ. We provide evidence that patients do not modulate their behaviour in response to social context information (i.e. top-down processing) and positive behavioural cues of others (i.e. bottom-up processing). Our results point to a general reduced sensitivity to social information on a behavioural and neural level. The findings suggest that increasing trust may facilitate social engagement in patients. Also, indices of real-life social functioning seem to be associated with lower neural activity in reward- and context processing (i.e. cognitive control) brain areas. The current study suggests that improving social interaction in SZ requires a multi-faceted approach in clinical practice, which considers both bottom-up and top-down processing of social information. An example of an intervention that incorporates these two facets is the Social Cognition and Interaction Training (SCIT) (Penn et al., 2007), targeting social skills bottom-up and social cognition top-down, which shows promising effects on social

functioning. Clinical practice will benefit from newly developed treatments, building on the SCIT for instance, targeting these facets in daily life by implementing an ecological momentary intervention by means of a smartphone based treatment. In addition, future studies are warranted to investigate how this is related to fine grained assessments of functional capacity and real world social functioning.

### **Acknowledgements**

We would like to thank the participants of the DECOP project and the NIHR CROs, as well as Farah Yakub, Marieke Helmich, Marina Volguraki, Katie Wong, Mathew Harvey and Tracy Bobin for their support with testing and recruitment.

### **References**

- Baez S, Herrera E, Villarin L, et al. (2013) Contextual social cognition impairments in schizophrenia and bipolar disorder. *PLoS One* 8: e57664.
- Bakker JM, Goossens L, Kumar P, et al. (2019) From laboratory to life: associating brain reward processing with real-life motivated behaviour and symptoms of depression in non-help-seeking young adults. *Psychological medicine* 49: 2441-2451.
- Barch DM and Ceaser A. (2012) Cognition in schizophrenia: core psychological and neural mechanisms. *Trends in cognitive sciences* 16: 27-34.
- Berg J, Dickhaut J and McCabe K. (1995) Trust, reciprocity, and social history. *Games and economic behavior* 10: 122-142.
- Bickart KC, Wright CI, Dautoff RJ, et al. (2011) Amygdala volume and social network size in humans. *Nature neuroscience* 14: 163.

- Björklund A and Dunnett SB. (2007) Dopamine neuron systems in the brain: an update. *Trends in neurosciences* 30: 194-202.
- Campellone TR, Fisher AJ and Kring AM. (2016) Using social outcomes to inform decision-making in schizophrenia: Relationships with symptoms and functioning. *Journal of abnormal psychology* 125: 310.
- Carrington SJ and Bailey AJ. (2009) Are there theory of mind regions in the brain? A review of the neuroimaging literature. *Human brain mapping* 30: 2313-2335.
- Catalano LT, Heerey EA and Gold JM. (2018) The valuation of social rewards in schizophrenia. *Journal of abnormal psychology* 127: 602.
- Chung YS, Mathews JR and Barch DM. (2010) The effect of context processing on different aspects of social cognition in schizophrenia. *Schizophrenia bulletin* 37: 1048-1056.
- Corrigan PW and Phelan SM. (2004) Social support and recovery in people with serious mental illnesses. *Community mental health journal* 40: 513-523.
- Couture SM, Penn DL and Roberts DL. (2006) The functional significance of social cognition in schizophrenia: a review. *Schizophrenia bulletin* 32: S44-S63.
- Declerck CH, Boone C and Emonds G. (2013) When do people cooperate? The neuroeconomics of prosocial decision making. *Brain and cognition* 81: 95-117.
- Delespaul PA. (1995) Assessing schizophrenia in daily life: The experience sampling method. Maastricht university.
- Delgado MR, Frank RH and Phelps EA. (2005) Perceptions of moral character modulate the neural systems of reward during the trust game. *Nature neuroscience* 8: 1611.
- Euston DR, Gruber AJ and McNaughton BL. (2012) The role of medial prefrontal cortex in memory and decision making. *Neuron* 76: 1057-1070.
- Fett A-KJ, Mouchlianitis E, Gromann PM, et al. (2019) The neural mechanisms of social reward in early psychosis. *Social cognitive and affective neuroscience* 14: 861-870.
- Fett A-KJ, Shergill SS, Joyce DW, et al. (2012) To trust or not to trust: the dynamics of social interaction in psychosis. *Brain* 135: 976-984.
- Fett AK, Shergill SS and Krabbendam L. (2015) Social neuroscience in psychiatry: unravelling the neural mechanisms of social dysfunction. *Psychol Med* 45: 1145-1165.
- Friston KJ, Frith CD, Turner R, et al. (1995) Characterizing evoked hemodynamics with fMRI. *Neuroimage* 2: 157-165.
- Green MF, Horan WP and Lee J. (2015) Social cognition in schizophrenia. *Nature Reviews Neuroscience* 16: 620.
- Gromann P, Heslenfeld DJ, Fett AK, et al. (2013) Trust versus paranoia: abnormal response to social reward in psychotic illness. *Brain* 136: 1968-1975.
- Gromann P, Shergill S, de Haan L, et al. (2014) Reduced brain reward response during cooperation in first-degree relatives of patients with psychosis: an fMRI study. *Psychological medicine* 44: 3445-3454.
- Hanssen E, Krabbendam L, Robberegt S, et al. (2020) Social and non-social reward learning reduced and related to a familial vulnerability in schizophrenia spectrum disorders. *Schizophrenia research* 215: 256-262.

- Hooker CI, Tully LM, Verosky SC, et al. (2011) Can I trust you? Negative affective priming influences social judgments in schizophrenia. *Journal of abnormal psychology* 120: 98.
- Howes OD and Kapur S. (2009) The dopamine hypothesis of schizophrenia: version III—the final common pathway. *Schizophrenia bulletin* 35: 549-562.
- Juckel G, Schlagenhauf F, Koslowski M, et al. (2006a) Dysfunction of ventral striatal reward prediction in schizophrenic patients treated with typical, not atypical, neuroleptics. *Psychopharmacology* 187: 222-228.
- Juckel G, Schlagenhauf F, Koslowski M, et al. (2006b) Dysfunction of ventral striatal reward prediction in schizophrenia. *Neuroimage* 29: 409-416.
- Kanai R, Bahrami B, Duchaine B, et al. (2012) Brain structure links loneliness to social perception. *Current Biology* 22: 1975-1979.
- Kapur S, Mizrahi R and Li M. (2005) From dopamine to salience to psychosis—linking biology, pharmacology and phenomenology of psychosis. *Schizophrenia research* 79: 59-68.
- Kay SR, Fiszbein A and Opfer LA. (1987) The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophrenia bulletin* 13: 261.
- King-Casas B, Tomlin D, Anen C, et al. (2005) Getting to know you: reputation and trust in a two-person economic exchange. *Science* 308: 78-83.
- Kluge A, Kirschner M, Hager OM, et al. (2018) Combining actigraphy, ecological momentary assessment and neuroimaging to study apathy in patients with schizophrenia. *Schizophr Res* 195: 176-182.
- Krach S, Paulus FM, Bodden M, et al. (2010) The rewarding nature of social interactions. *Frontiers in behavioral neuroscience* 4: 22.
- Lee J, Jimenez AM, Reavis EA, et al. (2018) Reduced Neural Sensitivity to Social vs Nonsocial Reward in Schizophrenia. *Schizophrenia bulletin*.
- Lee K-H, Farrow T, Spence S, et al. (2004) Social cognition, brain networks and schizophrenia. *Psychological medicine* 34: 391-400.
- Lemmers-Jansen IL, Fett A-KJ, Hanssen E, et al. (2018) Learning to trust: social feedback normalizes trust behavior in first-episode psychosis and clinical high risk. *Psychological medicine*: 1-11.
- Lewicki RJ and Wiethoff C. (2006) Trust, trust development, and trust repair. *The handbook of conflict resolution: Theory and practice* 2: 92-119.
- Lewis PA, Rezaie R, Brown R, et al. (2011) Ventromedial prefrontal volume predicts understanding of others and social network size. *Neuroimage* 57: 1624-1629.
- MacDonald AW and Carter CS. (2003) Event-related fMRI study of context processing in dorsolateral prefrontal cortex of patients with schizophrenia. *Journal of abnormal psychology* 112: 689.
- MacDonald AW, Cohen JD, Stenger VA, et al. (2000) Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science* 288: 1835-1838.
- Moran EK, Culbreth AJ, Kandala S, et al. (2019) From neuroimaging to daily functioning: A multimethod analysis of reward anticipation in people with schizophrenia. *J Abnorm Psychol* 128: 723-734.



- Murray G, Corlett P, Clark L, et al. (2008) Substantia nigra/ventral tegmental reward prediction error disruption in psychosis. *Molecular psychiatry* 13: 267.
- Myin-Germeys I, Kasanova Z, Vaessen T, et al. (2018) Experience sampling methodology in mental health research: new insights and technical developments. *World Psychiatry* 17: 123-132.
- Nielsen MO, Rostrup E, Wulff S, et al. (2012) Improvement of brain reward abnormalities by antipsychotic monotherapy in schizophrenia. *Archives of general psychiatry* 69: 1195-1204.
- Niendam TA, Lesh TA, Yoon J, et al. (2014) Impaired context processing as a potential marker of psychosis risk state. *Psychiatry Research: Neuroimaging* 221: 13-20.
- Palmier-Claus JE, Myin-Germeys I, Barkus E, et al. (2011) Experience sampling research in individuals with mental illness: reflections and guidance. *Acta Psychiatrica Scandinavica* 123: 12-20.
- Penn DL, Roberts DL, Combs D, et al. (2007) Best practices: the development of the social cognition and interaction training program for schizophrenia spectrum disorders. *Psychiatric services* 58: 449-451.
- Penn DL, Sanna LJ and Roberts DL. (2008) Social cognition in schizophrenia: an overview. *Schizophrenia bulletin* 34: 408-411.
- Pinkham AE, Hopfinger JB, Ruparel K, et al. (2008) An investigation of the relationship between activation of a social cognitive neural network and social functioning. *Schizophrenia bulletin* 34: 688-697.
- Pinkham AE and Penn DL. (2006) Neurocognitive and social cognitive predictors of interpersonal skill in schizophrenia. *Psychiatry research* 143: 167-178.
- Potheegadoo J, Cuervo-Lombard C, Berna F, et al. (2012) Distorted perception of the subjective temporal distance of autobiographical events in patients with schizophrenia. *Consciousness and cognition* 21: 90-99.
- Ruz M, Moser A and Webster K. (2011) Social expectations bias decision-making in uncertain inter-personal situations. *PLoS One* 6: e15762.
- Sanfey AG. (2007) Social decision-making: insights from game theory and neuroscience. *Science* 318: 598-602.
- Schlagenhauf F, Juckel G, Koslowski M, et al. (2008) Reward system activation in schizophrenic patients switched from typical neuroleptics to olanzapine. *Psychopharmacology* 196: 673-684.
- Schurz M, Radua J, Aichhorn M, et al. (2014) Fractionating theory of mind: a meta-analysis of functional brain imaging studies. *Neuroscience & Biobehavioral Reviews* 42: 9-34.
- StataCorp. (2015) Stata Statistical Software: Release 14. College Station, TX: StataCorp LP.
- Velthorst E, Fett A-KJ, Reichenberg A, et al. (2016) The 20-year longitudinal trajectories of social functioning in individuals with psychotic disorders. *American Journal of Psychiatry* 174: 1075-1085.
- Wechsler D. (1999) *Wechsler Abbreviated Scale of Intelligence*, New York, NY: The Psychological Corporation: Harcourt Brace & Company.

WHO. (1992) *The ICD-10 classification of mental and behavioural disorders: clinical descriptions and diagnostic guidelines*: World Health Organization.