**Supplementary Material**

Additional sample characteristics

Table S1:Diagnoses for FEP and SZ



Note: FEP = first episode psychosis; ICD 10 = International Statistical Classification of Diseases and Related Health Problems, 10th revision; SZ = schizophrenia

Table S2: Medication for FEP and SZ



Note: CPZ\_equiv = Chlorpromazine equivalents; FEP = first episode psychosis; IM = intramuscular; SZ = schizophrenia

**Tasks :**

Trust game

The trust game was a modified version of a previous multi-round trust game (King-Casas, 2005). Participants played the role of investor (first player) in a multi-round trust game consisting of 20 experimental and 20 control trials (see Figure S1). At the beginning of each round, they received the same starting budget of £10. Participants were asked to decide how much money they wanted to share with the other player (the trustee). Any amount between £0 and £10 could be shared with the trustee. The shared amount was tripled, and the second player had to decide how much to repay to the first player.

Participants were told that they were playing against a computer. The computer algorithm was programmed in a probabilistic way to respond in a cooperative style, where the return was 100, 150, or 200% of the invested amount. Each of these possible first repayments occurred with a probability of 33%. Subsequent repayment increased with 10% if the current investment reflected an increase in trust relative to the previous investment, however, remained stable in all other situations. Hence, with each increase in trust from the side of the investor, the chance of a repayment of 200% increased (Gromann et al., 2013; Lemmers-Jansen, Krabbendam, Veltman, & Fett, 2017).

A single trial lasted 18.5 s and was set up as follows (Figure S1 top row): Every trial started with an investment cue of £10 and the investment period of the participant (6 s). The invested amount was shown (2 s), followed by waiting period with a bar slowly filling itself with dots (2–4 s), and a fixation cross (0.5 s). The partner’s response was displayed (3 s), followed by the totals (2.5–4.5 s, depending on the length of the earlier waiting period, both adding up to 6.5 s). Each trial ended with a fixation cross (0.5 s). The design and duration of each event within the control trials (Figure S1 bottom row) was identical to the experimental trials. The control trials were included as a baseline condition for the functional MRI analysis. The task had a total duration of approximately 13 min.



Figure S1. Graphical overview of the trust game. Note: Top row represents the visual stimuli in the experimental trials; middle row are the separate phases of the trust game, including durations; bottom row represents the visual stimuli in the control trials. Printed with permission of Lemmers-Jansen et al. (2017).

Faces task

Participants were asked to choose one of two simultaneously presented faces and learn to identify which face was associated with a higher probability of reward. They were given the following instructions: “On each trial in this task you will be presented with two faces. You will have to select one of the faces. Press the top button to select the top face, the bottom button to select the lower face. Your task is to try to figure out which face in each block has the highest probability of winning and pick that face as many times as possible. You will be told when the block switches, and at each switch the faces will be associated with new probabilities of winning.” (Evans, Fleming, Dolan, & Averbeck, 2011).

The task consisted of 4 blocks with 30 trials each. In the two ‘emotional’ blocks (block 1 and 3), participants chose between angry and happy facial expressions, and in two ‘neutral’ blocks (block 2 and 4) participants chose between two neutral faces of different identities (see Figure S2). In each block, one face was associated with a higher probability of reward (60% vs. 40% contingency), where every correct choice (referred to as ‘wins’) was rewarded with 10p and every incorrect choice (referred to as ‘losses’) resulted in reward omission. The face stimuli and probabilities of winning were kept constant throughout each block.

Each trial began with a period of 1000 ms during which a white central fixation cross was presented against a dark background. This was followed by two faces being presented to the right and left of the fixation cross for 4500 ms. Within this time window, subjects were required to select one of the faces by pressing the corresponding button with their right hand. The selected face was highlighted by a yellow square surrounding it. Feedback was then presented on the screen for 1500 ms. The task had a total duration of approximately 15 min. (Vanes, Mouchlianitis, Collier, Averbeck, & Shergill, 2018).





Figure S2. Graphical overview of the task sequence (emotional block) and duration. The chosen face was highlighted with a red box surrounding it. Division of blocks is represented, and an example of the faces in a neutral block is shown. Printed and adapted with permission of Horne et al. (2021).

Stroop Vanes 2019 Supplement

Participants performed a verbal Stroop paradigm. On each trial, a single colour word was presented on the screen (“BLUE”, “RED”, “GREEN”, or “YELLOW”) against a black background. The word was printed in one of four possible font colours (blue, red, green, or yellow). Word meaning and font colour were either congruent or incongruent (see Figure S3), and subjects were instructed to respond verbally to the font colour and to ignore the word meaning. In addition, on fixation trials, a central white fixation cross was presented in the centre of the screen and no response was required. Thirty-three congruent, 33 incongruent, and 34 fixation trials were presented in randomised order, each with a duration of 700ms and inter-stimulus-interval of 2300ms. Responses were recorded via a microphone to assess reaction times. Only responses between 200 and 2700 ms after stimulus presentation were recorded (Vanes et al., 2019).



Figure S3. Graphical representation of the Stroop task, showing congruent trials on the left, where the font colour matches the word meaning; and incongruent trials on the right, where colour and meaning do not match, and participants have to respond to the font colour and ignore the word meaning.

SEM analyses

To investigate the associations between tasks, a structural equation model analysis (SEM) was performed, including Stroop accuracy and reaction time (level 1), faces task proportion ideal choices and emotion bias (level 2) and trust game mean trust and slope (level 3) (see Figure S4).



Figure S4. Graphical representation of the SEM model. Dashed lines show the significant associations for healthy controls, dotted lines for the FEP group. In SZ, no significant associations were found.

Table S3: Group means of principal components

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | HCN = 38 | FEPN = 99 | SZN = 41 | F | p-value | Group difference | p-value |
| pc1 M | 0.74 | -0.27 | -0.03 | 3.44 | 0.03 | HC > FEP |  0.01\* |
| (SD) | 1.32 | 2.18 | 2.18 |  |  | HC > SZ |  0.09 |
| pc2 M | -0.25 | -0.07 | 0.41 | 1.98 | 0.14 | HC < SZ | 0.06 |
| (SD) | 1.29 | 1.54 | 1.80 |  |  | FEP < SZ | < 0.1 |

Note: HC = healthy controls; FEP = first episode psychosis; SZ = schizophrenia; M = mean; SD = standard deviation; pc = principal component.

\* significant after Bonferroni correction

Table S4: Trust game slope by principal component scores.

|  |  |
| --- | --- |
|  |  |
|  | Slope |
|  | b/p |
| Scores for component 1 (HC) | 0.022 (0.126) |
| FEP | -0.030 (0.217) |
|  |  |
| SZ | -0.029 (0.310) |
|  |  |
| FEP # Scores for component 1 | -0.027 (0.075) |
|  |  |
| SZ # Scores for component 1 | 0.000 (0.982) |
|  |  |
| Scores for component 2 (HC) | -0.025 (0.092) |
|  |  |
| FEP # Scores for component 2 | 0.033\*  (0.044) |
|  |  |
| SZ # Scores for component 2 | 0.028 (0.110) |
|  |  |
| Constant | 0.055\* (0.012) |
|  |  |
| R2 | 0.108 |
| N | 178 |

\* p < 0.05.

Note: HC = healthy controls; FEP = first episode psychosis; SZ = schizophrenia.

Evans, S., Fleming, S. M., Dolan, R. J., & Averbeck, B. (2011). Effects of emotional preferences on value-based decision-making are mediated by mentalizing and not reward networks. *Journal of cognitive neuroscience, 23*(9), 2197-2210.

Gromann, P., Heslenfeld, D., Fett, A.-K. J., Joyce, D., Shergill, S., & Krabbendam, L. (2013). Trust versus paranoia: abnormal response to social reward in psychotic illness. *Brain*, awt076. doi:<http://doi.org/10.1093/brain/awt076>

Horne, C. M., Vanes, L. D., Verneuil, T., Mouchlianitis, E., Szentgyorgyi, T., Averbeck, B., . . . Shergill, S. S. (2021). Cognitive control network connectivity differentially disrupted in treatment resistant schizophrenia. *NeuroImage: Clinical, 30*, 102631.

King-Casas, B. (2005). Getting to know you: reputation and trust in a two-person economic exchange. *Science, 308*, 78-83. Retrieved from <http://dx.doi.org/10.1126/science.1108062>

Lemmers-Jansen, I. L., 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*, 235-245.

Vanes, L. D., Mouchlianitis, E., Collier, T., Averbeck, B. B., & Shergill, S. S. (2018). Differential neural reward mechanisms in treatment-responsive and treatment-resistant schizophrenia. *Psychological Medicine, 48*(14), 2418-2427.

Vanes, L. D., Mouchlianitis, E., Patel, K., Barry, E., Wong, K., Thomas, M., . . . Shergill, S. (2019). Neural correlates of positive and negative symptoms through the illness course: an fMRI study in early psychosis and chronic schizophrenia. *Scientific reports, 9*(1), 1-10.