Article

The Influence of Trial-By-Trial Feedback on Trust in Health, First-Episode and Chronic Psychosis

Imke L. J. Lemmers-Jansen [1,2,*] , Rune J. Wichmann [1], Sophie Perizonius [2] and Sukhi S. Shergill [1,3]

1   Institute of Psychiatry, Psychology and Neuroscience, King's College London, de Crespigny Park, London SE5 8AF, UK
2   Department of Clinical, Neuro and Developmental Psychology, Faculty of Behavioural and Movement Sciences, Vrije Universiteit Amsterdam, 1081 BT Amsterdam, The Netherlands
3   Kent and Medway Medical School, University of Kent, Canterbury CT2 7FS, UK
*   Correspondence: imke.jansen@vu.nl

Abstract: Trust is crucial to establishing reciprocal, positive social interactions and seems to be compromised in psychosis. The trust game offers methods to assess an individual's trust responses to trust-reciprocating, positive feedback. Various computational techniques have been implemented to measure trust responsiveness, mostly based on investments. Here, we propose a new method, focusing on feedback response. Psychosis patients show social dysfunction and reduced trust during early and more progressed illness stages. The present study inspects differences in feedback responsiveness of 102 first-episode psychosis patients (FEPs), 43 chronic psychosis patients (CPs), and 39 healthy controls (HCs) by adopting a novel assessment approach. Additionally, baseline trust, the trust exerted without any prior knowledge of the partner's trustworthiness, and mean trust were examined. Participants performed a multi-round trust game, playing the investor role, and were paired with a computer, programmed to return at least the invested amount, representing a trustworthy partner. The new method detected group differences, more distinguished than the former methods. Contrary to our expectations, baseline trust was intact in patients. Relative to HCs, patients were less responsive to feedback, failing to integrate the positive information into their decision-making process. The magnitude of returns was not associated with increases in trust. This novel method showed promising results and confirmed patients' deficits within the social interactional domain.

Keywords: trust game; novel technique; feedback responsiveness; chronic psychosis; first-episode psychosis; reward magnitude; baseline trust; psychotic symptoms

1. Introduction

Trust is essential for social functioning: it counteracts loneliness [1], promotes the generation of interpersonal relationships [2] and allows for mutually beneficial, reciprocal interactions [3]. Vice versa, reciprocating the trust shown by others also improves the relationship [4]. In general terms, an individual (trustor) relies on the benevolent action of another individual (trustee). The trustor's expectation that the trustee behaves in a way that promises a positive outcome for the trustor is regarded as trust [5], which can either be increased or decreased over the course of repeated social encounters [6].

Social interactions with others are compromised in patients with psychosis [7]. Examples are manifold: patients form incorrect judgments about social cues [8]; fail to identify others' emotions and mental states [9,10], and ascribe malicious intentions to other people [11]. In fact, paranoia, which is a prominent element of psychosis, depicts this suspicion or mistrust [12]. Furthermore, negative symptoms, that reflect a loss or decrease in normal functioning, such as reduced feelings of pleasure or social withdrawal, are directly associated with social problems [13]. The relationship between illness severity and social deficits may be bidirectional [14]. Initial social maladjustment partly triggers psychotic symptoms.

In turn, symptoms reinforce social difficulties, which may aggravate symptoms, creating a negative spiral. Positive symptoms (i.e., hallucinations and delusions), but especially negative symptoms, such as asociality (the lack of interest in social interactions) and blunted affect, are closely associated with social dysfunction [15]. These social problems exist throughout the psychosis spectrum, in individuals with chronic psychosis patients (CPs); in people who only experienced a single episode, referred to as first-episode psychosis patients (FEPs); and even in individuals with clinical high-risk for psychosis, or relatives of psychosis patients [15–19].

The most widely used paradigm to investigate trust is the classical trust game [20]. The trust game is recognised as a reliable and valid, direct, behavioural measure of trust [21–23]. Two anonymous players are required to reciprocate trust in order to achieve the best possible gains for both. Trusting involves risk, since the trustee earns the largest amount by not repaying the investor. In a multi-round setting, however, this might be counterproductive, as the investor potentially decides to invest less during the following rounds. By investing, the investor places trust in the trustee to be reimbursed. The amount of invested money represents a quantification of trust, with larger investments resembling more trust. Traditionally, baseline trust (e.g., the initial trust exerted without any knowledge of the trustworthiness of the trustee) and the development of trust during repeated interactions are investigated.

In previous studies in psychosis patients these methods have been used, and patients showed a tendency of lower baseline trust compared to HCs over the entire psychosis spectrum: Regarding CPs, the majority of studies found that the baseline trust was lower compared to HCs [24,25]; however, in one study, the results only bordered significance [26]. FEPs exhibited lower baseline trust [27], as well as early psychosis patients (i.e., teenagers), individuals at clinical high-risk (i.e., individuals with subthreshold psychotic symptoms), or first-degree relatives of psychosis patients [24,27–30]. In terms of symptom profiles, both increased positive and negative symptoms have been associated with reduced baseline trust [24,25,27,28].

In the current version of the trust game, participants were paired with a non-human, cooperative counterpart, returning at least the initial invested sum. As the trials progress, the trustor receives continuous, positive feedback in the form of repayments by the trustee. The trustor reacts adequately by increasing the investment over the course of the trials; this strategy maximises the monetary gains, showing feedback learning. Reciprocating the trustworthiness of the trustee by investing larger sums is regarded as responsive to feedback and acting trustfully. Decreasing investments is considered unresponsive to feedback and distrustful.

In psychosis patients, numerous studies showed that CPs were unresponsive to positive feedback, not showing an increase in investment over trials [26,31–34]; furthermore, self-ratings confirmed this lower propensity of placing trust in someone [33]. Illness duration might play a role in feedback learning, since FEPs still increased their investments after positive feedback, pointing towards intact feedback learning [27]. In support of that, early psychosis patients and the vast majority of at-risk individuals were sensitive to (social) feedback [24,27,28,30,31,35]. On the grounds of the current state of knowledge, especially patients with progressed psychosis expressions suffer from impaired feedback learning. Early psychosis stages or at-risk individuals are presumably not impacted. Similar to baseline trust, both positive and negative symptoms are associated with impaired feedback responsiveness [26–28,30,34].

Various techniques have been implemented to measure feedback responsiveness within the trust game context. A multilevel, mixed-effects regression analysis on a patient group-by-trial number interaction has been applied, with the trial number reflecting the consecutive rounds of the trust game [26,27,31,33,34]. In a less sensitive way, feedback responsiveness has been evaluated by dividing one game into four blocks of five trials and carrying out a multilevel, random regression analysis on the group-by-block interaction [30], or by comparing the mean investments of the first and the last block [32]. Lastly, mean

investments of all trials have been compared between groups [25,28,29], yet this yields even less informative results.

The current paper sets out to investigate the feedback responsiveness of HCs, FEPs and CPs within a multi-round trust game, applying a novel computational approach, using a trial-by-trial assessment of feedback. A detailed description of the procedure can be found in the Methods section. We tested if utilising this novel, detailed approach contributes to a more accurate understanding of feedback learning and the development of trust in psychosis, offering new perspectives on possible reasons underlying disrupted social functioning in patients. We compared the outcomes of these new analyses with earlier methods. Additionally, we hypothesised that (1) CPs would show reduced feedback learning, with HCs and FEPs showing similar learning. Secondly, patients presumably require stronger stimuli to incorporate the feedback information into their decision-making [25]. To test this, the magnitude of reward (i.e., percentage of return) was additionally considered in the analyses. We further hypothesised that (2) patients would show more response when the return magnitude would be higher. Moreover, (3) we expected reduced baseline trust in patients. Lastly, we explored the association between symptoms, feedback responsiveness and baseline trust, expecting that (4) symptoms are negatively associated with feedback responsiveness and with the amount of the first investments in patient groups.

2. Results

2.1. Sample Characteristics

Table 1 displays sample characteristics per group. Gender distribution differed between groups (χ^2(2183) = 6.49, p = 0.04). Specifically, the CP group consisted of less females than FEPs (p = 0.046) and HCs (p = 0.01). No difference between FEPs and HCs was observed (p = 0.32). Groups differed in age (F(2179) = 50.62, p > 0.001), with CPs being older than FEPs and HCs, and FEPs being younger than HCs (all p's > 0.001). Group differences in WASI were apparent (F(2168) = 26.99, p > 0.001), with patient groups scoring lower than HCs (both p's > 0.001), but not differing from each other (p = 0.21). Patient groups reported less years of education than HCs (both p's > 0.001), and groups differed concerning ethnical background composition (χ^2(4180) = 41.03, p > 0.001). Regarding patient characteristics, CPs reported higher positive symptoms (t(139) = 2.57, p = 0.01), negative symptoms (t(139) = 3.34, p = 0.001) and total symptom score (t(139) = 2.06, p = 0.04) than FEPs. No differences in general symptoms were observed (t(139) = $-$0.09, p = 0.94). There was a group difference in terms of medication (χ^2(3145) = 13.37, p = 0.004): CPs were more often prescribed with typical antipsychotics than FEPs (p = 0.001); atypical antipsychotics (p = 0.20) and no medication intake (p = 0.14) were similar among groups.

2.2. Trust Game Behaviour

Trust game behaviour is displayed in Table 2, showing similar first investments across groups (F(2181) = 2.29, p = 0.10). Adding the covariates age, gender and WASI to the model did not change the outcome, as the group remained nonsignificant (p = 0.17). Both patient groups showed lower mean investments compared to controls (F(2181) = 6.76, p = 0.001). The slope of individual regression lines bordered significance between groups, (F(2181) = 2.75, p = 0.067), with FEPs showing close to significant differences in increase of slope compared to HCs (p = 0.75). Since mean investment, and even slopes are a coarse measure of feedback learning, we investigated trial-by-trial feedback response.

Table 1. Participant and Patient Characteristics.

	HCs (*n* = 39)	FEPs (*n* = 102)	CPs (*n* = 43)
Gender, *n* male (%)	24 (61.5%)	71 (70.3%)	37 (86.0%) *
Age, *M* (*SD*)	34.05 (9.76) **	26.55 (5.89) **	41.05 (10.52) **
WASI, *M* (*SD*)	118.11 (11.46) **	99.28 (16.46)	94.45 (15.63)
Years of Education, *M* (*SD*)	18.25 (3.03) **	15.07 (3.73)	13.79 (3.64)
Ethnical Background, *n* (%)			
White/Caucasian	24 (61.5%) **	27 (26.5%)	14 (32.5%)
Black/African	2 (5.1%) **	54 (52.9%)	27 (62.8%)
Other/Mixed	13 (33.3%)	19 (18.6%)	-
Missing	-	2 (2.0%)	2 (4.7%)
PANSS			
Positive Symptoms		13.16 (5.49) *	15.77 (5.62)
Negative Symptoms		12.97 (5.45) **	16.40 (6.00)
General Symptoms		29.64 (7.93)	29.51 (9.34)
Total Score		55.78 (16.22) *	62.12 (18.52)
Diagnoses, *n* (%)			
Schizophrenia		14 (13.7%)	43 (100%)
Schizoaffective Disorder		21 (20.6%)	-
Other Psychotic Disorders		53 (52.0%)	-
Missing		14 (13.7%)	-
Medication, *n* (%)			
Atypical Antipsychotics		91 (89.2%)	35 (81.4%)
Typical Antipsychotics		3 (2.9%) **	8 (18.6%)
No Medication		5 (4.9%)	-
Missing		3 (2.9%)	-

M: mean; *SD*: standard deviation; WASI: Wechsler Abbreviated Scale of Intelligence; PANNS: Positive and Negative Syndrome Scale; HCs: healthy controls; FEPs: first-episode psychosis patients; CPs: chronic psychosis patients. * this group differs from both other groups at $p < 0.05$; ** this group differs from both other groups at $p < 0.001$.

Table 2. Trust Game Results.

	HCs (*n* = 39)	FEPs (*n* = 102)	CPs (*n* = 43)
Feedback Response			
Total over trials *M* (*SD*)	12.97 (12.32)	5.63 (10.75) **	7.37 (11.86) +
Mean per trial *M* (*SD*)	0.68 (0.65)	0.30 (0.57) **	0.39 (0.62) +
First Investment *M* (*SD*)	6.97 (2.47)	6.00 (2.42)	6.30 (2.38)
Mean Investment *M* (*SD*)	7.68 (1.57)	6.58 (1.61) *	6.68 (1.64) *

M: mean; *SD*: standard deviation; HCs: healthy controls; FEPs: first-episode psychosis patients; CPs: chronic psychosis patients. * this group differs from HCs at $p < 0.05$; ** this group differs from HCs at $p < 0.01$; + this group differs from HCs at trend level, $p = 0.07$.

2.3. Feedback Response

Considering the mean feedback response per trial over all groups, some values were exceptionally high or low, particularly within the patient groups. Correcting these outlier values to a value 2 *SD* from the mean resolved this problem; however, this yielded highly similar results to the uncorrected values, and it was decided to use the uncorrected values for the analyses. Group differences in feedback response were found ($F(2181) = 5.92$, $p = 0.003$), with FEPs responding less than HCs ($p = 0.002$), CPs responding less to feedback than HCs at trend level ($p = 0.07$), and CPs and FEPs displaying similar results ($p = 0.68$, see Table 2). Adding age, gender and WASI to the model did not impact the findings, as the group difference remained significant ($p = 0.02$). Figure 1a shows the development of the mean trial-by-trial feedback response values divided by groups over trials 2 to 20, and Figure 1b shows the development of the mean investments over trials per group, representing a similar shape of change, since increases of investments represent feedback responsiveness, and decreases represent feedback non-response.

(**a**)   (**b**)
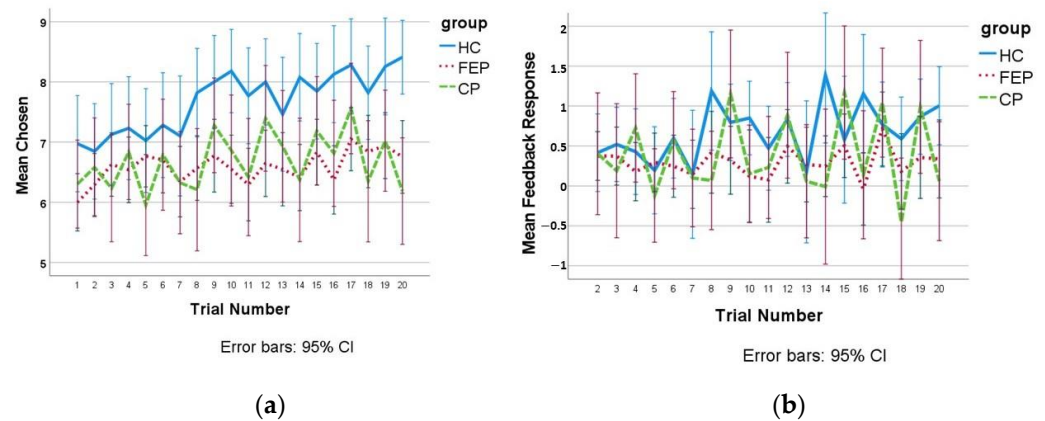
Figure 1. Development of (**a**) Mean Investments per Group over Trials, and (**b**) Mean Trial-by-Trial Feedback Response, with Error Bars. HCs: healthy controls; FEPs: first-episode psychosis patients; CPs: chronic psychosis patients.

As a data check, and a justification for the method, we inspected the individual data for investment patterns, and analysed changing behaviour, dividing the 20 trials in 4 blocks of 5 trials (analogous to [35]). There were differences in the number of changes in direction (from increase to decrease and vice versa, disregarding staying at the same number) between groups ($F(2181) = 4.08$, $p = 0.018$), with FEPs showing significantly more jumps ($p = 0.014$) than HCs. Furthermore, the times the number 10 was chosen differed per group ($F(2181) = 3.90$, $p = 0.022$), with FEPs choosing 10 significantly less than HCs ($p = 0.03$). CPs did not differ significantly from either group. Trial-by-trial analysis including staying behaviour, divided into 4 blocks of 5 trials (see Figure 2) showed that increasing investments only differed marginally between groups (blocks 1 and 3: HCs < FEPs $p = 0.05$ and 0.07, respectively), but differences were significant for decreasing and staying behaviour, however, only between HCs and FEPs. CPs did not significantly from either group in any condition. In the second block of 5 trials, FEPs decreased more ($p = 0.002$) and stayed less ($p = 0.02$) than HCs. In the third block, FEPs stayed less ($p = 0.02$) than HCs, and in the last block, FEPs decreased more and stayed less than HCs (both p's = 0.04).

2.4. Reward Magnitude Sensitivity

The interaction of group-by-reward magnitude (i.e., 100%, 150%, and 200%) was not associated with feedback response ($\beta = 0.02$, $p = 0.89$). After removing the interaction term from the model, reward magnitude was nonsignificant ($\beta = 0.16$, $p = 0.15$), whereas group ($\beta = -0.14$, $p = 0.03$) showed (as described above) a significant main effect on feedback response. An additional regression analysis examining all three reward magnitudes confirmed the absence of significant relationships with feedback response (all p's > 0.23).

2.5. Associations with Symptoms

The model including positive, negative and general symptoms predicted feedback response ($F(3137) = 3.13$, $p = 0.03$). The predictors explained 6.4% of the proportion of variance ($R^2 = 0.064$). Surprisingly, positive symptoms were positively associated with feedback response ($B = 0.44$, $p = 0.046$), and general symptoms were negatively associated with feedback response ($B = -0.48$, $p = 0.003$). Negative symptoms did not show such an association ($B = 0.16$, $p = 0.42$). No significant association between symptoms and first investment was found ($F(3137) = 0.65$, $p = 0.59$).
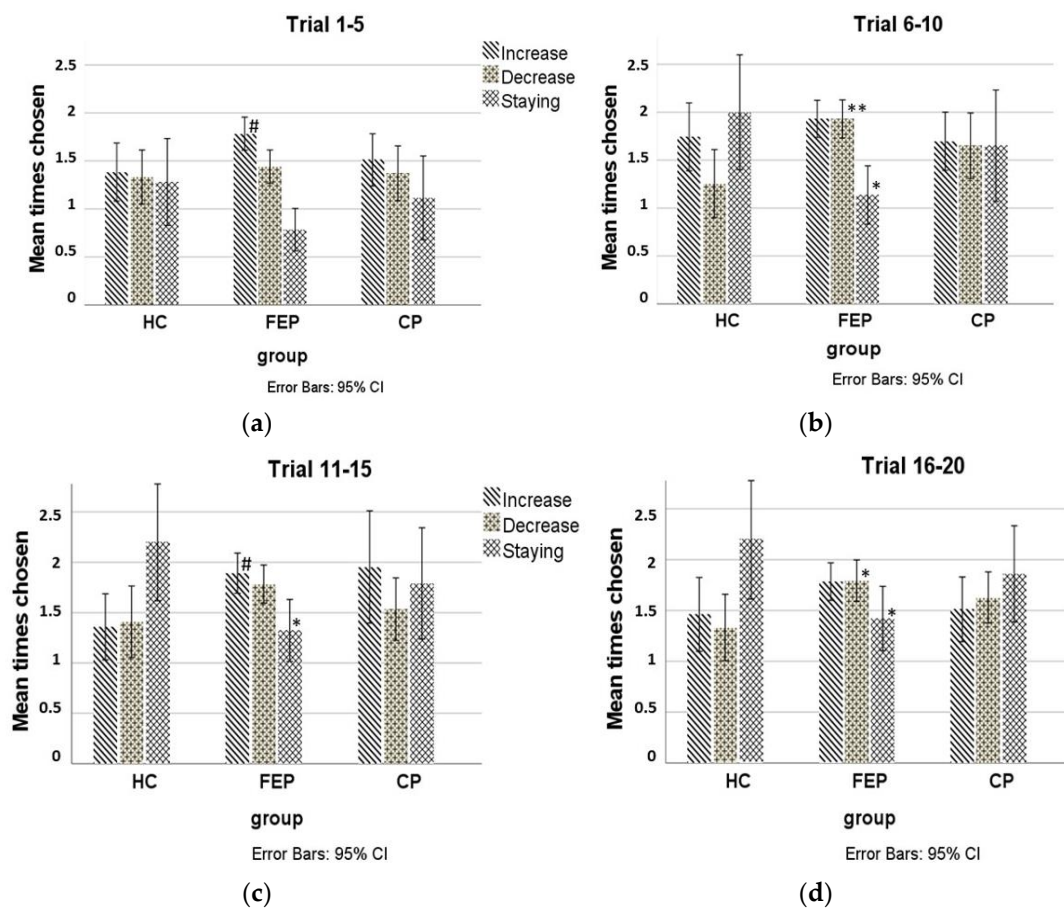
Figure 2. Investment Behaviour by Group During the First Block of 5 Trials (**a**), the Second Block (**b**), the Third (**c**) and Last Block of 5 Trials (**d**). HCs: healthy controls; FEPs: first-episode psychosis patients; CP: chronic psychosis patients. All group differences indicated are between FEPs and HCs, with * = $p < 0.05$; ** = $p < 0.01$; # = $p < 0.1$, at trend level. CPs did not differ significantly from either group.

3. Discussion

The study analysed changes in trust in response to direct, positive feedback (i.e., feedback response) of HCs, FEPs and CPs, during a multi-round, interactive trust game, by using a novel, trial-by-trial, proportional assessment method. We demonstrated that the new method reporting feedback response showed similar results to the investment analysis, however more detailed, and that magnitude of the feedback did not affect decision-making. We found that first investments were similar across groups, but feedback learning abilities were compromised in patients relative to HCs, but not different between patient groups.

The novel, trial-by-trial, proportional assessment method revealed results that were similar to the method using investments, yet, resulted to be more fine-grained, and allowed for the inclusion of response magnitude in the process of investing. Results resembled the analyses using mean investments, rather than the individual slope analyses. It was hypothesised that CPs would be less responsive to feedback than HCs. We could only tentatively accept this hypothesis, since group differences in feedback response only bordered significance ($p = 0.07$). Prior research uniformly emphasised CPs' feedback unresponsiveness [24–26,31–34]. Considering the results at trend level in combination with the number of prior sources, there is evidence to argue for impaired feedback learning abilities in CPs, nonetheless, it is not possible to reach definite conclusions.

The hypothesis that FEPs and HCs exhibit equal levels of feedback response was rejected. FEPs did not adjust their trust by increasing the investments over the course of repeated interactions, suggesting impaired feedback learning in early illness stages. The

availability of research on FEPs is sparse. At present, one other study explored feedback response in FEPs showing intact feedback learning in FEPs [27]. Apart from a different FEPs sample make-up (regarding age and diagnoses), Lemmers-Jansen, Fett [27] included a smaller sample of 22 patients. Here, 102 FEPs were utilised, providing more statistical power when drawing conclusions. In addition, the computational approach to assessing feedback response varied; however, the previously found increase of investment in FEP is not present in the current study, possibly suggesting a more random investment behaviour. A strength of the current paper is its trial-by-trial, proportional measuring technique, reflecting on feedback response from a novel angle and thus yielding new, possibly more precise insights. Still, considering the discrepancies with previous literature, no firm conclusions should be drawn. The heterogenous compositions of the patient groups may have influenced outcomes. For example, affective disorders (20% of the FEP sample) are known for abnormal trust behaviour and social decision-making [36,37].

The magnitude of the reward was irrelevant for the feedback response: participants were equally responsive to feedback after receiving a 100%, 150% or 200% return of investment in the previous round, leading to a rejection of the hypothesis. This association was similar for all groups. When deciding to trust or distrust, the overall direction of the previous feedback (trust-reciprocating in this case) seems to matter more than the magnitude of it. The data do not support the assumption that a reduced sensitivity to rewards underlies the insufficient feedback responsiveness in patients, even though prior research inferred such a connection [24,25]. Arguably, the amplitude of the rewards needs to be increased to identify any differences. Earlier studies, indeed, point out an undervaluation of reward magnitude in psychosis [38,39], highlighting the necessity of a larger range of rewards. Instead of implementing stimuli of 100% to 200% return of investment, effects might only arise between 100% to 500%, for instance. In the same line, investing £ 1 to £ 10 might be a too narrow spectrum of choices. Research with a wider dimension of incentives should be conducted to test for possible variations.

In contradiction with our hypothesis, baseline trust seemed intact in CPs and FEPs, meaning their capacity to invest in an unknown counterpart, without information about their trustworthiness, is not disturbed. The majority of earlier research challenges this finding, where patients start with a lower investment than HCs [24–27]. Disclosing the non-human nature of the game partner might have played a role in this incongruity, since both HCs and patients tend to invest more when playing a non-human counterpart [31]. This information was shared before the game, however, was concealed in most earlier studies, where participants thought they were playing a human counterpart. Initial suspicion could be specific towards humans, and maybe participants adopted a more strategic, less emotional approach on their first investment. This assumption should be tested experimentally by comparing computer and human counterparts directly.

Taken together, participants had a similar starting point (baseline trust). As the trials progressed, patients failed (CPs at trend level) to integrate the positive information received throughout the game into their decision-making process. As a result, the initial trust was not increased. Compared to other studies, where control participants reached a mean level of trust between 6.7 and 8 [25,27,29–31,40,41], HCs increased their investments in a similar way. For both patient groups, a general tendency was recognised: they shifted their behaviour from trusting to relatively less trusting after repeated interactions, suggesting an inadequate reaction to feedback. Translating this into real-life social implications, the social impairments in psychosis may at least partly stem from interactional deficits. When initiating or forming social relationships, patients perceive their partner as liable and trustworthy. The obstacle seems to be successfully maintaining those relationships, as their interactional deficiencies lead to a misinterpretation of benevolent, reciprocal social situations, therefore acting suspiciously towards al trustful social partner. There is well-established literature on interactional deficits in psychosis [42,43], as well as on the falsely malevolent interpretation of social information [44,45]. The current research underlines these findings. Distrust is associated with fewer friendships [46] and increased

loneliness [47]. Importantly, social support predicts recovery [48], stressing the need for more research on trust interventions in patients, to promote recovery.

This paper is the first directly comparing the two patient groups, arguably shining a more accurate light on potential differences. On investment measures and feedback response, CPs performed in-between FEPs and HCs; this is an unexpected pattern, we predicted that advanced psychosis would exhibit more impairments. Pre-existing group differences might account for these discrepancies in trust responses; however, analyses have shown that, age and gender, as well as intelligence, were of no influence. CPs displayed higher positive, negative and total symptoms than FEPs, ruling out symptom severity as an explanation. One possible explanation is the variation in diagnoses. All CPs were diagnosed with schizophrenia. FEPs consisted of a broader range of diagnoses, including schizoaffective disorders. In other studies, the sample make-up was more homogenous.

Only general symptoms showed a negative association with feedback response, suggesting an interference with the capacity to adequately learn from feedback. Against the hypothesis, more positive symptoms in CPs and FEPs were associated with increased feedback response. Earlier studies found associations in the expected, opposite direction [26–28]. Contrary to the hypothesis, negative symptoms were not associated with feedback response; however, social dysfunction is a prominent attribute of negative symptoms [15], and these outcomes are in dissonance with other research [26–28,30,34]. Symptoms were not associated with baseline trust, while other studies did find reduced baseline trust with increased symptoms [24,25,27,28]. Average total scores on the PANSS corresponded to a categorisation of mildly ill [49]. Symptoms may have been dampened by medication, reducing the associations with behaviour. Plausibly, the relationships remained nonsignificant as a consequence of generally low symptoms. Higher symptoms might be required to show any significant effects. Future research with separate symptom profiles, and in unmedicated patients could be valuable to disentangle specific effects on feedback response.

Several limitations should be considered. The disclosure of the non-human game partner, the varying diagnoses and the overall low symptoms have been discussed. Additionally, some methodological issues need to be addressed. Participants were not directly monetarily rewarded based on their trust game gains. Some propose that hypothetical payment impacts decision-making [50,51], while others suggest that real or hypothetical reward makes no difference [52]. Possibly, only the strength of the effect is amplified but not the direction [53], and all participant groups should react similarly to the effect of payment. By incorporating at least a 100% of investment return, it was intended to create a cooperative game environment, resembling positive interactions; it is debatable whether the size of a 100% return is indeed perceived as cooperative. While the participant does not loose, he does not gain either, and the partner gains twice the investment. One could evaluate this contribution as unfair. Economic theories (e.g., relative thinking theory by Azar [54]) put forward that individuals are more concerned about the relative standing to others than the absolute amount of money. In the game context, participants may have invested less after a 100% return of investment, and thus were appraised as less responsive to feedback. A further limitation is that a testing session lasted for about four hours, which could have been mentally exhaustive, causing fatigue and reduced attention during the trust game, which was presented at the end of the testing session.

While employing a proportional computation of feedback response presents many advantages, also with this approach a ceiling effect occurred. For increased investments, the assessment could be slightly distorted by overestimating low, more unresponsive investments and underestimating high, more responsive investments. An increase within the high, actual feedback responsive spectrum is evaluated as less feedback responsive than an increase within the low, actual feedback unresponsive spectrum (e.g., increasing from 1 to 2 = 2 vs. increasing from 9 to 10 = 1.11). Future research may test this possible limitation. In terms of decreasing investments, this distortion is not existent. The assignment of trial-by-trial feedback response values for not altered investments was somewhat arbitrary. Based on the *M* and *SD* of all observations, an appropriate estimate was decided. Other im-

plementations might have affected feedback responses differently. When analysing reward sensitivity, the group effect of receiving 100%, 150% or 200% return on feedback response was examined; furthermore here, a relative viewpoint could have been implemented. For instance, a return of 100% in the previous round and 150% in the current is an increase, whereas a return of 200% in the previous round and 150% in the current is a decrease. In this paper, that relative perspective was disregarded, as a large number of trials may have balanced this disparity out.

4. Materials and Methods

4.1. Participants

Forty-five HCs, 107 FEPs and 47 CPs were recruited for the study. After exclusion due to missing and invalid data or non-compliance with the recruitment criteria, the final sample consisted of 39 HCs, 102 FEPs, and 43 CPs. CPs were diagnosed with schizophrenia, as outlined by the International Classification of Diseases (ICD)-10 criteria. FEPs had experienced the first episode of psychosis within the last five years. FEPs' diagnoses varied, and were categorised as schizophrenia, schizoaffective disorder and other psychotic disorders. Patients were prescribed different antipsychotic medication over the course of the entire study. Only participants between 18 and 60 years were recruited.

Exclusion criteria for all subjects were a history of neurological illness, learning disabilities and drug or alcohol dependency/abuse within six months preceding the testing, as well as current major physical illness, suicidal risk, significant motor, hearing or speech difficulties, pregnancy and clozapine usage. For HCs, a history of psychiatric illness and having a first-degree relative with a psychotic disorder was an additional exclusion criterion.

Patients were recruited from the South London and Maudsley (SLaM) NHS Foundation Trust, the Oxleas NHS Foundation Trust and the North East London NHS Foundation Trust. The majority of SLaM participants has been contacted through a "consent for contact" scheme, offering a pool of patients that previously agreed to be consulted for potential research involvements. The remaining subjects were referred by consultant health professionals employed within the mentioned institutions. HCs of the same regional area as patients were recruited via Gumtree and Facebook advertisements.

4.2. Trust Game

A modified multi-round version of the trust game was used [27,28], during functional magnetic resonance imaging (fMRI) scanning. After receiving oral instructions, and before entering the scanner, practice trials were carried out. Subjects played the role of the trustor against a pre-programmed computer; they were informed that they played against a program beforehand. The game consisted of an experimental condition and a control condition, each consisting of 20 trials, presented in random order. At the beginning of each experimental trial, participants were endowed with a budget of £10. Next, they could invest any amount between £0 and £10 in the trustee; this investment was tripled and forwarded to the trustee, who repaid the investor with a certain sum. The computer's algorithm was programmed in a cooperative manner: 100%, 150% or 200% of the initial investment was repaid to the investor. The subjects hence gained experimental money on every trial, or at least played even. With every increased investment relative to the last played the experimental trial, the likelihood of a higher percentage of return increased (see Gromann, Heslenfeld [25] for the exact probabilistic computation). The gained sum after each round only served game purposes and was not awarded in real monetary value; however, at the end of the experiment, all participants were rewarded with £30 for participation and a maximum of £15 based on task performance. The set-up of control trials was identical to the experimental trials, with regard to visual stimuli, motor behaviour, and timings; however, instead of choosing to invest an amount, participants were required to move the cursor to the number indicated with an arrow. Since they do not include the element of investment, and, therefore lack of the possibility of learning, control trials were disregarded in this paper. Responses to these trials (correct/incorrect) did not influence the algorithm

of the experimental trials. A graphic illustration with corresponding stimuli is provided in Figure 3. According to the developers of the trust game, an investment represents trust, providing both players with the opportunity to increase the initial amount given [20]. A meta-analysis on one shot trust games also reports the first investment as trust [51]. In research using an iterative trust game, the first investment is thought to represent baseline trust, as opposed to mean trust or the development of trust over trials [25,29,36].
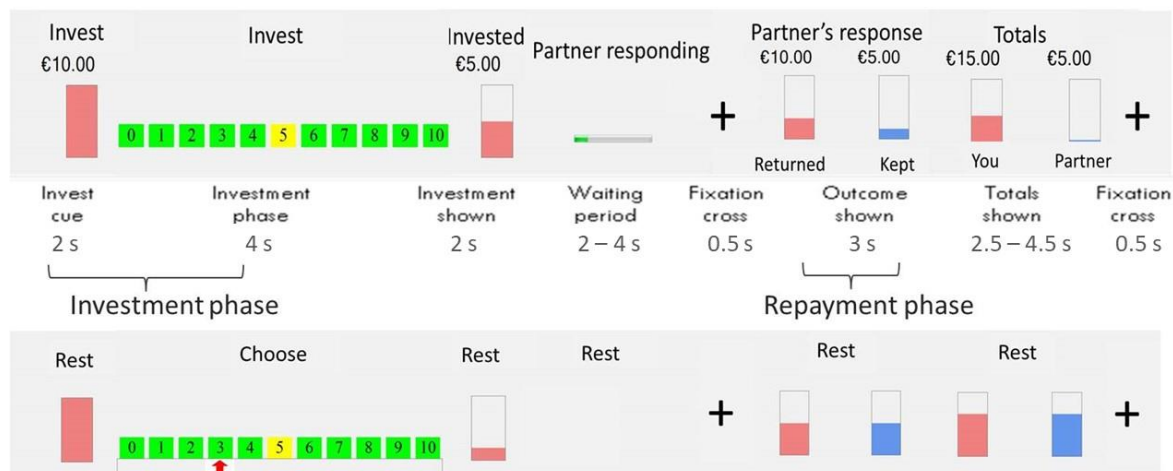
Figure 3. Visual representation of the trust game. Note. A depiction of the stimuli given in the trust game is shown, with above the experimental trials, and below the control trials, which are disregarded in this paper. Used with the permission of Lemmers-Jansen, Fett [27], who performed this study in the Netherlands, hence the currency in Euro's (€) and not in Pounds (£).

To evaluate feedback responsiveness, a proportional value relative to the investment of the previous round was created; this was done for rounds 2 to 20 of the experimental game trials. Whenever the investment was increased compared to the prior round, a trial-by-trial, proportional feedback response value was calculated by dividing the investment of the current round by the investment of the previous round. Whenever the investment was decreased, the trial-by-trial feedback response was derived by dividing the investment of the previous round with the investment of the current round. Since a reduction of investment reflects feedback unresponsiveness, a minus sign was added to the value. For instance, increasing one's investment from £7 to £9 provides a trial-by-trial feedback response of 9/7 = 1.29. Decreasing one's investment from £9 to £7 equals 9/7 = 1.29, adding a minus sign equals −1.29. Higher values indicate greater feedback responsiveness. Proportional values take into consideration the magnitude of the increase/decrease in investment. As an illustration, an increase from £2 to £9 is higher than from £2 to £4. The increase of trust is therefore greater in the first observation and thus receives a larger value (4.5 vs. 2); this approach also takes into account the proportional differences of absolutely equal steps. To demonstrate, in absolute terms, an increase from £1 to £2 and from £9 to £10 is the same. In proportional terms, the former is larger than the latter, therefore represented by a larger value (2 vs. 1.11). Since a dividend or divisor of zero (i.e., investing £0) is not feasible for calculating proportions, the investments were recoded from £1 to £11; furthermore, if the investment was not altered between two rounds, the trial-by-trial feedback responses were recoded. Initially, not changing an investment offered a value of one (e.g., from £8 to £8; 8/8 = 1); however, an unadjusted investment could either signal an adequate or an inadequate reaction, as staying at a large investment is more responsive than staying at a small investment. On account of a 0.40 *M* and a 2.23 *SD* for the trial-by-trial feedback responses, it was decided that staying equal on the following investments translated into subsequent related values: 10 = 2; 9 = 1.6; 8 = 1.2; 7 = 0.8; 6 = 0.4; 5 = 0; 4 = −0.4; 3 = −0.8; 2 = −1.2; 1 = −1.6; 0 = −2. Based on all trial-by-trial feedback responses, a total sum feedback response value for each participant was created, henceforth referred to

Lemmers-Jansen, Imke L. J., Wichmann, Rune J., Perizonius, Sophie and Shergill, Sukhi S. (2022) *The Influence of Trial-By-Trial Feedback on Trust in Health, First-Episode and Chronic Psychosis.* Games, 13 (5). p. 59. ISSN 2073-4336.

as feedback response. Baseline trust, the general inclination to trust an unknown person, was determined by the investment of the first round.

4.3. Other Measures

To evaluate symptom severity, the Positive and Negative Syndrome Scale (PANSS; [55]) was integrated in the study; this semi-structured interview consists of three sub-scales: positive symptoms (seven items, e.g., delusions, hallucinations, suspiciousness); negative symptoms (seven items, e.g., blunted affect, social and emotional withdrawal) and general symptoms (sixteen items). General symptoms depict more global psychological deficits not directly linked to psychosis but to other potential problem areas (e.g., anxiety, depression, poor attention, poor cooperativeness). In the current paper, the sum scores of the three subscales was used. Combined, these additionally generated a total PANSS score. Scores per item fall between one and seven, with higher scores indicating higher symptom intensity. Ratings of three and above suggest clinical relevance [56]. The scale has been proven to be a reliable and valid measurement tool [55]. Both patient groups took the clinical interview. The interviews and the ratings were conducted by trained professionals and lasted approximately 30 to 45 min.

To control for intelligence, the Wechsler Abbreviated Scale of Intelligence (WASI; [57]), an estimate for general intellectual ability, was administered. The sum scores of the Vocabulary and the Matrix Reasoning subscales were calculated, providing a total WASI score.

4.4. Procedure

After initiating contact with participants, an information sheet was shared. Testing took place at the Institute of Psychiatry, Psychology and Neuroscience of King's College London. After an explanation of the testing procedure, subjects provided written informed consent, and were made aware that withdrawal was possible at any given stage. First, general details were gathered (e.g., demographics and medication usage), including the PANSS and WASI. As part of a larger research project, additional measures (including neuroimaging) irrelevant to the present paper were obtained. Before performing the trust game, several practice rounds were played under the supervision of the researchers, to ensure understanding of the task. The entire session lasted about four hours. Subjects were compensated monetarily for their time and travel cost. Ethics approval for the study was provided by the London Camberwell St. Giles Research and Ethics Committee. The experiment was in compliance with the Declaration of Helsinki regulations.

4.5. Analysis

To test for group differences in participant and patient characteristics, chi-square tests and ANOVAs with post hoc tests applying a Tukey HSD correction were performed in SPSS (version 27). To examine group differences in terms of feedback response and baseline trust, a one-way ANOVA was used. Controlling for confounders, an additional one-way ANCOVA with group as independent, mean feedback response as dependent variable and age, gender and WASI as covariates was performed. To inspect if the reward magnitude of the previous round had an influence on feedback response on the following round and if this association differed between groups, a multilevel, mixed-effects regression analysis was carried out in Stata version 16 (xtmixed; feedback response per trial [level 1]; within participants [level 2]), to account for the repeated measurements (i.e., the dependence of observations within participants of trial 2 to 20), with group and percentage of return and their respective interaction as independent variables. To analyse whether symptoms are associated with feedback response and baseline trust, two multiple linear regression analyses with PANSS subscales as independent variables and feedback response or first investment as dependent variables were exercised. The transformation of the raw data into feedback response values was carried out in RStudio (version 4.1.2).

5. Conclusions

Summarising, this novel, exhaustive assessment approach gave a more detailed insight in feedback learning behaviour than previous methods; it was demonstrated that CPs and FEPs were less responsive to feedback than HCs, showing a more random investment behaviour. Over repeated interactions, patients failed to incorporate the positive feedback into their decision-making, and thus behaved distrustingly towards an in fact trustworthy partner. Baseline trust was intact in patients, but reduced mean trust in patients reflects the lack of feedback response. The results further emphasise the social and interactional deficiencies of psychosis patients. The novel approach might be further improved, but showed promising results for investigating feedback response in a more detailed manner.

References

1. Yang, J.; Moorman, S.M. Beyond the individual: Evidence linking neighborhood trust and social isolation among community-dwelling older adults. *Int. J. Aging Hum. Dev.* **2021**, *92*, 22–39. [CrossRef] [PubMed]
2. Wang, G.; Hu, W. Peer relationships and college students' cooperative tendencies: Roles of interpersonal trust and social value orientation. *Front. Psychol.* **2021**, *12*, 656412. [CrossRef] [PubMed]
3. Vilares, I.; Dam, G.; Kording, K. Trust and reciprocity: Are effort and money equivalent? *PLoS ONE* **2011**, *6*, e17113. [CrossRef]
4. van den Bos, W.; van Dijk, E.; Westenberg, M.; Rombouts, S.A.; Crone, E.A. What motivates repayment? Neural correlates of reciprocity in the Trust Game. *Soc. Cogn. Affect. Neurosci.* **2009**, *4*, 294–304. [CrossRef] [PubMed]
5. Ferrin, D.L.; Bligh, M.C.; Kohles, J.C. Can I trust you to trust me? A theory of trust, monitoring, and cooperation in interpersonal and intergroup relationships. *Group Organ. Manag.* **2007**, *32*, 465–499. [CrossRef]
6. Lewicki, R.J.; Tomlinson, E.C.; Gillespie, N. Models of interpersonal trust development: Theoretical approaches, empirical evidence, and future directions. *J. Manag.* **2006**, *32*, 991–1022. [CrossRef]
7. Green, M.F.; Horan, W.P.; Lee, J. Social cognition in schizophrenia. *Nat. Rev. Neurosci.* **2015**, *16*, 620–631. [CrossRef]
8. Corrigan, P.W.; Green, M.F. Schizophrenic patients' sensitivity to social cues: The role of abstraction. *Am. J. Psychiatry* **1993**, *150*, 589–594.
9. Edwards, J.; Jackson, H.J.; Pattison, P.E. Emotion recognition via facial expression and affective prosody in schizophrenia: A methodological review. *Clin. Psychol. Rev.* **2002**, *22*, 789–832. [CrossRef]
10. Langdon, R.; Connors, M.H.; Still, M.; Ward, P.B.; Catts, S. Theory of mind and neurocognition in early psychosis: A quasi-experimental study. *BMC Psychiatry* **2014**, *14*, 316. [CrossRef]
11. Randall, F.; Corcoran, R.; Day, J.; Bentall, R. Attention, theory of mind, and causal attributions in people with persecutory delusions: A preliminary investigation. *Cogn. Neuropsychiatry* **2003**, *8*, 287–294. [CrossRef] [PubMed]
12. Trotta, A.; Kang, J.; Stahl, D.; Yiend, J. Interpretation bias in paranoia: A systematic review and meta-analysis. *Clin. Psychol. Sci.* **2021**, *9*, 3–23. [CrossRef]
13. Mitra, S.; Mahintamani, T.; Kavoor, A.R.; Nizamie, S.H. Negative symptoms in schizophrenia. *Ind. Psychiatry J.* **2016**, *25*, 135. [PubMed]
14. Velthorst, E.; Froudist-Walsh, S.; Stahl, E.; Ruderfer, D.; Ivanov, I.; Buxbaum, J.; Banaschewski, T.; Bokde, A.L.; Büchel, C.; Quinlan, E.B. Genetic risk for schizophrenia and autism, social impairment and developmental pathways to psychosis. *Transl. Psychiatry* **2018**, *8*, 204. [CrossRef]

15. Corcoran, C.; Kimhy, D.; Parrilla-Escobar, M.; Cressman, V.; Stanford, A.; Thompson, J.; David, S.B.; Crumbley, A.; Schobel, S.; Moore, H. The relationship of social function to depressive and negative symptoms in individuals at clinical high risk for psychosis. *Psychol. Med.* **2011**, *41*, 251–261. [CrossRef]

16. Yung, A.R.; Phillips, L.J.; Yuen, H.P.; Francey, S.M.; McFarlane, C.A.; Hallgren, M.; McGorry, P.D. Psychosis prediction: 12-month follow up of a high-risk ("prodromal") group. *Schizophr. Res.* **2003**, *60*, 21–32. [CrossRef]

17. Cornblatt, B.A.; Auther, A.M.; Niendam, T.; Smith, C.W.; Zinberg, J.; Bearden, C.E.; Cannon, T.D. Preliminary findings for two new measures of social and role functioning in the prodromal phase of schizophrenia. *Schizophr. Bull.* **2007**, *33*, 688–702. [CrossRef]

18. Baas, D.; Van't Wout, M.; Aleman, A.; Kahn, R. Social judgement in clinically stable patients with schizophrenia and healthy relatives: Behavioural evidence of social brain dysfunction. *Psychol. Med.* **2008**, *38*, 747–754. [CrossRef]

19. Keshavan, M.S.; Kulkarni, S.R.; Bhojraj, T.; Francis, A.; Diwadkar, V.; Montrose, D.M.; Seidman, L.; Sweeney, J. Premorbid cognitive deficits in young relatives of schizophrenia patients. *Front. Hum. Neurosci.* **2010**, *3*, 62. [CrossRef]

20. Berg, J.; Dickhaut, J.; McCabe, K. Trust, reciprocity, and social history. *Games Econ. Behav.* **1995**, *10*, 122–142. [CrossRef]

21. Alós-Ferrer, C.; Farolfi, F. Trust games and beyond. *Front. Neurosci.* **2019**, *13*, 887. [CrossRef] [PubMed]

22. Banerjee, S.; Galizzi, M.M.; Hortala-Vallve, R. Trusting the trust game: An external validity analysis with a UK representative sample. *Games* **2021**, *12*, 66. [CrossRef]

23. Brülhart, M.; Usunier, J.-C. Does the trust game measure trust? *Econ. Lett.* **2012**, *115*, 20–23. [CrossRef]

24. Fett, A.-K.J.; Shergill, S.S.; Joyce, D.; Riedl, A.; Strobel, M.; Gromann, P.; Krabbendam, L. To trust or not to trust: The dynamics of social interaction in psychosis. *Brain* **2012**, *135*, 976–984. [CrossRef] [PubMed]

25. Gromann, P.; Heslenfeld, D.; Fett, A.-K.J.; Joyce, D.; Shergill, S.; Krabbendam, L. Trust versus paranoia: Abnormal response to social reward in psychotic illness. *Brain* **2013**, *136*, 1968–1975. [CrossRef] [PubMed]

26. Hanssen, E.; van Buuren, M.; Van Atteveldt, N.; Lemmers-Jansen, I.L.; Fett, A.-K.J. Neural, behavioural and real-life correlates of social context sensitivity and social reward learning during interpersonal interactions in the schizophrenia spectrum. *Aust. N. Z. J. Psychiatry* **2021**, *56*, 59–70. [CrossRef]

27. Lemmers-Jansen, I.L.; Fett, A.-K.J.; Hanssen, E.; Veltman, D.J.; Krabbendam, L. Learning to trust: Social feedback normalizes trust behavior in first-episode psychosis and clinical high risk. *Psychol. Med.* **2018**, *49*, 780–790. [CrossRef]

28. Fett, A.-K.J.; Mouchlianitis, E.; Gromann, P.; Vanes, L.; Shergill, S.S.; Krabbendam, L. The neural mechanisms of social reward in early psychosis. *Soc. Cogn. Affect. Neurosci.* **2019**, *14*, 861–870. [CrossRef]

29. Gromann, P.; Shergill, S.; De Haan, L.; Meewis, D.; Fett, A.-K.J.; Korver-Nieberg, N.; Krabbendam, L. Reduced brain reward response during cooperation in first-degree relatives of patients with psychosis: An fMRI study. *Psychol. Med.* **2014**, *44*, 3445–3454. [CrossRef]

30. Fett, A.-K.J.; Shergill, S.S.; Korver-Nieberg, N.; Yakub, F.; Gromann, P.; Krabbendam, L. Learning to trust: Trust and attachment in early psychosis. *Psychol. Med.* **2016**, *46*, 1437–1447. [CrossRef]

31. Hanssen, E.; Krabbendam, L.; Robberegt, S.; Fett, A.-K.J. Social and non-social reward learning reduced and related to a familial vulnerability in schizophrenia spectrum disorders. *Schizophr. Res.* **2020**, *215*, 256–262. [CrossRef] [PubMed]

32. Sutherland, C.A.; Rhodes, G.; Williams, N.; Connaughton, E.; Ewing, L.; Caruana, N.; Langdon, R. Appearance-based trust processing in schizophrenia. *Br. J. Clin. Psychol.* **2019**, *59*, 139–153. [CrossRef] [PubMed]

33. Campellone, T.R.; Fisher, A.J.; Kring, A.M. Using social outcomes to inform decision-making in schizophrenia: Relationships with symptoms and functioning. *J. Abnorm. Psychol.* **2016**, *125*, 310. [CrossRef] [PubMed]

34. Campellone, T.R.; Truong, B.; Gard, D.; Schlosser, D.A. Social motivation in people with recent-onset schizophrenia spectrum disorders. *J. Psychiatr. Res.* **2018**, *99*, 96–103. [CrossRef]

35. Fett, A.K.; Gromann, P.M.; Giampietro, V.; Shergill, S.S.; Krabbendam, L. Default distrust? An fMRI investigation of the neural development of trust and cooperation. *Soc. Cogn. Affect. Neurosci.* **2014**, *9*, 395–402. [CrossRef]

36. Kim, S.-S.; Chung, Y.; Perry, M.J.; Kawachi, I.; Subramanian, S. Association between interpersonal trust, reciprocity, and depression in South Korea: A prospective analysis. *PLoS ONE* **2012**, *7*, e30602. [CrossRef]

37. Zhang, H.j.; Sun, D.; Lee, T.M. Impaired social decision making in patients with major depressive disorder. *Brain Behav.* **2012**, *2*, 415–423. [CrossRef]

38. Albrecht, M.A.; Waltz, J.A.; Frank, M.J.; Gold, J.M. Probability and magnitude evaluation in schizophrenia. *Schizophr. Res. Cogn.* **2016**, *5*, 41–46. [CrossRef] [PubMed]

39. Hernaus, D.; Frank, M.J.; Brown, E.C.; Brown, J.K.; Gold, J.M.; Waltz, J.A. Impaired expected value computations in schizophrenia are associated with a reduced ability to integrate reward probability and magnitude of recent outcomes. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* **2019**, *4*, 280–290. [CrossRef]

40. Fett, A.-K.J.; Shergill, S.S.; Gromann, P.; Dumontheil, I.; Blakemore, S.-J.; Yakub, F.; Krabbendam, L. Trust and social reciprocity in adolescence—A matter of perspective-taking. *J. Adolesc.* **2014**, *37*, 175–184. [CrossRef]

41. Seres, I.; Unoka, Z.; Keri, S. The broken trust and cooperation in borderline personality disorder. *Neuroreport* **2009**, *20*, 388–392. [CrossRef] [PubMed]

42. Langdon, R.; Connors, M.H.; Connaughton, E. Social cognition and social judgment in schizophrenia. *Schizophr. Res. Cogn.* **2014**, *1*, 171–174. [CrossRef] [PubMed]

43. Sparks, A.; McDonald, S.; Lino, B.; O'Donnell, M.; Green, M.J. Social cognition, empathy and functional outcome in schizophrenia. *Schizophr. Res.* **2010**, *122*, 172–178. [CrossRef] [PubMed]

44. An, S.K.; Kang, J.I.; Park, J.Y.; Kim, K.R.; Lee, S.Y.; Lee, E. Attribution bias in ultra-high risk for psychosis and first-episode schizophrenia. *Schizophr. Res.* **2010**, *118*, 54–61. [CrossRef]

45. Combs, D.R.; Penn, D.L.; Michael, C.O.; Basso, M.R.; Wiedeman, R.; Siebenmorgan, M.; Tiegreen, J.; Chapman, D. Perceptions of hostility by persons with and without persecutory delusions. *Cogn. Neuropsychiatry* **2009**, *14*, 30–52. [CrossRef]

46. Harley, E.W.-Y.; Boardman, J.; Craig, T. Friendship in people with schizophrenia: A survey. *Soc. Psychiatry Psychiatr. Epidemiol.* **2012**, *47*, 1291–1299. [CrossRef]

47. Lamster, F.; Nittel, C.; Rief, W.; Mehl, S.; Lincoln, T. The impact of loneliness on paranoia: An experimental approach. *J. Behav. Ther. Exp. Psychiatry* **2017**, *54*, 51–57. [CrossRef]

48. Corrigan, P.W.; Phelan, S.M. Social support and recovery in people with serious mental illnesses. *Community Ment. Health J.* **2004**, *40*, 513–523. [CrossRef]

49. Leucht, S.; Kane, J.M.; Kissling, W.; Hamann, J.; Etschel, E.; Engel, R.R. What does the PANSS mean? *Schizophr. Res.* **2005**, *79*, 231–238. [CrossRef]

50. Johnson, N.D.; Mislin, A.A. Trust games: A meta-analysis. *J. Econ. Psychol.* **2011**, *32*, 865–889. [CrossRef]

51. Vlaev, I. How different are real and hypothetical decisions? Overestimation, contrast and assimilation in social interaction. *J. Econ. Psychol.* **2012**, *33*, 963–972. [CrossRef]

52. Locey, M.L.; Jones, B.A.; Rachlin, H. Real and hypothetical rewards. *Judgm. Decis. Mak.* **2011**, *6*, 552. [PubMed]

53. Derks, J. Adolescent Social Cognition. Ph.D. Thesis, Faculty of Psychology and Educational sciences, Vrije Universiteit, Amsterdam, The Netherlands, 2015.

54. Azar, O.H. Relative thinking theory. *J. Socio-Econ.* **2007**, *36*, 1–14. [CrossRef]

55. Kay, S.R.; Fiszbein, A.; Opler, L.A. The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophr. Bull.* **1987**, *13*, 261–276. [CrossRef] [PubMed]

56. Hermes, E.D.; Sokoloff, D.; Stroup, T.S.; Rosenheck, R.A. Minimum clinically important difference in the Positive and Negative Syndrome Scale with data from the Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE). *J. Clin. Psychiatry* **2012**, *73*, 11983. [CrossRef]

57. Wechsler, D. Wechsler Abbreviated Scale of Intelligence; APA PsycTests: 1999. Available online: https://psycnet.apa.org/doiLanding?doi=10.1037%2Ft15171-000 (accessed on 1 July 2022).