



## City Research Online

### City, University of London Institutional Repository

---

**Citation:** Lemmers-Jansen, I. L. J., Fett, A-K., Hanssen, E., Veltman, D. J. & Krabbendam, L. (2018). Learning to trust: social feedback normalizes trust behavior in first-episode psychosis and clinical high risk. *Psychological Medicine*, 49(5), pp. 780-790. doi: 10.1017/s003329171800140x

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/19935/>

**Link to published version:** <https://doi.org/10.1017/s003329171800140x>

**Copyright:** 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.

**Reuse:** Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

---

City Research Online:

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

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

---

1 **Title: Learning to trust: Social feedback normalizes trust behavior in first episode psychosis**  
2 **and clinical high-risk**

3 **Authors:** Imke L.J. Lemmers-Jansen <sup>a</sup> , Anne-Kathrin J. Fett <sup>a,b,c,d</sup>, Esther Hanssen <sup>a</sup>, Dick J. Veltman <sup>e,f</sup>,  
4 & Lydia Krabbendam <sup>c,d</sup>

5 **Affiliations:**

6 <sup>a</sup> Department of Educational and Family studies, Faculty of Behavioral and Movement Sciences, and  
7 Institute for Brain and Behavior Amsterdam, Vrije Universiteit Amsterdam, Van der  
8 Boechorststraat 1, 1081 BT Amsterdam, The Netherlands

9 <sup>b</sup> Department of Psychology, City, University of London, Northampton Square, London EC1V 0HB,  
10 United Kingdom

11 <sup>c</sup> Department of Clinical, Neuro and Developmental Psychology, Faculty of Behavioral and  
12 Movement Sciences, and Institute for Brain and Behavior Amsterdam, Vrije Universiteit  
13 Amsterdam, Van der Boechorststraat 1, 1081 BT Amsterdam, The Netherlands

14 <sup>d</sup> Department of Psychosis Studies, King's College London, Institute of Psychiatry, Psychology and  
15 Neuroscience, 16 De Crespigny Park, London SE5 8AF, Great Britain

16 <sup>e</sup> Department of Psychiatry, VU Medical Center, Van der Boechorststraat 7 , 1081 BT Amsterdam,  
17 The Netherlands

18 <sup>f</sup> Neuroscience Campus Amsterdam, De Boelelaan 1117, 1081 HV Amsterdam, The Netherlands

19

20 **Email addresses:** [imke.jansen@vu.nl](mailto:imke.jansen@vu.nl), [Anne-Kathrin.Fett@city.ac.uk](mailto:Anne-Kathrin.Fett@city.ac.uk), [esther.hanssen@vu.nl](mailto:esther.hanssen@vu.nl),  
21 [dj.veltman@vumc.nl](mailto:dj.veltman@vumc.nl), [lydia.krabbendam@vu.nl](mailto:lydia.krabbendam@vu.nl)

22 Manuscript consist of:

23 Words: 6010

24 Tables: 2

25 Graphs: 3

26

## Trust in CHR and psychosis

27 Abbreviations

28 FEP = First episode psychosis

29 CHR = Clinical high-risk

30 CAARMS = Comprehensive Assessment of At-Risk Mental States

31 PANSS = Positive and Negative Syndromes Scale

32 WAIS = Wechsler Adult Intelligence Scale

33 ROI = Region of interest

34 TPJ = Temporo-parietal junction

35 mPFC = medial prefrontal cortex

36

37 Abstract

38 Background: Psychosis is characterized by problems in social functioning that exist well before  
39 illness onset, and in individuals at clinical high-risk (CHR) for psychosis. Trust is an essential  
40 element for social interactions that is impaired in psychosis. In the trust game, chronic patients  
41 showed reduced baseline trust, impaired response to positive social feedback, and attenuated brain  
42 activation in reward and mentalizing areas. We investigated whether first episode psychosis  
43 patients (FEP) and CHR show similar abnormalities in the neural and behavioral mechanisms  
44 underlying trust.

45 Methods: Twenty-two FEP, 17 CHR, and 43 healthy controls performed two trust games, with a  
46 cooperative and an unfair partner in the fMRI scanner. Region of interest analyses were performed  
47 on mentalizing and reward processing areas, during the investment and outcome phases of the  
48 games.

49 Results: Compared to healthy controls, FEP and CHR showed reduced baseline trust, but like  
50 controls, learned to trust in response to cooperative and unfair feedback. Symptom severity was not  
51 associated with baseline trust, however in FEP associated with reduced response to feedback. The  
52 only group differences in brain activation were that CHR recruited the temporo-parietal junction  
53 (TPJ) more than FEP and controls during investment in the unfair condition. This hyper-activation  
54 in CHR was associated with greater symptom severity.

55 Conclusions: Reduced baseline trust may be associated with risk for psychotic illness, or generally  
56 with poor mental health. Feedback learning is still intact in CHR and FEP, as opposed to chronic  
57 patients. CHR however show distinct neural activation patterns of hyper-activation of the TPJ.

58 Keywords:

59 First episode psychosis; clinical high-risk; trust; social feedback; fMRI

60 1 Introduction

61 Psychosis is characterized by problems in social functioning (Couture et al, 2006; Fett et al, 2012).  
62 Lower social functioning is already present in childhood in individuals who continue to develop  
63 psychosis and has also been reported in individuals at high-risk for psychosis (Ballon et al, 2007;  
64 Corcoran et al, 2011; Cornblatt et al, 2007; Velthorst et al, 2016a; Velthorst et al, 2016b; Yung et al,  
65 2003). Clinical high-risk patients (CHR) are already in care for other psychopathology, reporting  
66 psychotic-like symptoms, but have not yet experienced – and maybe never will – full-blown  
67 psychosis. In CHR, the developmental course of social functioning is predictive of the conversion to  
68 psychosis (Cannon et al, 2008; Jang et al, 2011; Niendam et al, 2007). Understanding the  
69 mechanisms underlying deficits in social functioning in at-risk states and first-episode psychosis  
70 (FEP) is crucial for understanding transition and outcome prognosis. Intervening at these early  
71 stages targeting social functioning can improve outcome and possibly delay (or prevent) transition.

72 Social functioning relates to establishing relationships, both vocational and private  
73 (Velthorst et al, 2016a; Velthorst et al, 2016b). Patients show a steep decline in these domains  
74 starting about five years before illness onset. The basis of social functioning is the ability to interact  
75 in an appropriate way with other people. Previous research studying online social interactions in  
76 psychosis has suggested two possible explanatory mechanisms for impairments in social  
77 interactions; these are a reduced sensitivity to rewarding effects of social contact (Campellone et al,  
78 2016; Fett et al, 2012; Gromann et al, 2013), and an impaired social cognitive ability (Csukly et al,  
79 2011; Horat et al, 2017), including impaired mentalizing (Green et al, 2015). Social cognitive skills  
80 (Couture et al, 2006; Green & Leitman, 2008) are necessary for the formation and maintenance of  
81 relationships and for building trust in other people. Like patients with psychosis, CHR show deficits  
82 in a variety of these skills (Bora & Pantelis, 2013; Lavoie et al, 2013; McCleery et al, 2014), albeit to  
83 a lesser degree. Research has mainly focused on off-line cognitive skills, without investigating them

84 in real interactions. In the last decade, interactive designs have been widely used, that have the  
85 strength to capture social cognitive skills, as well as the rewarding effects of social behavior in an  
86 on-line setting. We therefore investigated cooperative and unfair social interactions and the neural  
87 correlates of trust, directly comparing FEP and CHR to controls, using an interactive trust game to  
88 test whether these groups display similar underlying mechanisms of reduced social interactions.

89         The trust game investigates real-time social interactions (Berg et al, 1995). In the game, the  
90 first player (investor), receives a certain endowment, e.g., €10. He or she can give any amount  
91 between €0 and €10 to the second player, the trustee. The given amount is tripled and the trustee  
92 then can return any part of this amount to the investor. The best pay-off for the trustee is reached  
93 by keeping the money. Thus, investing requires trust that a fair repayment will be made. The  
94 iterative game allows for the investigation of baseline trust (i. e. first investment), and the  
95 development of trust based on cooperative and unfair social feedback. Key-processes involved in  
96 the trust game are thought to be mentalizing (Declerck et al, 2013; Frith & Frith, 2006; Gallagher &  
97 Frith, 2003) and reward processing (Fehr & Camerer, 2007; King-Casas et al, 2005; Rilling & Sanfey,  
98 2011). Mentalizing appears to be important during both the investment and repayment phase,  
99 where estimations of the other's behavior are made. Reward learning signals have been shown to  
100 shift from the repayment phase to the investment phase in an iterative trust game (King-Casas et al,  
101 2005). Hence, we investigated both the investment and repayment phase (Figure 1).

102         Research in healthy subjects has shown that participants initially invest more than half of  
103 their endowment (Berg et al, 1995; Johnson & Mislin, 2011). Studies from our lab have shown that  
104 baseline trust tends to be lower in patients than controls (Fett et al, 2016; Gromann et al, 2013).  
105 Both positive (Fett et al, 2012) and negative (Fett et al, 2016) symptoms have been associated with  
106 lower baseline trust, suggesting that reduced trust may reflect either paranoia or a lack of social  
107 motivation. The ability to learn from social feedback seems to depend on context (cooperative or

108 unfair partner's responses) and illness duration: Early psychosis patients were able to adjust their  
109 trust to similar levels as controls, whereas chronic patients showed an insensitivity to positive  
110 feedback. In unfair interactions, early and chronic psychosis patients responded adequately to  
111 negative feedback (Campellone et al, 2016; Fett et al, 2015; Fett et al, 2016; Fett et al, 2012;  
112 Gromann et al, 2013). Understanding the mechanisms of trust in early psychosis stages may  
113 provide insights in focal points to target in social functioning interventions.

114 At the neural level, reduced caudate activation in chronic patients has been reported in  
115 cooperative interactions. Relatives of patients with psychosis, despite behavioral outcomes similar  
116 to controls, also showed reduced recruitment of the caudate and insula. These results possibly  
117 reflect reduced sensitivity to social reward processing mechanisms in both patients and relatives  
118 (Gromann et al, 2013; Gromann et al, 2014), which could account for social impairments.  
119 Associations of neural activity with positive symptoms have been reported (Gromann et al, 2013).

120 This study set out to investigate whether CHR and FEP patients, similar to chronic patients,  
121 show reduced baseline trust and to explore the neural mechanisms underlying trust behavior in  
122 these patient groups. Based on the existing trust game literature, we hypothesized that similar to  
123 relatives (Fett et al, 2012) and (chronic) patients (Fett et al, 2015) 1) FEP and CHR will show lower  
124 levels of baseline trust, and 2) CHR and FEP are able to learn from positive and negative feedback  
125 given by the counterpart and adjust their levels of trust accordingly (Fett et al, 2015; Gromann et al,  
126 2013). In both cases, we expected CHR to perform in between FEP and controls. In addition, the  
127 associations of symptoms with baseline trust and changes in trust in FEP and CHR were examined.  
128 On the neural level we hypothesized to find 3) attenuated activation in brain areas associated with  
129 mentalizing and reward (learning) in FEP compared to controls. Based on the trust literature in  
130 relatives and imaging research in CHR (Smieskova et al, 2013), we expected intermediate activation  
131 in CHR. Based on the findings by Gromann et al (2013), we hypothesized to find 4) positive



132 symptoms related positively to brain activation in mentalizing and negatively to reward areas in  
133 both patient groups. In addition, associations of brain activations with negative symptoms were  
134 investigated.

135

## 136 2 Methods

### 137 2.1 Subjects

138 Twenty-six FEP patients with non-affective psychosis, aged 16-22, and 17 CHR, aged 16-31, were  
139 recruited in the Amsterdam and The Hague area. Forty-nine healthy control participants (aged 16-  
140 31) were recruited to match both patient populations on age and gender. Patients were contacted  
141 through their caregivers at the academic medical center Amsterdam (AMC), the Amsterdam early  
142 intervention team psychosis and PsyQ, The Hague. FEP were diagnosed at the AMC, according to the  
143 DSM-IV criteria (Association & Association, 2000). FEP were included within 18 months of the  
144 diagnosis (mean 5.6 months). CHR were help seeking individuals that were referred to PsyQ by  
145 their general practitioners or other mental health institutions. After an initial diagnosis based on  
146 their complaints, all new admissions (between age 14-35) were screened for an “at-risk mental  
147 state” (ARMS) with the Comprehensive Assessment of At-Risk Mental States (CAARMS) (Yung et al,  
148 2005), a semi-structured interview that assesses psychotic experiences in the last year before  
149 assessment. Only the positive symptoms sub-scale was used, and both intensity and frequency of  
150 the symptoms were assessed. Patients met ARMS criteria for attenuated psychotic symptoms (a)  
151 with subthreshold intensity, when scoring 3-5 on severity and 3-6 on frequency of symptoms (all  
152 CHR participants), (b) or based on subthreshold frequency, when scoring 5-6 on severity and 3 or  
153 above on frequency (N = 2, in combination with intensity), or based on vulnerability, i.e.,  
154 schizotypal personality disorder, familial history of psychosis, or a drop in social functioning (N = 3,

155 in combination with intensity) (for details, see (Yung et al, 2005), p. 966). Patients were diagnosed  
156 as having a full-blown psychosis when scoring higher than 6. Additionally, patients had to display  
157 marked problems in socially useful activities (work and study), relationships, and self-care,  
158 indicated by a score below 55 on the Social and Occupational Functioning Assessment Scale  
159 (SOFAS; mean score 46.9) (Goldman et al, 1992; Morosini et al, 2000), see also (Rietdijk et al,  
160 2012). CHR were included within one year after CAARMS assessment (mean 4.8 months). Fourteen  
161 CHR patients also took part in a larger study (EU-GEI), with post-measurements at 6, 12 and 24  
162 months using the CAARMS. Of two CHR participants, follow-up data are missing. Symptoms of  
163 depression and anxiety are often the primary presenting complaints of CHR patients, rather than  
164 the attenuated psychotic symptoms (Modinos et al, 2014). Similar to other CHR samples (Fusar-Poli  
165 et al, 2014; Kelleher et al, 2012; Morrison et al, 2012; Wigman et al, 2012; Woods et al, 2009), the  
166 current CHR sample had comorbid diagnoses of anxiety (5), personality (3), eating (2) and mood  
167 (2) disorders, trauma (2), and ADHD (3). Exclusion criteria for all participants were an IQ < 80 and  
168 contraindications for scanning. For FEP additional exclusion criteria were a primary diagnosis of a  
169 mood disorder, and comorbidity with autism spectrum disorder. Healthy control participants were  
170 excluded if they had a (family) history of psychopathology, which was assessed with self-report,  
171 and by a systematic interview with questions regarding past and present mental help seeking,  
172 depressed and psychotic symptoms, and intake of medication.

173 We excluded four FEP and six controls due to invalid behavioral data (1 FEP and 2 controls),  
174 unusable or missing imaging data (3 FEP and 4 controls). The analysis sample consisted of 22 FEP,  
175 17 CHR and 43 controls.

## 176 2.2 Measures

### 177 2.2.1 *Trust game*

178 Participants played the role of investor in two multi-round trust games. They were told that they  
179 were connected to their anonymous counterpart via the Internet. In reality, they played against a  
180 computer, programmed to respond either in a cooperative or in an unfair way. In the cooperative  
181 condition, the return was 100%, 150% or 200% of the invested amount, with increasing likelihood  
182 of a 200% repayment after each increase of investment. In the unfair condition the return was 75%  
183 or 50%, with increasing likelihood of a 50% repayment after increase of investment. The two  
184 games were presented in counterbalanced order. Each game consisted of 20 experimental and 20  
185 control trials (Figure 1). For a detailed description of the paradigm see Lemmers-Jansen et al  
186 (2017). After the trust game, a questionnaire to investigate participants' opinions on the behavior  
187 of their counterpart was administered, to check if participants believed that they were playing a  
188 real person.

### 189 *2.2.2 Positive and Negative Syndrome Scale (PANSS; Kay et al, 1987)*

190 The 30-item PANSS semi-structured interview was used for rating symptoms in the two weeks  
191 prior to testing. The PANSS distinguishes between positive, negative, and general symptoms (Kay et  
192 al, 1987). Items are scaled on a 7-point Likert scale, ratings 3 and higher indicating clinical values.  
193 All FEP and 13 CHR completed the interview. Interviews were taken by four researchers, and audio  
194 tapes (and if consented video tapes) were made. Responses were rated on the basis of the  
195 recordings and notes taken during the interview, by two researchers. Based on the first participants  
196 an interrater reliability was calculated ( $r = .85$ ). All PANSS data were rated by the same two  
197 researchers.

### 198 *2.2.3 Wechsler Adult Intelligence Scale (WAIS; Wechsler, 1997)*

199 To control for confounding effects of intelligence, the vocabulary subscale of the WAIS was  
200 included, a measure of verbal comprehension, consisting of 33 words that had to be defined by the

201 participants. Answers were coded as fully correct (2 points), partially correct (1) or wrong (0),  
202 resulting in a maximum score of 66.

#### 203 2.2.4 Trust Manipulation Check Questionnaire

204 The trust game was followed by a questionnaire to investigate participants' opinions on the  
205 behavior of their counterpart, and to check if they believed that they were playing a real person  
206 ("What do you think this task was about?", "What were the main causes that influenced your  
207 behavior during the task?", "Did you use a strategy during the tasks? If so, which strategy?", and  
208 "Did you think that the counterparts made fair choices?"). The last questions required answers  
209 specified per counterpart. If a participant referred to the two counterparts as persons in his/her  
210 responses, we regarded the manipulation as successful. If participants reported on any of the  
211 questions that they had doubts, or did not believe that the counterpart was real, the manipulation  
212 was coded as failed. Four controls, three FEP and one CHR did not believe the manipulation.

213

#### 214 2.3 Procedure

215 This research was approved by the Medical Ethical Committee of the VU Medical Center  
216 Amsterdam. All participants signed an informed consent, and completed several questionnaires  
217 that are not included in the current paper. We then administered the PANSS to FEP and CHR  
218 patients. Prior to scanning, participants received oral and written instructions for the trust game,  
219 illustrated with screenshots of the game to illustrate the task. Participants played several practice  
220 rounds on a computer, accompanied by additional feedback, to ensure understanding of the task.  
221 Subsequently participants were scanned for about an hour. First, participants performed the trust  
222 game, followed by the structural scan. After this period of relative rest, they performed a second  
223 task (the Social Mindfulness Paradigm, to be reported in a separate paper (in preparation)),

224 followed by a resting state scan. Immediately after scanning, a manipulation check for the trust  
225 game was administered (see 2.2.4). Participants received an image of their brain, 25€ for  
226 participation and reimbursement of their travel costs.

## 227 2.4 fMRI Data Acquisition.

228 Imaging data were obtained at the Spinoza Center Amsterdam, using a 3.0 T Philips Achieva whole  
229 body scanner (Philips Healthcare, Best, The Netherlands) equipped with a 32 channel head coil. A  
230 T2\* EPI sequence (TR = 2.31, TE = 27.63, FA = 76.1°, FOV 240mm, voxel size 2.5 x 2.5 x 2.5, 40  
231 slices, 0.3 mm gap) was used, which resulted in 325 images per condition. A T1-weighted scan was  
232 obtained for anatomical reference (TR = 8.2, TE = 3.8, FA = 8°, FOV 240\*188mm, voxel size 1 x 1 x 1,  
233 220 slices).

## 234 2.5 Data Analysis

### 235 2.5.1 Behavioral data

236 Demographic and behavioral data were analyzed using Stata 13 (StataCorp, 2013) with regression  
237 analyses and chi-square tests. We analyzed group differences in first investment (baseline trust),  
238 and the development of investments (changes in trust) across repeated interactions (indicated by  
239 trial number) in each game (cooperative and unfair). We used multilevel random regression  
240 analyses to account for multiple observations [investments (level 1); within participants (level 2)].  
241 All analyses were controlled for age and WAIS score.

### 242 2.5.2 Imaging data

243 Imaging data were analyzed using Statistical Parametric Mapping 8 (SPM, 2009). Functional images  
244 for each participant were preprocessed as follows: realign and unwarp, coregistration with  
245 individual structural images, segmented for normalization to an MNI template and smoothing with  
246 a 6 mm Gaussian kernel (FWHM). At fist-level, a general linear model was used to construct

247 individual time courses for the investment and repayment phase per condition, using an event-  
248 related design. For each trial we defined the investment as the period of stimulus onset to the  
249 moment of investment, and the repayment phase as the period during which the partner's return  
250 was displayed (Figure 1). Trials from both the cooperative and unfair conditions were contrasted  
251 with control trials. Additionally, cooperative trials were contrasted with unfair trials, to directly  
252 compare the differences in response to cooperative and unfair feedback.

253 A priori ROI analyses were performed. ROIs were derived from Gromann et al (2014).  
254 Talairach coordinates were converted to MNI space (tal2mni under MatLab), resulting in the  
255 following ROIs: right caudate (MNI coordinates 16, 17, 7), superior temporal sulcus (STS; 62, -58, 5)  
256 and TPJ (51, -57, 26), left insula (-33, 14, 0), and medial prefrontal cortex (mPFC; -3, 65, 25). We  
257 tested group differences using MarsBaR (version0.43; <http://marsbar.sourceforge.net>). An adjusted  
258 *p*-value was calculated, taking the correlation between the  $\beta$ -values into account by using the  
259 Simple Interactive Statistical Analysis Bonferroni tool  
260 (<http://www.quantitativeskills.com/sisa/calculations/bonfer.htm>), resulting in adjusted *p*-values  
261 (see Table 2) (Li et al, 2014; Woudstra et al, 2013). Additional whole-brain analyses were  
262 performed, to investigate activation outside the predefined ROIs.

### 263 2.5.3 Associations with symptoms

264 Group differences in the association of first investment and development of investments with  
265 symptoms (paranoia item, positive and negative PANSS subscales) were investigated. The  
266 persecution item (P6) and the depression item (G6) were used as an additional index for paranoid  
267 ideation, as previously reported (Gromann et al, 2013), and depression, based on CHR comorbidity.  
268 Second, beta weights of the ROIs (average over all voxels) were associated with symptoms. To  
269 further explore the association of symptoms with behavior and brain activation, additional post-hoc  
270 analyses within patient groups were performed.

271 3 Results

272 3.1 Participant characteristics

273 Participant characteristics and baseline trust are displayed in Table 1. CHR were significantly older  
274 than controls ( $\beta = .40, p < .001$ ) and FEP ( $\beta = .57, p < .001$ ), and FEP scored significantly lower on  
275 the WAIS vocabulary scale than controls ( $\beta = -.39, p = .003$ ) and CHR ( $\beta = -.31, p = .01$ ). The time  
276 between diagnosis and inclusion in the study did not differ significantly between FEP and CHR ( $\beta =$   
277  $-.35, p = .21$ ).

278 3.2 Behavioral results

279 Based on the group differences, all analyses were controlled for age and WAIS vocabulary score.  
280 Four controls, three FEP and one CHR did not believe they played against a human counterpart.  
281 Analyses without these subjects yielded similar results. Below the results of the complete sample  
282 are reported.

283 Group differences in baseline trust – the first investment of the first game – were found ( $\beta =$   
284  $-.27, p = .02$ ; Table 1), with FEP and CHR showing lower baseline trust than controls (FEP:  $\beta = -0.24,$   
285  $p = .04$ ; CHR:  $\beta = -0.25, p < .05$ ). CHR and FEP did not differ significantly from each other ( $p = 0.8$ ).

286 To investigate the development of trust over trials we performed a three-way interaction  
287 “trial number-by-group-by-condition” on investment. This interaction was not significant and  
288 therefore removed from the model. Significant trial number-by-group, condition-by-group and trial  
289 number-by-condition interactions on investment were found ( $b = .03, 95\% \text{ CI } [.031, .05], p = .001$ ;  $b$   
290  $= .49, 95\% \text{ CI } [.28, .71], p < .001$ ;  $b = -.20, 95\% \text{ CI } [-.24, -.17], p < .001$ , respectively), indicating that  
291 the development of trust differed between groups (Figure 2).

292 In the *cooperative condition* there was a significant group-by trial number interaction on  
293 investment ( $b = .03, 95\% \text{ CI } [.006, .05], p = .01$ ), with FEP showing significantly stronger increase  
294 than controls ( $b = .08, 95\% \text{ CI } [-.07, .08], p = .03$ ). Controls and CHR did not differ significantly from  
295 each other ( $p = .9$ ). Analysis by group showed that all groups increased investments significantly

296 (all  $p$ 's  $<.01$ ; see Figure 2a). In the *unfair condition*, analysis revealed a significant group-by trial  
297 number interaction on investment ( $b = .03$ , 95% CI [.01, .06],  $p = .005$ ), with FEP showing  
298 significantly less decrease than the other groups ( $b = -.97$ , 95% CI [-1.89, -.06],  $p = .04$ ). All groups  
299 decreased investments significantly (all  $p$ 's  $<.001$ ; see Figure 2b).

300 All analyses were also conducted with medication type (no medication; atypical anti-  
301 psychotics; combination of typical and atypical; other psychotropic medication) as a grouping  
302 variable. No differences in baseline trust, nor in adjustment of trust were found between the  
303 medication groups, and no interactions with symptoms on trust were found.

304

### 305 3.3 Imaging results

306 ROI analyses revealed significant group differences in the right TPJ only, showing more  
307 activation in CHR compared to controls and FEP during the investment phase in the unfair  
308 condition (Table 2). Furthermore, CHR activated the TPJ and mPFC more than controls, when  
309 investing in an unfair partner compared to a cooperative partner. Other ROIs showed no group  
310 differences. During cooperative investment, cooperative and unfair repayment no significant group  
311 differences in ROI activation were found. All ROI analyses were also conducted between medication  
312 groups. No significant differences in activation were found.

313 Additional exploratory whole-brain analyses, based on a significance level of  $p < .05$  FWE  
314 cluster corrected, did not reveal significant group differences. Results with a more lenient threshold  
315 are presented in Supplementary Table S1.

### 316 3.4 Symptoms

317 CHR and FEP did not differ significantly from each other in terms of overall and positive symptoms,  
318 and on the depression item. Only on the paranoia item, CHR scored significantly higher than FEP ( $\beta$   
319 = .37,  $p = .03$ ). There was a trend towards significance indicating that FEP had higher negative  
320 symptoms than CHR ( $p < .07$ ).



321 3.4.1 Associations between behavioral outcomes and symptoms

322 No group-by-symptoms interactions on first investment were found (all  $ps > .24$ ). After removing  
323 the interaction from the model, no main effects of symptoms on baseline trust were found (all  $ps >$   
324  $.16$ ).

325 In the *cooperative condition*, the group-by-trial number-by-symptoms models showed a  
326 significant interaction for negative symptoms only ( $b = -.02$ , 95% CI  $[-.03, -.001]$ ,  $p = .03$ ), indicating  
327 that negative symptoms impacted upon the development of trust differentially in the three groups.  
328 Post-hoc analyses showed a significant association between symptoms and changes in investments  
329 over trials in FEP ( $b = -.01$ , 95% CI  $[-.02, -.003]$ ,  $p = .004$ ), but not in CHR. To visualize this  
330 association, we divided the negative symptoms in three levels (Figure 3). Analysis indicated that the  
331 only highest level of negative symptoms interfered with increasing investments. No significant  
332 interactions with positive symptoms, paranoia, or depression were found.

333 In the *unfair condition* the group-by-trial number-by-symptoms models did not show  
334 significant interactions. After removing the 3-way interaction from the model, the interaction of  
335 positive symptoms with trial number became significant ( $b = .01$ , 95% CI  $[.0002, .012]$ ,  $p = .04$ ).  
336 Higher positive symptoms were associated with less decrease in investments in FEP and CHR.  
337 Associations of decreasing investment with negative symptoms, paranoia, and depression showed  
338 no significant group differences. Analyses within each patient group revealed a significant  
339 association between depression and investment over trials in FEP ( $b = .03$ , 95% CI  $[.007, .056]$ ,  $p =$   
340  $.01$ ), showing that FEP with a more severe depression score adjusted their investment less to the  
341 negative feedback than FEP with milder depression scores.

342 3.4.2 Associations of ROI beta weights and symptoms

343 Beta weights of the ROIs showing group differences (Table 2) were correlated with the positive and  
344 negative subscales, and PANSS paranoia and depression score. No group differences were found in  
345 the association between symptoms and ROI activation. After removing the group-by-symptoms  
346 interaction from the model there was a positive association at trend level for TPJ activation and  
347 paranoia in the unfair investment phase ( $\beta = .32, p = .07$ ), indicating that in both patient groups the  
348 TPJ was increasingly activated in patients with higher paranoia.

349 Beta weights of the TPJ were unevenly distributed. Therefore, Spearman rank correlation  
350 was used for the exploratory analyses per group. This analysis revealed significant associations  
351 between symptoms and beta weights of the TPJ for CHR, but not for FEP. Specifically, CHR showed a  
352 significant positive association between the paranoia item, positive, and negative symptoms, and  
353 TPJ activation during unfair investments ( $\rho = .57, p = .04$ ;  $\rho = .70, p = .008$ ;  $\rho = .64, p = .02$ ,  
354 respectively). No associations between symptoms and mPFC activation were found.

355

#### 356 4 Discussion

357 This study investigated the behavioral and neural mechanisms associated with trust and the  
358 association with symptoms in a high-risk (CHR) and first episode psychosis (FEP) sample using an  
359 interactive trust game. Participants played two trust games, with a cooperative and an unfair pre-  
360 programmed partner. Behaviorally, FEP and CHR only differed from controls, and not from each  
361 other, showing reduced basic trust, that is initial trust before partner feedback is revealed. No  
362 impairments in the development of trust in response to feedback over trials were found in either  
363 patient group, compared to controls. Only in FEP associations between trust development and  
364 symptoms were found. On the neural level, CHR recruited the TPJ more than the other groups

365 during investment in the unfair condition, suggesting differential processing as compared to  
366 healthy controls and FEP.

#### 367 4.1 Behavioral mechanisms of trust

368 Importantly, and in line with previous research both FEP and CHR showed reduced baseline trust  
369 toward unknown others (Fett et al, 2015; Fett et al, 2016; Fett et al, 2012; Gromann et al, 2013).

370 Reduced baseline trust has been found in individuals at genetic risk for psychosis, (Fett et al, 2012),  
371 but contradicted by another (Gromann et al, 2014). Contrary to previous research (Fett et al, 2016;  
372 Fett et al, 2012) baseline trust in FEP and CHR was not associated with symptom severity. The  
373 association with positive symptoms found by Fett et al (2012) was at trend level (.09), providing  
374 only tentative support. The fact that reduced baseline trust has also been found in individuals at  
375 genetic risk for psychosis and now in CHR in combination with the lack of an association with  
376 symptoms, tentatively suggests that reduced baseline trust is linked to the risk for psychosis trait,  
377 rather than a consequence of the illness (a state marker) that would be associated with (temporal  
378 fluctuations of) symptoms.

379 Feedback learning in cooperative and unfair interactions is still intact in CHR and FEP, as  
380 opposed to chronic psychosis (Campellone et al, 2016; Fett et al, 2015). The development of trust in  
381 response to positive feedback by the game partner showed, as predicted, that FEP and CHR  
382 increased their levels of trust significantly. The same pattern was found in response to negative  
383 feedback: over game rounds FEP and CHR decreased their trust to the same level as controls. FEP  
384 showed steeper increase in positive interactions than controls, possibly because they were more  
385 sensitive to the effects of the positive feedback given they initially had lower expectations, as  
386 reflected in lower baseline trust. Furthermore, a ceiling effect might result in a less steep increase  
387 for controls, with 23% of the control participants investing the maximum of 10 in 75% or more of

388 the trials. The slightly reduced response to negative feedback in FEP, resembling results of Fett et al  
389 (2016), might be explained by the differences in first investment (FEP starting significantly lower).

#### 390 4.1.1 Symptoms

391 Symptom severity on average was similar in the two patient groups. FEP showed substantial  
392 variability in symptom severity, which reflects the fact that we included both hospitalized patients  
393 and ambulant patients who were in a rehabilitation trajectory. CHR experienced more paranoia  
394 than FEP, and FEP tended to have more negative symptoms than CHR. These differences might be  
395 explained by medication effects: 64% of the FEP were on atypical antipsychotic medication,  
396 probably dampening positive symptoms, whereas 47% of the CHR was on other psychotropic  
397 medication (see Table 1). First episode patients with highest negative symptoms, as opposed to  
398 milder symptoms, showed almost no adjustment of trust in response to positive feedback. Intact  
399 feedback learning mechanisms in FEP were associated with milder negative symptom severity. The  
400 association between negative symptoms and problems in social functioning and responding to  
401 feedback has been well established in psychosis (Addington & Addington, 2005; Campellone et al,  
402 2016; Milev et al, 2005; Strauss et al, 2013; Voges & Addington, 2005; Waltz et al, 2011), possibly  
403 reflecting a lack of (social) motivation or depression. In the unfair condition positive symptoms  
404 were associated with less decrease in trust in both FEP and CHR. This suggests that positive  
405 symptoms interfere with learning from negative social feedback, contradicting earlier findings that  
406 found no associations of positive symptoms with learning to trust (Fett et al, 2016).

#### 407 4.2 Neural mechanisms of trust

408 On the neural level CHR activated the TPJ significantly more than the other two groups. The TPJ  
409 forms part of the mentalizing system (Fletcher et al, 1995; Frith & Frith, 2006; Van Overwalle,  
410 2009), and was previously found to be activated in the trust game (King-Casas et al, 2005; Krueger

411 et al, 2007; Saxe & Kanwisher, 2003; van den Bos et al, 2011). In chronic patients reduced TPJ  
412 activation was associated with more positive symptoms (Gromann et al, 2013). In our sample, TPJ  
413 activation did not differ between FEP and controls, suggesting a decline in TPJ response with longer  
414 illness duration. CHR however, showed increased activation in this area compared to FEP and  
415 controls during unfair investment. CHR also showed more TPJ and mPFC activation than controls  
416 during investments towards the unfair counterpart, as compared to cooperative counterpart. Since  
417 both areas form part of the mentalizing system, this could suggest that unfair interactions elicit  
418 increased mentalizing in CHR. Gromann et al (2013), in contrast, found the mPFC to be activated  
419 more in cooperative interactions in both patients and controls. Increased neural activation in  
420 patients at-risk for psychosis during mentalizing and emotion processing areas despite similar  
421 behavioral performance was previously found (Brüne et al, 2011; Derntl et al, 2015; Marjoram et al,  
422 2006). The elevated TPJ activation in CHR was associated with higher symptoms in all domains,  
423 thus associating higher illness severity with greater neural activity. This association was not found  
424 in FEP, possibly suggesting different underlying mechanisms between groups. In combination with  
425 the behavioral data, showing that CHR adapted adequately to negative social feedback, the  
426 increased TPJ activity could indicate a cognitive mechanism by which increased mentalizing helps  
427 to respond adequately to negative feedback, indicating more effort, or an inefficient use of the TPJ.  
428 The data do not point to compensating mechanisms, since they would suggest deficiencies or  
429 reduced processing in other parts of the brain. These were not found in the ROI analyses, nor in the  
430 additional whole brain analyses (see supplementary Table S1). The results show no evidence for  
431 reduced sensitivity to social reward in FEP and CHR, and suggest that altered mentalizing might be  
432 associated with reduced baseline trust. However, due to the small sample size, this result must be  
433 interpreted with caution.

#### 434 4.3 Clinical high-risk

435           Following the procedure of previous CHR investigations (Fusar-Poli et al, 2010a; McGorry &  
436 van Os, 2013; Phillips et al, 2009; Rietdijk et al, 2012; Shim et al, 2008; Thompson et al, 2012;  
437 Valmaggia et al, 2013; van der Gaag et al, 2012; Wood et al, 2011), we included participants  
438 assessed with the CAARMS, and with a score below 55 on the SOFAS. Our sample was comparable  
439 to other samples in terms of comorbidities (Corcoran et al, 2011; Fusar-Poli et al, 2014; Ising et al,  
440 2016; Modinos et al, 2014; Morrison et al, 2012; Woods et al, 2009). One year after testing, e.g.  
441 around two years after initial assessment, CHR participants were re-assessed with the CAARMS, to  
442 investigate their current status. One of the CHR had made the transition to psychosis. Of two CHR  
443 transition data were missing. In this aspect, our high-risk group differed from other high-risk  
444 groups. Variant transition rates have been reported in comparable samples with regard to  
445 assessment and age range (Broome et al, 2012; Broome et al, 2005; Demjaha et al, 2010; Fusar-Poli  
446 et al, 2010b; Nelson et al, 2011). Transition rates in similar referred samples are under 10%  
447 (Rietdijk et al, 2012; Yung et al, 2011). Patients already received treatment for their primary  
448 problems, including cognitive behavioral therapy (CBT) for their CHR status (psychotic symptoms).  
449 This has shown to be an adequate strategy to reduce symptoms, increasing their social functioning  
450 skills, to reduce the transition rates (by 46%), and to increase chances for remission (Cannon et al,  
451 2008; Ising et al, 2016; Jang et al, 2011; Niendam et al, 2007; van Os & Murray, 2013).

452           In a recent discussion on CHR it has been argued that the presence of psychotic symptoms is  
453 possibly more important than transition in the assessment of CHR (van Os & Reininghaus, 2016).  
454 Many patients in care for anxiety and depression report psychotic symptoms (van Os & Linscott,  
455 2012; van Os & Reininghaus, 2016; Velthorst et al, 2009; Wigman et al, 2012; Woods et al, 2009),  
456 but do not transition to psychosis. The current sample fits previous descriptions, making it a  
457 representative sample of patients with psychotic symptoms and generally poor mental health. The  
458 addition of psychotic symptoms renders these patients at-risk for developing psychopathology,  
459 without the direct consequence of developing a psychotic disorder (Fusar-Poli et al, 2013; Yung et

460 al, 2012). In many cases subclinical psychotic experiences are transitory (van Os & Reininghaus,  
461 2016). However, the presence of psychotic symptoms is associated with a poorer prognosis,  
462 showing that these patients are certainly in need of special care (McGorry & van Os, 2013;  
463 Ruhrmann et al, 2010; Valmaggia et al, 2013; van Os & Linscott, 2012; van Os & Reininghaus, 2016).

#### 464 4.4 Limitations and future directions

465 Several limitations should be considered. First, current results should be interpreted with caution,  
466 due to the small sample size, especially of the CHR sample. Our CHR results should therefore be  
467 considered as a first step in research on real-time social interaction in high-risk individuals, that  
468 warrants replication and extension in future research (see also (Broome et al, 2010; Fusar-Poli et  
469 al, 2011; Juckel et al, 2012). With a larger CHR sample, the number of converters will increase,  
470 allowing for an investigation of converters vs. non-converters. Additionally, a larger sample would  
471 provide the possibility to subdivide the CHR sample on the basis of different symptomatology  
472 (Fusar-Poli et al, 2014; Valmaggia et al, 2013), yielding more insight in the factors causing social  
473 problems, and explaining transition trajectories. Transition in these studies was not explained by  
474 comorbid anxiety and depression, but by the severity of CAARMS score, social dysfunction and  
475 increased negative symptoms. Direct comparison between at-risk subjects with and without  
476 progression into a diagnosis of schizophrenia could also elucidate whether greater activation of the  
477 mentalizing network in CHR is serving as a compensatory mechanism, or could also be linked to  
478 transition to psychosis. Larger samples could have revealed group differences that were not  
479 apparent in this sample. Further, FEP symptom severity was rather mild, possibly due to  
480 responsiveness to antipsychotic treatment. Similar symptom severity has been found in stable and  
481 medicated patients (Möller et al, 2005). Including a broader range of symptom severity would  
482 increase the validity of the sample. Methodologically, a limitation is that participants were not paid  
483 on performance. There is some evidence that real payment has a different effect on decisions and

484 related brain activity than hypothetical payment (Hertwig & Ortmann, 2001; Johnson & Mislin,  
485 2011; Vlaev, 2012), but other studies have found no differences (Locey et al, 2011; Madden et al,  
486 2003). Plausibly, hypothetical payment may influence the strength, but not the direction of the  
487 effect (Derks, 2015). Furthermore, eight participants did not believe they were playing against a  
488 human counterpart. This might have influenced their behavior. However, analyses without these  
489 participants did not change the results. Adequate increase and decrease of investments in response  
490 to cooperative and unfair feedback showed that overall the experimental manipulation of the  
491 counterpart was effective. Additionally, including an online mentalizing task in the scanner could  
492 establish direct links with the current outcomes, pointing in the direction of differential mentalizing  
493 processes.

494

## 495 5 Conclusion

496 Summarizing, baseline trust is impaired in FEP and CHR, indicating that reduced baseline  
497 trust is associated with the risk for psychotic illness trait, or with poor mental health in general,  
498 rather than a consequence of psychotic disorder. CHR performed in between controls and FEP  
499 (Pukrop et al, 2006; Thompson et al, 2012), resembling most the patient group. In contrast to  
500 chronic patients (Fett et al, 2015), reward learning was not impaired in FEP and CHR (see also  
501 (Juckel et al, 2012). This suggests that with positive social feedback, the lack of initial trust in FEP  
502 and CHR can be restored, at least in the context of the trust game. The neural results pointed to  
503 globally intact neural mechanisms associated with trust, except for changes in brain areas  
504 associated with mentalizing in CHR. However, these findings should be considered as preliminary,  
505 and more research in the field is needed to replicate and extend our findings.

506



507 **Acknowledgements:** We would like to thank Prof. Lieuwe de Haan and his team at the Amsterdam  
508 Medical Center (AMC), Prof. Mark van der Gaag and Dr. Helga Ising at PsyQ, the Hague, Onno  
509 Ackema and the department of early psychosis of the AMC, and the VIP team for helping with  
510 recruitment of both patient groups, and Tinka Beemsterboer and colleagues at the Spinoza Centre  
511 for Neuroimaging, Roeterseiland Amsterdam for their help during scanning, and all participants for  
512 completing the testing session and providing us with valuable material.

513

514 **Funding:** This work was supported by funding of the Hersenstichting Nederland [KS2011(1)-75], a  
515 VIDI and VICI grant from the Netherlands Organization for Scientific Research (NWO) [452-07-  
516 007,453-11-005]; and a ERC Consolidator grant (648082 SCANS) to Prof. Lydia Krabbendam. Anne-  
517 Kathrin Fett was supported by a Netherlands Organization for Scientific Research (NWO) VENI  
518 grant (451-13-035) and a NARSAD Young Investigator Grant from the Brain & Behaviour research  
519 Foundation (24138).

520

521

- 522 Addington, J. & Addington, D. (2005) Patterns of premorbid functioning in first episode psychosis:  
 523 relationship to 2-year outcome. *Acta Psychiatrica Scandinavica*, 112(1), 40-46.
- 524 Association, A. P. & Association, A. P. (2000) DSM-IV-TR: Diagnostic and statistical manual of mental  
 525 disorders, text revision. *Washington, DC: American Psychiatric Association*, 75, 78-85.
- 526 Ballon, J. S., Kaur, T., Marks, I. I. & Cadenhead, K. S. (2007) Social functioning in young people at risk for  
 527 schizophrenia. *Psychiatry research*, 151(1), 29-35.
- 528 Berg, J., Dickhaut, J. & McCabe, K. (1995) Trust, reciprocity, and social history. *Games and economic  
 529 behavior*, 10(1), 122-142.
- 530 Bora, E. & Pantelis, C. (2013) Theory of mind impairments in first-episode psychosis, individuals at ultra-  
 531 high risk for psychosis and in first-degree relatives of schizophrenia: systematic review and meta-  
 532 analysis. *Schizophrenia research*, 144(1), 31-36.
- 533 Broome, M. R., Day, F., Valli, I., Valmaggia, L., Johns, L., Howes, O., Garety, P. & McGuire, P. (2012)  
 534 Delusional ideation, manic symptomatology and working memory in a cohort at clinical high-risk for  
 535 psychosis: a longitudinal study. *European Psychiatry*, 27(4), 258-263.
- 536 Broome, M. R., Fusar-Poli, P., Matthiasson, P., Woolley, J. B., Valmaggia, L., Johns, L. C., Tabraham, P.,  
 537 Bramon, E., Williams, S. C. & Brammer, M. J. (2010) Neural correlates of visuospatial working memory in  
 538 the 'at-risk mental state'. *Psychological medicine*, 40(12), 1987-1999.
- 539 Broome, M. R., Woolley, J. B., Johns, L. C., Valmaggia, L. R., Tabraham, P., Gafoor, R., Bramon, E. &  
 540 McGuire, P. K. (2005) Outreach and support in south London (OASIS): implementation of a clinical  
 541 service for prodromal psychosis and the at risk mental state. *European Psychiatry*, 20(5), 372-378.
- 542 Brüne, M., Özgürdal, S., Ansorge, N., von Reventlow, H. G., Peters, S., Nicolas, V., Tegenthoff, M., Juckel,  
 543 G. & Lissek, S. (2011) An fMRI study of "theory of mind" in at-risk states of psychosis: comparison with  
 544 manifest schizophrenia and healthy controls. *Neuroimage*, 55(1), 329-337.
- 545 Campellone, T. R., Fisher, A. J. & Kring, A. M. (2016) Using social outcomes to inform decision-making in  
 546 schizophrenia: Relationships with symptoms and functioning. *Journal of abnormal psychology*, 125(2),  
 547 310.
- 548 Cannon, T. D., Cadenhead, K., Cornblatt, B., Woods, S. W., Addington, J., Walker, E., Seidman, L. J.,  
 549 Perkins, D., Tsuang, M. & McGlashan, T. (2008) Prediction of psychosis in youth at high clinical risk: a  
 550 multisite longitudinal study in North America. *Archives of general psychiatry*, 65(1), 28-37.
- 551 Corcoran, C., Kimhy, D., Parrilla-Escobar, M., Cressman, V., Stanford, A., Thompson, J., David, S. B.,  
 552 Crumley, A., Schobel, S. & Moore, H. (2011) The relationship of social function to depressive and  
 553 negative symptoms in individuals at clinical high risk for psychosis. *Psychological medicine*, 41(02), 251-  
 554 261.
- 555 Cornblatt, B. A., Auther, A. M., Niendam, T., Smith, C. W., Zinberg, J., Bearden, C. E. & Cannon, T. D.  
 556 (2007) Preliminary findings for two new measures of social and role functioning in the prodromal phase  
 557 of schizophrenia. *Schizophrenia bulletin*, 33(3), 688-702.
- 558 Couture, S. M., Penn, D. L. & Roberts, D. L. (2006) The functional significance of social cognition in  
 559 schizophrenia: a review. *Schizophrenia bulletin*, 32(suppl 1), S44-S63.
- 560 Csukly, G., Polgár, P., Tombor, L., Réthelyi, J. & Kéri, S. (2011) Are patients with schizophrenia rational  
 561 maximizers? Evidence from an ultimatum game study. *Psychiatry research*, 187(1), 11-17.
- 562 Declerck, C. H., Boone, C. & Emonds, G. (2013) When do people cooperate? The neuroeconomics of  
 563 prosocial decision making. *Brain and Cognition*, 81(1), 95-117.
- 564 Demjaha, A., Valmaggia, L., Stahl, D., Byrne, M. & McGuire, P. (2010) Disorganization/cognitive and  
 565 negative symptom dimensions in the at-risk mental state predict subsequent transition to psychosis.  
 566 *Schizophrenia bulletin*, 38(2), 351-359.
- 567 Derks, J. (2015) Adolescent Social Cognition (doctoral dissertation).

- 568 Derntl, B., Michel, T. M., Prempeh, P., Backes, V., Finkelmeyer, A., Schneider, F. & Habel, U. (2015)  
569 Empathy in individuals clinically at risk for psychosis: brain and behaviour. *The British Journal of*  
570 *Psychiatry*, bjp. bp. 114.159004.
- 571 Fehr, E. & Camerer, C. F. (2007) Social neuroeconomics: the neural circuitry of social preferences. *Trends*  
572 *in cognitive sciences*, 11(10), 419-427.
- 573 Fett, A.-K., Gromann, P., Shergill, S. & Krabbendam, L. (2015) Trust vs. Paranoia: The Dynamics of Social  
574 Interaction in Early and Chronic Psychosis, *SCHIZOPHRENIA BULLETIN*. OXFORD UNIV PRESS GREAT  
575 CLARENDON ST, OXFORD OX2 6DP, ENGLAND.
- 576 Fett, A.-K., Shergill, S., Korver-Nieberg, N., Yakub, F., Gromann, P. M. & Krabbendam, L. (2016) Learning  
577 to trust: trust and attachment in early psychosis. *Psychological medicine*, 46(7), 1437.
- 578 Fett, A.-K., Shergill, S. S., Joyce, D. W., Riedl, A., Strobel, M., Gromann, P. M. & Krabbendam, L. (2012) To  
579 trust or not to trust: the dynamics of social interaction in psychosis. *Brain*, 135(3), 976-984.
- 580 Fletcher, P. C., Happe, F., Frith, U., Baker, S. C., Dolan, R. J., Frackowiak, R. S. & Frith, C. D. (1995) Other  
581 minds in the brain: a functional imaging study of "theory of mind" in story comprehension. *Cognition*,  
582 57(2), 109-128.
- 583 Frith, C. D. & Frith, U. (2006) The neural basis of mentalizing. *Neuron*, 50(4), 531-534.
- 584 Fusar-Poli, P., Borgwardt, S., Bechdolf, A., Addington, J., Riecher-Rössler, A., Schultze-Lutter, F.,  
585 Keshavan, M., Wood, S., Ruhrmann, S. & Seidman, L. J. (2013) The Psychosis High-Risk State: A  
586 Comprehensive State-of-the-Art Review. *JAMA psychiatry*, 70(1), 107.
- 587 Fusar-Poli, P., Broome, M. R., Matthiasson, P., Woolley, J. B., Johns, L., Tabraham, P., Bramon, E.,  
588 Valmaggia, L., Williams, S. & McGuire, P. (2010a) Spatial working memory in individuals at high risk for  
589 psychosis: longitudinal fMRI study. *Schizophrenia Research*, 123(1), 45-52.
- 590 Fusar-Poli, P., Broome, M. R., Woolley, J. B., Johns, L. C., Tabraham, P., Bramon, E., Valmaggia, L.,  
591 Williams, S. & McGuire, P. (2011) Altered brain function directly related to structural abnormalities in  
592 people at ultra high risk of psychosis: longitudinal VBM-fMRI study. *Journal of psychiatric research*,  
593 45(2), 190-198.
- 594 Fusar-Poli, P., Byrne, M., Valmaggia, L., Day, F., Tabraham, P., Johns, L., McGuire, P. & Team, O. (2010b)  
595 Social dysfunction predicts two years clinical outcome in people at ultra high risk for psychosis. *Journal*  
596 *of psychiatric research*, 44(5), 294-301.
- 597 Fusar-Poli, P., Nelson, B., Valmaggia, L., Yung, A. R. & McGuire, P. K. (2014) Comorbid depressive and  
598 anxiety disorders in 509 individuals with an at-risk mental state: impact on psychopathology and  
599 transition to psychosis. *Schizophrenia bulletin*, 40(1), 120-131.
- 600 Gallagher, H. L. & Frith, C. D. (2003) Functional imaging of 'theory of mind'. *Trends in cognitive sciences*,  
601 7(2), 77-83.
- 602 Goldman, H. H., Skodol, A. E. & Lave, T. R. (1992) Revising axis V for DSM-IV: a review of measures of  
603 social functioning. *Am J Psychiatry*, 149(9), 1148-1156.
- 604 Green, M. F., Horan, W. P. & Lee, J. (2015) Social cognition in schizophrenia. *Nature Reviews*  
605 *Neuroscience*, 16(10), 620.
- 606 Green, M. F. & Leitman, D. I. (2008) Social cognition in schizophrenia. *Schizophrenia Bulletin*, 34(4), 670-  
607 672.
- 608 Gromann, P. M., Heslenfeld, D. J., Fett, A.-K., Joyce, D. W., Shergill, S. S. & Krabbendam, L. (2013) Trust  
609 versus paranoia: abnormal response to social reward in psychotic illness. *Brain*, awt076.
- 610 Gromann, P. M., Shergill, S., de Haan, L., Meewis, D., Fett, A.-K., Korver-Nieberg, N. & Krabbendam, L.  
611 (2014) Reduced brain reward response during cooperation in first-degree relatives of patients with  
612 psychosis: an fMRI study. *Psychological medicine*, 44(16), 3445-3454.
- 613 Hertwig, R. & Ortmann, A. (2001) Experimental practices in economics: A methodological challenge for  
614 psychologists? *Behavioral and Brain Sciences*, 24(03), 383-403.

- 615 Horat, S. K., Favre, G., Prévot, A., Ventura, J., Herrmann, F. R., Gothuey, I., Merlo, M. C. & Missonnier, P.  
 616 (2017) Impaired social cognition in schizophrenia during the Ultimatum Game: An EEG study.  
 617 *Schizophrenia Research*.
- 618 Ising, H. K., Kraan, T. C., Rietdijk, J., Dragt, S., Klaassen, R. M., Boonstra, N., Nieman, D. H., Willebrands-  
 619 Mendrik, M., van den Berg, D. P. & Linszen, D. H. (2016) Four-year follow-up of cognitive behavioral  
 620 therapy in persons at ultra-high risk for developing psychosis: the Dutch early detection intervention  
 621 evaluation (EDIE-NL) trial. *Schizophrenia bulletin*, 42(5), 1243-1252.
- 622 Jang, J. H., Shin, N. Y., Shim, G., Park, H. Y., Kim, E., Jang, G.-E., Kwon, S. J., Hur, J.-W., An, S. K. & Kwon, J.  
 623 S. (2011) Longitudinal patterns of social functioning and conversion to psychosis in subjects at ultra-high  
 624 risk. *Australian and New Zealand Journal of Psychiatry*, 45(9), 763-770.
- 625 Johnson, N. D. & Mislin, A. A. (2011) Trust games: A meta-analysis. *Journal of Economic Psychology*,  
 626 32(5), 865-889.
- 627 Juckel, G., Friedel, E., Koslowski, M., Witthaus, H., Özgürdal, S., Gudlowski, Y., Knutson, B., Wrase, J.,  
 628 Brüne, M. & Heinz, A. (2012) Ventral striatal activation during reward processing in subjects with ultra-  
 629 high risk for schizophrenia. *Neuropsychobiology*, 66(1), 50-56.
- 630 Kay, S., Fiszbein, A. & Opler, L. (1987) The positive and negative syndrome scale (PANSS) for  
 631 schizophrenia. *Schizophrenia bulletin*, 13(2), 261.
- 632 Kelleher, I., Keeley, H., Corcoran, P., Lynch, F., Fitzpatrick, C., Devlin, N., Molloy, C., Roddy, S., Clarke, M.  
 633 C. & Harley, M. (2012) Clinicopathological significance of psychotic experiences in non-psychotic young  
 634 people: evidence from four population-based studies. *The British Journal of Psychiatry*, 201(1), 26-32.
- 635 King-Casas, B., Tomlin, D., Anen, C., Camerer, C. F., Quartz, S. R. & Montague, P. R. (2005) Getting to  
 636 know you: reputation and trust in a two-person economic exchange. *Science*, 308(5718), 78-83.
- 637 Krueger, F., McCabe, K., Moll, J., Kriegeskorte, N., Zahn, R., Strenziok, M., Heinecke, A. & Grafman, J.  
 638 (2007) Neural correlates of trust. *Proceedings of the National Academy of Sciences*, 104(50), 20084-  
 639 20089.
- 640 Lavoie, M.-A., Lacroix, J. B., Godmaire-Duhaime, F., Jackson, P. L. & Achim, A. M. (2013) Social cognition  
 641 in first-degree relatives of people with schizophrenia: a meta-analysis. *Psychiatry Research*, 209(2), 129-  
 642 135.
- 643 Lemmers-Jansen, I. L. J., Krabbendam, L., Veltman, D. J. & Fett, A.-K. J. (2017) Boys vs. girls: Gender  
 644 differences in the neural development of trust and reciprocity depend on social context. *Developmental*  
 645 *Cognitive Neuroscience*.
- 646 Li, W., Tol, M. J., Li, M., Miao, W., Jiao, Y., Heinze, H. J., Bogerts, B., He, H. & Walter, M. (2014) Regional  
 647 specificity of sex effects on subcortical volumes across the lifespan in healthy aging. *Human brain*  
 648 *mapping*, 35(1), 238-247.
- 649 Locey, M. L., Jones, B. A. & Rachlin, H. (2011) Real and hypothetical rewards. *Judgment and decision*  
 650 *making*, 6(6), 552.
- 651 Madden, G. J., Begotka, A. M., Raiff, B. R. & Kastern, L. L. (2003) Delay discounting of real and  
 652 hypothetical rewards. *Experimental and clinical psychopharmacology*, 11(2), 139.
- 653 Marjoram, D., Job, D. E., Whalley, H. C., Gountouna, V.-E., McIntosh, A. M., Simonotto, E., Cunningham-  
 654 Owens, D., Johnstone, E. C. & Lawrie, S. (2006) A visual joke fMRI investigation into Theory of Mind and  
 655 enhanced risk of schizophrenia. *Neuroimage*, 31(4), 1850-1858.
- 656 McCleery, A., Horan, W. P. & Green, M. F. (2014) Social Cognition During the Early Phase of  
 657 Schizophrenia. *Social Cognition and Metacognition in Schizophrenia*, 49.
- 658 McGorry, P. & van Os, J. (2013) Redeeming diagnosis in psychiatry: timing versus specificity. *The Lancet*,  
 659 381(9863), 343-345.
- 660 Milev, P., Ho, B.-C., Arndt, S. & Andreasen, N. C. (2005) Predictive values of neurocognition and negative  
 661 symptoms on functional outcome in schizophrenia: a longitudinal first-episode study with 7-year follow-  
 662 up. *American Journal of Psychiatry*, 162(3), 495-506.

- 663 Modinos, G., Allen, P., Frascarelli, M., Tognin, S., Valmaggia, L., Xenaki, L., Keedwell, P., Broome, M.,  
 664 Valli, I. & Woolley, J. (2014) Are we really mapping psychosis risk? Neuroanatomical signature of  
 665 affective disorders in subjects at ultra high risk. *Psychological medicine*, 44(16), 3491-3501.
- 666 Möller, H.-J., Llorca, P.-M., Sacchetti, E., Martin, S. D., Medori, R., Parellada, E. & Group, S. S. (2005)  
 667 Efficacy and safety of direct transition to risperidone long-acting injectable in patients treated with  
 668 various antipsychotic therapies. *International clinical psychopharmacology*, 20(3), 121-130.
- 669 Morosini, P., Magliano, L., Brambilla, L., Ugolini, S. & Pioli, R. (2000) Development, reliability and  
 670 acceptability of a new version of the DSM-IV Social and Occupational Functioning Assessment Scale  
 671 (SOFAS) to assess routine social functioning. *Acta Psychiatrica Scandinavica*, 101(4), 323-329.
- 672 Morrison, A. P., French, P., Stewart, S. L., Birchwood, M., Fowler, D., Gumley, A. I., Jones, P. B., Bentall, R.  
 673 P., Lewis, S. W. & Murray, G. K. (2012) Early detection and intervention evaluation for people at risk of  
 674 psychosis: multisite randomised controlled trial. *Bmj*, 344, e2233.
- 675 Nelson, B., Yuen, K. & Yung, A. (2011) Ultra high risk (UHR) for psychosis criteria: are there different  
 676 levels of risk for transition to psychosis? *Schizophrenia research*, 125(1), 62-68.
- 677 Niendam, T. A., Bearden, C. E., Zinberg, J., Johnson, J. K., O'Brien, M. & Cannon, T. D. (2007) The course  
 678 of neurocognition and social functioning in individuals at ultra high risk for psychosis. *Schizophrenia  
 679 bulletin*, 33(3), 772-781.
- 680 Phillips, L. J., Nelson, B., Yuen, H. P., Francey, S. M., Simmons, M., Stanford, C., Ross, M., Kelly, D., Baker,  
 681 K. & Conus, P. (2009) Randomized controlled trial of interventions for young people at ultra-high risk of  
 682 psychosis: study design and baseline characteristics. *Australian & New Zealand Journal of Psychiatry*,  
 683 43(9), 818-829.
- 684 Pukrop, R., Schultze-Lutter, F., Ruhrmann, S., Brockhaus-Dumke, A., Tendolkar, I., Bechdolf, A.,  
 685 Matuschek, E. & Klosterkötter, J. (2006) Neurocognitive functioning in subjects at risk for a first episode  
 686 of psychosis compared with first-and multiple-episode schizophrenia. *Journal of Clinical and  
 687 Experimental Neuropsychology*, 28(8), 1388-1407.
- 688 Rietdijk, J., Klaassen, R., Ising, H., Dragt, S., Nieman, D., Van de Kamp, J., Cuijpers, P., Linszen, D. & Van  
 689 der Gaag, M. (2012) Detection of people at risk of developing a first psychosis: comparison of two  
 690 recruitment strategies. *Acta psychiatrica scandinavica*, 126(1), 21-30.
- 691 Rilling, J. K. & Sanfey, A. G. (2011) The neuroscience of social decision-making. *Annual review of  
 692 psychology*, 62, 23-48.
- 693 Ruhrmann, S., Schultze-Lutter, F. & Klosterkötter, J. (2010) Probably at-risk, but certainly ill—advocating  
 694 the introduction of a psychosis spectrum disorder in DSM-V. *Schizophrenia research*, 120(1), 23-37.
- 695 Saxe, R. & Kanwisher, N. (2003) People thinking about thinking people: the role of the temporo-parietal  
 696 junction in “theory of mind”. *Neuroimage*, 19(4), 1835-1842.
- 697 Shim, G., Kang, D.-H., Sun Chung, Y., Young Yoo, S., Young Shin, N. & Soo Kwon, J. (2008) Social  
 698 functioning deficits in young people at risk for schizophrenia. *Australian and New Zealand Journal of  
 699 Psychiatry*, 42(8), 678-685.
- 700 Smieskova, R., Marmy, J., Schmidt, A., Bendfeldt, K., Riecher-Rössler, A., Walter, M., E Lang, U. &  
 701 Borgwardt, S. (2013) Do subjects at clinical high risk for psychosis differ from those with a genetic high  
 702 risk?-A systematic review of structural and functional brain abnormalities. *Current medicinal chemistry*,  
 703 20(3), 467-481.
- 704 SPM (2009) *Statistical Parametric Mapping*. London, UK: Wellcome Trust Centre for Neuroimaging.  
 705 Available online: [www.fil.ion.ucl.ac.uk/spm](http://www.fil.ion.ucl.ac.uk/spm) [Accessed].
- 706 StataCorp (2013) *Stata Statistical Software* Release 13. TX: StataCorp LP.
- 707 Strauss, G. P., Waltz, J. A. & Gold, J. M. (2013) A review of reward processing and motivational  
 708 impairment in schizophrenia. *Schizophrenia bulletin*, 40(Suppl\_2), S107-S116.

- 709 Thompson, A., Papas, A., Bartholomeusz, C., Allott, K., Amminger, G. P., Nelson, B., Wood, S. & Yung, A.  
 710 (2012) Social cognition in clinical “at risk” for psychosis and first episode psychosis populations.  
 711 *Schizophrenia research*, 141(2), 204-209.
- 712 Valmaggia, L., Stahl, D., Yung, A., Nelson, B., Fusar-Poli, P., McGorry, P. & McGuire, P. (2013) Negative  
 713 psychotic symptoms and impaired role functioning predict transition outcomes in the at-risk mental  
 714 state: a latent class cluster analysis study. *Psychological medicine*, 43(11), 2311-2325.
- 715 van den Bos, W., van Dijk, E., Westenberg, M., Rombouts, S. A. & Crone, E. A. (2011) Changing brains,  
 716 changing perspectives the neurocognitive development of reciprocity. *Psychological Science*, 22(1), 60-  
 717 70.
- 718 van der Gaag, M., Nieman, D. H., Rietdijk, J., Dragt, S., Ising, H. K., Klaassen, R. M., Koeter, M., Cuijpers,  
 719 P., Wunderink, L. & Linszen, D. H. (2012) Cognitive behavioral therapy for subjects at ultrahigh risk for  
 720 developing psychosis: a randomized controlled clinical trial. *Schizophrenia Bulletin*, 38(6), 1180-1188.
- 721 van Os, J. & Linscott, R. J. (2012) Introduction: the extended psychosis phenotype—relationship with  
 722 schizophrenia and with ultrahigh risk status for psychosis. *Schizophrenia Bulletin*, 38(2), 227-230.
- 723 van Os, J. & Murray, R. M. (2013) Can we identify and treat “schizophrenia light” to prevent true  
 724 psychotic illness. *BMJ*, 346(jan18 1), f304-f304.
- 725 van Os, J. & Reininghaus, U. (2016) Psychosis as a transdiagnostic and extended phenotype in the  
 726 general population. *World Psychiatry*, 15(2), 118-124.
- 727 Van Overwalle, F. (2009) Social cognition and the brain: a meta-analysis. *Human brain mapping*, 30(3),  
 728 829-858.
- 729 Velthorst, E., Fett, A.-K. J., Reichenberg, A., Perlman, G., van Os, J., Bromet, E. J. & Kotov, R. (2016a) The  
 730 20-Year Longitudinal Trajectories of Social Functioning in Individuals With Psychotic Disorders. *American*  
 731 *Journal of Psychiatry*, appi. ajp. 2016.15111419.
- 732 Velthorst, E., Nieman, D. H., Becker, H. E., van de Fliert, R., Dingemans, P. M., Klaassen, R., de Haan, L.,  
 733 van Amelsvoort, T. & Linszen, D. H. (2009) Baseline differences in clinical symptomatology between ultra  
 734 high risk subjects with and without a transition to psychosis. *Schizophrenia Research*, 109(1), 60-65.
- 735 Velthorst, E., Reichenberg, A., Kapara, O., Goldberg, S., Fromer, M., Fruchter, E., Ginat, K., de Haan, L.,  
 736 Davidson, M. & Weiser, M. (2016b) Developmental Trajectories of Impaired Community Functioning in  
 737 Schizophrenia. *JAMA psychiatry*, 73(1), 48-55.
- 738 Vlaev, I. (2012) How different are real and hypothetical decisions? Overestimation, contrast and  
 739 assimilation in social interaction. *Journal of Economic Psychology*, 33(5), 963-972.
- 740 Voges, M. & Addington, J. (2005) The association between social anxiety and social functioning in first  
 741 episode psychosis. *Schizophrenia research*, 76(2), 287-292.
- 742 Waltz, J. A., Frank, M. J., Wiecki, T. V. & Gold, J. M. (2011) Altered probabilistic learning and response  
 743 biases in schizophrenia: behavioral evidence and neurocomputational modeling. *Neuropsychology*,  
 744 25(1), 86.
- 745 Wechsler, D. (1997) WAIS-III Dutch Translation. Lisse: Swets & Zeitlinger.
- 746 Wigman, J. T., van Nierop, M., Vollebergh, W. A., Lieb, R., Beesdo-Baum, K., Wittchen, H.-U. & van Os, J.  
 747 (2012) Evidence that psychotic symptoms are prevalent in disorders of anxiety and depression,  
 748 impacting on illness onset, risk, and severity—implications for diagnosis and ultra-high risk research.  
 749 *Schizophrenia Bulletin*, sbr196.
- 750 Wood, S. J., Yung, A. R., McGorry, P. D. & Pantelis, C. (2011) Neuroimaging and treatment evidence for  
 751 clinical staging in psychotic disorders: from the at-risk mental state to chronic schizophrenia. *Biological*  
 752 *psychiatry*, 70(7), 619-625.
- 753 Woods, S. W., Addington, J., Cadenhead, K. S., Cannon, T. D., Cornblatt, B. A., Heinssen, R., Perkins, D.  
 754 O., Seidman, L. J., Tsuang, M. T. & Walker, E. F. (2009) Validity of the prodromal risk syndrome for first  
 755 psychosis: findings from the North American Prodrome Longitudinal Study. *Schizophrenia bulletin*, 35(5),  
 756 894-908.

- 757 Woudstra, S., van Tol, M.-J., Bochdanovits, Z., van der Wee, N. J., Zitman, F. G., van Buchem, M. A.,  
758 Opmeer, E. M., Aleman, A., Penninx, B. W. & Veltman, D. J. (2013) Modulatory effects of the piccolo  
759 genotype on emotional memory in health and depression. *Plos one*, 8(4), e61494.
- 760 Yung, A. R., Phillips, L. J., Nelson, B., Francey, S. M., Yuen, H. P., Simmons, M. B., Ross, M. L., Kelly, D.,  
761 Baker, K. & Paul Amminger, G. (2011) Randomized controlled trial of interventions for young people at  
762 ultra high risk for psychosis: 6-month analysis. *Journal of Clinical Psychiatry*, 72(4), 430.
- 763 Yung, A. R., Phillips, L. J., Yuen, H. P., Francey, S. M., McFarlane, C. A., Hallgren, M. & McGorry, P. D.  
764 (2003) Psychosis prediction: 12-month follow up of a high-risk ("prodromal") group. *Schizophrenia*  
765 *research*, 60(1), 21-32.
- 766 Yung, A. R., Woods, S. W., Ruhrmann, S., Addington, J., Schultze-Lutter, F., Cornblatt, B. A., Amminger, G.  
767 P., Bechdolf, A., Birchwood, M. & Borgwardt, S. (2012) Whither the attenuated psychosis syndrome?  
768 *Schizophrenia bulletin*, 38(6), 1130-1134.
- 769 Yung, A. R., Yung, A. R., Pan Yuen, H., Mcgorry, P. D., Phillips, L. J., Kelly, D., Dell'olio, M., Francey, S. M.,  
770 Cosgrave, E. M. & Killackey, E. (2005) Mapping the onset of psychosis: the comprehensive assessment of  
771 at-risk mental states. *Australian and New Zealand Journal of Psychiatry*, 39(11-12), 964-971.

772

773

774 Table 1

775 Participant Characteristics and Baseline Trust in the Trust Game

776

	Controls <i>N</i> = 43	CHR <i>N</i> = 17	FEP <i>N</i> = 22	
Gender, <i>N</i> male (%)	22 (51.16 %)	7 (41.18 %)	14 (63.64 %)	777
Age, mean ( <i>SD</i> )	21.06 (2.74)	23.78 (2.42) **	19.88 (1.54)	778
WAIS vocabulary, mean ( <i>SD</i> )	41.77 (11.39)	41.71 (12.16)	32.5 (10.53) **	779
PANSS total ( <i>SD</i> )		58.92 (11.84)	61.32 (14.51)	
mean ( <i>SD</i> )		1.95 (.37)	2.09 (.51)	780
- Positive total ( <i>SD</i> )		13.38 (2.59)	13.23 (5.72)	781
mean ( <i>SD</i> )		1.91 (.39)	1.89 (.82)	
- Negative total ( <i>SD</i> )		13.69 (3.73)	17.18 (5.85)	782
mean ( <i>SD</i> )		1.96 (.53)	2.45 (.84) * <sup>^</sup>	783
- General total ( <i>SD</i> )		31.85 (6.07)	30.91 (7.66)	784
mean ( <i>SD</i> )		2.0 (.38)	1.93 (.48)	785
- Paranoia mean ( <i>SD</i> )		3.38 (1.50)	2.23 (1.41) *	
- Depression mean ( <i>SD</i> )		4 (1.73)	2.89 (1.49)	786
Medicated <i>N</i> (%)		8 (47%) *	16 (73%)	787
- Atypical anti-psychotics (%)			13 (81.5%)	788
- Typical and atypical antipsychotics (%)			1 (6.25%)	789
- Anti-depressant (%)		3 (37.5%)	-	
- SSRI (%)		3 (37.5%)	-	790
- Benzodiazepine (%)		2 (25%)	1 (6.25%)	791
- Sertraline (%)		-	1 (6.25%)	792
Baseline trust, mean ( <i>SD</i> )	7.02 (1.81) **	5.82 (2.32)	5.52 (2.02)	793

794 \* = significant difference between FEP and CHR at  $p < .05$ 795 \*\* = significantly different from both other groups at  $p < .05$ 796 \*<sup>^</sup> = FEP > CHR at  $p < .07$ 

797 Note: CHR = clinical high-risk; FEP = first episode psychosis; SD = standard deviation; WAIS =

798 Wechsler Adult Intelligence Scale; PANSS = Positive and Negative Syndrome Scale; SSRI = selective

799 serotonin reuptake inhibitors. The paranoia item forms part of the positive subscale (P6), the

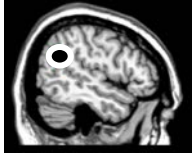



## Trust in CHR and psychosis

800 depression item forms part of the general subscale (G6). For analyses, these item were investigated  
801 separately.  
802

803 Table 2

804 Region of Interest (ROI) Analyses Outcome

Condition	ROI	$p$	Contrast	Location
Unfair Investment *	TPJ	.019	CHR > FEP	
	TPJ	<.001	CHR > Controls	
Unfair investment > cooperative investment **	TPJ	.005	CHR > Controls	
	mPFC	.007	CHR > Controls	

805 *Note:* \* = adjusted significance level for multiple comparisons of  $p = .027$ .806 \*\* = adjusted significance level of  $p = .012$ .

807 Montreal Neurological Institute (MNI) coordinates: TPJ = temporo-parietal junction, 51, -57, 26;

808 mPFC = medial prefrontal cortex, -3, 65, 25. CHR = clinical high-risk; FEP = first episode psychosis

809

810