

Twin differences in the Minnesota Trust Game relate to neural mechanisms of suspiciousness

Rebecca Kazinka^{1,2}, Anita N.D. Kwashie³, Danielle N. Pratt⁴, Iris Vilares³, William Iacono³, Sylvia Wilson⁵, & Angus W. MacDonald III⁵

1. University of Minnesota Medical School, Psychiatry Department, Minneapolis, MN
2. University of Minnesota, Biomedical Engineering Department, Minneapolis, MN
3. University of Minnesota, Psychology Department, Minneapolis, MN
4. Northwestern University, Psychology Department, Evanston, IL
5. University of Minnesota, Institute for Child Development, Minneapolis, MN

Corresponding author: angus@umn.edu

Abstract

Spite sensitivity, or the fear that a person is willing to intentionally take a loss to ensure that another person will as well, may be a key component in understanding persecutory ideation (the belief that others want to harm you). We implemented a co-twin control design to examine potentially causal relationships among persecutory ideation, spite sensitivity, and neural activity and connectivity. Sixty-nine participants (23 monozygotic twin pairs and an additional 23 unpaired monozygotic twins) completed the Minnesota Trust Game, a social decision-making game played asynchronously with an anonymous partner that targets spite sensitivity by varying the incentives of the partner. Participants with more self-reported persecutory ideation (relative to those with lower persecutory ideation) trusted less even when the partner was incentivized to be fair. Similarly, computational modeling showed that increased persecution was associated with greater beliefs of a partner's spitefulness. Twins with greater beliefs of a partner's spitefulness (relative to their co-twins) also reported higher persecution. In addition, twin differences in left lateral OFC activation during the task was associated with spite sensitivity. These results point towards a potentially causal role of the lateral OFC on spite sensitivity, and in turn causal impacts of spite sensitivity on persecutory ideation.

Introduction

Although debilitating psychotic disorders affect a relatively low proportion of the population (~1%) (1), attenuated psychotic-like experiences are far more common (2). Persecutory delusions, beliefs that others are conspiring against oneself, exist at the extreme; however many people endorse questions suggestive of paranoia, such as “I often feel that others have it in for me” (3–5). In individuals with psychosis, persecutory ideation has been associated with theory-of-mind deficits such that they have difficulty understanding another’s intentions, leading to conflicts in relationships (6,7). Even in nonclinical populations, persecutory ideation is associated with social anxiety and fear of others’ negative evaluations, which make social relationships more difficult (8). In the present paper, we aimed to 1) assess social decision-making and related neural functioning in individuals who report elevated levels of suspiciousness yet are not diagnosed with a psychotic disorder and 2) leverage differences between participants and their monozygotic co-twins to provide insight into the potentially causal impacts of persecutory ideation on interpersonal behaviors and related brain activation.

Social decision-making paradigms, in which choices affect not only the player but also others, can reveal not just an individual’s preferences about a tangible reward like money, but also their sense of fairness and beliefs about a partner’s intentions (9). Traditional trust games have sequential choices between two players, such that the first mover’s monetary offer is multiplied when given to the second mover, and the second mover decides how much money to return. Importantly, the second mover is disincentivized to return any money (i.e., earn the most winnings), but often participants pay some amount back (10). However, trust games often only assess situations when the second mover has reasons to be

untrustworthy. Our study uses the Minnesota Trust Game (MTG), which expands upon traditional trust games by controlling the amount of money each player can receive to set up two different conditions based on the second mover's incentives (11). Like the traditional trust game, in one condition the second mover is offered two choices, and taking the higher amount leads to a loss for the first mover (Rational Mistrust); we extend the traditional game by including a condition in which the second mover receives less money if they choose to give the first mover less money as well, incentivizing the second mover to be trustworthy (Suspiciousness). This small change in the incentives in the MTG allows us to target spite sensitivity, or the fear that the second mover will purposefully choose a poor outcome for themselves to ensure that the first mover suffers as well. The incentive structures in the MTG can uniquely provide insights into concerns that another person may behave spitefully in the Suspiciousness condition; we leverage this feature to target spite sensitivity's relationship with persecutory ideation.

Prior research in the MTG has found that individuals higher in persecutory ideation are less trusting in the Suspiciousness condition, even though they are incentivized to trust (11,12). In addition, there is mixed evidence that persecutory ideation is associated with computational measures of spite sensitivity (13,14). Neuroimaging results from the first mover's perspective while playing the MTG match similar results seen in traditional trust games related to mistrust and expand our understanding of neural correlates of spite sensitivity. Individuals with psychosis have shown reduced caudate nucleus activity during a Rational Mistrust condition, increasing the player's perceived risk of losing money, similar to results in the trust game (15). However, in the Suspiciousness condition, there was greater lateral orbitofrontal cortex (OFC) activity associated with increased risk of

losing money for the first mover (14). In addition, specific functional connectivity patterns correspond to suspicious beliefs in individuals with psychosis, namely decreased connectivity between frontoparietal regions (FPN) and the ventromedial prefrontal cortex (vmPFC)/OFC (Wisner et al., 2021). Kazinka et al. (2024) further found that disconnectivity between the frontoparietal network and the OFC/salience network was associated with spite sensitivity in individuals with psychosis. These findings indicated different neural correlates when predicting rational mistrust versus suspicious mistrust, suggesting that they are separate constructs.

In this study, we had the unique opportunity to implement the co-twin control design (16), which is a common method used in twin research (17,18). While most past research on spite sensitivity has focused on cross-sectional designs that are limited to correlational relationships, the co-twin control design can provide additional insights into potentially causal mechanisms. Monozygotic twins have shared genetics and many shared experiences; by comparing an individual against their co-twin, we can control for these factors. This approach allows us to compare differences in traits within a pair, which represent unshared experiences. When these differences also predict differences in measured outcomes, then we can more confidently assert that the differences indicate a causal relationship with the measured outcomes. For example, the co-twin control design has suggested potential causal influences linking paranoia to symptoms such as hallucinations, cognitive disorganization (19) and depression (20). Of note, the MTG has previously been studied in a small sample of monozygotic twins without psychosis using the co-twin control design (12). This study replicated the relationship between FPN-vmPFC/OFC disconnectivity and persecutory ideation

seen in individuals with psychosis and further found that co-twin differences in persecutory ideation were associated with differences in trust in the Suspiciousness condition, suggesting that persecutory ideation may be causally linked to trust in the Suspiciousness condition (i.e., spite sensitivity). The co-twin control model can provide stronger evidence for the role of spite sensitivity in persecutory ideation.

The current study examines spite sensitivity in a community sample selected to match the range of self-reported suspiciousness found in an early psychosis sample previously reported (14). We additionally recruited a subsample of their monozygotic co-twins to examine potentially causal associations of spite sensitivity with persecutory ideation. We aimed to replicate previous results, but we were most interested in examining twin differences to further tease apart the potentially causal role of spite sensitivity and neural activation during the MTG on persecutory ideation. The co-twin control design provides stronger evidence that *a priori* brain regions are associated with spite sensitivity and that spite sensitivity is predictive of persecutory ideation, providing valuable insights for future research on treating persecutory ideation.

Methods

Participants

Seventy-two individual monozygotic twins were recruited from the ES and MS cohorts from the Minnesota Center for Twin and Family Research (MCTFR) at the University of Minnesota (21–23). Participants were between 18 and 45 years. Exclusion criteria included DSM-IV diagnosis of a mental disorder with psychosis, diagnosis of a substance use disorder in the past 3 months, current intoxication,

health issues contraindicating MR scanning, history of neurological or developmental disorders or traumatic brain injury, or an estimated IQ below 70 as determined by the Wechsler Test of Adult Reading (WTAR) (24). Forty-nine of these individuals (referred to as probands) were selected based on having a score on the Multidimensional Personality Questionnaire (MPQ) Alienation subscale that could be near-matched to scores for individuals with psychosis previously described by Kazinka et al. (2024), providing us with a community sample with elevated MPQ-Alienation but not diagnosed with psychosis. In total, three probands were excluded (for Autism Spectrum Disorder, Bipolar Disorder, or claustrophobia), leaving 46 eligible probands (see Table 1). We recruited a subsample of 23 co-twins of the probands (referred to as “co-twins”), not selected for their scores on MPQ Alienation. Procedures for this study were approved by the University of Minnesota Institutional Review Board.

INSERT TABLE 1 HERE

Personality & Clinical Measures

Persecutory ideation was measured using the MPQ-Alienation subscale (25,26), which assesses feelings that others will harm them, that others cannot be trusted, and of betrayal. To screen for mental health symptoms on the assessment day, trained research assistants administered the Depression, Mania, Alcohol Use, Substance Use, and Psychosis sections of the Mini International Neuropsychiatric Interview 5th edition (M.I.N.I., interrater ICC-2 = .80) (27). Handedness laterality was assessed using the Edinburgh Handedness Inventory (28), in which positive values reflect right-handedness and negative values reflect left-handedness.

Decision-Making Task

The Minnesota Trust Game (MTG) is a social decision-making game that depends on the decisions of the player and an anonymous partner (see Figure 1). There are two sequential roles: the first mover decides between two choices (take a small amount of money or trust the second mover to decide); if given the opportunity, the second mover makes a decision between two different ways of allocating money to the first mover and to themselves. Importantly, there is no immediate feedback in the game, so these roles were played asynchronously. Participants played as the second mover first and then reversed roles. Our results focused on the perspective of the first mover. This paradigm is designed to distinguish between sources of distrust, depending on the partner's motivations (either incentivized or disincentivized to betray the first mover). First movers make a decision to either take a small, safe amount of money for both players (\$10) or choose to let the second mover decide between two options that are known to the first mover. The second mover's choices would then be between a mutual option with an equal, larger amount of money for each player (\$20) or unequal amounts where the second mover receives some temptation (T) and the first mover receives some adverse payoff (AD). There are two conditions based on the temptation: when $T = \$25$, the second mover is incentivized to betray the first mover to get a larger reward (because T is greater than \$20), which we refer to as the Rational Mistrust condition (RMT25); when $T = \$15$, the second mover is *disincentivized* to betray the player (receiving less than \$20), which we refer to as the Suspiciousness condition (SUS15). ADs vary in each trial from losing \$15 to winning \$22 to parametrically identify a boundary between trusting and not trusting (referred to as thresholds).

Payment incentives were determined by a random trial from each game based on the participant's choices paired with the choices of a random previous participant.

Data Analysis

Behavioral Analyses

Analyses were conducted in RStudio (v. 4.2.2) unless otherwise specified.

Behavior on the MTG was assessed by calculating a threshold for each condition (RMT25 and SUS15) using a Heaviside function to identify where participants switch between trusting and not trusting in the MTG. Higher thresholds indicate greater mistrust. To test the hypothesis that individuals with greater persecutory ideation (compared to those with less) were less trusting in the SUS15 condition but not the RMT25 condition, we conducted a mixed-effects logistic regression. The model predicted the choice to let the second mover decide (i.e., trust the second mover) or take the safe payoff (\$10), based on condition (RMT25 or SUS15), persecutory ideation (MPQ-Alienation) and adverse payoff (each AD value offered) nested within family (i.e., twin pairs). To examine spite sensitivity, we calculated the first mover's percentage of choices to not trust the second mover during the SUS15 condition (referred to as suspicious mistrust). We conducted similar analyses from the second mover's decision, which can be found in the supplementary materials (Figure S2).

We additionally applied a computational model previously tested and validated on two separate samples (13), using principles from the Fehr-Schmidt inequity aversion model (29). This model calculates four parameters: 1) envy (α), or the dislike of a partner's larger outcome; 2) estimated partner spite-guilt (β' ; referred to as spite-guilt beliefs, a measure of spite sensitivity), which measures

the estimation of the partner's guilt about a greater outcome (for positive values) or spitefulness towards the player (for negative values); 3) risk aversion (R), which assesses the difference of the adverse payoff to the safe amount; and 4) lambda (λ), which calculates inverse temperature (i.e., noisiness in responses). We used the `fmincon` function in MATLAB (2018 edition) to calculate the best parameter values for each individual using maximum likelihood estimation. We hypothesized that spite-guilt beliefs would be more closely related to SUS15 condition thresholds, while risk aversion would be more associated with the RMT25 thresholds. Finally, we used a random effects linear regression to test the relationship of spite sensitivity (measured by suspicious mistrust and the spite-guilt beliefs parameter) and persecutory ideation, with family as a random effect. Model equations and model recovery are reported in the supplemental materials.

Co-twin Control Analysis

To further examine the potentially causal effect of spite sensitivity on persecutory ideation, we conducted co-twin control analyses with 23 twin pairs, which accounts for shared familial confounds (30). In brief, we conducted linear mixed models using the twin pair average score (between-pair) and the twin pair difference score (within-pair) on spite sensitivity. The between-pair score controls what is shared by twins in a twin pair (i.e., shared genetic and environmental factors), while the within-pair score signifies what is unique to each twin (i.e., nonshared environmental factors). Our analyses focused on the within-pair results for behavioral results such as thresholds and parameters, personality traits such as MPQ Alienation, and z-scored beta values from neuroimaging regions of interest (ROIs). We hypothesized that there would be a negative relationship, in which twins who believed the second mover would be more spiteful (negative spite-guilt

beliefs) compared to their co-twins' beliefs about spitefulness would be more likely to also have higher persecutory ideation.

Neuroimaging Analyses

We conducted three task-fMRI scans of the first mover in the MTG. We additionally collected structural scans (T1) and used TOPUP (31,32) to implement distortion correction. Two of the probands were excluded for poor resolution of the T1 structural scan. For another two of the probands, only two of the three runs were used for neuroimaging results due to movement $>.5$ mm for more than 20% of the run. One co-twin was excluded from the neuroimaging data collection because of pregnancy; another co-twin only completed 2 of the 3 scans.

Neuroimaging analyses were conducted using BOLD GLMs and functional connectivity using FSL version 6.0.4 (33). The general linear model compared conditions and risk (no, low, and high risk) based on the adverse payoff. "No risk" was any adverse payoff above \$10 ($\geq \$$), "low risk" was an adverse payoff between \$3 and \$10, and "high risk" was an adverse payoff below \$3. We examined contrasts only between no and high risk, the two conditions, and their interaction. We additionally assessed individual differences by using *a priori* ROIs from the interaction seen in individuals with psychosis (14), which included the bilateral caudate nucleus, left and right lateral OFC, dorsomedial prefrontal cortex (dmPFC), and ventromedial prefrontal cortex (vmPFC) (Figure 3A & S3). We used these ROIs to compare previous results from individuals with psychosis to this community sample. GLM models additionally included an explanatory variable for missed responses, standard motion parameters, and confound variables for timepoints in which motion was greater than .5 mm. No participants met these

exclusion criteria. We used FSL's `easythresh` function with a threshold of $Z = 3.1$ and $p_{\text{brainwise}} = .05$.

To assess functional connectivity, we performed dual regression using FSL (34,35) applied to components defined from a sample of healthy community members ($N=218$) previously reported (36; Figure S5). Five networks of interest were selected: OFC, OFC/insula/dmPFC network, bilateral FPN, and caudate nucleus, based on previous studies (12). Time courses from the intrinsic connectivity networks selected were correlated to calculate interconnectivity. We examined how interconnectivity between these regions related to individual difference measures (MPQ-Alienation, suspicious mistrust, and spite-guilt beliefs), and additionally examined potential causal relationships using the co-twin control design described above.

Data Availability

The code is available on Open Science Foundation at https://osf.io/3k2z9/?view_only=2fd26bc8877e4a4a955d57d53030f70c. Due to the sensitive nature of the participants' data, it is not publicly available, but it can be made available upon reasonable request. These hypotheses were derived from previous work (12,14), but this study was not pre-registered.

Results

Sample demographics and their relationship to MPQ-Alienation and spite-guilt beliefs are reported in Table 1. We found that those with greater MPQ-Alienation also had fewer years of education. We therefore statistically controlled for this factor when conducting the analyses including MPQ-Alienation.

Comparisons between probands and co-twins, as well as probands whose twins

were recruited versus not, are reported in Table S1. Of note, there were more left-handed paired probands versus unpaired probands.

Behavioral Performance and Computational Modeling

MTG performance showed that participants had a similar pattern of behavior to previous MTG performance, in which they were more trusting in the SUS15 condition than the RMT25 condition. Logistic regression showed a significant interaction between condition and adverse payoff (Estimate = .442, SE = .33, $Z = 13.5$, $p < .001$), in which participants trusted more in the SUS15 condition, compared to the RMT25 condition, as the adverse payoff decreased. There were also significant main effects of condition and adverse payoff (p 's < .001). These results did not change when controlling for education, which was also a significant, independent predictor (Estimate = .55, $Z = 3.30$, $p = .001$).

We further tested the role of persecutory ideation on MTG performance. When adding persecutory ideation to the model, we found a significant three-way interaction (condition \times adverse payoff \times Alienation: Estimate = .099, SE = .036, $Z = 2.71$, $p = .007$; Figure 1B). We replicated a significant interaction of condition and Alienation, such that those with higher persecutory ideation were less trusting in the SUS15 condition (Estimate = .141, SE = .034, $Z = 4.20$, $p < .001$). There was a main effect of persecutory ideation as well (Estimate = -.92, SE = .054, $Z = 16.8$, $p < .001$). When controlling for education, the interaction of condition and persecutory ideation held. The main effect of education was also significant (Estimate = .488, $Z = 2.84$, $p = .005$).

Spite sensitivity was tested using computational modeling of the first mover. Participants showed good replication of the group data after simulating the data

based on the estimated variables (Figure S1). There was a strong negative relationship between spite-guilt beliefs and SUS15 condition thresholds (Estimate = -3.71, SE = .528, $t = 7.04$, $p < .001$; Figure 1C), suggesting that participants trusted more when they believed the second mover would feel more guilty. We also found a strong negative correlation between spite-guilt beliefs and the RMT25 condition thresholds (Estimate = -4.85, SE = .433, $t = 11.22$, $p < .001$). In terms of risk aversion, we found this parameter correlated with both the RMT25 condition thresholds (Estimate = .145, SE = .027, $t = 5.30$, $p < .001$) and the SUS15 condition thresholds (Estimate = .098, SE = .029, $t = 3.42$, $p = .001$). Finally, we examined individual differences related to persecutory ideation, which showed a significant negative correlation with spite-guilt beliefs (Estimate = -1.45, SE = .48, $t = 3.06$, $p = .005$; Figure 2A) and a positive correlation with suspicious mistrust (Estimate = .011, SE = .004, $t = 2.97$, $p = .005$). These results did not differ when controlling for education.

Twin differences in behavior

To further assess the potentially causal role of spite sensitivity on persecutory ideation, we compared the selected community sample to their monozygotic co-twins in a subsample of 23 pairs. A strong negative correlation was found between twin differences in spite-guilt beliefs and persecutory ideation (Estimate = -2.18, SE = .63, $t = 3.44$, $p = .002$; Figure 2B), so that twins with greater spite-guilt beliefs (compared to their co-twins) also had higher persecutory ideation. Results held when controlling for education. There was also a strong negative correlation between twin differences in spite-guilt beliefs and suspicious mistrust (Estimate = -.08, SE = .02, $t = 4.11$, $p < .001$), such that twins with higher spite-guilt beliefs also had less trust during the SUS15 condition. However,

we did not identify a sufficiently strong relationship between twin differences in suspicious mistrust and persecutory ideation (Estimate = 8.98, SE = 6.13, $t = 1.46$, $p = .158$).

Neuroimaging Results

Main effects of risk and condition

We compared neural activation across the two conditions (RMT25 and SUS15) and two risk levels (high and no risk), based on the adverse payoff (Figure 3). *A priori* ROIs included bilateral caudate nucleus, bilateral OFC, dmPFC, and vmPFC (Figure 3A & S3). The interaction between Risk and Condition showed a small region of positive activation in the caudate. When we examined the main effect of Risk, we found a significant reduction in activation in bilateral caudate during high risk trials, consistent with previous findings. In addition, there was a significantly greater activation in the no risk trials than high risk trials during RMT25 in the bilateral caudate nucleus. Activation comparing no risk>high risk during RMT25 was negatively correlated with the risk aversion parameter in the left caudate (Estimate = -.007, SE = .003, $t = 2.58$, $p = .013$). The main effect of condition showed a significant correlation between the right lateral OFC activation and suspicious mistrust (Estimate = .084, SE = .037, $t = 2.26$, $p = .034$). When looking at high risk>no risk trials in the SUS15 condition, this contrast was negatively correlated with suspicious mistrust in the left lateral OFC (Estimate = -.14, SE = .04, $t = 3.72$, $p < .001$), right lateral OFC (Estimate = -.096, SE = .035, $t = 2.80$, $p = .010$), and dmPFC (Estimate = -.091, SE = .027, $t = 3.35$, $p = .003$); Figure 3A). This result suggests that lateral OFC played a role in decision-making in the SUS15 condition specifically, even if it was not significant in the overall contrast. Instead, the caudate nucleus had a greater involvement in decision-

making in the MTG in this sample. Whole brain analyses are reported in the supplemental results (Figure S4).

Functional connectivity

We explored the influence of task performance on functional connectivity analyses. There was a significant positive relationship between right FPN-OFC and suspicious mistrust (Estimate = .612, SE = .226, $t = 2.71$, $p = .009$; Figure 4); spite-guilt beliefs showed a non-significant negative relationship for the FPN-OFC connection (Estimate = -3.18, SE = 1.65, $t = 1.93$, $p = .058$). In other studies, this connection was positively correlated with paranoid thoughts in individuals with psychosis during rest (37).

Twin differences in neural activation

We finally examined the potentially causal role of spite sensitivity on neural activations. In the SUS15 condition high-risk>no-risk analyses, we found that twin differences in left lateral OFC activation predicted suspicious mistrust (Estimate = -.14, SE = .06, $t = 2.53$, $p = .019$; Figure 3C) so that twins with more negative left lateral activation also trusted less during the SUS15 condition. These results were not seen in twin differences of the right lateral OFC or dmPFC, nor when examining relationships with spite-guilt beliefs or MPQ-Alienation.

Discussion

This study contributes to the growing literature supporting spite sensitivity as a construct underlying suspiciousness. Forty-six community participants (probands) were selected because their MPQ-Alienation scores closely matched those of a group of individuals with psychosis. In addition, 23 of these probands were compared to their monozygotic co-twins. MTG behavior in the community

sample replicated previous results for both the First and Second Mover Games. Spite-guilt beliefs and twin differences in spite-guilt beliefs predicted persecutory ideation. In the Rational Mistrust condition, we replicated the relationship with bilateral caudate; when comparing high-risk vs. no-risk trials, left caudate was also associated with risk aversion. Furthermore, right lateral OFC activation when comparing the two conditions was correlated with suspicious mistrust. Bilateral OFC and dmPFC activation comparing high- and no-risk trials in the Suspiciousness condition predicted suspicious mistrust as well. When examining twin differences, the left lateral OFC was also associated with suspicious mistrust. These results suggest a potentially causal relationship between lateral OFC activation and spite sensitivity and, in turn, between spite sensitivity and persecutory ideation.

Similar to previously published work (11–14), we found a relationship between the Suspiciousness condition and persecutory ideation. A recent meta-analysis showed that individuals with psychosis and clinically high-risk participants, but not relatives of individuals with psychosis, had lower baseline trust than healthy controls in the standard trust game (38). This meta-analysis also examined the relationship between positive and negative symptoms and trust but found mixed evidence that positive symptoms (like delusions) were specifically related to distrust. A key component of the MTG is distinguishing situations in which the second mover is or is not incentivized to betray the first mover, which may explain why we find more consistent results associating persecutory ideation and suspicious mistrust. For example, reported persecutory ideation in a large online sample was associated with increased beliefs that a second mover in the standard trust game was operating with a harmful intent rather than a selfish

intent (39), consistent with the distinctions of the Suspiciousness and Rational Mistrust conditions. This further suggests that spite sensitivity is a valuable construct, measurable across several samples, that can contribute to our understanding of specific influences on persecutory ideation.

Furthermore, computational modeling linked spite-guilt beliefs and MPQ-Alienation more strongly using the co-twin control approach. Of note, several other persecutory ideation models based on behavioral paradigms exist. Using a modified Social Value Orientation paradigm, Barnby, Raihani, and Dayan (2022) developed a model integrating Bayesian modeling and the Fehr-Schmidt inequity aversion model; like our model, guilt was allowed to be negative and was used to predict beliefs of a partner (40). Barnby and colleagues (2022) found that individuals with greater paranoia were more likely to want to earn more than their partners, believe that their partners had harmful intent, and have more rigidity in their belief-updating. Similarly, individuals playing a modified Dictator Game showed that greater paranoia was associated with beliefs that their partners had greater harmful intent and more rigidity in these beliefs (41). In an advice-taking task in a probabilistic lottery game, individuals with persecutory ideation were less likely to trust advice and believed that it was intentionally misleading; hierarchical Bayesian modeling showed that these individuals had more rigid belief-updating (42,43). In contrast, research using a probabilistic reversal learning paradigm, without any social cues, also showed that paranoia is associated with an increased prior belief of volatility in the game in individuals with psychosis, in online samples, and in rats with chronic methamphetamine exposure (44,45). Given that there was no feedback provided in our game, we measure prior beliefs about a partner's intentions. The use of the co-twin control model provides stronger

support that there is a potentially causal link between persecutory ideation and prior beliefs that a partner might not have good intentions. Overall, these models converge on the idea that prior beliefs are more negative and that they remain more rigid despite new information. However, it is unclear if these findings are driven by specifically social priors or more general beliefs about volatility in the environment. Previous work in the MTG showed that spite-guilt beliefs were not associated with performance against a nonhuman random partner, suggesting a privileged role for social information processing.

To further elucidate the mechanisms of spite sensitivity, we examined the interaction of risk and condition. Our community sample had a more tempered response during decision-making than did individuals with psychosis (14). Most notably, the whole-brain analysis did not show significant differences between high and no risk trials in the Suspiciousness condition; however, when we examined *a priori* ROIs more specifically, those with higher spite sensitivity had greater activation in the no-risk trials compared to the high-risk trials. This result suggests that these regions are still relevant for processing perceived risks seen in the community sample, although to a lesser degree than the individuals with psychosis. Instead, we found that the caudate nucleus activation was greater in the no risk trials compared to high risk trials, similar to the individuals with psychosis previously reported. The caudate nucleus is associated with reward processing (46), and this result may reflect expected processing—in no-risk trials, the participants are more likely to trust their partner and would receive a larger amount of money than the safe amount. Previous studies of the standard trust game found reduced activation in the right caudate in individuals with psychosis and non-affected siblings compared to healthy controls when making investments

(15,47). Our results suggest intact reward processing in the caudate, although we did not directly compare these findings to another control group. Unlike Gromann and colleagues (2013), the caudate was not associated with persecutory ideation. The MTG sees more specific relationships with persecutory ideation in the Suspiciousness condition instead, while the left caudate activation in the Rational Mistrust condition was associated with risk aversion. In comparison to the previously reported individuals with psychosis, our community sample may be able to compensate more readily, showing a more tempered neural response.

We additionally found supporting evidence for neural correlates of spite sensitivity. Suspicious mistrust was associated with lateral OFC when comparing conditions, as well as comparing risk in the Suspiciousness condition. This region was identified as important for suspiciousness decision-making previously in a group of individuals with psychosis and therefore may suggest that altered activation may still occur depending on individual beliefs about spite (14). This brain region is associated with assessment of punitive outcomes and may be related to fear of spite (48). There was also an association between twin differences in the left lateral OFC and suspicious mistrust. Previous research has also shown twin differences in these regions were associated with twin differences in the severity of stressful life events (49), which may be involved in the changes in spite sensitivity, although we cannot directly test this hypothesis. These overlapping findings would support the stress-vulnerability model (50). These results suggest that these regions are generally implicated in spite sensitivity.

There are several limitations to this study. Most notably, we had a relatively small sample of monozygotic twin pairs, limiting the power we had to identify potentially causal relationships. Nonetheless, we believe that these preliminary

results warrant further research in the causal influences of persecutory ideation. In addition, by controlling for shared genetic and experiential influences, the co-twin model offers stronger support for causal associations between variables, but without longitudinal data, the direction of these relationships is unclear. It is also possible that associations may be driven by unshared experiences that influence both persecutory ideation and spite sensitivity but were not included here.

Our unique opportunity to investigate persecutory ideation in monozygotic twins demonstrated that spite sensitivity is potentially causal for persecutory ideation, strengthening our understanding of this relationship in individuals with psychosis and in undergraduate students (13,14). Furthermore, we found that twin differences in left lateral OFC activation also predicted spite sensitivity; this region of interest was previously identified in individuals with psychosis, providing stronger evidence that it may be important in processing spite sensitivity and persecutory ideation. Given these individual differences in spite sensitivity, future research may benefit from examining longitudinal factors such as previous life stress and social support, which might explain the development of persecutory ideation and psychosis. Spite sensitivity, measured by the Minnesota Trust Game, may therefore be an important component in understanding persecutory ideation in research in psychosis and psychosis risk.

Acknowledgements

Thank you to all of the participants, without whom this work would not have been possible, and to Saaraa Ameri and Naomi Skarsgard for their assistance with data collection. Finally, we appreciate the support and critical feedback from the University of Minnesota Institute for Engineering Medicine Manuscript Writing Workshop.

Funding

This project was funded by U.S. National Institutes of Mental Health grants to AM (R21MH112918) and RK (T32-MH115688). RK was also supported by funding from the U.S. National Institutes of Biomedical Imaging and Bioengineering (T32-EB031512). The National Institutes of Health grants additionally supported WI (R01DA013240, R01DA036216, and R37DA005147) and SW (K01DA037280). This work derives from a portion of RK's dissertation.

Author Contributions

Rebecca Kazinka: Methodology, investigation, formal analysis, writing - original draft, and writing-review & editing.

Anita Kwashie: Investigation.

Danielle Pratt: Methodology, investigation, and writing-review & editing.

Iris Vilares: Writing - review & editing and supervision.

William Iacono: Funding acquisition, review & editing, and resources.

Sylia Wilson: Writing - review & editing, methodology, resources, and supervision.

Angus MacDonald III: Conceptualization, methodology, writing - review & editing, funding acquisition, resources, and supervision.

References

1. Kendler KS, Gallagher TJ, Abelson JM, Kessler RC. Lifetime prevalence, demographic risk factors, and diagnostic validity of nonaffective psychosis as assessed in a US community sample. The National Comorbidity Survey. *Arch Gen Psychiatry*. 1996 Nov;53(11):1022-31.
2. Johns LC, Van Os J. The continuity of psychotic experiences in the general population. *Clin Psychol Rev*. 2001;21(8):1125-41.
3. Bebbington PE, McBride O, Steel C, Kuipers E, Radovanovic M, Brugha T, et al. The structure of paranoia in the general population. *Br J Psychiatry*. 2013 Jun;202:419-27.
4. Verdoux H, Van Os J. Psychotic symptoms in non-clinical populations and the continuum of psychosis. *Schizophr Res*. 2002;54(1-2):59-65.
5. van Os J, Reininghaus U. Psychosis as a transdiagnostic and extended phenotype in the general population. *World Psychiatry*. 2016 Jun;15(2):118-24.
6. Mehl S, Rief W, Lüllmann E, Ziegler M, Kesting ML, Lincoln TM. Are theory of mind deficits in understanding intentions of others associated with persecutory delusions? *J Nerv Ment Dis*. 2010 Jul;198(7):516-9.
7. Langdon R, Siegert RJ, McClure J, Harrington L. Schizophrenia, theory of mind, and persecutory delusions. *Cognit Neuropsychiatry*. 2005 May 1;10(2):87-104.
8. Martin JA, Penn DL. Attributional style in schizophrenia: An investigation in outpatients with and without persecutory delusions. *Schizophr Bull*. 2002;28(1):131-41.
9. Alos-Ferrer C, Farolfi F. Trust Games and Beyond. *Front Neurosci*. 2019 Sep 10;13:887.

10. Berg J, Dickhaut J, McCabe KA. Trust, Reciprocity, and Social History. *Games Econ Behav.* 1995;10:122-42.
11. Johnson MK, Rustichini A, MacDonald III AW. Suspicious personality predicts behavior on a social decision-making task. *Personal Individ Differ.* 2009 Jul;47(1):30-5.
12. Wisner KM, Johnson MK, Porter JN, Krueger RF, MacDonald AW. Task-related neural mechanisms of persecutory ideation in schizophrenia and community monozygotic twin-pairs. *Hum Brain Mapp.* 2021;42(16):5244-63.
13. Kazinka R, Vilares I, MacDonald III AW. A Computational Model of Non-optimal Suspiciousness in the Minnesota Trust Game. *Comput Psychiatry.* 2022;6(1):60-78.
14. Kazinka R, Kwashie AND, Pratt DN, Vilares I, MacDonald AW. Value Representations of Spite Sensitivity in Psychosis on the Minnesota Trust Game. *Biol Psychiatry Cogn Neurosci Neuroimaging.* 2024 Apr 1;9(4):429-36.
15. Gromann PM, Heslenfeld DJ, Fett AK, Joyce DW, Shergill SS, Krabbendam L. Trust versus paranoia: Abnormal response to social reward in psychotic illness. *Brain J Neurol.* 2013;136(6):1968-75.
16. McGue M, Osler M, Christensen K. Causal Inference and Observational Research: The Utility of Twins. *Perspect Psychol Sci J Assoc Psychol Sci.* 2010 Sep;5(5):546-56.
17. Boomsma D, Busjahn A, Peltonen L. Classical twin studies and beyond. *Nat Rev Genet.* 2002 Nov;3(11):872-82.
18. Neale MC, Cardon LR. Methodology for Genetic Studies of Twins and Families [Internet]. 1992 [cited 2024 Apr 25]. Available from: <https://link.springer.com/book/10.1007/978-94-015-8018-2>

19. Zavos HMS, Freeman D, Haworth CMA, McGuire P, Plomin R, Cardno AG, et al. Consistent Etiology of Severe, Frequent Psychotic Experiences and Milder, Less Frequent Manifestations: A Twin Study of Specific Psychotic Experiences in Adolescence. *JAMA Psychiatry*. 2014 Sep 1;71(9):1049-57.
20. Zavos HMS, Eley TC, McGuire P, Plomin R, Cardno AG, Freeman D, et al. Shared Etiology of Psychotic Experiences and Depressive Symptoms in Adolescence: A Longitudinal Twin Study. *Schizophr Bull*. 2016 Sep;42(5):1197-206.
21. Keyes MA, Malone SM, Elkins IJ, Legrand LN, McGue M, Iacono WG. The Enrichment Study of the Minnesota Twin Family Study: Increasing the Yield of Twin Families at High Risk for Externalizing Psychopathology. *Twin Res Hum Genet Off J Int Soc Twin Stud*. 2009 Oct;12(5):489-501.
22. Wilson S, Elkins IJ, Bair JL, Oleynick VC, Malone SM, McGue M, et al. Maladaptive personality traits and romantic relationship satisfaction: A monozygotic co-twin control analysis. *J Abnorm Psychol*. 2018;127(4):339-47.
23. Iacono WG, McGue M, Krueger RF. Minnesota Center for Twin and Family Research. *Twin Res Hum Genet*. 2006 Dec;9(6):978-84.
24. Wechsler D. Wechsler Test of Adult Reading: WTAR. San Antonio, TX: The Psychological Corporation; 2001.
25. Patrick CJ, Curtin JJ, Tellegen A. Development and validation of a brief form of the Multidimensional Personality Questionnaire. *Psychol Assess*. 2002;14(2):150-63.
26. Tellegen A, Waller NG. Exploring personality through test construction: Development of the Multidimensional Personality Questionnaire. In: Boyle GJ, Matthews G, Saklofske DH, editors. *The SAGE handbook of personality theory and*

assessment, Vol 2 Personality measurement and testing. Sage Publications, Inc.; 2008. p. 261-92.

27. Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, et al. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J Clin Psychiatry*. 1998;59 Suppl(20):22-33.

28. Oldfield RC. The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*. 1971;9(1):97-113.

29. Fehr E, Schmidt KM. A theory of fairness, competition, and cooperation. *Q J Econ*. 1999;(August):817-68.

30. Begg MD, Parides MK. Separation of individual-level and cluster-level covariate effects in regression analysis of correlated data - Begg - 2003 - *Statistics in Medicine* - Wiley Online Library [Internet]. [cited 2024 Apr 25]. Available from: <https://onlinelibrary.wiley.com/doi/10.1002/sim.1524>

31. Andersson JLR, Skare S, Ashburner J. How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging. *NeuroImage*. 2003;20(2):870-88.

32. Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TEJ, Johansen-Berg H, et al. Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*. 2004 Jan 1;23:S208-19.

33. Jenkinson M, Beckmann CF, Behrens TEJ, Woolrich MW, Smith SM. FSL. *NeuroImage*. 2012;62(2):782-90.

34. Beckmann CF, Mackay CE, Filippini N, Smith SM. Group comparison of resting-state fMRI data using multi-subject ICA and dual regression. *Organ Hum Brain Mapp*. 2009;

35. Nickerson L, Smith SM, Öngür D, Beckmann CF. Using Dual Regression to Investigate Network Shape and Amplitude in Functional Connectivity Analyses. *Front Neurosci.* 2017;11:115–115.
36. Rueter AR, Abram SV, MacDonald III AW, Rustichini A, DeYoung CG. The Goal Priority Network as a Neural Substrate of Conscientiousness. *Hum Brain Mapp.* 2018;39(9):3574–85.
37. Yu L, Kazinka R, Pratt D, Kwashie A, Macdonald AW. Resting-state networks associated with behavioral and self-reported measures of persecutory ideation in psychosis. *Brain Sci.* 2021;11(11).
38. Prasannakumar A, Kumar V, Rao NP. Trust and psychosis: a systematic review and meta-analysis. *Psychol Med.* 2022 Aug 17;1–9.
39. Greenburgh A, Bell V, Raihani N. Paranoia and conspiracy: group cohesion increases harmful intent attribution in the Trust Game. *Peerj.* 2019 Aug 16;7:e7403.
40. Barnby JM, Raihani N, Dayan P. Knowing me, knowing you: Interpersonal similarity improves predictive accuracy and reduces attributions of harmful intent. *Cognition.* 2022 Aug 1;225:105098.
41. Barnby JM, Mehta MA, Moutoussis M. The computational relationship between reinforcement learning, social inference, and paranoia. *PLOS Comput Biol.* 2022 Jul 25;18(7):e1010326.
42. Diaconescu AO, Wellstein KV, Kasper L, Mathys C, Stephan KE. Hierarchical bayesian models of social inference for probing persecutory delusional ideation. *J Abnorm Psychol.* 2020;129(6):556–69.

43. Wellstein KV, Diaconescu AO, Bischof M, Rüesch A, Paolini G, Aponte EA, et al. Inflexible social inference in individuals with subclinical persecutory delusional tendencies. *Schizophr Res.* 2020;215:344–51.
44. Reed EJ, Uddenberg S, Suthaharan P, Mathys CD, Taylor JR, Groman SM, et al. Paranoia as a deficit in non-social belief updating. *eLife.* 2020;9:1–55.
45. Sheffield JM, Suthaharan P, Leptourgos P, Corlett PR. Belief Updating and Paranoia in Individuals With Schizophrenia. *Biol Psychiatry Cogn Neurosci Neuroimaging.* 2022 Apr;S2451902222000799.
46. Bartra O, McGuire JT, Kable JW. The valuation system: A coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *NeuroImage.* 2013;76:412–27.
47. Gromann PM, Shergill SS, Haan LD, Meewis DGJ, Fett AKJ. Reduced brain reward response during cooperation in first-degree relatives of patients with psychosis: an fMRI study. *Psychol Med.* 2014;44:3445–54.
48. Rolls ET. Limbic structures, emotion, and memory. *The Curated Reference Collection in Neuroscience and Biobehavioral Psychology.* 2016. 1 p.
49. Godinez DA, McRae K, Andrews-Hanna JR, Smolker H, Banich MT. Differences in frontal and limbic brain activation in a small sample of monozygotic twin pairs discordant for severe stressful life events. *Neurobiol Stress.* 2016 Dec 1;5:26–36.
50. Ingram RE, Luxton DD. Vulnerability-Stress Models. In: *Development of psychopathology: A vulnerability-stress perspective.* Thousand Oaks, CA, US: Sage Publications, Inc; 2005. p. 32–46.

Table 1. Sample Demographics

Category	Participants	Relationship with MPQ-Alienation	Relationship with guilt beliefs
N	69	69	
Age	30.9 (5.3)	Estimate = -.15, SE = .197, t = .80, p = .426	Estimate = .09, SE = .197, t = .948, p = .343
% Male	43.5%	t(44) = 0.76, p = .451	t(44) = 1.20, p = .234
Years of Education	16.5 (2.3)	Estimate = -1.34, SE = .402, t = 3.34, p = .002	Estimate = -.10, SE = .100, t = .64, p = .521
Parental Years of Education (average)	14.8 (3.2)	Estimate = -.495, SE = .310, t = 1.59, p = .119	Estimate = -.073, SE = .310, t = .44, p = .659
Handedness Laterality	73.3 (55.8)	Estimate = -.020, SE = .017, t = 1.19, p = .240	Estimate = -.004, SE = .017, t = .11, p = .911
% Racial Minority	2.9%	W = 91, p = .400	W = 28, p = .611
% Hispanic or Latinx	1.4%	W = 19, p = .466	W = 49, p = .001
MPQ-Alienation	34.5 (8.0)	-	Estimate = .48, t = 3.60, p = .001

Note. Parentheses indicate standard deviation. Demographics for all participants were correlated (Spearman) with MPQ-Alienation and spite-guilt beliefs. Degrees of freedom were reduced to match the unique number of families in the study (46). For racial minority and hispanic or Latinx comparisons, the Wilcoxon rank sum test was used.

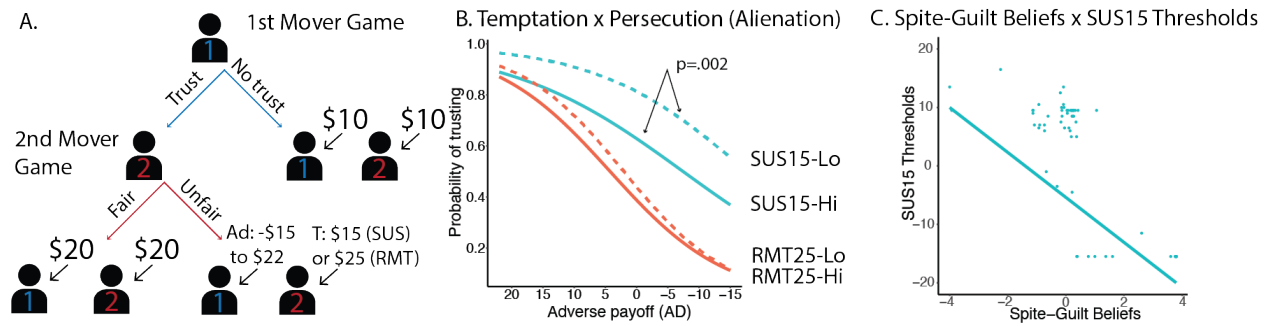


Figure 1. Schematic of Minnesota Trust Game. A) Decision tree for the two roles in the Minnesota Trust Game: First Mover and Second Mover. The first mover decided whether to take a small sum of money ($S=\$10$) or let the second mover decide between two options: one where both partners got a mutually beneficial larger amount of money ($M=\$20$) or an uneven choice. Temptation T was manipulated to either be greater than M (Rational Mistrust - RMT25) or less than M (Suspiciousness - SUS15). Adapted from Kazinka et al., 2024, and used with permission. B) Aggregate first-mover behavior across all community participants, divided by condition. Overall, trust decreases as the adverse payoff decreases, yet participants are generally more trusting in the Suspiciousness condition. Individuals who were higher on MPQ-Alienation were less likely to trust particularly during the Suspiciousness condition (shown as a median split for clarity). C) We replicated a significant negative relationship between spite-guilt beliefs and thresholds in the Suspiciousness condition. Line fit represents the predicted values from the regression model.

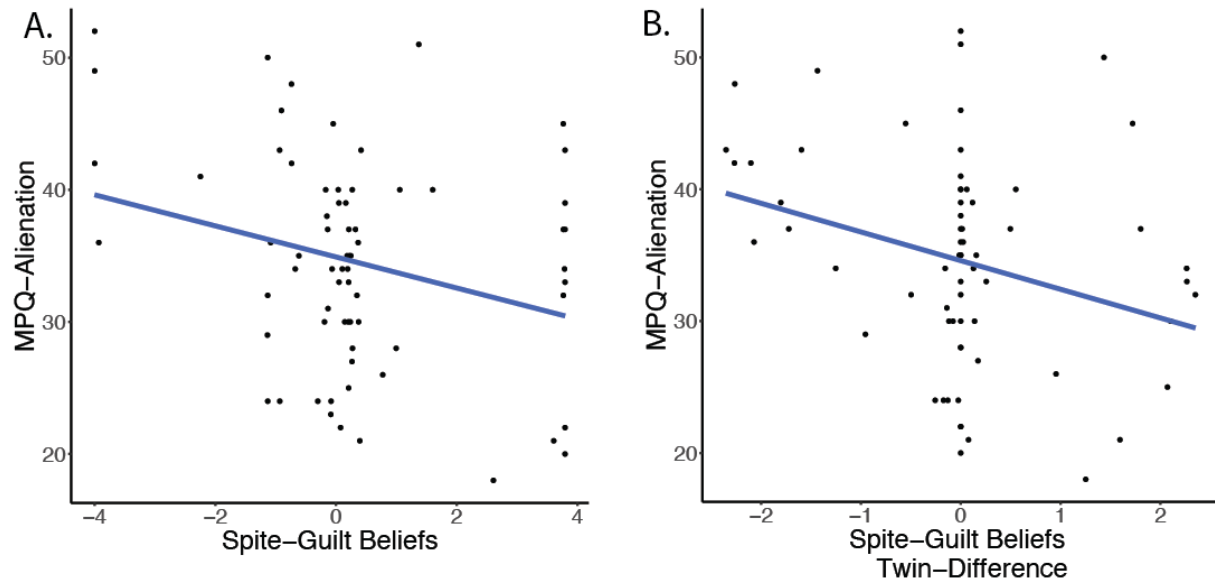


Figure 2. Spite-Guilt Beliefs predicts MPQ-Alienation. A) Spite-guilt beliefs had a strong negative relationship with MPQ-Alienation, in which more beliefs that others were spiteful was associated with reporting more alienation. B) The co-twin control model further supported this result by showing that twin differences in spite-guilt beliefs predicted MPQ-Alienation, suggesting a potentially causal role of spite-guilt beliefs on MPQ-Alienation. Line fit represents the predicted values from the regression model.

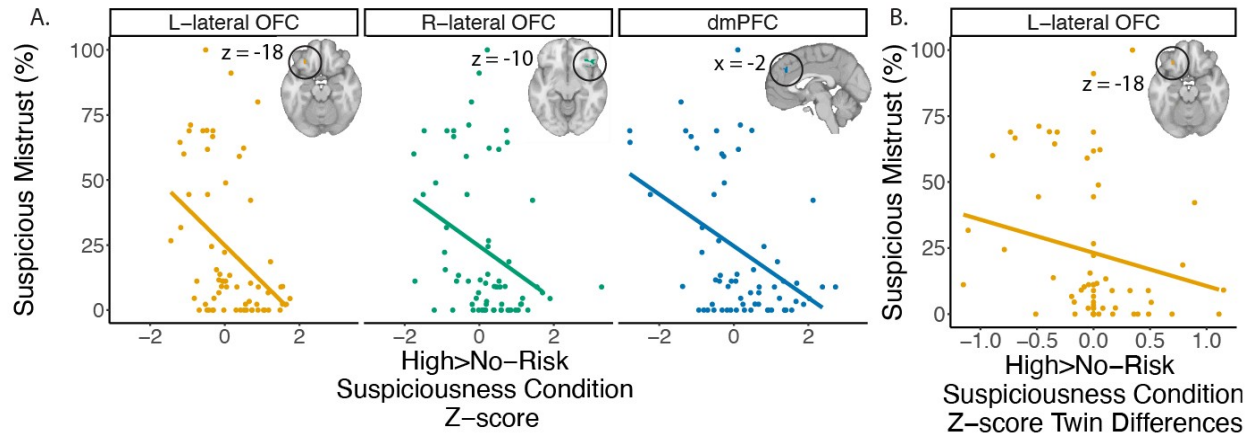


Figure 3. Task-based BOLD response predicts suspicious mistrust. A. the percentage of mistrust in the Suspiciousness condition was related to activation in the bilateral OFC and dmPFC when comparing High-Risk>No-Risk trials in the Suspiciousness condition. B. We further showed that twin differences in the left lateral OFC contrast of High-Risk>No-Risk trials in the Suspiciousness conditions predicted the percentage of mistrust in the Suspiciousness condition (suspicious mistrust), strengthening our understanding of this relationship. Line fit represents the predicted values from the regression model.

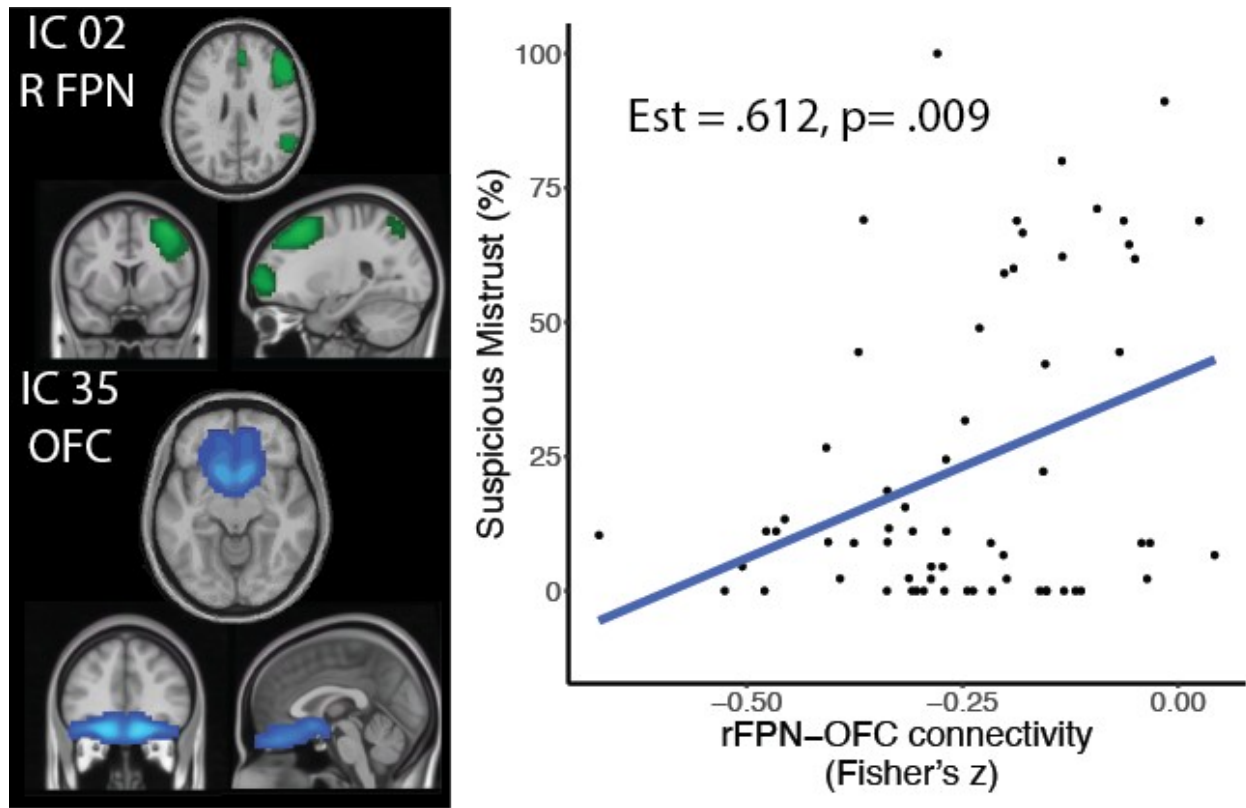


Figure 4. Functional connectivity. Suspicious mistrust was related to the connectivity between the right frontoparietal network and OFC.