NIH-PA Author Manuscript

Functional neuroimaging of mentalizing during the trust game in social anxiety disorder

Chandra Sehkar Sripada[a], **Mike Angstadt**[a], **Sarah Banks**[c], **Pradeep J. Nathan**[d], **Israel Liberzon**[a,b], and **K. Luan Phan**[a,b]

[a]Psychiatry Affective Neuroimaging Laboratory, Department of Psychiatry, University of Michigan

[b]Mental Health Service, VA Ann Arbor Healthcare System, Ann Arbor, Michigan, USA

[c]Neuropsychology and Cognitive Neuroscience Unit, Montreal Neurological Institute, Montreal, Quebec, Canada

[d]Brain Mapping Unit, Department of Psychiatry, University of Cambridge, Cambridge, UK

Abstract

Individuals with generalized social anxiety disorder tend to make overly negative and distorted predictions about social events, which enhance perceptions of threat and contribute to excessive anxiety in social situations. Here, we coupled functional magnetic resonance imaging and a multiround economic exchange game ('trust game') to probe mentalizing, the social-cognitive ability to attribute mental states to others. Relative to interactions with a computer, those with human partners ('mentalizing') elicited less activation of medial prefrontal cortex in generalized social anxiety patients compared with matched healthy control participants. Diminished medial prefrontal cortex function may play a role in the social-cognitive pathophysiology of social anxiety.

Keywords

Introduction

Generalized social anxiety disorder is a highly prevalent and disabling psychiatric illness that is characterized by excessive fear of public scrutiny and negative evaluation across a variety of social situations [1,2]. The underlying cause of this exaggerated social fear is unknown, but could partly be because of deficits in social cognition, which manifest as a tendency toward inaccurate and distorted interpretations of the beliefs and intentions of others during interpersonal interactions [3]. Identification of a neural mechanism that explains these social-cognitive deficits in generalized social anxiety disorder (hereafter referred to as 'social anxiety') remains elusive.

So far, evidence for brain-based dysfunction in social anxiety has focused on exaggerated reactivity of the amygdala in response to facial signals that convey criticism and/or negative feedback (angry, contemptuous/disgusted, and/or fearful faces) (for meta-analysis, see Ref.

Correspondence to K. Luan Phan, MD, Department of Psychiatry, University of Michigan, Rachel Upjohn Building, Room 2751, PO Box 67654250 Plymouth Road, Ann Arbor, MI 48109-2700, USA, Tel: +734 936 9275; fax: +734 936 7868; e-mail: luan@umich.edu.

[4]). However, the usefulness of static face stimuli in elucidating social-cognitive deficits in social anxiety is likely to be limited, as these stimuli primarily engage perception of emotional signals, typically do not engage prefrontal cortex, and do not reflect real-world social interactions that are inherently dynamic and interactive.

Functional magnetic resonance imaging (fMRI) studies of social cognition have recently used a novel approach called 'social neuroeconomics' [5,6], which uses economic game theory to dynamically model fundamental aspects of social interaction such as cooperation, trust, and social signaling [7-11]. Collectively, these studies have begun to characterize the neural correlates of social interactions, including processes that encode the motives and intentions of social partners to guide appropriate behaviors, referred to as 'theory of mind' or 'mentalizing' [12,13]. The medial prefrontal cortex has been implicated as a key brain region that implements mentalizing during the social interactions by these studies [12-16], including those studies that have specifically used economic exchange games [7,10].

As a prototypical interactive economic exchange game, the 'trust game' serves as a potent probe of mentalizing abilities because it sets up the need to make inferences about the mental state of others [7], including partners' beliefs, desires, intentions, and certain dispositions (e.g. trustworthiness, reliability, approachability, etc.). Such information is crucial for predicting partners' responses; as recently noted by Frith and Frith [14], 'It is important for us to be able to read the minds of others because it is their mental states that determine their actions'. Therefore, functional neuroimaging of the trust game [17] may shed light on a novel behavioral and neural mechanism that explains the tendency of individuals with social anxiety to misinterpret partner motives, which may contribute to their reticence to enter into or continue social interactions. In this study, we used a version of the trust game previously coupled with fMRI [7,11] to probe the neural correlates of mentalizing in social anxiety. Given that individuals with social anxiety tend to form inaccurate impressions about others [18], we predicted that relative to healthy controls, prefrontal areas such as the medial prefrontal cortex would exhibit less engagement during mentalizing in individuals with social anxiety.

Methods

Participants

Twenty-six patients with generalized social anxiety disorder, based on *Diagnostic and Statistical Manual of Mental Disorders (fourth edition)* criteria as confirmed by the Structured Clinical Interview for *Diagnostic and Statistical Manual of Mental Disorders (fourth edition)* with additional probes from the Liebowitz Social Anxiety Scale and the Social Phobic Situations Interview, and 26 healthy control individuals (hereafter referred to as 'controls') participated. None of the social anxiety patients had a current depressive episode or recent substance abuse/dependence (within 6 months of study entry), or a lifetime history of bipolar disorder, psychotic disorder, any autism/pervasive developmental disorder, or mental retardation. All participants were right-handed and free of current or past major medical or neurologic illness, as confirmed by a medical exam by a physician. All of the participants were free of psychoactive medications at the time of the study, and none of the participants tested positive on a urine toxicology screen or alcohol breathalyzer on the day of scanning. Demographic and clinical characteristics are reported in Table 1. All participants provided written informed consent for this study, as approved by the local institutional review board.

Social-cognitive task

The fMRI task involved an event-related design. Participants were instructed that they are assigned to be 'Decision Maker 1' (DM1) in a 'decision-making' game. They were also told that they would be playing with other people (HUMAN partners) who had previously

participated in the same game as 'Decision Maker 2' (DM2) and whose responses were previously recorded and now serve as DM2 'reactions' to their (DM1's) decisions. Participants were further instructed to imagine as if they were playing the game with DM2 partners in real-time. As DM1, the participant was informed of the task in the following way (Fig. 1): (i) for each trial, you are given 20 monetary units (to be converted into actual money after the end of the experiment); (ii) for each trial, you must make a decision to keep, and thus, equally divide the 20 units between yourself and your partner (KEEP) or to invest the 20 units (INVEST); (iii) if you choose to KEEP, then the actual outcome of this trial is complete (e.g. you will receive 10 units, and your partner will receive 10 units); (iv) if you choose to INVEST, then the amount doubles to 40 monetary units, and the actual outcome of the trial is to be decided by DM2, who can decide to split the money equally with you (e.g. you will receive 20 units, and DM2 will receive 20 units) or to keep the entire amount (e.g. you will receive 0 units, and DM2 will receive 40 units) (SPLIT). Participants were told that they would play with three types of DM2 players who were classified based on their previously recorded actions as: (i) type 1: 'tended to split the money more than 50% of the time'; (ii) type 2: 'tended to split the money about 50% of the time'; and (iii) type 3: 'tended to split the money less than 50% of the time.' They were also told that they were playing with a computer partner (COMPUTER) who 'split the money 50% of the time.' Unbeknownst to the participants, the frequency to 'split' (e.g. reciprocate) an investment was actually fixed at the following frequencies: (i) type 1 $\lambda=\lambda 75\%$; (ii) type 2 $\lambda=\lambda 50\%$; (iii) type 3 $\lambda=\lambda 25\%$. To maximize economic gain, participants had to decode and predict DM2's actions (e.g. operationalized as the likelihood that DM2 would reciprocate), and thus engage mentalizing.

At the start of the trial (Fig. 1), participants viewed three different face photographs representing each HUMAN DM2 type but the 'face' was obscured by a colored oval (purple, yellow, brown) so as to convey a sense of anonymity and to obviate neural responses to facial features, or they viewed an image of a computer screen (COMPUTER). The DM2 image appeared for 4λs during which the participants were instructed to make their choice (KEEP or INVEST) by button press. In real-time and based on the participant's own decision/choice, feedback was provided immediately in the form of a DM2 image reappearing for 2λs along with information about the participant's choice, and the DM2's actual (in instances of INVEST) or hypothetical (in instances of KEEP) response. Each trial was separated by an intertrial interval (blank grey-scale screen with fixation cross-hair), jittered from 0 to 12λs. There were a total of 80 trials equally representing the four DM2 types, which were semirandomly presented and distributed evenly across four fMRI runs. After the experimental session was complete, participants were paid according to the actual outcomes during the fMRI trust game.

Behavioral analysis

We analyzed the data pertaining to the choice [frequency (i.e. % of trials) INVEST] the participants made during fMRI scanning to confirm the manipulations of DM2 split frequencies would affect participants' behavior as intended; in contrast, it was expected that over time, participants would accurately associate a DM2 type with the corresponding split frequency and modify their INVEST versus SPLIT choice accordingly. We conducted a repeated-measures analysis of variance using group (patients, controls) as the between-subject factor and the four types of DM2 as the within-subject factor; significant main effects and interactions were clarified by *t*-tests.

Neuroimaging

All scanning was performed with BOLD-sensitive whole-brain fMRI on a 3.0 Tesla GE Signa System (General Electric; Milwaukee, Wisconsin, USA) using a standard radiofrequency coil and associated software (LX 8.3, Neuro-optimized gradients). Whole-brain functional scans were acquired using a T2*-weighted reverse spiral sequence (echo timeλ=λ25λms, repetition

timeλ=λ2000λms, 64 × 64 matrix, flip angleλ=λ77°, field of viewλ=λ24λcm, 30 contiguous 5λmm axial slices per volume, aligned with the anterior commissure posterior commissure line). A high-resolution T1 scan (3D-MPRAGE; repetition timeλ=λ25λms; min echo time; field of viewλ=λ24λcm; slice thicknessλ=λ1.5λmm) was also acquired.

Neuroimaging analysis

Data from all 52 participants met criteria for high quality and scan stability with minimum motion correction (<1λmm displacement in any one direction) and were subsequently included in the data analyses. The first four volumes from each run were discarded to allow for T1 equilibration effects. Preprocessing steps, implemented using Statistical Parametric Mapping software (SPM2; Wellcome Department of Cognitive Neurology, London; www.fil.ion.ucl.ac.uk/spm) were as follows: (i) spatial realignment; (ii) normalization to the Montreal Neurologic Institute (MNI) template through the use of nonlinear warping algorithm; (iii) spatial smoothing through the use of a Gaussian 8λmm full-width-half-maximum kernel. Statistical analyses were performed at the individual and group level using the general linear model and Gaussian random fields theory as implemented in SPM2 for event-related fMRI analyses [19] with covariates representing each DM2 type 'event' (three HUMAN DM2s, one COMPUTER DM2). In the first-level analysis, covariates were convolved with a canonical hemodynamic response function, using the temporal derivative to account for intersubject variability in BOLD signal time to peak. In the second-level analysis, we conducted a random-effects comparison of the linear contrast between HUMAN and COMPUTER trials within and between the two groups (patients, controls). Statistical maps were first created using a threshold of $P_{uncorrected}$ less than 0.005 with a cluster volume of at least 120λmm^3, consistent with earlier studies of mentalizing/impression formation [10,20] and the trust game with fictive partners [7,11]. As this was the first study of this kind, we set a liberal significance threshold to maximize sensitivity for group differences in the whole-brain search; to obviate bias, we report all activations surviving this threshold. These statistical maps were superimposed on normalized high-resolution T1-weighted images and their locations interpreted by using known neuroanatomical landmarks with the brain atlas by Tzourio-Mazoyer and colleagues [21].

Results

Behavioral results

The behavioral results confirmed that the reciprocity manipulation (of DM2's split frequency) as implemented in this modified multiround trust game affected participants' behavior, and did so similarly in both control and social anxiety groups. There was a significant main effect of DM2 partner type on frequency (% of trials) of INVEST decisions showing that participants in both groups chose to INVEST rather than KEEP more frequently when playing against DM2 types who exhibited a higher likelihood of reciprocation [main effect of DM2 type: F(3,150) λ=λ55.71, $P<0.001$]; subsequent t-tests revealed that percentage of INVEST trials had the following pattern (type 1>type 2λ=λCOMPUTER>type 3; all $P<0.05$) (Fig. 2). However, the main effect of group and group × DM2 type interactions were not significant (all $P>0.50$) (Fig. 2).

Neuroimaging results

In control participants, the contrast of HUMAN DM2s (relative to COMPUTER DM2) activated a diverse network including dorsal medial prefrontal cortex, right inferior frontal gyrus, calcarine cortex, inferior and middle occipital gyrus, middle temporal gyrus, putamen (ventral striatum), and postcentral gyrus (Table 2). In social anxiety participants, a similar, though less distributed, set of regions was activated including calcarine cortex, inferior occipital gyrus, fusiform gyrus, and hippocampus (Table 2). Notably, there was an absence of engagement of medial prefrontal cortex in social anxiety patients (Fig. 3). The difference in

medial prefrontal cortex activation to HUMAN (>COMPUTER) partners was confirmed by between-group comparisons (controls>patients) (Fig. 3). To ensure that this group difference was not driven by differential activation to COMPUTER partners between the control and patient groups, we compared activations to the COMPUTER partner (>fixation) across the two groups, and observed no group differences in medial prefrontal cortex.

To further clarify the group differences noted above, we extracted parameter estimates [β weights (units) of the amplitude of the BOLD fMRI response] from each participant for the main contrast (HUMAN vs. COMPUTER) from a 10λmm sphere surrounding the voxel of peak activation in the group difference map within the medial prefrontal cortex [MNI coordinates: (10,22,50)] using the Plot Function in SPM2. Extracted BOLD response from these regions showed that the control group activated medial prefrontal cortex, whereas the patient group exhibited 'deactivations' in this area (mean β±SEM: medial prefrontal cortex, control: 0.128±0.046; patient: −0.084±0.053; tλ=λ3.02; Pλ=λ0.004, two-tailed) (Fig. 3).

Discussion

To begin to test a brain-based model of social-cognitive dysfunction in generalized social anxiety, this study probed the neural correlates of mentalizing during the trust game in patients with social anxiety and healthy controls. The online invest decisions made by participants during this multiround trust game confirmed our intended experimental manipulation involving three different person partner types, with differing motives/intentions as operationalized as their likelihood of reciprocation ('SPLIT'). At the behavioral level, we did not observe group differences in INVEST frequency across the three different HUMAN partner types. At the brain level, the groups did differ in their response to HUMAN partners versus a COMPUTER partner. Unlike healthy controls, individuals with social anxiety failed to activate the medial prefrontal cortex, a key region engaged by mentalizing and impression formation of other people. These findings add to our understanding of current neurobiological models of social anxiety that implicate a social threat-sensitive amygdala and expand the model to include the medial prefrontal cortex in the context of mentalizing and social cognition. More generally, these results show that social neuroeconomic approaches that use tasks derived from game theory can be used to better reflect the dynamic nature of social interactions, and can illuminate the neural basis of social impairments observed in a number of mental illnesses (e.g. autism, schizophrenia, personality disorders) [22,23].

Our finding of activation in medial prefrontal cortex in control participants during interactions with HUMAN (>COMPUTER) partners is consistent with earlier studies that reliably activated medial prefrontal cortex in neuroeconomic exchange games similar to the one used in this study, in which participants interact with the same social partners over time and use information about partners' prior actions to learn about their motives, intentions, and reputations [7,10]. The dorsal regions of the medial wall prefrontal cortex is increasingly understood to be a central node in the brain's mentalizing network with processing in this region uniquely specialized for storing and reasoning about desires, intentions, and (potentially false) beliefs of others [12, 13,17]. Consistent with this view, several fMRI studies that attempt to selectively engage 'pure' mental state attributions without engaging processing of faces (similar to our approach here), eye gaze or bodily motion have specifically found activation in medial prefrontal cortex as opposed to other regions in the brain's mentalizing network [7,24]. Furthermore, many studies find that medial prefrontal cortex is preferentially engaged during conditions that require inferences about stable underlying character traits and behavioral dispositions compared with control conditions in which temporary intentions or goals must be inferred [16,20,25].

Conclusion

This is the first study to use game theory to model the neural activity associated with mentalizing in a simulated social exchange in participants with social anxiety. We found that compared with healthy controls, individuals with social anxiety exhibit impaired responses in medial prefrontal cortex, a region increasingly implicated by social neuroscience as part of the brain's social-cognitive network, and particularly critical to mentalizing and forming impressions about others.
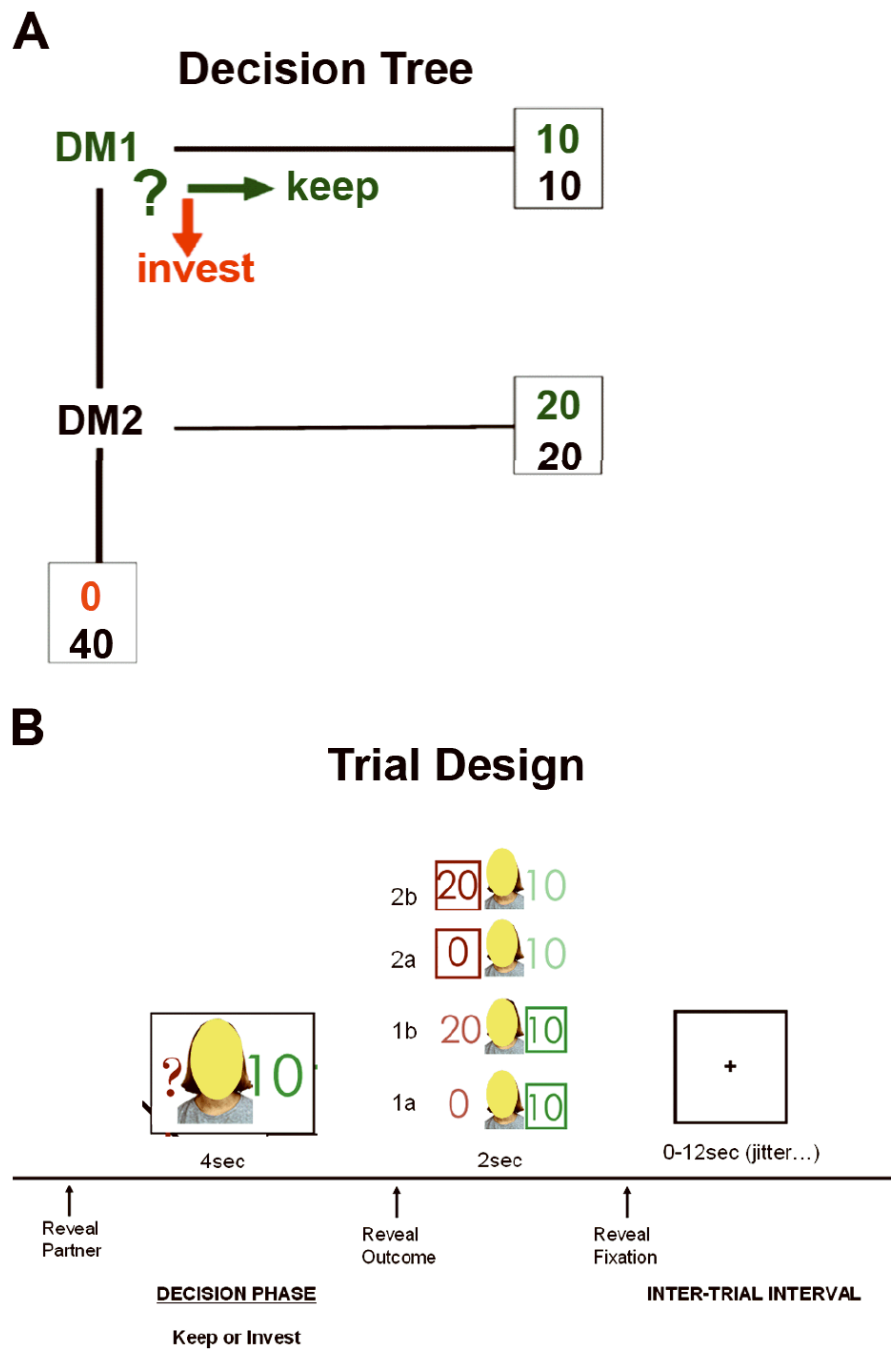
Acknowledgments

References

1. Stein MB, Stein DJ. Social anxiety disorder. Lancet 2008;371:1115–1125. [PubMed: 18374843]

2. Schneier FR. Clinical practice. Social anxiety disorder. N Eng J Med 2006;355:1029–1036.

3. Hirsch CR, Clark DM. Information-processing bias in social phobia. Clin Psychol Rev 2004;24:799–825. [PubMed: 15501557]

4. Etkin A, Wager TD. Functional neuroimaging of anxiety: a meta-analysis of emotional processing in PTSD, social anxiety disorder, and specific phobia. Am J Psychiatry 2007;164:1476–1488. [PubMed: 17898336]

5. Fehr E, Camerer CF. Social neuroeconomics: the neural circuitry of social preferences. Trends Cogn Sci 2007;11:419–427. [PubMed: 17913566]

6. Loewenstein G, Rick S, Cohen JD. Neuroeconomics. Ann Rev Psychol 2008;59:647–672. [PubMed: 17883335]

7. McCabe K, Houser D, Ryan L, Smith V, Trouard T. A functional imaging study of cooperation in two-person reciprocal exchange. Proc Natl Acad Sci U A 2001;98:11832–11835.

8. King-Casas B, Tomlin D, Anen C, Camerer CF, Quartz SR, Montague PR. Getting to know you: reputation and trust in a two-person economic exchange. Science (New York, NY) 2005;308:78–83.

9. Sanfey AG, Rilling JK, Aronson JA, Nystrom LE, Cohen JD. The neural basis of economic decision-making in the Ultimatum Game. Science (New York, NY) 2003;300:1755–1758.

10. Rilling JK, Sanfey AG, Aronson JA, Nystrom LE, Cohen JD. The neural correlates of theory of mind within interpersonal interactions. Neuroimage 2004;22:1694–1703. [PubMed: 15275925]

11. Delgado MR, Frank RH, Phelps EA. Perceptions of moral character modulate the neural systems of reward during the trust game. Nat Neurosci 2005;8:1611–1618. [PubMed: 16222226]

12. Mitchell JP. Mentalizing and Marr: an information processing approach to the study of social cognition. Brain Res 2006;1079:66–75. [PubMed: 16473339]

13. Gallagher HL, Frith CD. Functional imaging of theory of mind. Trends Cogn Sci 2003;7:77–83. [PubMed: 12584026]

14. Frith CD, Frith U. The neural basis of mentalizing. Neuron 2006;50:531–534. [PubMed: 16701204]

15. Amodio DM, Frith CD. Meeting of minds: the medial frontal cortex and social cognition. Nat Rev Neurosci 2006;7:268–277. [PubMed: 16552413]

16. Van Overwalle F. Social cognition and the brain: a meta-analysis. Hum Brain Mapp 2008;M:M–M.

17. Krueger F, Grafman J, McCabe K. Neural correlates of economic game playing. Philos Trans R Soc Lond 2008;363:3859–3874. [PubMed: 18829425]

18. Stopa L, Clark DM. Social phobia and interpretation of social events. Behav Res Ther 2000;38:273–283. [PubMed: 10665160]

19. Friston KJ, Fletcher P, Josephs O, Holmes A, Rugg MD, Turner R. Event-related fMRI: characterizing differential responses. NeuroImage 1998;7:30–40. [PubMed: 9500830]

20. Mitchell JP, Neil Macrae C, Banaji MR. Forming impressions of people versus inanimate objects: social-cognitive processing in the medial prefrontal cortex. Neuroimage 2005;26:251–257. [PubMed: 15862225]

21. Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, et al. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage 2002;15:273–289. [PubMed: 11771995]

22. King-Casas B, Sharp C, Lomax-Bream L, Lohrenz T, Fonagy P, Montague PR. The rupture and repair of cooperation in borderline personality disorder. Science 2008;321:806–810. [PubMed: 18687957]

23. Seres I, Unoka Z, Keri S. The broken trust and cooperation in borderline personality disorder. Neuroreport 2009;20:388–392. [PubMed: 19218873]

24. Gallagher HL, Jack AI, Roepstorff A, Frith CD. Imaging the intentional stance in a competitive game. Neuroimage 2002;16:814–821. [PubMed: 12169265]

25. Mitchell JP, Heatherton TF, Macrae CN. Distinct neural systems subserve person and object knowledge. Proc Natl Acad Sci U S A 2002;99:15238–15243. [PubMed: 12417766]

A

Decision Tree

DM1 ———————— | 10 |
| 10 |

? → keep

↓ invest

DM2 ———————— | 20 |
| 20 |

| 0 |
| 40 |

B

Trial Design

2b 20 😊 10

2a 0 😊 10

1b 20 😊 10

1a 0 😊 10

? 😊 10   +

4sec   2sec   0-12sec (jitter…)

↑ Reveal Partner   ↑ Reveal Outcome   ↑ Reveal Fixation

DECISION PHASE   **INTER-TRIAL INTERVAL**

Keep or Invest

Fig. 1.
Schematic representation of the trust game. (a) Choice options and potential outcomes for Decision Maker 1 (DM1). (b) Exemplar trial structure with probe during decision phase and four potential outcomes based on participant (DM1) choice.

Online Performance
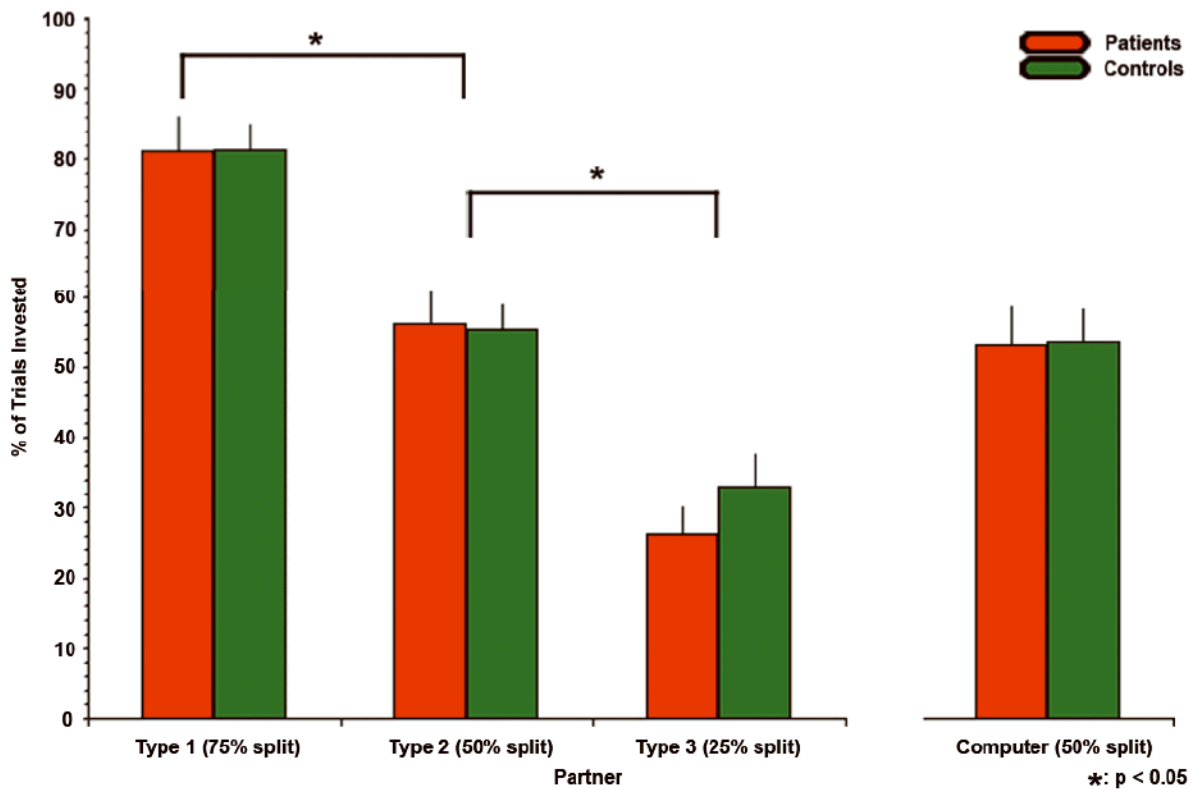
Fig. 2.
Bar graph showing percent of trials (mean±SEM) in which participants made a decision to invest across the four different partner types (three HUMAN partners, one COMPUTER partner) for social anxiety and healthy control groups. *$P<0.05$.
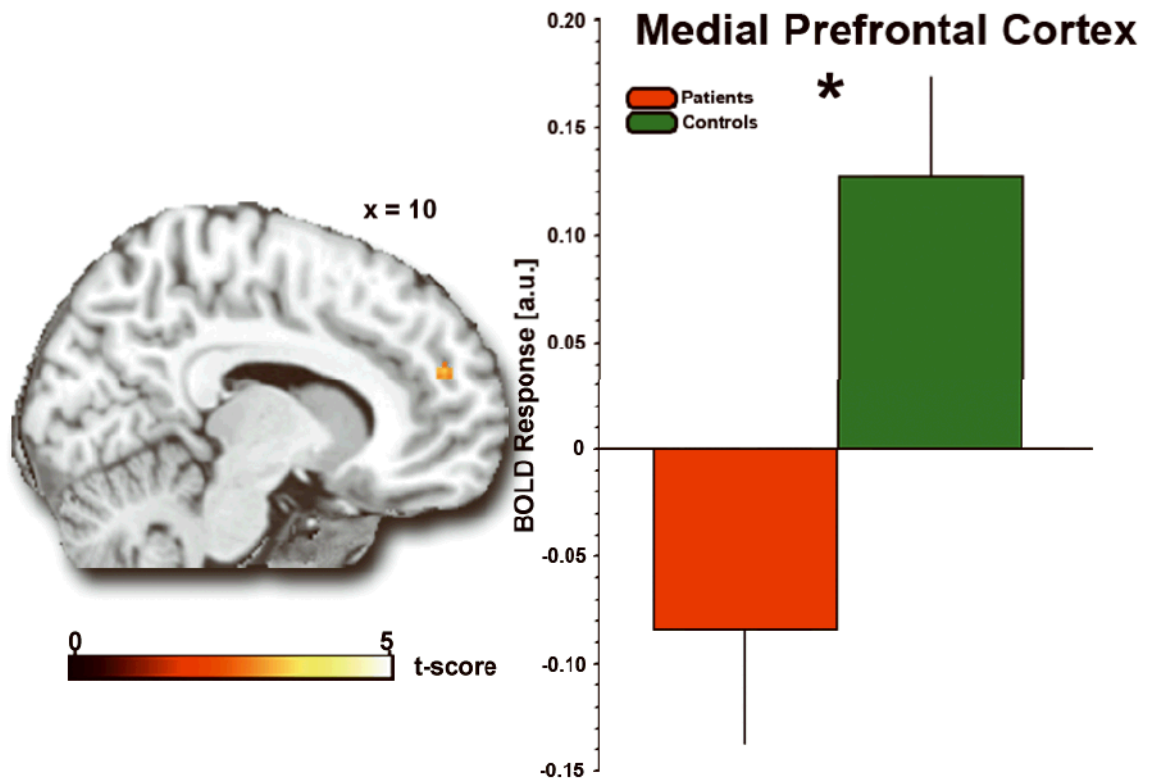
Fig. 3.
Group differences in brain response during mentalizing. Statistical *t*-maps (*P*<0.005) rendered on a canonical brain image and bar graph of extracted BOLD response showing less activation to HUMAN (>COMPUTER) partners in social anxiety patients compared with healthy controls in medial prefrontal cortex. **P*<0.05. a.u., arbitrary units.

Table 1

Demographic and clinical characteristics of participants

	Patients (*n*λ=λ26)	Controls (*n*λ=λ26)	*P* value[a]
λAge (years, mean±SD)	32.7±8.1	28.0±8.2	NS
Sex (*n*, male/female)	13/12	9/16	NS
Education (years, mean±SD)	15.6±1.6	15.7±1.4	NS
Social anxiety severity[b] (mean±SD)	77.3±18.35	15.6±11.9	1.52e-14
Trait anxiety level[c] (mean±SD)	48.8±9.64	29.6±5.3	9.53e-9
Depression severity[d] (mean±SD)	10.3±6.23	1.9±2.6	1.24e-6

NS, nonsignificant.

[a]Between-group comparisons by *t*-test or χ^2, as appropriate.

[b]Severity of social anxiety as measured by the Liebowitz Social Anxiety Scale.

[c]Trait anxiety level as measured by the Spielberger State-Trait Anxiety Inventory.

[d]Depression severity as measured by the Beck Depression Inventory.

Table 2

Activation foci in response to HUMAN (>COMPUTER) partners within each group and differences between groups

Group and brain region	Cluster size (mm³)	MNI coordinates (x,y,z)	Z-score
λControls			
vCalcarine	10λ600	0, −80, 0	5.62
vInferior occipital gyrus	2632	−46, −80, −6	5.47
vInferior occipital gyrus	1464	46, −80, −4	4.90
vInferior frontal gyrus	176	50, 26, 28	3.91
vMiddle temporal gyrus	408	48, −70, 8	3.83
vPutamen	120	−14, 10, −8	3.50
vPostcentral gyrus	160	−64, −2, 20	3.28
vMiddle occipital gyrus	232	−24, −86, 12	3.15
vSuperior medial frontal gyrus	120	8, 52, 24	2.91
Patients			
vCalcarine	8376	0, −82, −2	5.20
vInferior occipital gyrus	2000	48, −76, −8	4.39
vFusiform gyrus	472	42, −48, −24	4.16
vInferior occipital gyrus	984	−46, −78, −4	4.09
vHippocampus	128	−16, −14, −12	2.92
Controls>patients			
vPostcentral gyrus	920	52, −14, 24	3.54
vMiddle occipital gyrus	216	−24, −86, 12	3.43
vSuperior medial frontal gyrus	200	10, 50, 22	3.12
vCuneus	144	−12, −84, 28	3.08
vInferior frontal gyrus	128	48, 6, 16	2.87
vInferior frontal gyrus	144	52, 14, 22	2.78
Patients>controls			
vSupplementary motor area	312	−8, 18, 74	3.56
vSupramarginal gyrus	128	−62, −52, 32	3.41
vMiddle frontal gyrus	120	44, 44, 32	2.91

MNI, Montreal Neurologic Institute.