



Frontoparietal and Default Mode Network Contributions to Self-Referential Processing in Social Anxiety Disorder

Matthew L. Dixon¹ · Craig A. Moodie¹ · Philippe R. Goldin² · Norman Farb³ · Richard G. Heimberg⁴ · Jinxiao Zhang¹ · James J. Gross¹

Accepted: 9 July 2021
© The Psychonomic Society, Inc. 2021

Abstract

Social anxiety disorder (SAD) is characterized by negative self-referential processing, which triggers excessive emotional reactivity. In healthy individuals, positive self-views typically predominate and are supported by regions of the default mode network (DMN) that represent self-related information and regions of the frontoparietal control network (FPCN) that contribute to metacognitive awareness and emotion regulation. The current study used functional magnetic resonance imaging (fMRI) to examine patterns of DMN and FPCN activation during positive and negative self-referential judgments in SAD patients ($N = 97$) and controls ($N = 34$). As expected, SAD patients demonstrated a striking difference in self-beliefs compared with non-anxious healthy controls, endorsing fewer positive traits and more negative traits. However, SAD patients and controls demonstrated largely similar patterns of DMN and FPCN recruitment during self-referential judgements. No significant group differences were observed. However, equivalence testing identified numerous regions demonstrating effect sizes that were not small enough to conclude that they were practically equivalent to zero, despite the nonsignificant null hypothesis test. These regions may be key targets to investigate in future studies using larger samples.

Keywords Emotion · Regulation · Prefrontal · Beliefs · Clinical

Introduction

To effectively navigate a complex social world, the human brain generates models of the self and others based on features that are useful for predicting behavior (Denny et al., 2012; Frith & Frith, 2010; Hassabis et al., 2013). These features include traits, such as “outgoing,” that specify how an individual is likely to behave across different contexts. These traits are generalizations about behavior that allow people to predict how social interactions are likely to unfold and to make

decisions that maximize the chances of experiencing positive states and avoiding potentially harmful states.

Notably, most traits carry a positive or negative valence. A “courageous” person is associated with a greater tendency to do the right thing in the face of threat, whereas a “cowardly” person will likely try to escape the situation. As such, one’s self-attribution of traits has powerful implications for one’s sense of self-esteem. Several lines of evidence suggest that healthy individuals are biased to select traits that promote a positive self-image, for example, updating self-evaluations to a greater extent when receiving positive versus negative feedback about character traits (Koban et al., 2017; Korn et al., 2012), and taking more responsibility for successes than failures (Leary, 2007). In contrast, vulnerability to psychiatric disorders involves predominantly negative self-beliefs. For example, a central component of social anxiety disorder (SAD) is negative self-beliefs related to socially relevant traits (e.g., seeing oneself as boring and awkward) (Clark & Wells, 1995; Heimberg et al., 2014; Moscovitch, 2009). Moreover, SAD patients update self-evaluations to a greater extent than controls when receiving negative feedback (Koban et al., 2017). Cognitive models of SAD suggest that these maladaptive beliefs promote excessive

✉ Matthew L. Dixon
dixonm@stanford.edu

✉ James J. Gross
gross@stanford.edu

¹ Department of Psychology, Stanford University, Stanford, CA, USA

² University of California Davis, Davis, CA, USA

³ University of Toronto, Toronto, ON, Canada

⁴ Temple University, Philadelphia, PA, USA

negative emotional reactivity and interfere with emotion regulation (Clark & Wells, 1995; Heimberg et al., 2014; Jazaieri et al., 2015; Moscovitch, 2009).

Neuroscientific studies have examined the neural basis of self-referential processing using tasks in which participants view and rate trait adjectives in terms of self-descriptiveness. These studies have consistently reported activation in regions of the default mode network (DMN), including the medial prefrontal cortex, pregenual anterior cingulate cortex, posterior cingulate cortex, and temporopolar cortex (D'Argembeau et al., 2005; Davey et al., 2016; Denny et al., 2012; Farb et al., 2007; Northoff et al., 2006; Schmitz & Johnson, 2007). Consistent with a role in representing self-relevant information, the DMN is activated when individuals recall autobiographical memories (Andrews-Hanna et al., 2014), plan future personal activities (Spreng et al., 2010), reflect on personal goals and values (D'Argembeau et al., 2010; D'Argembeau et al., 2012), and engage in self-evaluative and narrative-based self-evaluation (Dixon et al., 2017; Farb et al., 2007). Although the DMN is also engaged when making judgments about other individuals, meta-analyses have shown that several regions in this network, such as the pregenual anterior cingulate cortex, are preferentially engaged during self-related judgments compared to other-related judgments (Denny et al., 2012; Murray et al., 2012; van der Meer et al., 2010).

Critically, alterations in DMN function have been noted as a key factor in a number of disorders that involve disrupted self-referential processing, including Alzheimer's disease (Buckner et al., 2009), depression (Kaiser et al., 2015; Sheline et al., 2009), anxiety (Sylvester et al., 2012), and schizophrenia (Whitfield-Gabrieli et al., 2009). SAD is associated with altered DMN activation across a range of tasks, including processing facial expressions, receiving criticism, and attempting to regulate emotions (Blair et al., 2008; Bruhl et al., 2014; Dixon et al., 2020). However, only a few studies have examined neural activity in SAD during explicit reflection on personality traits (Blair et al., 2011; Goldin et al., 2012; Goldin, Ramel, & Gross, 2009a). One study found that healthy controls demonstrated greater MPFC activation when reflecting on their own traits from a first person perspective (e.g., "I'm ugly") compared with reflecting on their own traits as if from a second person perspective (e.g., "you are ugly"), whereas SAD patients showed the opposite pattern (Blair et al., 2011). This finding is consistent with the notion that the self-referential neural circuitry in SAD is hypersensitive to the opinions of other people (Blair et al., 2011). The conclusions from this study are somewhat limited, however, by the small sample size ($N = 15$ patients). Additionally, it is unclear whether potential neural dysfunction in SAD during self-reflection is dependent on the valence of the traits being reflected upon. This is a key issue given that negative self-views and positive self-views are separable components of

SAD and show distinct relationships to symptom reduction following cognitive behavioral therapy and mindfulness-based stress reduction (Goldin, Jazaieri, et al., 2013a; Thurston et al., 2017).

Another large-scale network that influences self-referential processing and emotional reactivity is the frontoparietal control network (FPCN). The FPCN includes parts of the lateral prefrontal cortex, pre-supplementary motor area, intraparietal sulcus, and posterior middle temporal gyrus (Cole et al., 2014; Dixon et al., 2018; Dosenbach et al., 2007; Spreng et al., 2013; Vincent et al., 2008; Yeo et al., 2011). The FPCN is believed to support adaptive functioning through its roles in metacognitive awareness and top-down modulation of cognitive and affective processes (Buhle et al., 2014; Bunge, 2004; Cole et al., 2013; Dosenbach et al., 2006; Fleming & Dolan, 2012; Goldin, Ziv, et al., 2013b; Kaiser et al., 2015; Ochsner & Gross, 2005; Spreng et al., 2010). Although the FPCN has received less attention than the DMN in the context of self-referential processing, it is consistently engaged during tasks that involve self-reflection (Farb et al., 2007; Schmitz et al., 2004), metacognitive judgments (i.e., reflecting on one's knowledge) (Baird et al., 2013; Fleming & Dolan, 2012; McCaig et al., 2011), and autobiographical planning (Spreng et al., 2010). Moreover, growing evidence suggests that the FPCN interacts with the DMN to monitor and regulate self-related thinking in line with goals and task demands (Dixon et al., 2018; Dobbins et al., 2002; Smallwood et al., 2012; Spreng et al., 2010).

Although prior studies have examined FPCN activation in SAD in the context of emotion regulation tasks (e.g., Goldin, Manber, et al., 2009b), it has not been examined in the context of self-referential processing. A key question is whether individuals with SAD recruit the FPCN differently during self-reflection given their tendency towards negative self-referential thinking and maladaptive patterns of emotional reactivity and regulation (Jazaieri et al., 2015). One possibility is that SAD patients will show less FPCN engagement during self-referential processing, reflecting a loss of regulatory capacity over self-related thought content. Alternatively, SAD patients may show excessive FPCN engagement reflecting the effort to regulate emotional responses that may be elicited during self-referential judgments.

The goal of the present study was to examine whether SAD patients exhibit altered activity in DMN and FPCN regions during self-referential processing. We scanned SAD patients and healthy controls with functional magnetic resonance imaging (fMRI) as they viewed social trait words and made self-referential judgments (self-descriptive: yes/no) or case judgments (upper-case/lower-case). There were separate blocks of positively and negatively valenced words. We used whole-brain analyses and *a priori* defined regions of interest (ROIs) to examine patterns of activation in DMN and FPCN regions. Our first hypothesis was that we would observe DMN

and FPCN activation during self-referential judgments in both groups, but that SAD patients would show aberrant activation in select regions within these networks. Given that few studies have examined self-referential processing in SAD, we thought an open-ended rather than directional hypothesis was more appropriate. Our second hypothesis was that group differences would depend on trait valence, with different regions differentiating controls from SAD patients during reflection on positive trait words and reflection on negative trait words. Our third hypothesis was that activation patterns within the DMN and FPCN during self-referential processing would correlate with social anxiety severity in the SAD group.

Methods

Participants

Participants who met criteria for inclusion in the fMRI study (Table S1) included 114 SAD patients ($\text{mean}_{\text{age}} = 32.9$, $\text{SD}_{\text{age}} = 7.92$; $n = 61$ females) and 36 healthy controls ($\text{mean}_{\text{age}} = 32.1$, $\text{SD}_{\text{age}} = 8.70$; $n = 22$ females) who provided informed consent in accordance with the Institutional Review Board at Stanford University, passed MRI safety screening, were 22–55 years of age, were fluent in English, and right-handed as assessed by the Edinburgh Handedness Inventory (Oldfield, 1971). Controls had no history of psychiatric disorders. SAD patients met criteria for a primary diagnosis of generalized SAD based on the Anxiety Disorders Interview Schedule for the DSM-IV-Lifetime version (ADIS-IV-L) (Di Nardo et al., 1994). Clinical interviews were conducted by doctoral clinical psychologists and doctoral students in clinical psychology trained on the ADIS-IV-L. Patients met criteria for the generalized subtype of SAD if they endorsed greater than moderate social fear in five or more distinct social situations assessed by the ADIS-IV-L. Patients also had a score greater than 60 on the Liebowitz Social Anxiety Scale Self-Report (LSAS-SR) (Fresco et al., 2001; Liebowitz, 1987), the cutoff score for the generalized subtype of SAD as determined by receiver operator characteristics analysis (Rytwinski et al., 2009). Participants were excluded for comorbid diagnoses of current major depressive disorder, posttraumatic stress disorder, or obsessive-compulsive disorder; pharmacotherapy or psychotherapy during the past year; any history of neurological disorders or head trauma; cardiovascular disorders, thought disorders, or bipolar disorder; current substance/alcohol abuse or dependence.

Self-Referential Encoding Task

Participants performed a self-referential encoding task (Fig. 1) (Derry & Kuiper, 1981) programmed in EPrime software. Stimuli were 25 positive and 25 negative social trait adjectives

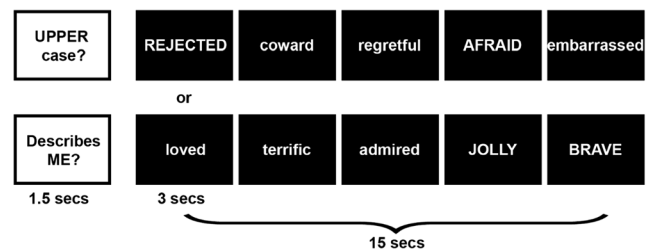


Fig. 1 Trial structure of the self-referential encoding task. Top panel illustrates a case judgment block with a set of negatively-valenced words. Bottom panel illustrates a self-referential judgment block with positively valenced words. There were 5 blocks of each trial type

(see Supplementary Information Appendix) from the Affective Norms for English Words database (Bradley & Lang, 1999) and were balanced (all $ps > 0.51$) on word frequency (positive adjectives = 40.5, negative adjectives = 33.6), number of letters (positive adjectives = 6.9, negative adjectives = 7.2), arousal (positive adjectives = 5.54, negative adjectives = 5.43; on a scale of 1 = low to 9 = high), and valence (deviation from neutral: positive adjectives = 2.66, negative adjectives = 2.58; on a scale of 1 = most negative, 5 = neutral, 9 = most positive) based on the nine-point Self-Assessment Manikin rating system (Lang, 1980). Participants viewed the trait words and made a yes/no judgment indicating whether the trait was self-descriptive (“self” judgment condition), or a yes/no judgment indicating whether the trait was written in all upper-case letters (“case” judgement condition). There were four trial types (self vs. case judgment \times positive vs. negative valence). Each adjective was presented twice: once in the self-referential judgment condition, and once in the case judgment condition. The self-referential condition required participants to engage in self-focused attention and access their self-schema in relation to social-evaluative processing. The case identification condition required attention to the concrete sensory features of the words and was used as a comparison condition to control for reading negative and positive adjectives.

There were five blocks of each of the four trial types. Each block started with a question screen (either “Describes ME?” or “UPPER case?”) for 1.5 seconds, and then five positive or five negative adjectives were presented one at a time for 3 seconds each. Participants made a yes/no response using a button pad during presentation of each of the five stimuli. Thus, in total, participants made 25 judgments for each of the four trial types. At the end of the run there was a 3-second fixation cross and a 3-second blank screen. Stimulus order was pseudo-randomized in terms of block sequence, with no more than two blocks of the same condition in a row. The sequence of words as well as whether they were upper or lower case were randomized within each block. All trials were presented within a single functional run lasting 5.6 minutes. The short length of the task was justified by the simple task design—unlike many cognitive tasks that involve

multicomponent trials (e.g., fixation, cue, delay, target, feedback), our task had a single component (make a self-referential judgment or a letter case judgment on each word). Moreover, the simplicity of making self-referential judgments was expected to yield consistent activation across participants within each group, even with data from just a single run. It is worth noting that this task was part of a larger battery that also included two emotion regulation tasks. Task order was fixed, with the self-referential encoding task always being the third task.

Behavioral Data Analysis

For each participant, we computed the percentage of positive trait words that were endorsed (i.e., assigned a yes response), and the percentage of negative trait words that were endorsed. We also computed median reaction times (RT) for positive and negative trait word trials. Trials in which participants responded in under 250 ms were excluded from the analysis. We also computed accuracy on the case judgment trials. Only participants who scored 70% correct or higher were included in further behavioral and fMRI analyses to ensure that the results were based on participants who understood the task and were paying attention. This criterion led to the exclusion of 2 control participants (5.6% of sample) and 17 patients (14.9% of sample), resulting in final samples of $N = 34$ and $N = 97$, respectively. The proportion of excluded subjects did not differ by group, $\chi^2(1, 150) = 2.17, p = 0.14$.

MRI Data Acquisition

fMRI data were collected using a General Electric 3T Signa magnet with a T2*-weighted gradient-echo spiral-in/out pulse sequence (Glover & Law, 2001). Head movement was restricted using a bite-bar and foam padding. Twenty-four ascending interleaved axial slices were acquired (4.5-mm slice thickness; single shot; repetition time (TR) = 1.5 s; echo time (TE) = 28.5 ms; flip angle (FA) = 65°; field of view (FOV) = 220 mm; matrix size = 64 × 64, voxel resolution 3.438 mm² × 4.5 mm). Each participant completed one functional run during which 230 functional volumes were acquired. Data collected during the first 4 TRs were discarded to allow for equilibration effects. Before functional imaging, a high resolution T1-weighted structural image was acquired using fast spin-echo spoiled gradient recall (132 slices; TR = 3 s; TE = 68 ms; FOV = 220 mm; matrix size: 256 × 256; voxel size = 1 × 1 × 1.2 mm).

fMRI Preprocessing

Using Statistical Parametric Mapping (SPM 12), the data were reoriented to approximate Montreal Neurological Institute (MNI) space, corrected for motion via realignment to the first

volume (using a 6-parameter rigid body transformation), and slice-time corrected (to the middle slice). Each subject's T1 image was bias-corrected and segmented using a nonlinear deformation field to map it onto template (ICBM) tissue probability maps for gray/white matter and CSF. Parameters obtained from this step were subsequently applied to the functional data (resampled to 3 mm³ voxels) during normalization to MNI space. The data were spatially smoothed using an 8-mm³ full-width at half-maximum Gaussian kernel to reduce the impact of inter-subject variability in brain anatomy.

First-Level Analysis

Brain activation was assessed in each participant with linear multiple regression using the following regressors: (i) instruction (question) cue; (ii) negative self judgment; (iii) positive self judgment; (iv) negative case judgment; (v) positive case judgment. Each regressor was convolved with a canonical hemodynamic response function. The model included regressors of no interest to account for subject motion: six parameters from realignment and framewise displacement time course (computed based on Power et al., 2012). Framewise displacement showed a low correlation with the six motion parameters (mean $r = 0.18$) and was, therefore, included in the model to account for residual noise. A high-pass filter (128-s cutoff) was included to remove low-frequency drifts.

Second-Level (Group) Analysis

Single-subject contrast images (e.g., positive self > case) were entered into second-level random effects analyses to assess group-level significance using independent samples *t*-tests (controls vs. SAD patients). Controlling for multiple comparisons ($p < 0.05$ family-wise-error corrected) in whole-brain analyses was achieved through threshold-free cluster enhancement (TFCE) (Smith & Nichols, 2009) (see *Supplementary Methods* for details). TFCE controls familywise error rate based on a permutation testing approach and determines optimal voxel-wise cluster-forming thresholds using an automated algorithmic method. It therefore eliminates the need for the experimenter to arbitrarily choose thresholds to correct for multiple comparisons.

Relationship between Brain Activation and Social Anxiety Severity

To explore potential heterogeneity within the SAD group, we performed a second-level regression analysis to examine whether the strength of activation during self-referential processing was associated with individual differences in the severity of social anxiety within the SAD group. This group-level regression analysis used LSAS-SR total severity scores as a predictor variable and voxel-wise responses from the

positive self > case or negative self > case contrast as the outcome variable.

As a complementary approach to the whole-brain analysis, we also examined correlations between social anxiety severity and mean activation (beta values) in three functionally defined ROIs representing classic self-referential processing regions from a prior meta-analysis (Northoff et al., 2006) (see *Supplementary Methods*). Given that these regions have been consistently associated with self-referential processing, they were key targets to evaluate individual differences in brain activation in the SAD group.

Equivalence Testing

Null results by themselves do not provide evidence in favor of zero effect, because it could be the case that the analyses were underpowered for detecting meaningful effect sizes. A complementary approach is to use equivalence tests, which determine whether observed effect sizes are small enough to be considered practically equivalent to zero (Lakens, 2017). To be considered as practically equivalent to zero, the 90% confidence intervals (CIs) around observed effect sizes should not overlap with equivalence bounds, usually defined as the smallest effect size of interest (SESOI). In other words, observed effect sizes should be significantly different from the smallest effect size of interest (Lakens, 2017). In our case, the minimum effect size we had power to detect with our sample size was a medium effect size of ~ 0.5 , so we used this as the smallest effect size of interest (i.e., lower and upper bounds were Hedge's $g = -0.5$ and 0.5 , respectively). Although equivalence testing has yet to become widespread in fMRI research, recent work suggests that it provides important complementary information to standard null hypothesis testing (Vidal-Ribas et al., 2021).

We conducted equivalence tests on group comparisons (i.e., we looked for equivalence of mean differences). If effects are not statistically different from zero, but the 90% CI includes one of the equivalence bounds, the effect falls within an undecidability range (i.e., it cannot be presumed that the effect is small enough to be considered equivalent to zero for practical purposes) (Lakens, 2017). To compute effect sizes and CIs on the fMRI data, we used analysis code made available by Gerchen et al. (2021) at https://github.com/Fungisai/g_ci_spm.

Exploratory Analyses

As a complementary *hypothesis generating* procedure, we also performed exploratory analyses using a liberal threshold of $p < 0.005$, with a 5-voxel cluster extent (Lieberman & Cunningham, 2009). For the liberal exploratory analysis, we restricted our search volume to a single region of interest (ROI) consisting of the DMN and FPCN. The DMN was

derived from the brain network parcellation of Yeo et al. (2011) and the FPCN mask was custom made (see *Supplementary Methods* for details).

Results

Preliminary Results

A 2 group (SAD, controls) \times 2 valence (positive, negative) ANOVA on trait endorsement responses revealed a significant interaction ($F_{1,129} = 145.58, p < 0.001, \eta_p^2 = 0.53$) (Fig. 2). Independent samples t -tests revealed significant differences between groups in the percentage of positive traits endorsed ($t_{129} = 10.74, p < 0.001, \eta_p^2 = 0.47$) and the percentage of negative traits endorsed ($t_{129} = 10.80, p < 0.001, \eta_p^2 = 0.48$). SAD patients were far less likely to endorse positive traits (41.07% vs. 85.76%) and far more likely to endorse negative traits (54.43% vs. 6.82%), consistent with a robust negativity bias. A 2 (group) \times 2 (valence) ANOVA on reaction time (RT) revealed a main effect of group ($F_{1,129} = 43.38, p < 0.001, \eta_p^2 = 0.25$), but no interaction ($F < 1$). Patients responded slower when making judgments on positive words (1406.93 ms vs. 1100.10 ms; $t_{129} = 6.20, p < 0.001, \eta_p^2 = 0.23$) and negative words (1405.38 ms vs. 1115.44 ms; $t_{129} = 5.85, p < 0.001, \eta_p^2 = 0.21$).

Positive Self Judgment Versus Case Judgment

In a whole-brain analysis corrected for multiple comparisons ($p < 0.05$ family-wise error corrected), both groups demonstrated significant activation for the positive self > case contrast in DMN regions (medial prefrontal cortex, posterior cingulate cortex, posterior inferior parietal lobule, superior frontal sulcus, lateral temporal cortex, temporopolar cortex) and FPCN regions (inferior frontal sulcus, inferior frontal junction, middle frontal gyrus, pre-supplementary motor area, and anterior mid-cingulate cortex). Both groups also activated the anterior insula, striatum, thalamus, and visual regions (Fig. 3).

There were no regions showing significant activation in the positive case judgment > self judgment contrast. The whole-brain analysis revealed no significant group differences in activation at the corrected statistical threshold.

Negative Self Judgment Versus Case Judgment

In a whole-brain analysis corrected for multiple comparisons ($p < 0.05$ family-wise error corrected), both groups demonstrated significant activation for the negative self > case contrast in DMN regions (left inferior frontal gyrus, left lateral temporal cortex, left temporopolar cortex, left retrosplenial cortex), as well as visual regions and left anterior insula (Fig. 4). For both groups, the negative case judgment > self

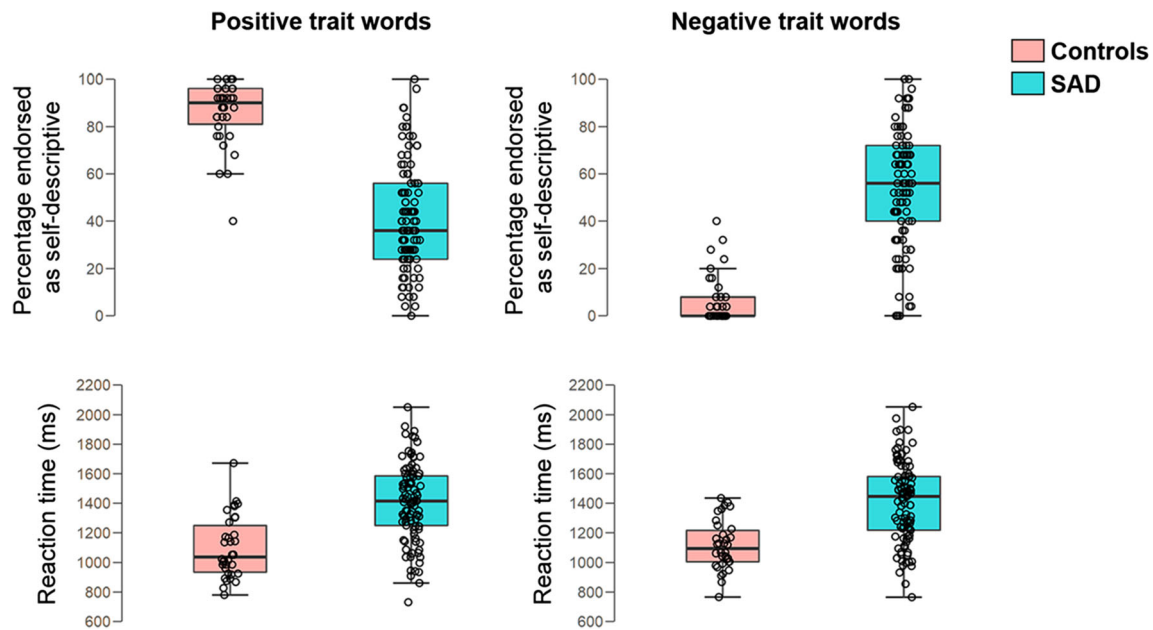


Fig. 2 Ratings of trait words. Top panel: percentage of positive and negative trait words endorsed as self-descriptive by each participant. Bottom panel: median reaction time in milliseconds (ms) during trait judgments. The plots include individual data points with the median

(black line), the interquartile range (shaded area), and $1.5\times$ interquartile range (whiskers). The figure was created with JASP computer software Version 0.9 (JASP Team, 2018)

judgment contrast revealed significant activation in dorsal attention network regions (frontal eye fields, anterior intraparietal sulcus, superior parietal lobule), somatomotor

cortex, dorsal posterior cingulate cortex/precuneus, posterior superior temporal sulcus, right posterior inferior parietal lobule, and superior frontal sulcus (Fig. 4). There were no group

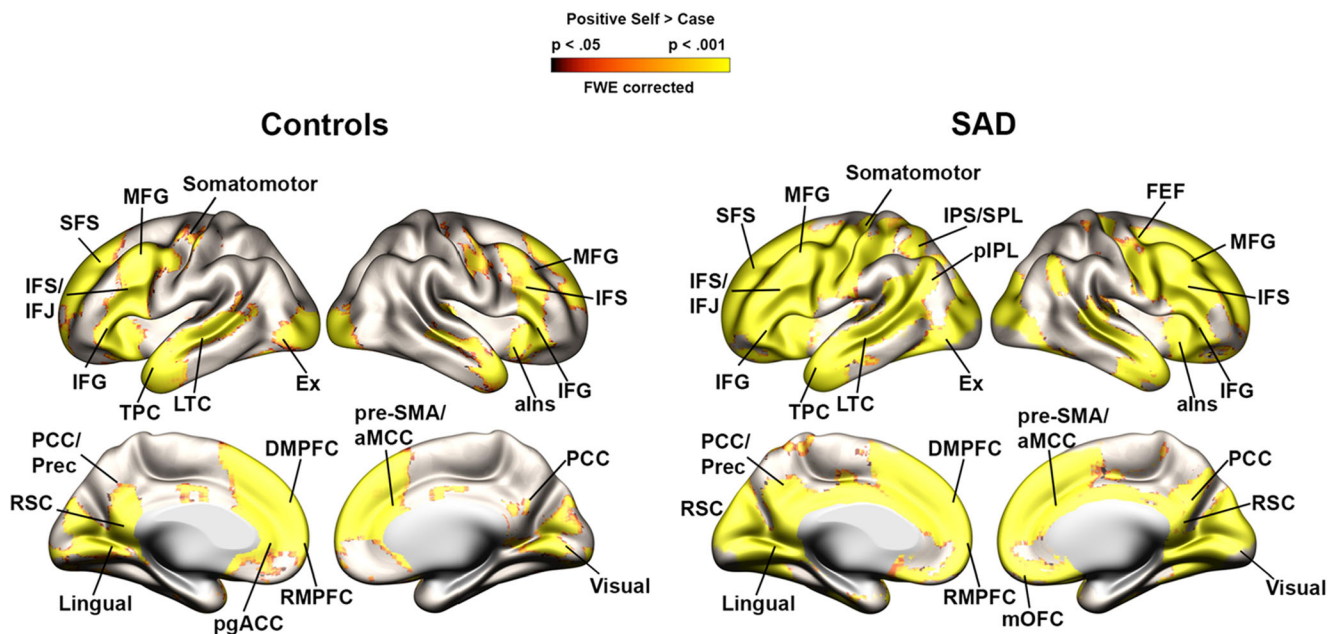


Fig. 3 Surface rendering of the positive self judgment versus case judgment contrast map. Both groups demonstrated robust activation of default mode network regions and frontoparietal regulation regions for the positive self > case judgement contrast. Maps are $p < 0.05$ family-wise error corrected for the whole-brain volume. alns, anterior insula; DMPFC, dorsomedial prefrontal cortex; Ex, extrastriate cortex; FEF, frontal eye fields; IFG, inferior frontal gyrus; IFS/IFJ, inferior

frontal sulcus/inferior frontal junction; IPS/SPL, intraparietal sulcus/superior parietal lobule; LTC, lateral temporal cortex; MFG, middle frontal gyrus; PCC/Prec, posterior cingulate cortex/precuneus; pIPL, posterior inferior parietal lobule; pgACC, pregenual anterior cingulate cortex; pre-SMA/aMCC, pre-supplementary motor area/anterior mid cingulate cortex; RMPFC, rostromedial prefrontal cortex; RSC, retrosplenial cortex; SFS, superior frontal sulcus; TPC, temporopolar cortex

differences observed in the whole-brain analysis at the corrected statistical threshold.

Positive Versus Negative Self Judgments

In a whole-brain analysis corrected for multiple comparisons ($p < 0.05$ family-wise error corrected), both groups demonstrated significant activation for the positive > negative self contrast in DMN regions (rostral and dorsal medial prefrontal cortex, pregenual anterior cingulate cortex, superior frontal gyrus, posterior inferior parietal lobule), FPCN regions (inferior frontal gyrus, posterior middle frontal gyrus, pre-supplementary motor cortex), as well as the striatum, mediodorsal thalamus, hippocampus, somatomotor cortex, cerebellum, and visual cortex (Fig. 5). There were no significant activations for the negative > positive self contrast and no significant group differences at the corrected threshold.

Relationship between Social Anxiety Severity and Brain Activation

To test hypothesis 3, we examined whether the strength of activation during self-referential processing was associated with individual differences in the severity of social anxiety within the SAD group, as measured with the LSAS-SR. There were no significant associations detected using a whole-brain analysis corrected for multiple comparisons.

As a complementary approach, we also examined the relationship between social anxiety severity and ROIs derived from a highly cited meta-analysis of self-referential processing (Northoff et al., 2006). These ROIs were 10-mm spheres centered on peak coordinates for the medial prefrontal cortex, posterior cingulate cortex, and pre-supplementary motor area. Greater social anxiety severity was associated with less mean activation in the pre-supplementary motor area ROI during positive self-judgments ($r = -0.20$, $p = 0.055$), but this association was not statistically significant when correcting for multiple comparisons (corrected $p = 0.33$). There was no relationship between symptom severity and any other ROI during positive or negative self-judgements (all $r < 0.15$, corrected $p > 0.78$).

Equivalence Tests

Although we did not observe significant group differences using null hypothesis testing, this does not by itself provide support for the absence of an effect. To provide complementary information, we performed equivalence tests. The results demonstrated that for numerous regions—including parts of the DMN, FPCN, striatum, amygdala, and orbitofrontal cortex—the confidence intervals around observed effect sizes for the group comparison overlapped with a minimum effect size of interest (in this case, a medium effect size of Hedge's $g = 0.5$) (Fig. 6). Thus, the effect size in these regions *cannot* be

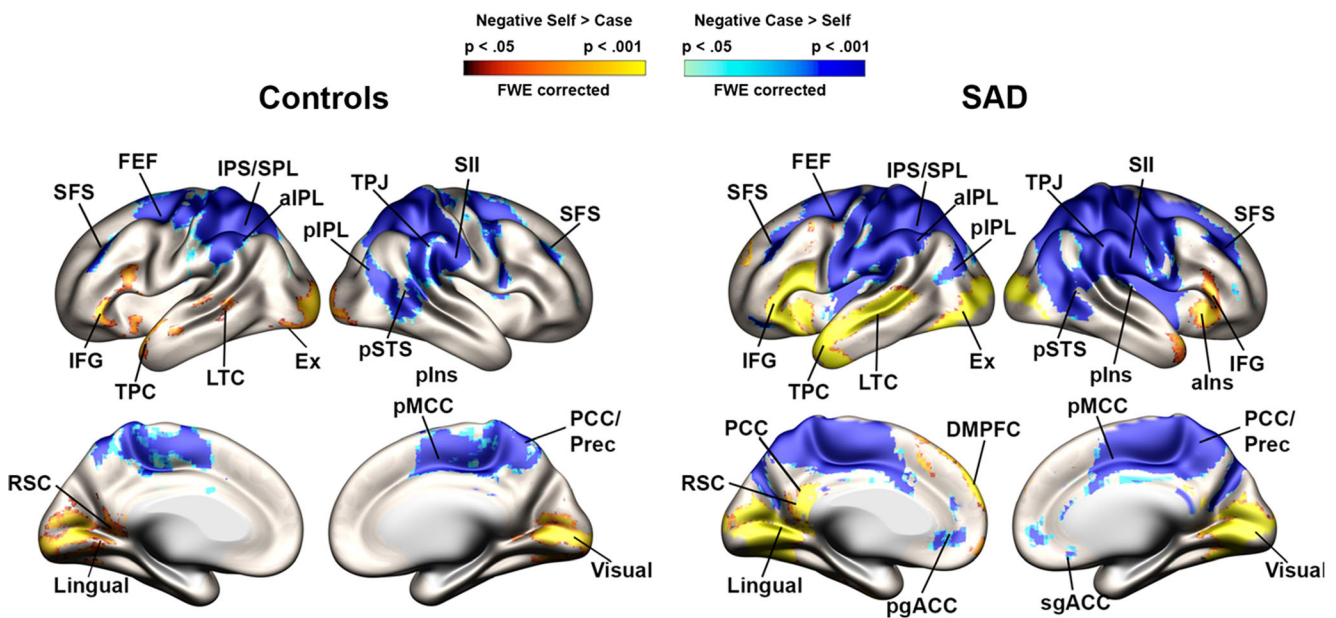


Fig. 4 Surface rendering of the negative self judgment versus case judgment contrast map. Maps are $p < 0.05$ family-wise error corrected for the whole-brain volume. aIPL, anterior inferior parietal lobule; alns, anterior insula; DMPFC, dorsomedial prefrontal cortex; Ex, extrastriate cortex; FEF, frontal eye fields; IFG, inferior frontal gyrus; IFS/IFJ, inferior frontal sulcus/inferior frontal junction; IPS/SPL, intraparietal sulcus/superior parietal lobule; LTC, lateral temporal cortex; PCC/Prec,

posterior cingulate cortex/precuneus; plns, posterior insula; pIPL, posterior inferior parietal lobule; pgACC, pregenual anterior cingulate cortex; pre-SMA/aMCC, pre-supplementary motor area/anterior mid cingulate cortex; pSTS, posterior superior temporal sulcus; RSC, retrosplenial cortex; SII, secondary somatosensory cortex; SFS, superior frontal sulcus; sgACC, subgenual anterior cingulate cortex; TPC, temporopolar cortex; TPJ, temporoparietal junction

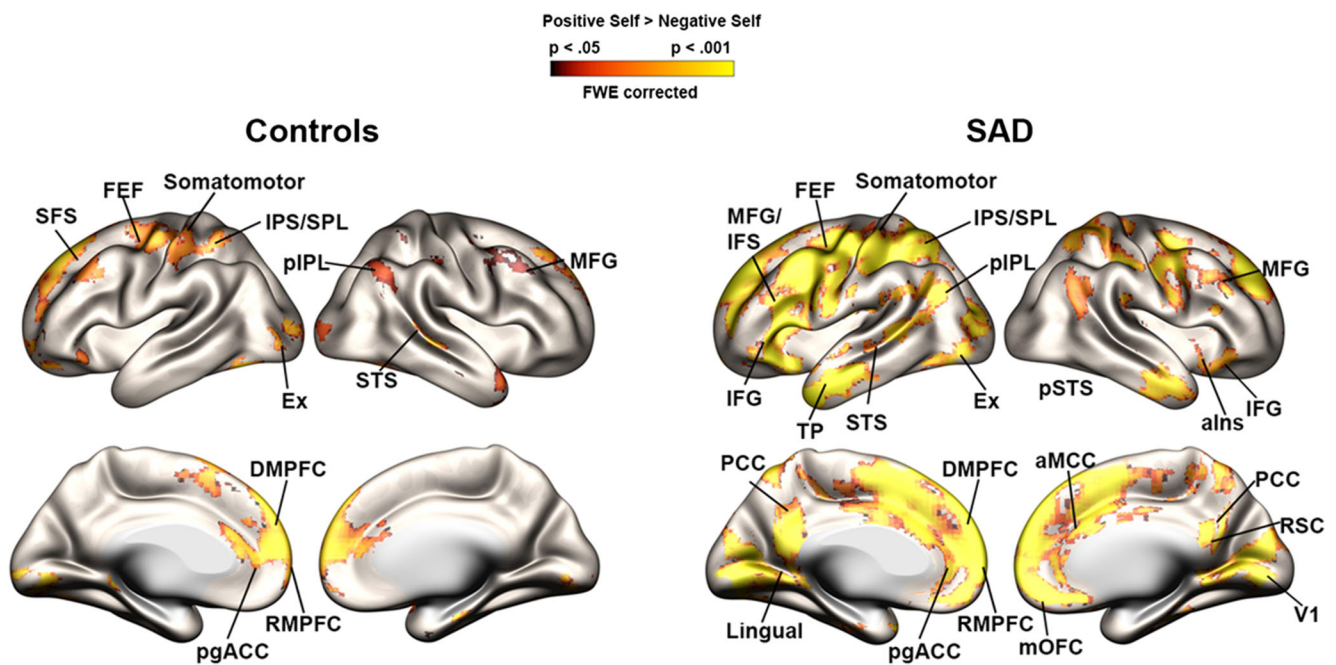


Fig. 5 Surface rendering of the positive self judgment versus negative self judgment contrast map. Maps are $p < 0.05$ family-wise error corrected for the whole-brain volume. aIns, anterior insula; aMCC, anterior mid cingulate cortex; DMPFC, dorsomedial prefrontal cortex; Ex, extrastriate cortex; FEF, frontal eye fields; IFG, inferior frontal gyrus;

IFS, inferior frontal sulcus; IPS/SPL, intraparietal sulcus/superior parietal lobe; MFG, middle frontal gyrus; PCC, posterior cingulate cortex; pIPL, posterior inferior parietal lobule; pgACC, pregenual anterior cingulate cortex; RSC, retrosplenial cortex; SFS, superior frontal sulcus; TP, temporopolar cortex

interpreted as practically equivalent to zero. Therefore, no conclusions can be drawn about group differences in these regions. However, they may serve as regions of interest to explore in future studies that are better powered.

Exploratory Analyses

Finally, we used an exploratory approach with a liberal statistical threshold ($p < 0.005$ uncorrected, voxel extent ≥ 5) to identify regions that may be targets to explore in future work. Notably, the regions demonstrating a group difference using a more liberal threshold generally overlapped with the regions of undecidability identified via equivalence testing.

For the positive self > case contrast, controls relative to patients exhibited greater activation of the pregenual anterior cingulate cortex within the DMN (Figure S1; Table S2). Conversely, patients relative to controls demonstrated greater activation in several regions within the boundaries of the FPCN and DMN, including the inferior frontal sulcus, posterior middle frontal gyrus/premotor cortex, and intraparietal sulcus (Figure S1; Table S2).

For the negative self > case contrast, controls relative to SAD patients demonstrated greater activation in the pregenual anterior cingulate cortex and left anterior prefrontal cortex within the DMN (Figure S2; Table S3). The pregenual anterior cingulate cortex cluster overlapped with cluster showing a group difference in the positive self > case judgment contrast (Figure S3). Patients relative to controls demonstrated greater

activation of the left lateral temporal cortex and left inferior frontal gyrus within the DMN ROI for the negative self > case judgment contrast (Figure S2; Table S3).

For the positive > negative self contrast, controls relative to SAD patients exhibited greater activation in the lateral temporal cortex.

Finally, we explored associations between social anxiety severity in the SAD group and brain activation. Several regions within the DMN and FPCN ROIs correlated with with social anxiety severity (Figure S4; Table S4).

Discussion

The current study examined the neural basis of self-referential processing in SAD. As expected, we found a striking group difference in the self-beliefs that were endorsed, with the SAD patients endorsing significantly more negative traits and significantly fewer positive traits. However, this difference in self-referential processing was not mirrored in brain activation patterns. Both patients and controls demonstrated similar patterns of brain activation, including recruitment of canonical DMN and FPCN self-referential regions. Whole-brain analyses corrected for multiple comparisons did not identify significant group differences. However, equivalence tests indicated that for many regions, the effect size of the group contrast was not small enough to be considered practically equivalent to zero. In other words, the current data do not provide evidence

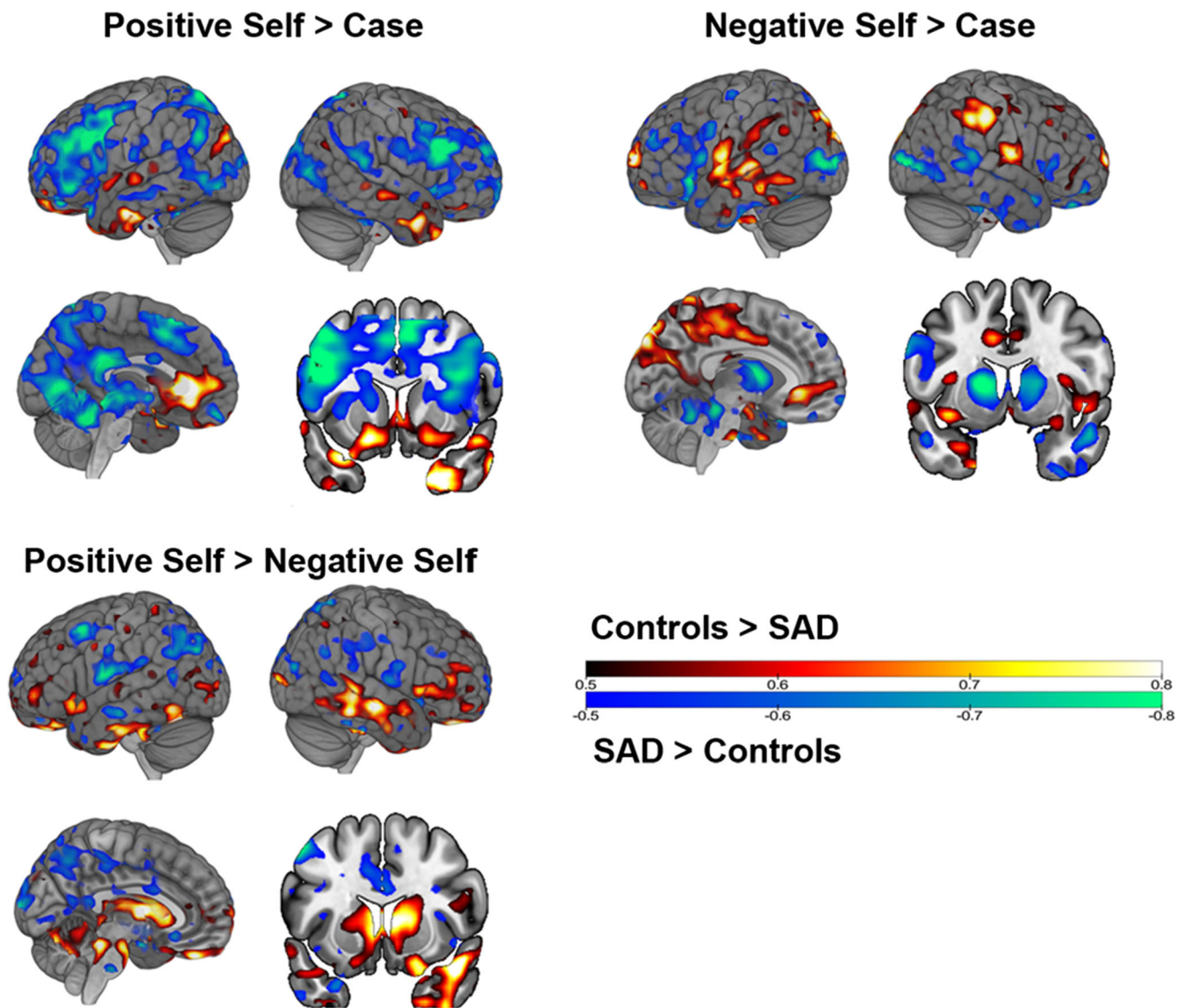


Fig. 6 Equivalence testing. For each contrast, we illustrate the values of the 90% confidence interval around the observed effect size at each voxel for the group comparison. Because our minimum effect size of interest was Hedge's $g = \pm 0.5$, we illustrate voxels with confidence intervals that are equal to or exceed this value. These illustrated voxels indicate regions of undecidability in which we cannot conclude that the effect size for the

group contrast was small enough to be practically equivalent to zero. Although we did not observe significant group differences using null hypothesis testing, the fairly widespread colored voxels in this figure suggest that our study was potentially underpowered for detecting potentially meaningful group differences.

for an absence of a group difference in brain activation during self-referential processing. It could be the case that a study with a larger sample size could detect significant differences. The results of this study may therefore point to regions of interest to explore in future work and highlight the necessity of very large sample sizes to have sufficient power to detect differences between controls and SAD patients in brain activation during self-referential processing.

Self-referential processing engages the DMN and FPCN in SAD patients and controls

Consistent with prior studies (D'Argembeau et al., 2005; Denny et al., 2012; Farb et al., 2007; Murray et al., 2012;

Northoff et al., 2006; Schmitz et al., 2004; Schmitz & Johnson, 2007; van der Meer et al., 2010), we found that both groups demonstrated robust engagement of DMN regions including the medial prefrontal cortex, posterior cingulate cortex, posterior inferior parietal lobule, and temporopolar cortex when making self-referential judgments relative to non-self-referential (case) judgments. This collection of regions may represent the content and valence of one's self-schema by integrating information from episodic and semantic memory with affective signals from limbic regions (Andrews-Hanna et al., 2014; D'Argembeau, 2013; Dixon et al., 2017). It is possible that the content and valence of one's self-schema

may be reflected in multivariate activity patterns spanning multiple regions in the DMN and beyond—patterns that would not be revealed using standard univariate analyses, as was done here.

Although the FPCN has received less attention than the DMN in self-referential processing, we found that both groups demonstrated extensive activation of FPCN regions (e.g., inferior frontal sulcus, middle frontal gyrus, pre-SMA) during self versus case judgments. This is consistent with prior reports of FPCN activation during self-referential processing (Farb et al., 2007; Schmitz et al., 2004). Mounting evidence indicates that the FPCN plays a critical role in metacognitive awareness—the ability to accurately report on one’s knowledge, including knowledge related to incoming sensory signals (Fleming & Dolan, 2012), retrieved memories (Baird et al., 2013), value signals (De Martino et al., 2013), and the arising of spontaneous thoughts (Ellamil et al., 2016; Fox et al., 2015). During self-referential judgments, this network may allow individuals to consciously monitor and compare self-knowledge to the trait in question. Thus, while the DMN may represent the content of self-knowledge and automatic self-related feelings, the FPCN may operate on this information to make task-related judgments and to regulate emotional reactivity when necessary.

Limitations

Several limitations should be noted when interpreting these results. First, we conducted only univariate analyses. It remains possible that distinct signatures of self-referential processing in SAD could be observed using other methods (e.g., multivoxel pattern analysis). Our task design was not optimized for multivoxel pattern analysis, so this will be a key issue to address in future studies. Second, despite using a relatively large clinical sample, our study may still have been underpowered to detect small to moderate effects. In particular, our sample included far fewer controls than SAD patients, which could have been a factor in the lack of observed group differences. In line with this, equivalence tests identified a number of brain regions with effect sizes that could not be statistically rejected as significantly different from a minimal effect size of interest. The relevance of activation in these regions cannot be interpreted in the current study. However, these regions may be useful regions of interest in future work using larger samples. Third, our study may have been limited by the fact that the task included only 5 blocks (25 self-referential judgments) per condition. It is possible that a longer task could result in less variability in brain responses and will be important to consider with future work. Fourth, the task we employed required participants to make a categorical yes/no judgment about the self-descriptiveness of the trait words. Additional insights could potentially be obtained in future work by using a Likert-type scale which would allow for a

parametric analysis that can isolate neural activity that precisely corresponds to the extent to which traits are self-relevant. Fifth, the present study only examined explicit reflection on self-relevant stimuli. It could be informative to compare neural responses during explicit reflection versus more spontaneous self-referential processing in SAD to dissociate the neural structures involved in the conscious representation and evaluation of self-relevant material from those that may be involved in the automatic detection of self-relevant stimuli and the enduring representation of self-views. Finally, in the current study, we treated the SAD patients as a single group that was compared against the control group. This may have masked differences, given that there may be heterogeneity within the SAD sample (Talmon et al., 2021).

Conclusions

The current study used positive and negative trait words to examine the neural basis of self-referential processing in a large sample of SAD patients. While the overall pattern of DMN and FPCN engagement during self-referential judgments was similar in patients and controls, there were specific regions within these networks that showed a group difference and correlations with SAD symptom severity when using a liberal statistical threshold. These findings suggest that there is still much to be learned about the neural basis of diminished positive and increased negative self-views in SAD.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.3758/s13415-021-00933-6>.

Acknowledgments This research was supported by National Institute of Mental Health Grant R01 MH76074 awarded to James Gross.

References

- Andrews-Hanna, J. R., Saxe, R., & Yarkoni, T. (2014). Contributions of episodic retrieval and mentalizing to autobiographical thought: Evidence from functional neuroimaging, resting-state connectivity, and fMRI meta-analyses. *Neuroimage*, *91*, 324–335.
- Baird, B., Smallwood, J., Gorgolewski, K. J., & Margulies, D. S. (2013). Medial and lateral networks in anterior prefrontal cortex support metacognitive ability for memory and perception. *The Journal of Neuroscience*, *33*(42), 16657–16665.
- Blair, K., Geraci, M., Devido, J., McCaffrey, D., Chen, G., Vythilingam, M., et al. (2008). Neural response to self- and other referential praise and criticism in generalized social phobia. *Archives of General Psychiatry*, *65*(10), 1176–1184.
- Blair, K. S., Geraci, M., Otero, M., Majestic, C., Odenheimer, S., Jacobs, M., et al. (2011). Atypical modulation of medial prefrontal cortex to self-referential comments in generalized social phobia. *Psychiatry Research*, *193*(1), 38–45.
- Bradley, M. M., & Lang, P. J. (1999). Affective norms for English words (ANEW): Instruction manual and affective ratings. Citeseer.
- Bruhl, A. B., Delsignore, A., Komossa, K., & Weidt, S. (2014). Neuroimaging in social anxiety disorder—a meta-analytic review

- resulting in a new neurofunctional model. *Neuroscience and Biobehavioral Reviews*, 47, 260-280.
- Buckner, R. L., Sepulcre, J., Talukdar, T., Krienen, F. M., Liu, H., Hedden, T., et al. (2009). Cortical hubs revealed by intrinsic functional connectivity: Mapping, assessment of stability, and relation to Alzheimer's disease. *The Journal of Neuroscience*, 29(6), 1860-1873.
- Buhle, J. T., Silvers, J. A., Wager, T. D., Lopez, R., Onyemekwu, C., Kober, H., et al. (2014). Cognitive reappraisal of emotion: A meta-analysis of human neuroimaging studies. *Cerebral Cortex*, 24(11), 2981-2990.
- Bunge, S. A. (2004). How we use rules to select actions: A review of evidence from cognitive neuroscience. *Cognitive, Affective, & Behavioral Neuroscience*, 4(4), 564-579.
- Clark, D. M., & Wells, A. (1995). A cognitive model of social phobia. In M. R. L. R. G. Heimberg, D. A. Hope, & F. R. Schneier (Ed.), *Social phobia: Diagnosis, assessment, and treatment* (pp. 69-93). Guilford Press.
- Cole, M. W., Reynolds, J. R., Power, J. D., Repovs, G., Anticevic, A., & Braver, T. S. (2013). Multi-task connectivity reveals flexible hubs for adaptive task control. *Nature Neuroscience*, 16(9), 1348-1355.
- Cole, M. W., Repovs, G., & Anticevic, A. (2014). The frontoparietal control system: A central role in mental health. *The Neuroscientist*, 20(6), 652-664.
- D'Argembeau, A. (2013). On the role of the ventromedial prefrontal cortex in self-processing: The valuation hypothesis. *Frontiers in Human Neuroscience*, 7, 372.
- D'Argembeau, A., Collette, F., Van der Linden, M., Laureys, S., Del Fiore, G., Degueldre, C., et al. (2005). Self-referential reflective activity and its relationship with rest: A PET study. *Neuroimage*, 25(2), 616-624.
- D'Argembeau, A., Stawarczyk, D., Majerus, S., Collette, F., Van der Linden, M., Feyers, D., et al. (2010). The neural basis of personal goal processing when envisioning future events. *Journal of Cognitive Neuroscience*, 22(8), 1701-1713.
- D'Argembeau, A., Jedidi, H., Balteau, E., Bahri, M., Phillips, C., & Salmon, E. (2012). Valuing one's self: Medial prefrontal involvement in epistemic and emotive investments in self-views. *Cerebral Cortex*, 22(3), 659-667.
- Davey, C. G., Pujol, J., & Harrison, B. J. (2016). Mapping the self in the brain's default mode network. *Neuroimage*, 132, 390-397.
- De Martino, B., Fleming, S. M., Garrett, N., & Dolan, R. J. (2013). Confidence in value-based choice. *Nature Neuroscience*, 16(1), 105-110.
- Denny, B. T., Kober, H., Wager, T. D., & Ochsner, K. N. (2012). A meta-analysis of functional neuroimaging studies of self- and other judgments reveals a spatial gradient for mentalizing in medial prefrontal cortex. *Journal of Cognitive Neuroscience*, 24(8), 1742-1752.
- Derry, P. A., & Kuiper, N. A. (1981). Schematic processing and self-reference in clinical depression. *Journal of Abnormal Psychology*, 90(4), 286.
- Di Nardo, P. A., Brown, T. A., & Barlow, D. H. (1994). *Anxiety Disorders Interview Schedule for DSM-IV: Lifetime version (ADIS-IV-L)*. Oxford University Press.
- Dixon, M. L., Thiruchselvam, R., Todd, R., & Christoff, K. (2017). Emotion and the prefrontal cortex: An integrative review. *Psychological Bulletin*, 143(10), 1033-1081.
- Dixon, M. L., De La Vega, A., Mills, C., Andrews-Hanna, J., Spreng, R. N., Cole, M. W., et al. (2018). Heterogeneity within the frontoparietal control network and its relationship to the default and dorsal attention networks. *Proceedings of the National Academy of Sciences of the United States of America*, 115(7), E1598-E1607.
- Dixon, M. L., Moodie, C. A., Goldin, P. R., Farb, N., Heimberg, R. G., & Gross, J. J. (2020). Emotion regulation in social anxiety disorder: Reappraisal and acceptance of negative self-beliefs. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, 5(1), 119-129.
- Dobbs, I. G., Foley, H., Schacter, D. L., & Wagner, A. D. (2002). Executive control during episodic retrieval: Multiple prefrontal processes subservise source memory. *Neuron*, 35(5), 989-996.
- Dosenbach, N. U., Visscher, K. M., Palmer, E. D., Miezin, F. M., Wenger, K. K., Kang, H. C., et al. (2006). A core system for the implementation of task sets. *Neuron*, 50(5), 799-812.
- Dosenbach, N. U., Fair, D. A., Miezin, F. M., Cohen, A. L., Wenger, K. K., Dosenbach, R. A., et al. (2007). Distinct brain networks for adaptive and stable task control in humans. *Proceedings of the National Academy of Sciences of the United States of America*, 104(26), 11073-11078.
- Ellamil, M., Fox, K. C., Dixon, M. L., Pritchard, S., Todd, R. M., Thompson, E., et al. (2016). Dynamics of neural recruitment surrounding the spontaneous arising of thoughts in experienced mindfulness practitioners. *Neuroimage*, 136, 186-196.
- Farb, N. A., Segal, Z. V., Mayberg, H., Bean, J., McKeon, D., Fatima, Z., et al. (2007). Attending to the present: Mindfulness meditation reveals distinct neural modes of self-reference. *Social Cognitive and Affective Neuroscience*, 2(4), 313-322.
- Fleming, S. M., & Dolan, R. J. (2012). The neural basis of metacognitive ability. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 367(1594), 1338-1349.
- Fox, K., Spreng, R. N., Ellamil, M., Andrews-Hanna, J. R., & Christoff, K. (2015). The wandering brain: Meta-analysis of functional neuroimaging studies of mind-wandering and related spontaneous thought processes. *Neuroimage*, 111, 611-621.
- Fresco, D. M., Coles, M. E., Heimberg, R. G., Liebowitz, M. R., Hami, S., Stein, M. B., et al. (2001). The Liebowitz social anxiety scale: A comparison of the psychometric properties of self-report and clinician-administered formats. *Psychological Medicine*, 31(6), 1025-1035.
- Frith, U., & Frith, C. (2010). The social brain: Allowing humans to boldly go where no other species has been. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 365(1537), 165-176.
- Gerchen, M. F., Kirsch, P., & Feld, G. B. (2021). Brain-wide inferiority and equivalence tests in fMRI group analyses: Selected Applications. *bioRxiv*.
- Glover, G. H., & Law, C. S. (2001). Spiral-in/out BOLD fMRI for increased SNR and reduced susceptibility artifacts. *Magnetic Resonance in Medicine*, 46(3), 515-522.
- Goldin, P., Ramel, W., & Gross, J. (2009a). Mindfulness meditation training and self-referential processing in social anxiety disorder: Behavioral and neural effects. *Journal of Cognitive Psychotherapy*, 23(3), 242-257.
- Goldin, P. R., Manber, T., Hakimi, S., Canli, T., & Gross, J. J. (2009b). Neural bases of social anxiety disorder: Emotional reactivity and cognitive regulation during social and physical threat. *Archives of General Psychiatry*, 66(2), 170-180.
- Goldin, P., Ziv, M., Jazaieri, H., & Gross, J. J. (2012). Randomized controlled trial of mindfulness-based stress reduction versus aerobic exercise: Effects on the self-referential brain network in social anxiety disorder. *Frontiers in Human Neuroscience*, 6, 295.
- Goldin, P. R., Jazaieri, H., Ziv, M., Kraemer, H., Heimberg, R., & Gross, J. J. (2013a). Changes in positive self-views mediate the effect of cognitive-behavioral therapy for social anxiety disorder. *Clinical Psychological Science: A Journal of the Association for Psychological Science*, 1(3), 301-310.
- Goldin, P. R., Ziv, M., Jazaieri, H., Hahn, K., Heimberg, R., & Gross, J. J. (2013b). Impact of cognitive behavioral therapy for social anxiety disorder on the neural dynamics of cognitive reappraisal of negative self-beliefs: Randomized clinical trial. *JAMA Psychiatry*, 70(10), 1048-1056.
- Hassabis, D., Spreng, R. N., Rusu, A. A., Robbins, C. A., Mar, R. A., & Schacter, D. L. (2013). Imagine all the people: How the brain creates

- and uses personality models to predict behavior. *Cerebral Cortex*, 24(8), 1979–1987.
- Heimberg, R. G., Brozovich, F. A., & Rapee, R. M. (2014). A cognitive-behavioral model of social anxiety disorder. In S.G. Hofmann & P. M. DiBartolo (Eds.), *Social anxiety: Clinical, developmental, and social perspectives* (3rd ed., pp. 705–728). Academic Press.
- JASP Team. (2018). *JASP (Version 0.9)*[Computer software].
- Jazaieri, H., Morrison, A.S., Goldin, P.R., & Gross, J.J. (2015). The role of emotion and emotion regulation in social anxiety disorder. *Current Psychiatry Reports*, 17(1), 531.
- Kaiser, R. H., Andrews-Hanna, J. R., Wager, T. D., & Pizzagalli, D. A. (2015). Large-scale network dysfunction in major depressive disorder: A meta-analysis of resting-state functional connectivity. *JAMA Psychiatry*, 72(6), 603–611.
- Koban, L., Schneider, R., Ashar, Y. K., Andrews-Hanna, J. R., Landy, L., Moscovitch, D. A., et al. (2017). Social anxiety is characterized by biased learning about performance and the self. *Emotion*, 17(8), 1144–1155.
- Korn, C. W., Prehn, K., Park, S. Q., Walter, H., & Heekeren, H. R. (2012). Positively biased processing of self-relevant social feedback. *The Journal of Neuroscience*, 32(47), 16832–16844.
- Lakens, D. (2017). Equivalence tests: A practical primer for t tests, correlations, and meta-analyses. *Social Psychological and Personality Science*, 8(4), 355–362.
- Lang, P. J. (1980). Behavioral treatment and bio-behavioral assessment: Computer applications. *Technology in Mental Health Care Delivery Systems*, 119–137.
- Leary, M. R. (2007). Motivational and emotional aspects of the self. *Annual Review of Psychology*, 58, 317–344.
- Lieberman, M. D., & Cunningham, W. A. (2009). Type I and Type II error concerns in fMRI research: Re-balancing the scale. *Social Cognitive and Affective Neuroscience*, 4(4), 423–428.
- Liebowitz, M. R. (1987). Social phobia. *Anxiety. Modern Problems of Pharmacopsychiatry*, 22, 141–173.
- McCaig, R. G., Dixon, M., Keramatian, K., Liu, I., & Christoff, K. (2011). Improved modulation of rostralateral prefrontal cortex using real-time fMRI training and meta-cognitive awareness. *Neuroimage*, 55(3), 1298–1305.
- Moscovitch, D. A. (2009). What is the core fear in social phobia? A new model to facilitate individualized case conceptualization and treatment. *Cognitive and Behavioral Practice*, 16(2), 123–134.
- Murray, R. J., Schaer, M., & Debbane, M. (2012). Degrees of separation: A quantitative neuroimaging meta-analysis investigating self-specificity and shared neural activation between self- and other-reflection. *Neuroscience and Biobehavioral Reviews*, 36(3), 1043–1059.
- Northoff, G., Heinzel, A., de Greck, M., Birmpohl, F., Dobrowolny, H., & Panksepp, J. (2006). Self-referential processing in our brain—a meta-analysis of imaging studies on the self. *Neuroimage*, 31(1), 440–457.
- Ochsner, K. N., & Gross, J. J. (2005). The cognitive control of emotion. *Trends in Cognitive Sciences*, 9(5), 242–249.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, 9(1), 97–113.
- Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage*, 59(3), 2142–2154.
- Rytwinski, N. K., Fresco, D. M., Heimberg, R. G., Coles, M. E., Liebowitz, M. R., Cissell, S., et al. (2009). Screening for social anxiety disorder with the self-report version of the Liebowitz Social Anxiety Scale. *Depression and Anxiety*, 26(1), 34–38.
- Schmitz, T. W., & Johnson, S. C. (2007). Relevance to self: A brief review and framework of neural systems underlying appraisal. *Neuroscience and Biobehavioral Reviews*, 31(4), 585–596.
- Schmitz, T. W., Kawahara-Baccus, T. N., & Johnson, S. C. (2004). Metacognitive evaluation, self-relevance, and the right prefrontal cortex. *Neuroimage*, 22(2), 941–947.
- Sheline, Y. I., Barch, D. M., Price, J. L., Rundle, M. M., Vaishnavi, S. N., Snyder, A. Z., et al. (2009). The default mode network and self-referential processes in depression. *Proceedings of the National Academy of Sciences of the United States of America*, 106(6), 1942–1947.
- Smallwood, J., Brown, K., Baird, B., & Schooler, J. W. (2012). Cooperation between the default mode network and the frontal-parietal network in the production of an internal train of thought. *Brain Research*, 1428, 60–70.
- Smith, S. M., & Nichols, T. E. (2009). Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. *Neuroimage*, 44(1), 83–98.
- Spreng, R. N., Stevens, W. D., Chamberlain, J. P., Gilmore, A. W., & Schacter, D. L. (2010). Default network activity, coupled with the frontoparietal control network, supports goal-directed cognition. *Neuroimage*, 53(1), 303–317.
- Spreng, R. N., Sepulcre, J., Turner, G. R., Stevens, W. D., & Schacter, D. L. (2013). Intrinsic architecture underlying the relations among the default, dorsal attention, and frontoparietal control networks of the human brain. *Journal of Cognitive Neuroscience*, 25(1), 74–86.
- Sylvester, C. M., Corbetta, M., Raichle, M. E., Rodebaugh, T. L., Schlaggar, B. L., Sheline, Y. I., et al. (2012). Functional network dysfunction in anxiety and anxiety disorders. *Trends in Neurosciences*, 35(9), 527–535.
- Talmon, A., Dixon, M.L., Goldin, P.R., Heimberg, R. G., & Gross, J.J. (2021). Neurocognitive heterogeneity in social anxiety disorder: The role of self-referential processing and childhood maltreatment. *Clinical Psychological Science* <https://doi.org/10.1177/21677026211004452>.
- Thurston, M. D., Goldin, P., Heimberg, R., & Gross, J. J. (2017). Self-views in social anxiety disorder: The impact of CBT versus MBSR. *Journal of Anxiety Disorders*, 47, 83–90.
- van der Meer, L., Costafreda, S., Aleman, A., & David, A. S. (2010). Self-reflection and the brain: A theoretical review and meta-analysis of neuroimaging studies with implications for schizophrenia. *Neuroscience and Biobehavioral Reviews*, 34(6), 935–946.
- Vidal-Ribas, P., Janiri, D., Doucet, G. E., Pompattananangkul, N., Nielson, D. M., Frangou, S., & Stringaris, A. (2021). Multimodal neuroimaging of suicidal thoughts and behaviors in a US population-based sample of school-age children. *American Journal of Psychiatry*, 178(4), 321–332.
- Vincent, J. L., Kahn, I., Snyder, A. Z., Raichle, M. E., & Buckner, R. L. (2008). Evidence for a frontoparietal control system revealed by intrinsic functional connectivity. *Journal of Neurophysiology*, 100(6), 3328–3342.
- Whitfield-Gabrieli, S., Thermenos, H. W., Milanovic, S., Tsuang, M. T., Faraone, S. V., McCarley, R. W., et al. (2009). Hyperactivity and hyperconnectivity of the default network in schizophrenia and in first-degree relatives of persons with schizophrenia. *Proceedings of the National Academy of Sciences of the United States of America*, 106(4), 1279–1284.
- Yeo, B. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., et al. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology*, 106(3), 1125–1165.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.