Spatiotemporal dissociation of brain activity underlying threat and reward in social anxiety disorder

John A. Richey,1 Merage Ghane,1 Andrew Valdespino,1 Marika C. Coffman,1 Marlene V. Strege,1 Susan W. White,1,2 and Thomas H. Ollendick1,2

1Department of Psychology, Virginia Tech., 109 Williams Hall, MC0436 Blacksburg, VA 24061, USA and 2Virginia Tech Child Study Center, Suite 207, Turner St, Blacksburg, VA 24061, USA

Correspondence should be addressed to: John A. Richey, Department of Psychology, Virginia Tech, 109 Williams Hall, MC0436, Blacksburg, VA 24061, USA. Email: richey@vt.edu.

Abstract

Social anxiety disorder (SAD) involves abnormalities in social motivation, which may be independent of well-documented differences in fear and arousal systems. Yet, the neurobiology underlying motivational difficulties in SAD is not well understood. The aim of the current study was to spatiotemporally dissociate reward circuitry dysfunction from alterations in fear and arousal-related neural activity during anticipation and notification of social and non-social reward and punishment. During fMRI acquisition, non-depressed adults with social anxiety disorder (SAD; N = 21) and age-, sex- and IQ-matched control subjects (N = 22) completed eight runs of an incentive delay task, alternating between social and monetary outcomes and interleaved in alternating order between gain and loss outcomes. Adults with SAD demonstrated significantly reduced neural activity in ventral striatum during the anticipation of positive but not negative social outcomes. No differences between the SAD and control groups were observed during anticipation of monetary gain or loss outcomes or during anticipation of negative social images. However, consistent with previous work, the SAD group demonstrated amygdala hyper-activity upon notification of negative social outcomes. Degraded anticipatory processing in bilateral ventral striatum in SAD was constrained exclusively to anticipation of positive social information and dissociable from the effects of negative social outcomes previously observed in the amygdala. Alterations in anticipation-related neural signals may represent a promising target for treatment that is not addressed by available evidence-based interventions, which focus primarily on fear extinction and habituation processes.

Key words: social anxiety disorder; fMRI; reward; monetary incentive delay; nucleus accumbens; threat

Introduction

Social anxiety disorder (SAD) is a common and debilitating psychiatric disorder predominantly characterized by persistent fear of one or more social or performance situations (American Psychiatric Association and American Psychiatric Association and DSM-5 Task Force, 2013). Previous behavioral and neuroimaging studies of SAD have highlighted a central role for negative affects and threat-related neural circuits in symptom expression, primarily including a limbic–medial prefrontal circuit that involves enhanced processing of threat stimuli (Phan et al., 2006; Goldin et al., 2009; Etkin, 2010; Schmidt et al., 2010; Hattingh et al., 2013). Additionally, recent work has also suggested that SAD may be uniquely characterized by diminished positive affect (Brown et al., 1998; Kashdan, 2007; Alden et al., 2008; Morrison and Heimberg, 2013; Weeks, 2015), which may mechanistically relate to its development and maintenance (Caouette and Guyer, 2014; Haller et al., 2015) and cannot be accounted for by depressive symptomatology (Kashdan, 2007; Eisner et al., 2009; Weeks, 2015). Despite the increasing interest in characterizing psychiatric disorders in terms of positively and negatively valenced motivational systems (Insel et al., 2010; Casey et al., 2014; Insel, 2014) and the potential to inform treatment development, few studies have examined the...
mechanisms of positive affect deficits in SAD within the functional neurobiology of reward.

Reward processing in typically developing humans and non-human primates is mediated by dense dopaminergic projections originating from the ventral tegmental area (VTA) that project to the striatum, the orbitofrontal cortex (OFC), the ventromedial prefrontal cortex (vmPFC) and the anterior cingulate cortex (Haber and Knutson, 2010). Collectively, these regions form a mesolimbic dopamine pathway that is sensitive to both the magnitude and the probability of reward (Schultz, 1998; Ikemoto and Panksepp, 1999; Schultz, 2000; Berdidge et al., 2009; Saddoris et al., 2015). In particular, patterned firing of dopaminergic neurons in the nucleus accumbens (NAc) is thought to encode incentive motivation related to approach behaviors toward salient goals (Knutson et al., 2001; Knutson and Cooper, 2005; Kim et al., 2006; Bjork and Hommer, 2007; Forbes et al., 2009). Comparative research further suggests that dopamine-mediated responses in NAc are increased during anticipation of learned cue-outcome associations (Martin and Ono, 2000; Melendez et al., 2002), which are thought to be broadly reflective of the motivational relevance of upcoming events (Carter et al., 2009) and related to the modulation and planning of complex motivated behavior (Mogenson et al., 1980; Roesch et al., 2009). Social interaction mobilizes the same mesolimbic network that is active while processing non-social rewards such as food, money, sex and drugs of addiction (Koob and LeMoal, 1997; Izuma et al., 2008; Billing et al., 2008; Spreckelmeyer et al., 2009; Rademacher et al., 2010; Trezza et al., 2011). Such reward network responses toward social information are also present during both anticipatory and outcome periods (Hayden et al., 2007; Winston et al., 2007; Rademacher et al., 2010), suggesting that behavior is strongly guided by both the motivation to attain social rewards and the enjoyment of such rewards once received (Ruff and Fehr, 2014).

While convergent evidence suggests that in typically developing individuals, social cohesion and affiliation are associated with increase in striatal responses to social cues, this perspective also suggests that reduced striatal activation during the assignment of values to social stimuli ought to be associated with degraded social affiliation and perceptions of weaker social bond formation, both of which are observed in SAD (Mathew et al., 2001; Fox and Kalin, 2014; Haller et al., 2015). Consistent with this conceptualization, relatively early PET imaging studies demonstrated altered striatal functions in SAD, which may be rooted in abnormal central dopamine function, linked to dopamine D2 receptor and dopamine transporter (DAT) availability in the striatum (Tiihonen et al., 1997; Schneier et al., 2000; Schneier et al., 2008) (also see Schneier et al., 2009). In behavioral terms, when compared with non-anxious counterparts, adults with SAD report that positive events are less likely to occur (Gilboa-Schechtman et al., 2000), report greater difficulty expressing positive emotions (Turk et al., 2005) and demonstrate a reduced tendency to sustain or savor positive affect once experienced (Eisner et al., 2009).

A small number of neuroimaging studies have probed the neurobiological substrates of dampened positive affect in SAD. For example, Sripada et al. (2013) reported that the experience of cooperative social reciprocity in an economic exchange (‘trust’) game was associated with increased striatal activity in control subjects but absent in SAD, potentially reflecting diminished valuation of positive partner feedback. Richey et al. (2014) observed diminished NAc activation in SAD compared to controls and adults with autism during anticipation of social but not monetary outcomes in the Monetary Incentive Delay (MID) task (Knutson et al., 2001, 2008). This work illustrated comparatively reduced striatal activity during anticipation of non-emotional (neutral) faces from the NimStim face set (Tottenham et al., 2009), which was interpreted as a social-specific effect because it was not observed during anticipation of non-social (i.e. monetary) gains. Consistent with these results, Maresh et al. (2014) reported no differences between adults with and without SAD during anticipation of monetary gains. Collectively, this pattern of results suggests that while percepts of non-social reward may be preserved in SAD, abnormalities in incentive motivation may be constrained exclusively to social stimuli.

Alterations in value-related neural signals may mechanistically relate to disorder development and maintenance by compromising the ability to select between social options that should be pursued or avoided. However, a lack of appetitive motivation toward social reward does not rule out the possibility that diminished ‘seeking’ of social incentives in SAD may in fact be attributable to enhanced avoidance motivation. Previous animal and human work has demonstrated a role for ventral striatum in aspects of aversive motivation and avoidance (Reis et al., 2004; Oleson et al., 2012), indicating that cues preceding stressors can also increase dopamine (DA) release and furthermore that DA depletion can impair active avoidance in rats (McCullough et al., 1993). Although this question has not been specifically examined in anxious samples, prior neuroimaging studies have used incentive tasks (including the MID) to determine whether cues predicting appetitive (gain) and aversive (loss) outcomes are encoded by separate systems. Knutson and Greer (2008) reported results from a meta-analysis of 21 studies using the MID, which indicated that gain anticipation contrasts showed greater activation in NAc than loss anticipation contrasts of the same magnitude. These authors framed this result in terms of an anticipatory affect model of NAc function, in which positive (gain) and negative (loss) outcomes recruit spatiotemporally distinct functions within the mesolimbic circuit (NAc and insular cortices, respectively). Under this anticipatory affect account, valence-specific effects for cue-driven striatal responses in the MID appear to be indexing a particular component of accumbens DA function in the form of activational or instrumental seeking behaviors, but not the preparation of an avoidance behavior. Thus, the differentiation between enhanced avoidance and diminished approach motivation appears to be testable within incentivized reaction time tasks such as the MID.

Yet, despite clear theoretical predictions the landscape of reward circuitry deficits in SAD remains largely unexplored. Accordingly, the primary purpose of this study was to determine whether SAD is characterized by deficiencies in reward processing that are not addressed by currently available evidence-based treatments, which focus primarily on fear extinction and habituation processes as their central mechanisms of action (Feske and Chambless, 1995; Blanco et al., 2010; Hofmann et al., 2012). By manipulating the valence (positive, negative) and type of incentives (social, non-social) in a sample of adults with SAD and healthy control subjects during an incentivized reaction time-task (Knutson et al., 2000, 2008), we sought to identify the dissociable contributions of reward- and threat-related neural systems to social incentive processing in SAD. Given previous findings of reduced NAc activation in response to social, but not monetary rewards in SAD (Maresh et al., 2014; Richey et al., 2014) and cognitive and behavioral research indicating that this disorder is characterized by circumscribed deficits in positive affects (Brown et al., 1998; Kashdan,
Table 1. Mean (SDs) age and symptom profiles.

<table>
<thead>
<tr>
<th></th>
<th>Control (n = 22)</th>
<th>SAD (n = 21)</th>
<th>t or (\chi^2) (P)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>26.50 (7.98)</td>
<td>25.67 (7.61)</td>
<td>0.35 (0.73)</td>
</tr>
<tr>
<td>Sex (F/M)</td>
<td>10/12</td>
<td>8/13</td>
<td></td>
</tr>
<tr>
<td>WASI (full-scale)</td>
<td>118.2 (7.93)</td>
<td>117.2 (10.9)</td>
<td>0.33 (0.73)</td>
</tr>
<tr>
<td>WASI (performance)</td>
<td>118 (8.89)</td>
<td>118.1 (12.3)</td>
<td>0.1 (0.99)</td>
</tr>
<tr>
<td>WASI (verbal)</td>
<td>114.7 (10)</td>
<td>113.1 (11)</td>
<td>0.46 (0.64)</td>
</tr>
<tr>
<td>LSAS total score</td>
<td>17.94 (11.62)</td>
<td>8.72 (8.05)</td>
<td>9.15 (0.00001)</td>
</tr>
<tr>
<td>LSAS fear subscale</td>
<td>12.44 (7.64)</td>
<td>32.0 (11.07)</td>
<td>8.21 (0.00001)</td>
</tr>
<tr>
<td>LSAS avoidance subscale</td>
<td>8.72 (8.05)</td>
<td>7.39 (0.00001)</td>
<td></td>
</tr>
<tr>
<td>BDI total score</td>
<td>5.65 (4.56)</td>
<td>7.2 (6.25)</td>
<td>0.91 (0.37)</td>
</tr>
</tbody>
</table>

Both groups completed: (1) The Wechsler Abbreviated Scale of Intelligence (WASI); (2) The Liebowitz Social Anxiety Scale (LSAS; Liebowitz, 1987), a measure designed to assess social fear and avoidance; (3) the Beck Depression Inventory-II (BDI; Beck, 1961; Beck Steer and Garbin, 1998), administered to assess the overall severity of depressive symptoms and to verify that the SAD group did not have significant mood symptoms.

In 2004, we predicted that SAD would be characterized by diminished activation in the ventral striatum during anticipation of positive social outcomes and not enhanced striatal activation during anticipation of negative social outcomes, in keeping with the anticipatory affect account of striatal functions (Knutson and Greer, 2008). We further predicted that these effects (degraded approach motivation and lack of enhanced avoidance motivation) would be constrained exclusively to social incentives and not observed during monetary incentive processing. Second, given the framework of the present study that reward network dysfunction may be dissociable from threat-related processing of social stimuli, we also predicted that hypo-activation of NAc during anticipatory processing of positive social incentives in NAc can be distinguished temporally from amygdala hyper-activation during negative social outcomes. We further hypothesized that individual estimates of NAc activity during positive social incentives would be related to symptom severity in the SAD group.

Method

Participants

The total sample comprised 43 adults (control N = 22, SAD N = 21) recruited from the community. All participants provided written informed consent, as reviewed and approved by the local Institutional Review Board (IRB). Participants were matched on age, sex and IQ. Information on group size, age and gender distributions are summarized in Table 1 and further described in the supplementary materials. Participants were right-handed and had normal or corrected to normal vision. All participants were assessed for preliminary exclusion criteria, including MRI contraindications, uncontrolled epilepsy (seizure within 6 months prior to consent), history of neurological injury (head trauma), current use of psychotropic medications (past 3 months) and color blindness as measured by the Ishihara Color Plate Test (Ishihara, 1990).

fMRI task

The fMRI task was modified from the Monetary Incentive Delay (MID) task as first described in Knutson et al. (2000). Task conditions and trial timings are summarized in Figure 1. Runs began with an instructional screen indicating the forthcoming run type. Events (money [or faces, gain [or loss]) were segregated into separate runs (36 trials per run), in order to minimize the number of cues to be memorized, resulting in four distinct run types: (1) Money ’gain’ [+ $1], (2) Money ’loss’ [− $1], (3) Face ’gain’ [happy] and (4) Face ’loss’ [angry]. Eight total runs, two of each type as depicted in Figure 2b were presented in counterbalanced order across subjects. Additional task details are provided in the supplementary materials.

Measures

Diagnostic assessment. All subjects completed the Anxiety Disorders Interview Schedule for DSM-IV (ADIS-IV) (Brown et al., 1994), administered by research reliable doctoral-level assessors and reviewed by a Ph.D. level clinical psychologist (J.A.R.).

Cognitive assessment. Subjects completed the Wechsler Abbreviated Scale of Intelligence-II (WASI-II) (Wechsler, 2011), a clinician-administered measure of general cognitive function and IQ.

Self-reported symptoms. All subjects completed the Liebowitz Social Anxiety Scale – LSAS (Liebowitz, 1987), which comprises 24 items and uses a 0 to 3 rating for both fear (ranging from ‘no fear’ to ‘severe fear’) and avoidance (ranging from ‘never’ to ‘usually’), for different social situations (Cronbach’s \(z\) for the current sample = 0.93). Participants completed the Beck Depression Inventory – BDI-II (Beck et al., 1996), a 21-item self-report measure of depression symptomatology (Cronbach’s \(z\) for the current sample = 0.77).

Picture ratings. After scanning, participants completed a picture-rating task outside of the scanner. In this task, participants provided a rating (0–8 point scale) reflecting valence and arousal ratings for each face stimulus using a self-assessment mannequin (Bradley and Lang, 1994).

fMRI data analysis


Consistent with previous studies using the MID in clinical samples (Guyer et al., 2006; Beck et al., 2009; Delmonte et al., 2012), the primary group-level (random effects) method of fMRI data analysis was a mixed model ANOVA (cf. Henson and Penny, 2005; equations (21)–(25)), with one between-subjects factor: group (SAD, Control) and three within-subjects factors: 2 (reward potential [win, non-win]), \(\times\) 2 (stimulus type [Social, Monetary]), \(\times\) 2 (outcome type [Reward, Punishment]), applied separately for the anticipatory and outcome phases of the task.
Fig. 1. Schematic depiction of the modified MID task. Participants alternated between ‘money’ and ‘face’ runs, and ‘win’ and ‘loss’ runs, denoted by an instructional screen at the start of each run. Each trial consisted of a cue (i.e. a gray triangle indicates a ‘win’ trial and a red triangle indicates a ‘loss’ trial), an anticipatory delay, a target and outcome feedback. Not shown: in all runs, a blue circle cue appeared in 50% of trials, indicating a non-incentivized trial in which a ‘hit’ resulted in a check mark and a ‘miss’ in an ‘X’.

Fig. 2. Average reaction times during face and money conditions for both potential reward (‘Rew’) and non-potential reward (‘Non’) trials. Error bars represent the standard errors of the mean.
Table 2. Clusters showing significant group differences in ANOVA models.

<table>
<thead>
<tr>
<th>Anticipation Phase</th>
<th>Reward Type [Money, Faces] X Valence [Gain, Loss] X Possible Win [Win, Non-win]</th>
<th>Control &gt; SAD; Clusters significant at $P &lt; 0.005$; FDR corrected at $P &lt; 0.05$</th>
<th>SAD &gt; control; No clusters significant at $P &lt; 0.005$, FDR corrected at $P &lt; 0.05$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Talairach–Tournoux atlas label</td>
<td>Brodmann area*</td>
<td>Size (mm³)</td>
<td>t mean</td>
</tr>
<tr>
<td>Right posterior cingulate</td>
<td>30</td>
<td>20,985</td>
<td>3.1495</td>
</tr>
<tr>
<td>Postcentral gyrus</td>
<td>7</td>
<td>5405</td>
<td>3.0387</td>
</tr>
<tr>
<td>Right orbital gyrus</td>
<td>—</td>
<td>4807</td>
<td>3.1262</td>
</tr>
<tr>
<td>Right superior temporal gyrus</td>
<td>22</td>
<td>4180</td>
<td>3.1934</td>
</tr>
<tr>
<td>Right subcallosal gyrus</td>
<td>47b</td>
<td>2656</td>
<td>3.1262</td>
</tr>
<tr>
<td>Right medial frontal gyrus</td>
<td>6</td>
<td>1935</td>
<td>3.0029</td>
</tr>
<tr>
<td>Right orbital gyrus</td>
<td>—</td>
<td>4807</td>
<td>3.1262</td>
</tr>
<tr>
<td>Right superior temporal gyrus</td>
<td>22</td>
<td>4180</td>
<td>3.1934</td>
</tr>
</tbody>
</table>

SAD > control; No clusters significant at $P < 0.005$, FDR corrected at $P < 0.05$

Outcome Phase

<table>
<thead>
<tr>
<th>Reward Type [Money, Faces] X Valence [Gain, Loss] X Possible Win [Win, Non-win]</th>
<th>Control &gt; SAD; Clusters significant at $P &lt; 0.005$; FDR corrected at $P &lt; 0.05$</th>
<th>SAD &gt; control; No clusters significant at $P &lt; 0.005$, FDR corrected at $P &lt; 0.05$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Right cingulate gyrus</td>
<td>8a</td>
<td>1040</td>
</tr>
<tr>
<td>Right medial frontal gyrus</td>
<td>9</td>
<td>745</td>
</tr>
<tr>
<td>Right orbital gyrus</td>
<td>11b</td>
<td>630</td>
</tr>
<tr>
<td>Right cerebellar tonsil</td>
<td>—</td>
<td>628</td>
</tr>
<tr>
<td>Left superior frontal gyrus</td>
<td>10</td>
<td>438</td>
</tr>
<tr>
<td>Right superior parietal lobule</td>
<td>7</td>
<td>433</td>
</tr>
<tr>
<td>Left superior frontal gyrus</td>
<td>32</td>
<td>335</td>
</tr>
<tr>
<td>Right postcentral gyrus</td>
<td>3</td>
<td>261</td>
</tr>
</tbody>
</table>

SAD > control; No clusters significant at $P < 0.005$, FDR corrected at $P < 0.05$

---

*BA label is indicated in table if the focus point of cluster within <5 mm of BA label boundary.

bIndicates clusters displayed in Figure 4a and b.

Indicates clusters displayed in Figure 5a.
In addition, we sought to characterize neural responses to social outcomes and to replicate previous findings in SAD, indicating amygdala activation when viewing angry faces (Phan et al., 2006; Evans et al., 2008). Because amygdala activity is most likely to be observed in SAD for social conditions only and thought to be less related to non-social (monetary) manipulations (cf. Maresh et al., 2014), we examined the outcome phase for social-only trials. In this planned contrast, we opted to code non-social conditions in this model, to avoid their inclusion in implicit baseline parameters, which would likely add noise to subsequent parameter estimates. The specific interaction of interest within this mixed model ANOVA was a 2 (group) × 2 (reward and punishment [R, P]) × 2 (hit, miss) for social (Soc) outcomes.

For all whole-brain results, statistical parameter maps were thresholded at $P < 0.005$ and false discovery rate (FDR) corrected at $P < 0.05$. Clusters that passed corrections for height and multiple comparisons were further scrutinized by extracting subject- and condition-specific signal intensity coefficients to evaluate simple effects. Relations between neural responses to rewards and social anxiety symptoms from the LSAS were assessed in SAD by using group-level activation maps to extract mean subject-specific parameters. The specific interaction of interest within this mixed model ANOVA was a 2 (group) × 2 (hit, miss) for social (Soc) outcomes. For all whole-brain results, statistical parameter maps were thresholded at $P < 0.005$ and false discovery rate (FDR) corrected at $P < 0.05$. Clusters that passed corrections for height and multiple comparisons were further scrutinized by extracting subject- and condition-specific signal intensity coefficients to evaluate simple effects. Relations between neural responses to rewards and social anxiety symptoms from the LSAS were assessed in SAD by using group-level activation maps to extract mean subject-specific parameters. The specific interaction of interest within this mixed model ANOVA was a 2 (group) × 2 (hit, miss) for social (Soc) outcomes.

Results

**MID reaction times**

In-scanner reaction times (RTs) to task targets are depicted in Figure 2. Full details regarding analysis and interpretation of RT data are provided in the supplementary materials.

**Picture ratings**

Valence and arousal ratings of faces are depicted in Figure 3. Full analysis and interpretation of face rating data are provided in the supplementary materials.

Cognitive functioning and self-reported symptoms

Results for self-reported symptoms and IQ measurements are reported in Table 1.

fMRI statistical comparisons

**Anticipatory phase.** A whole-brain mixed-model ANOVA compared voxel-wise activity of controls vs SAD for each stimulus category during the anticipatory period of the MID (i.e. subsequent to the cue period and prior to target presentation). Main effect images and corresponding activation tables are provided in the supplementary material. In the mixed-model ANOVA, a significant interaction was observed in bilateral NAc for the contrast controls > SAD during anticipation ($P < 0.005$ FDR corrected at $P < 0.05$; left $k = 302$, peak voxel MNI[$x$, $y$, $z$]: $-8$, $13$, $-13$, $t = 3.11$; right $k = 2,696$, peak voxel MNI: $22$, $25$, $-17$, $t = 4.17$), as well as in left and right temporal cortex, thalamus and insula (Figure 4a). No significant clusters were observed at a similar threshold for the contrast SAD > controls. Following upon significant clusters in a priori regions of interest (ROIs) in left and right NAc, we followed the procedure outlined in Knutson et al. (2008) by placing 8 mm spheres (e.g. 4 mm radius) with centroids at the following coordinates (MNI: ±10, 10, −2). We further constrained each ROI to include only voxels that demonstrated significant activation ($P < 0.005$; FDR at $P < 0.05$) in the group-level ANOVA model. Case- and event-wise parameter estimates were extracted for each ROI and aggregated into separate models for comparison of simple effects within gain and loss trials (Figure 4b). In line with initial predictions, planned post hoc tests of simple effects revealed that controls demonstrated significantly greater activation in bilateral NAc during anticipation of positive social images than SAD subjects (left NAc: $t(39) = 3.11$ and $P = 0.005$; right NAc: $t(39) = 3.22$ and $P = 0.002$), an effect that appeared to be exclusive to the anticipation of positive social images and was not observed between groups during anticipation of monetary gains. However, we did find evidence for comparatively reduced NAc activation in controls relative to SAD during anticipation of angry faces for right
Table 3. Clusters showing significant group differences in social-only trials.

<table>
<thead>
<tr>
<th>Outcome Phase</th>
<th>Emotion [Angry, Happy] X Outcome [Hit, Miss]</th>
<th>Talairach-Tournoux atlas label</th>
<th>Brodmann area</th>
<th>Size (mm$^3$)</th>
<th>t mean</th>
<th>t max</th>
<th>MNI coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Right middle occipital gyrus</td>
<td>19</td>
<td>1393</td>
<td>3.077</td>
<td>4.43</td>
<td>34  -70 4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Left precentral gyrus</td>
<td>4</td>
<td>513</td>
<td>3.064</td>
<td>3.514</td>
<td>-27 -27 55</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Right parahippocampal gyrus</td>
<td>36$^a$</td>
<td>512</td>
<td>3.046</td>
<td>4.108</td>
<td>15  -10 -28</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Right inferior temporal gyrus</td>
<td>21</td>
<td>435</td>
<td>2.989</td>
<td>3.619</td>
<td>46  -3  -20</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Left parahippocampal gyrus</td>
<td>36$^a$</td>
<td>262</td>
<td>3.059</td>
<td>3.854</td>
<td>-14 -8 -29</td>
</tr>
</tbody>
</table>

SAD > control
No clusters significant at P < 0.005, FDR corrected at P < 0.05

$^a$BA label is indicated in table if the focus point of cluster within <5 mm of BA label boundary.
$^b$Indicates clusters displayed in Figure 5b.

Given previous reports indicating an association between depression and abnormal striatal function (Pizzagalli et al., 2009), all analyses, for both anticipatory and outcome phases of the MID, were repeated with casewise scores on BDI covaried (see the supplementary materials for full details). In short, the

Direct comparison of positive vs negative social outcomes

To examine the outcome phase for social-only trials, we constructed another mixed model ANOVA by again modeling both social and non-social outcomes, but in this case examining only the 2 (group) × 2 (reward and punishment [R, P]) × 2 (hit, miss) interaction. The results of this analysis revealed significant activation in bilateral amygdala, which prompted follow-up pairwise t tests to determine the direction and magnitude of simple effects that produced this result (Figure 5b). Clusters identified in this analysis included voxels that fell within and also outside of predefined amygdala ROIs from the Harvard-Oxford subcortical atlas (Frazier et al., 2005; Makris et al., 2006). Plotted post hoc tests revealed that SAD subjects demonstrated greater bilateral activation in amygdala exclusively when viewing angry faces (left amygdala: t(39) = 3.49 and P = 0.001; right amygdala: t(39) = 3.31 and P = 0.009), whereas no other between-group differences for the remaining three event types were observed in either ROI (all Ps > 0.28).

Inclusion of Beck Depression Inventory (BDI) scores as a covariate

Given previous reports indicating an association between depression and abnormal striatal function (Pizzagalli et al., 2009), all analyses, for both anticipatory and outcome phases of the MID, were repeated with casewise scores on BDI covaried (see the supplementary materials for full details). In short, the

1 Note that the spatial extent of clusters observed bilaterally in the amygdala was further scrutinized in order to determine the amount of spatial overlap with pre-defined segmentations of left and right amygdala derived from standardized atlases of subcortical anatomy. We used the Harvard-Oxford subcortical atlas to calculate the number of voxels within significant clusters reported in Figure 6b that overlapped with left (k = 188) and right (k = 390) amygdala as defined by separate left and right amygdala segmentations within this atlas. ROI analyses reported here reflect BOLD signal activity extracted from the entire cluster, however supplemental analyses also scrutinized the extent to which event-wise patterns were similar when conducted only on voxels (1) within and (2) outside of Harvard-Oxford atlas labels. The results of this additional analysis indicated no differences in the patterns of observed effects for (1) or (2), suggesting that observed effects were consistent across the spatial extent of the cluster as defined by the Harvard-Oxford atlas (see the supplementary results for additional details).
principal results reported above remained unchanged, with the exception of clusters in the bilateral insula, which were regressed out entirely, indicating that the variance in these voxels could be substantially accounted for by the variance contained in the BDI.

Finite impulse response model: direct dissociation of reward-vs threat-related ROIs

To enhance the comparability of results across ROIs and to rule out the possibility that regional differences in the shape or temporal dynamics of evoked BOLD responses were driving between-group differences observed in the univariate GLM, we conducted an unbiased finite impulse response (FIR) model comparison of average signal intensity in the neural responses observed in ventral striatum and amygdala during anticipatory and outcome periods (respectively; Figures 4 and 5; full details provided in the supplementary materials). Overall results of this analysis indicated that relative to controls, SAD subjects showed a significantly diminished response in the ventral striatum during anticipation of social reward images, which was dissociable from significant hyper-activation in amygdala observed subsequently in the outcome phase while viewing an angry face (Figure 6a).

fMRI-symptom associations

The results of follow-up correlational analyses indicated that neither left nor right NAc2 demonstrated any relationship with the LSAS total scale score during monetary anticipation (within potential win trials only). However, we observed a significant correlation between right ($r = -0.48$ and $P = 0.04$) but not left

2 All fMRI-symptom correlations reported here reflect calculations that take into account casewise scores on the BDI-II as a covariate, to guard against the possibility that depressive symptoms could be inflating relationships between NAc activation and LSAS scores. The authors wish to thank an anonymous reviewer for this suggestion.
\(r = -0.04\) and \(P = 0.96\) NAc activation and LSAS total scale score, indicating that more severe SAD symptoms were associated with lower levels of right NAc activation exclusively during anticipation of positive faces (Figure 6b). No significant correlations were observed between LSAS scores and NAc parameter estimates during anticipation of potential loss trials for either social or monetary incentives (all Ps > 0.21).

**Discussion**

This study was designed to dissociate neural responses in social and non-social rewards and punishments in a sample of non-depressed adults with SAD compared with healthy control subjects. Results collectively indicated diminished anticipatory BOLD responses in ventral striatum that were social-specific as well as spatially and temporally independent of enhanced BOLD responses observed in threat and arousal systems. These findings provide evidence for a reward-centric mechanism of functional pathophysiology in SAD that is not currently incorporated into major treatment approaches, which instead focus primarily on fear extinction and habituation processes.

Because individuals with social anxiety have been found to report reduced social approach motivation and diminished positive affect, we predicted that diminished striatal function would be observed in the SAD group exclusively during anticipation of positive social incentives. Consistent with prediction, a significant decrement in NAc activity was observed during anticipation of positive social incentives, an effect that was distinguishable from responses to monetary incentives and from later stages of reward processing when notification of outcomes occurred. When considered together, the effect of positive and negative social images produced a dissociation in the spatial and temporal domains, such that the direction of striatal and amygdala responses in the SAD group depended on the emotional valence (positive/negative) and temporal period (anticipation/notification) under consideration. Interestingly, this specific combination of effects was observed against the backdrop of no between-groups differences in the vmPFC and dmPFC during outcome processing, suggesting that the subjective value of both monetary and social incentives are not neurally differentiable once achieved. Nor did we observe differences between groups in insula activation during incentive processing. Considering research linking insula activity to visceral and autonomic arousal (Eickhoff et al., 2006), this suggests that the intensity of internal states of arousal did not differ between groups during anticipation of social incentives regardless of valence. This finding is in contrast with a number of prior studies that have found enhanced insula activity in SAD samples while viewing social stimuli. However these studies have exclusively focused on direct viewing of social images (Straube et al., 2004; Shah et al., 2009) and social feedback (Peterburs et al., 2016) and not anticipation of such stimuli as was the case here, suggesting that the characteristics of anticipatory processing in SAD may differ substantially from outcome processing along dimensions of both motivation and arousal.
Moreover, the double dissociation of striatal and amygdala responses may offer novel insights into how the particular dimensions of affect and arousal may be separable and also mechanistically altered in SAD. Knutson et al. (2014) recently proposed a testable quantitative model for inferring affect from certain combinations of results in the MID. At the core of this model is theoretical space defined by orthogonal axes of ‘positive arousal’ and ‘negative arousal.’ Specific combinations of brain activity in NAc and insula appear to track closely with self-reported positive affect and overall arousal (respectively) during anticipatory processing in the MID. The scheme for affective inference contained in this circumplex model of positive and negative arousal, when applied to the results observed here, yields the inference that differences in positive affect (as indexed by NAc activation) but not general arousal (indexed by insular cortex) may uniquely characterize SAD during incentive processing, which in turn lends support to the hypothesis that social-motivational deficits in SAD may be due to trait-like differences in approach-related motivation, rather than general levels of anticipatory arousal.

We further sought to determine whether diminished pursuit of social incentives in SAD might be due to (1) diminished approach motivation or (2) enhanced avoidance-related motivation, as indexed by between-groups differences in striatal activation during anticipation of social reward and punishments, respectively. Relative to baseline, we found evidence for significant reductions in striatal activation bilaterally in controls during anticipation of negative social images (Figure 4b). However, SAD subjects did not display a similar pattern of diminished striatal activity relative to baseline and indeed parameter estimates were not different from zero for either trial type in the SAD group. In contrast to initial prediction, a direct group-wise comparison of responses in striatum during anticipation of negative social outcomes revealed that the SAD group demonstrated comparatively enhanced activity in right ($P = 0.03$) but not left ($P = 0.15$) striatum. Along these same lines, reaction time data in our study indicated significantly faster responses to social punishment trials in the SAD group. We thus found only partial support for the hypothesized lack of enhanced avoidance motivation. Instead, this pattern of results appears to support a model in which the striatum is disinhibited by cues preceding avoidable social punishment in SAD. Although unexpected, this finding is not without precedent. For example, animal research has indicated that DA neurons appear to respond more selectively to rewards and normally show weak responses or even inhibition to aversive stimuli (Mirenowicz and Schultz, 1996; Ungless et al., 2004). However, elevated DA levels have been observed in NAc in response to conditioned stimuli when associations are well learned (Samejima et al., 2005; Eldar et al., 2016). In relation to the current study, it is therefore possible that cue-outcome associations were learned more rapidly by the SAD group, thus bolstering striatal responses over time. Supporting this account, Levita et al. (2012) examined striatal responses during a well-learned conditioned avoidance task and found evidence for greater activation of NAc during active avoidance and furthermore found that individual differences in NAc activity positively correlated with individual differences in self-reported anxiety.
Although the MID is not a conditioned learning task and thus does not directly measure learning rate, we did select for individuals high and low in anxiety. Consequently, if SAD subjects learned the cue-outcome associations more rapidly than controls in social (but not non-social) punishment trials, this may explain the comparatively greater striatal responses exclusively during social avoidance. Thus, a potentially profitable line of future work might examine whether such striatal responses in SAD are directly tracking with an enhanced rate of avoidance learning during aversive social conditioning.

In total, results from this study also have a number of mechanistic implications for intervention. Prior treatment development efforts have largely fallen in line with major theoretical approaches that feature negative valence systems at their core (Schneider et al., 1999). Yet, current evidence-based treatments for SAD such as Cognitive Behavioral Therapy (CBT) and selective-serotonin reuptake inhibitors (SSRIs), which focus on fear extinction and serotonergic systems, respectively, produce only modest effects, with between 40 and 60% of SAD patients remaining symptomatic after treatment (Heimberg et al., 1998; Stein et al., 1999; Davidson et al., 2004; Liebowitz et al., 2005).

Results from our study point toward a separate pathophysiological mechanism that is not currently incorporated into CBT and related families of interventions. In light of these results, it is possible that targeting hedonic function in SAD may result in provable gains in treatment and reduce rates of dropout. Consistent with this possibility, treatment of hedonic deficits, separate from negative affect has proven successful in other psychiatric disorders characterized by reward circuitry dysfunction (Dimidjian et al., 2006; Johnson et al., 2011; Kuyken et al., 2015).

Results and conclusions presented herein should be evaluated in light of study limitations. First, although the MID is a reliable and well-validated index of reward-circuitry function, we are observing post hoc the generative model of payoffs, where prior beliefs are being described by the task. This allowed us to determine correlational relationships to symptom severity, but does not allow a definitive answer to the question of whether social-motivational deficits are developmentally linked to striatal abnormalities (see supplementary materials for additional discussion). Second, we note that in comparison to some clinical neuroimaging reports, our sample size was not large. However, we also note that studies employing the MID often report 20 participants or fewer and a recent review of this task by Stein et al., 2006; Johnson et al., 2011; Kuyken et al., 2015.

Results and conclusions presented herein should be evaluated in light of study limitations. First, although the MID is a reliable and well-validated index of reward-circuitry function, we are observing post hoc the generative model of payoffs, where prior beliefs are being described by the task. This allowed us to determine correlational relationships to symptom severity, but does not allow a definitive answer to the question of whether social-motivational deficits are developmentally linked to striatal abnormalities (see supplementary materials for additional discussion). Second, we note that in comparison to some clinical neuroimaging reports, our sample size was not large. However, we also note that studies employing the MID often report 20 participants or fewer and a recent review of this task by Stein et al., 2006; Johnson et al., 2011; Kuyken et al., 2015.


