Theory of Mind Impairments in Social Anxiety Disorder

Dianne M. Hezel
Richard J. McNally
Harvard University

Social anxiety disorder (SAD) is a common psychiatric disorder characterized by a persistent, excessive fear and avoidance of social and performance situations. Research on cognitive biases indicates individuals with SAD may lack an accurate view of how they are perceived by others, especially in social situations when they allocate important attentional resources to monitoring their own actions as well as external threat. In the present study, we explored whether socially anxious individuals also have impairments in theory of mind (ToM), or the ability to comprehend others’ mental states, including emotions, beliefs, and intentions. Forty socially anxious and 40 non-socially-anxious comparison participants completed two ToM tasks: the Reading the Mind in the Eyes and the Movie for the Assessment of Social Cognition. Participants with SAD performed worse on ToM tasks than did non-socially-anxious participants. Relative to comparison participants, those with SAD were more likely to attribute more intense emotions and greater meaning to what others were thinking and feeling. These group differences were not due to interpretation bias. The ToM impairments in people with SAD are in the opposite direction of those in people with autism spectrum conditions whose inferences about the mental states of other people are absent or very limited. This association between SAD and ToM may have important implications for our understanding of both the maintenance and treatment of social anxiety disorder.

Keywords: social anxiety disorder; theory of mind; social cognition; anxiety disorders

Social anxiety disorder (SAD) affects 6.8% of American adults annually, and its lifetime prevalence is 12.1% (Kessler, Berglund, et al., 2005; Kessler, Chiu, Demler, & Walters, 2005). SAD is characterized by a marked, persistent fear of scrutiny and humiliation and by avoidance of social and performance situations (American Psychiatric Association, 2000). Cognitive-processing biases figure in the onset and maintenance of SAD (Constans, Penn, Ihen, & Hope, 1999). Attentional (Amir, Foa, & Coles, 1998; Hope, Rapee, Heimberg, & Dombeck, 1990), memory (Foa, Gilboa-Schechtman, Amir, & Freshman, 2000; Kim, 2004), imagery (Hirsch, Meynen, & Clark, 2004; Makkar & Grisham, 2011; Wells & Papageorgiou, 1999), and interpretive biases (Amir et al., 1998; Constans et al., 1999; Stopa & Clark, 2000; Voncken, Bogels, & de Vries, 2003) for threat-relevant information are evident in SAD.

Studies on interpretive bias suggest that people with SAD construe neutral and ambiguous stimuli as more threatening than do non-socially-anxious individuals. For example, Niels-Christensen, Stein, and Means-Christensen (2003) found that socially anxious individuals evaluated themselves negatively, and erroneously believed that others judged them negatively in a social interaction, implying they seemingly lack an accurate view of how others view them, consistent with an interpretation bias. Other research has suggested that socially anxious people possess abnormal processing of positive stimuli, including fearful responses to favorable feedback (for a review, see Kashdan, Weeks, & Savostyanova, 2011) and the absence of a positive interpretation bias that nonanxious individuals possess (Hirsch...
& Mathews, 1997, 2000). It is also possible, however, that socially anxious individuals have difficulty comprehending the mental states of others irrespective of the valence or ambiguity of the stimulus. That is, they may be impaired in inferring and reasoning about others’ beliefs, emotions, and intentions, and hence in predicting their thoughts and actions, especially in social situations. The cognitive capacity to identify and reason about mental states in other people is called theory of mind (Premack & Woodruff, 1978).

A term coined by researchers studying the cognitive abilities of chimpanzees, theory of mind (ToM) is both a critical adaptation for social functioning and an important developmental milestone in humans. Sabbagh (2004) delineates two component processes of ToM: “(1) detecting or decoding others’ mental states based on immediately available observation information and (2) reasoning about those mental states in the service of explaining or predicting others’ actions” (p. 210). Decoding abilities refer to basic skills, such as identifying facial expressions or following eye gaze, whereas reasoning abilities require higher-order skills such as detecting sarcasm or inferring that someone is upset because they did poorly on a job review (Sabbagh, 2004; Washburn, 2012).

People with ToM deficits have difficulty evaluating others’ thoughts, and thereby experience social impairment that may contribute to functional impairment seen in autism (Baron-Cohen, 1995, 2005; Baron-Cohen, Leslie, & Frith, 1985; Frith, 1989) and schizophrenia (Brune, 2005; Corcoran, 2000; Couture, Penn, & Roberts, 2006). Likewise, it is possible that if individuals have trouble identifying and reasoning about others’ emotions and intentions that they may experience anxiety when in social situations. Social anxiety and ToM ability correlate inversely among people with schizophrenia spectrum disorders; one interpretation of this finding is that “… fully intact ToM capacities have a protective effect against paranoia or that high levels of social anxiety have a negative impact on ToM” (p. 84; Lysaker et al., 2010). Similarly, others (Samson, Lackner, Weiss, & Papousek, 2012) found that people with high levels of social anxiety rated cartoons requiring an understanding of others’ mental states (“ToM cartoons”), but not other cartoons, as less humorous than did people without social anxiety. However, only one study has examined ToM in individuals with diagnosed SAD (Washburn, 2012). This study found that nondepressed, socially anxious participants performed worse than nonanxious participants on one decoding measure of ToM, whereas individuals with comorbid depression and anxiety performed better on the task (one interpretation of enhanced decoding abilities in individuals with depression is that these individuals may be especially attentive to subtle social cues). However, the nondepressed, socially anxious group, which consisted of only nine participants, was rather small and hence perhaps underpowered to reveal group differences in the two other reasoning ToM tasks used in the study.

In the present study, we compared socially anxious and non-socially-anxious participants’ performance on socially relevant ToM tasks that require participants to both decode others’ emotions and reason about their mental states. We attempted to extend the findings of Washburn (2012) with the addition of a cognitive load condition. According to cognitive-behavioral models, people with SAD disproportionately allocate attentional resources to monitoring self-image, external threat, and personal expectations of how others will react to them in social situations (Rapee & Heimberg, 1997). This increased cognitive load, in turn, impairs performance on unrelated, complex cognitive tasks. Hope, Heimberg, and Klein (1990) found that socially anxious participants reported increased self-focused attention and performed less accurately than did nonanxious participants on a recall task following a social interaction. Hence, increased allocation of attentional resources may impair the processing of social information. If self-preoccupation impairs the ability to make accurate inferences about the mental states of other people, we would expect cognitive load to impair ToM ability in both non-socially-anxious comparison and socially anxious participants. To test this hypothesis, we gave half of the participants a memory task prior to completing the ToM measures. If, on the other hand, socially anxious participants have impairments in ToM irrespective of cognitive load, we would expect them to perform worse than non-socially-anxious participants in the no-load condition when they are presumably not self-monitoring.

Unlike interpretation bias paradigms, the tasks used in this study require that participants identify not only the emotions, but also the thoughts and intentions of others, irrespective of valence. By analyzing the errors that people make on these ToM tasks, one can discern whether participants’ errors are due to an interpretation bias (in which case we would expect that they would choose answers more negative in valence than the correct answer) as well as the extent to which they are taking the perspective of others. These ToM tasks are ecologically valid in that they require that participants make real-time assessments of what other people are thinking and feeling and why they
are acting a certain way. If individuals with social anxiety disorder do possess impairments in ToM, clinicians could potentially target these weaknesses in treatment. Research indicates that people with schizophrenia have benefited greatly from cognitive remediation, which aims to improve cognition and social cognition through different computerized exercises (McGurk, Twamley, Sitzer, McHugo, & Mueser, 2007; Wykes, Huddy, Cellard, McGurk, & Czobor, 2011).

Lastly, we examined the relation of IQ to ToM performance to ensure that differences in cognitive ability did not account for differential performance on the ToM tasks. Similarly, we measured depressive symptoms since some studies have identified less accurate ToM performance in depressed individuals (Lee, Harkness, Sabbagh, & Jacobson, 2005) whereas other studies have found enhanced ToM performance in dysthymic individuals (Harkness, Sabbagh, Jacobson, Chowdrey, & Chen, 2005).

Method

Participants

Participants were recruited from the Harvard University study pool, which consists of Harvard undergraduate students and adults living in metropolitan Boston, and from a Boston University job/volunteer site. The online postings included a brief description of the study and instructions that anyone, socially anxious or not, between the ages of 18 and 65 were welcome to participate. Individuals who signed up for the Harvard study pool have the option of completing a prescreen questionnaire. Three of the questions on the prescreen were specific to the current study and included items from the Mini Social Phobia Inventory, a self-report measure that assesses the extent to which a person fears and avoids social situations (Seeley-Wait, Abbott, & Rapee, 2009). In order to recruit additional socially anxious participants, we emailed individuals who scored six or higher on the scale, which is indicative of social anxiety, to invite them to take part in the study. All participants, irrespective of the results of the prescreen questionnaire, were assessed with the Mini International Neuropsychiatric Interview (Sheehan et al., 1998). The first author classified all participants into one of two groups: (1) those meeting DSM-IV criteria for SAD and (2) those without SAD (see Table 3 for information on the groups’ additional Axis I diagnoses). The SAD group comprised 40 participants (27 women) with a mean age of 26.5 years (SD = 11.9), and the non-SAD comparison group comprised 40 participants (34 women) with a mean age of 20.1 years (SD = 2.2). We excluded the data of 10 additional participants for various reasons, including suspected malingering and inability to complete the cognitive load task. With the remaining 80 participants, we had .94 power to detect large effects (Faul, Erdfelder, Lang, & Buchner, 2007). Harvard students received study pool credit, whereas others received $10 per hour for their participation. Data on race and ethnicity were not collected.

Materials

The Mini International Neuropsychiatric Interview (MINI) is a structured interview used to diagnose a range of current and lifetime Axis I disorders according to DSM-IV criteria (Sheehan et al., 1998). Administration takes approximately 15 minutes, and the MINI has good to very good concordance with the International Classification of Diseases and the Structured Clinical Interview for DSM-IV Diagnoses (SCID). Interrater reliability is excellent, with the majority of the scales having a kappa of .9 or higher; and test-retest reliability is very good, with most scales having a kappa of .75.

The Liebowitz Social Anxiety Scale Self-Report (LSAS) is a 24-item scale that accurately identifies the presence and severity of SAD (Fresco et al., 2001; Rytwinski et al., 2009). Participants indicate on a Likert Scale of zero to three the extent to which they fear and avoid 24 different social and performance situations (e.g., eating in public, speaking to an authority figure, etc.). Scores range from 0 to 144, with higher scores signifying greater social anxiety; a score of 60 indicates generalized SAD (Mennin et al., 2002; Rytwinski et al., 2009). The LSAS has high internal consistency (α = .95), strong convergent and discriminant validity, and good test-retest reliability (r = .83, p < .01) (Baker, Heinrichs, Kim, & Hofmann, 2002).

The Center for Epidemiologic Studies Depression Scale, Revised (CESD) scale consists of 20 items that assess the frequency at which individuals have experienced symptoms of depression over the prior week (Eaton, Smith, Ybarra, Muntaner, & Tien, 2004). Scores ranging from 0 to 60 are calculated by adding item responses (four items are reversed scored), and 16 is the suggested clinical cutoff score for depression. The CESD has high internal consistency (α = .90), acceptable test-retest reliability (r = .57), and good discriminant and concurrent validity as measured by correlations with self-report measures (r = .74) and clinical interviews of depression (r = .46; Radloff, 1977).

The American National Adult Reading Test (NART) requires participants to read aloud a list of 50 short words of irregular pronunciation. The number of pronunciation errors are tallied and
used to estimate aspects of IQ. The NART validly estimates general IQ, verbal IQ, and performance IQ (Crawford, Parker, Stewart, Besson, & De Lacy, 1989), and has very high test-retest reliability \((r = .98)\), very high interrater reliability (ranging from \(r = .96\) to \(r = .98\)), and high split-half reliability \((r = .90\) to \(.93\); Crawford et al., 1989). Administration takes approximately 2 to 3 minutes. NART scores were not used for eight participants (five SAD and three non-SAD) who were not native English speakers.

The Wechsler Adult Intelligence Scale–Fourth Edition (WAIS) measures the cognitive ability, or IQ, of adults. In the present study, we used Similarities, a verbal comprehension task, and Matrix Reasoning, a perceptual organization task (Wechsler, 2008). Both Similarities and Matrix Reasoning have good to excellent internal consistency \((r = .87\) and \(.90\), respectively), good test-retest stability \((r = .83\) and \(r = .76\), respectively), excellent interrater reliability \((r = .93\) and \(.98\), respectively), and high convergent and discriminant validity.

The Reading the Mind in the Eyes (MIE) is a decoding theory of mind task that consists of photographs of the eye-regions of actors and actresses (Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001). Individuals are asked to choose which of four words best describes what the person in the picture is thinking or feeling, as determined by the test-maker, earning one point per correct answer. This ToM test comprises two 18-question parts (Part A and Part B) and scores range from 0 to 18 (on each part), with low scores indicating impaired decoding abilities. The Mind in the Eyes is an advanced ToM task as individuals must identify complex emotions by relying on a limited amount of information (from the eyes only). The Mind in the Eyes has good test-retest reliability with an intraclass correlation coefficient of \(.83\) and good internal consistency with a Cronbach’s alpha of \(.61\) (Vallante et al., 2013), and can reliably detect subtle differences in social cognitive abilities (Baron-Cohen et al., 2001). Following others (Harkness et al., 2005), we classified all possible answers (both correct and incorrect) on this task in terms of accuracy and valence. Specifically, 12 members of our lab rated each emotion on a 1 (negative) to 7 (positive) scale, yielding a mean intraclass reliability correlation coefficient of \(.98\). We then conducted one-sample \(t\)-tests to determine what words were significantly different from neutral \((\alpha = .05)\).

Words that had a significantly higher mean than 4 were classified as positive, words that had a mean significantly lower than 4 were classified as negative, and words that did not differ from 4 were classified as neutral. We determined that 20 of the eyes depict a negative valence, 13 a positive valence, and 3 a neutral valence (see Table 1).

The Movie for the Assessment of Social Cognition (MASC) is a reasoning ToM task that consists of a 15-minute video depicting four actors and actresses interacting (Dziobek et al., 2006). Throughout the task, the film is paused and participants are asked to answer a total of 45 questions requiring them to identify the characters’ feelings, thoughts, and intentions (e.g., “What is Sandra feeling?”). Correct answers receive 1 point and incorrect answers are scored in one of three ways: (1) no ToM, meaning the answer was based on some physical aspect (e.g., “her hair does not look that nice”); (2) less ToM signifying an insufficient response that misses a crucial aspect of the social situation (i.e., a “half-right” answer [e.g., “she is pleased about his compliment”]); and (3) excessive ToM, signifying reading too much into the character’s state of mind (e.g., “she is exasperated about Michael coming on too strong”). Identifying the types of errors people make on this task enables one to evaluate the extent to which participants are taking the perspective of the characters in the film. The task also includes six additional control questions assessing understanding of non-socially-relevant aspects of the plot and characters; these questions are totaled separately from the overall MASC score. The MASC has high internal consistency \((\alpha = .84)\) and strong test-retest reliability \((ICC = .97)\) and high

<table>
<thead>
<tr>
<th>Word</th>
<th>(t(11))</th>
<th>(p)</th>
<th>Word</th>
<th>(t(11))</th>
<th>(p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upset</td>
<td>-16.58</td>
<td>&lt;.001</td>
<td>Doubtful</td>
<td>-5.00</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Insisting</td>
<td>-3.02</td>
<td>.01</td>
<td>Tentative</td>
<td>-2.35</td>
<td>.04</td>
</tr>
<tr>
<td>Worried</td>
<td>-11.00</td>
<td>&lt;.001</td>
<td>Defiant</td>
<td>-2.87</td>
<td>.02</td>
</tr>
<tr>
<td>Uneasy</td>
<td>-5.93</td>
<td>&lt;.001</td>
<td>Hostile</td>
<td>-16.58</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Despondent</td>
<td>-9.57</td>
<td>&lt;.001</td>
<td>Cautious</td>
<td>-2.57</td>
<td>.03</td>
</tr>
<tr>
<td>Preoccupied</td>
<td>-4.18</td>
<td>.002</td>
<td>Serious</td>
<td>-2.80</td>
<td>.02</td>
</tr>
<tr>
<td>Cautious</td>
<td>-2.57</td>
<td>.03</td>
<td>Distrustful</td>
<td>-17.23</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Regretful</td>
<td>-11.00</td>
<td>&lt;.001</td>
<td>Nervous</td>
<td>-7.34</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Skeptical</td>
<td>-3.55</td>
<td>.005</td>
<td>Suspicious</td>
<td>-11.00</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Accusing</td>
<td>-10.34</td>
<td>&lt;.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anticipating</td>
<td>1.82</td>
<td>.10</td>
<td>Concerned</td>
<td>-1.48</td>
<td>.17</td>
</tr>
<tr>
<td>Pensive</td>
<td>0</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Playful</td>
<td>12.54</td>
<td>&lt;.001</td>
<td>Friendly</td>
<td>10.58</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Desire</td>
<td>5.20</td>
<td>&lt;.001</td>
<td>Interested</td>
<td>4.73</td>
<td>.001</td>
</tr>
<tr>
<td>Fantasizing</td>
<td>4.53</td>
<td>&lt;.001</td>
<td>Reflective</td>
<td>3.92</td>
<td>.002</td>
</tr>
<tr>
<td>Contemplative</td>
<td>2.35</td>
<td>.04</td>
<td>Flirtatious</td>
<td>6.28</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Thoughtful</td>
<td>8.86</td>
<td>&lt;.001</td>
<td>Confident</td>
<td>12.54</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Decisive</td>
<td>2.97</td>
<td>.01</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
convergent validity with other measures of social cognition (Dziobek et al., 2006).

**Procedure**
Participants provided written informed consent as approved by Harvard’s Committee on the Use of Human Subjects and in compliance with the Helsinki Declaration of 1975. After being categorized as either socially anxious or not, participants were randomly assigned to either a cognitive load or no cognitive load condition. We achieved random assignment by using an Excel sheet that was programmed to randomly assign group membership; at no time were the investigators able to predict what the next assignment would be.

Participants first completed the CESD scale and the LSAS, followed by Part A of Mind in the Eyes, which was used as a baseline measure of participants’ performance. Participants completed these tasks, which took approximately 15 minutes combined, on a computer that was facing away from the examiner (as to reduce any anxiety participants may feel when completing the measures). The first author then administered the Similarities and Matrix Reasoning subscales of the WAIS-IV and the NART. Administration of these cognitive ability assessments took approximately 20 minutes. Finally, participants completed Part B of the MIE and the MASC on the same computer as the other measures.

Participants in the load condition received a memory task before completing Part B of the MIE and the MASC. Specifically, before the MIE task, they had 20 seconds to memorize a sequence of eight symbols (e.g., * ? = \ > & ! #) and were instructed to rehearse them aloud while completing the task (Gilbert & Osborne, 1989; van den Bos, Peters, Bobocel, & Ybema, 2006). When they finished the test, participants repeated the symbols back to the first author and then rated how difficult they found the memory task on a scale of 1 to 10.

**Results**

**Preliminary Results**
The groups did not differ significantly in number of men and women, $\chi^2(1) = 3.38, p = .07$, or in cognitive ability (Table 2). Relative to non-socia

<table>
<thead>
<tr>
<th></th>
<th>SAD (M, SD)</th>
<th>Non-SAD (M, SD)</th>
<th>t(df)</th>
<th>$p$</th>
<th>$r$</th>
</tr>
</thead>
<tbody>
<tr>
<td>LSAS (anxiety severity)</td>
<td>72.48 (22.35)</td>
<td>26.03 (16.37)</td>
<td>t(78) = 10.60</td>
<td>&lt;.001</td>
<td>.77</td>
</tr>
<tr>
<td>CESD (depression severity)</td>
<td>20.70 (13.88)</td>
<td>6.35 (6.63)</td>
<td>t(78) = 5.90</td>
<td>&lt;.001</td>
<td>.56</td>
</tr>
<tr>
<td>NART: full-scale IQ</td>
<td>118.23 (5.10)</td>
<td>118.38 (3.66)</td>
<td>t(70) = 14</td>
<td>.89</td>
<td>.02</td>
</tr>
<tr>
<td>NART: verbal IQ</td>
<td>118.63 (5.35)</td>
<td>118.97 (4.02)</td>
<td>t(70) = 31</td>
<td>.76</td>
<td>.04</td>
</tr>
<tr>
<td>NART: performance IQ</td>
<td>113.31 (3.7)</td>
<td>113.51 (2.71)</td>
<td>t(70) = 26</td>
<td>.79</td>
<td>.03</td>
</tr>
<tr>
<td>WAIS-IV: Similarities</td>
<td>10.73 (2.21)</td>
<td>11.33 (1.79)</td>
<td>t(78) = 13.4</td>
<td>.19</td>
<td>.15</td>
</tr>
<tr>
<td>WAIS-IV: Matrix Reasoning</td>
<td>10.75 (2.73)</td>
<td>11.58 (2.76)</td>
<td>t(78) = 1.34</td>
<td>.90</td>
<td>.15</td>
</tr>
</tbody>
</table>

*Note. $^* = p \leq .05$ criteria; $^1$ = scaled scores.

LSAS = Liebowitz Social Anxiety Scale; CESD = Center for Epidemiologic Studies – Depression; NART = National Adult Reading Test; WAIS-IV = Weschler Adult Intelligence Scale, Fourth Edition.
Five of the nine individuals in the non-socially anxious group remaining four individuals met criteria for recurrent MDD. The remaining 20 individuals met criteria for a single past major depressive episode; the action fell short of significance, $r = .22$, whereas the Group × Cognitive Load interaction fell short of significance, $F(1, 76) = 2.67, p = .11, r = .18$ (see Figure 1). Likewise, there were no significant main effects or interactions for the repeated measure (MIE at time one versus MIE at time two).

To evaluate the type of errors participants made on the Mind in the Eyes task, we used our classification of the MIE items to determine if valence was related to group performance. A repeated measure ANOVA showed an interaction effect between group and the valence of the eyes in question. Using follow-up $t$-tests, we found that participants with SAD made significantly more errors than did comparison participants on questions about negative, $t(78) = 3.40, p = .001, r = .36$, valenced sets of eyes, and this difference remained significant after we applied a Bonferroni correction for multiple comparisons ($p < .02$). The groups did not differ, however, in the number of errors they made in response to positive, $t(78) = .24, p = .81, r = .03$, or neutral, $t(78) = .72, p = .47, r = .08$, valenced expressions. We categorized each incorrect answer on the MIE as more negatively valenced, more positively valenced, or the same valence as the correct answer. For example, if the correct answer for a particular item was positive in valence, but the participant chose an incorrect answer that was either neutral or negative in valence, this would be coded as a “more negative error.” Alternatively, if a participant chose an incorrect answer that was more positive in valence than the correct answer, this was coded as a “more positive error.” Finally, if someone chose an incorrect answer that was the same valence as the correct answer (e.g., the correct answer is negative in valence and the person chose an incorrect answer that was also negative in valence), this was coded as a “same valence error.” After a repeated-measures ANOVA indicated a significant interaction between group and error type, we performed follow-up $t$-tests, which revealed no group differences in the number of more positive, $t(78) = .39, p = .70, r = .04$, or more negative answers, $t(78) = .61, p = .54, r = .07$. However, socially anxious participants chose significantly more incorrect answers that were the same valence as the correct answer, $t(78) = 3.5, p = .001, r = .37$. This difference remained significant after we corrected for multiple comparisons (Bonferroni corrected $p < .02$).

Analysis of MASC scores (with a second 2 × 2 ANOVA) showed a similar pattern of findings to the MIE, though performance on the two tasks was uncorrelated when we controlled for the presence of SAD (partial $r = .17, p = .13$). Participants with SAD were less accurate on the task than were non-socially-anxious participants, $F(1, 76) = 9.37, p = .003, r = .33$. Moreover, participants in the cognitive load condition performed worse overall on the MASC than did participants under no load.

### Table 3

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>Non-SAD n (%)</th>
<th>SAD n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Major Depressive Disorder Current</td>
<td>1 (2.5%)</td>
<td>0</td>
</tr>
<tr>
<td>Lifetime</td>
<td>9 (22.5%)**</td>
<td>26 (67.5%)*</td>
</tr>
<tr>
<td>Generalized Anxiety Disorder</td>
<td>1 (2.5%)</td>
<td>13 (32.5%)</td>
</tr>
<tr>
<td>Panic Disorder (with and without Agoraphobia)</td>
<td>1 (2.5%)</td>
<td>12 (30%)</td>
</tr>
<tr>
<td>Substance Abuse/Dependence</td>
<td>1 (2.5%)</td>
<td>6 (15%)</td>
</tr>
<tr>
<td>Bipolar Disorder (I or II)</td>
<td>2 (5%)</td>
<td>3 (7.5%)</td>
</tr>
<tr>
<td>Post Traumatic Stress Disorder</td>
<td>0</td>
<td>2 (5%)</td>
</tr>
<tr>
<td>Agoraphobia (without Panic Disorder)</td>
<td>7 (18%)</td>
<td>1 (2.5%)</td>
</tr>
<tr>
<td>Obsessive Compulsive Disorder</td>
<td>0</td>
<td>1 (2.5%)</td>
</tr>
<tr>
<td>Bulimia</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
| **Six of the 26 individuals in the SAD group met criteria for a single past major depressive episode; the remaining 20 individuals met criteria for recurrent MDD.**

**Five of the nine individuals in the non-socially anxious group met criteria for a single past major depressive episode; the remaining four individuals met criteria for recurrent MDD.

---

**Figure 1** Performance on Reading the Mind in the Eyes Revised. Note. Dotted line represents mean score of subjects with Asperger Syndrome or High Functioning Autism (Baron-Cohen et al., 2001).
there was an association between depressive symptoms, as measured by the CESD, and performance on the ToM. Depression was not significantly correlated with performance on either the MIE \( (r = -.03, p = .78) \) or the MASC \( (r = -.03, p = .82) \).

We compared participants’ difficulty ratings of the memory task (range: 1–10) to determine if the task successfully taxed working memory for both the SAD and the non-SAD groups. Collectively, the participants gave the first memory task a mean difficulty rating of 7.82 and the second task a mean rating of 7.70. The groups did not differ on ratings for task one, \( t(36) = .58, p = .57, r = .13 \), or task two, \( t(36) = .56, p = .58, r = .12 \), nor did the groups differ in how accurately they recalled the symbols of memory task one, \( t(39) = .58, p = .57, r = .12 \), or two, \( t(35) = .42, p = .26, r = .11 \).

Finally, to test the association between different factors that may be associated with ToM performance, we examined the relation between ToM accuracy and cognitive ability and sex. Scores on the NART and Similarities subtest predicted performance on the MIE \( (r = .23, p = .05 \) and \( r = .22, p = .06 \), respectively), but not on the MASC. Finally, men performed worse on the MASC than did women, \( t(78) = 3.14, p = .002, r = .33 \), though SAD still significantly predicted performance when we controlled for sex in a simultaneous multiple regression \( (\beta = -2.34, t(76) = 2.53, p = .014, r = .28) \). There was no difference between men and women’s performance on the MIE, \( t(78) = .27, p = .79, r = .09 \). Sex did not significantly interact with group (SAD vs. non-SAD) on either the MIE, \( F(1, 76) = .40, p = .53 \), or the MASC, \( F(1, 76) = .06, p = .82 \).

**Discussion**

Participants with SAD, irrespective of cognitive load, performed worse than did non-socally-anxious participants on two ToM tasks, partly replicating Washburn (2012), who found that nondepressed, socially anxious participants performed worse than nonanxious participants on the MIE (but not on the MASC). Though participants with SAD had higher scores than autism spectrum participants typically do (Figures 1 and 2; Baron-Cohen et al., 2001; Dziobek et al., 2006), they do have difficulty understanding others’ thoughts and emotions relative to individuals without SAD. That socially anxious and comparison participants did not differ in accuracy on the MASC’s control questions indicates that this impairment is specific to social information. The absence of group differences in cognitive ability further suggests that anxious participants did not have general difficulty completing laboratory tasks.
The fact that the socially anxious group performed worse than non-socially-anxious participants irrespective of condition suggests that this impairment is not due to increased cognitive load associated with self-monitoring processes. Although participants in the load condition performed worse on the MASC, the MASC control questions, and the MIE task than did individuals in the no load condition, there were no interactions between group and condition; this finding suggests that increasing cognitive load does not have differential effects on ToM performance for people with and without SAD. Hope, Heimberg, et al. (1990) found that memory impairments in socially anxious individuals occurred after social interaction and a subsequent increase in self-focused attention. Mellings and Alden (2000) likewise found that “social interaction tasks may lend themselves to encoding biases in socially anxious individuals whereas semantic memory tasks may not” (p. 256). Although these studies indicate social information interferes with memory recall, we found that a nonsocial memory task interfered with ToM accuracy. Alternatively, participants may have viewed the study itself as a performance situation because their accuracy on these tasks was evaluated. Any situation involving scrutiny by others may pose an evaluative threat to someone with SAD (Rapee & Heimberg, 1997; Washburn, 2012). If our participants were anxious during the experiment, then this may have impaired their performance on the ToM tasks. However, the only tasks that were administered by the examiner were the WAIS subscales and the NART, which were used to assess cognitive ability; all other tasks were completed on a computer that was facing away from the examiner in order to reduce any anxiety associated with evaluation of their answers. That socially anxious participants did not perform any worse on the cognitive ability tasks suggests that any anxiety they experienced did not adversely affect their general performance on study tasks.

Interestingly, socially anxious participants made more “excessive ToM” errors on the MASC than did non-socially-anxious participants. That is, socially anxious people were more likely to attribute more intense emotions and greater meaning to what the characters in the movie were thinking, feeling, and intending. Notably, their impairments in ToM are in the opposite direction of those in people with autism spectrum conditions whose inferences about the mental states of other people are limited, at best. Accordingly, individuals with SAD are neither unable nor unwilling to make inferences about the thoughts of others. Rather, they read too much into how others are feeling, thereby misunderstanding social situations. Importantly, excessive ToM responses are not simply more negative or catastrophic interpretations of the situations, which would suggest a negative interpretation bias. Rather, they indicate that participants are incorrectly attributing beliefs—negative, positive, and neutral—and intentions to others when it is not contextually appropriate (Sharp et al., 2001). Perhaps they imagined themselves in the social situations in the movie and (incorrectly) based their answers on how they would feel in those situations rather than on how the characters in the film actually felt. This interpretation is speculative and merits further investigation. However, Tibi-Elhanany and Shamay-Tsoory (2011) discovered that, compared to those with low levels of social anxiety, highly socially anxious people showed an elevated tendency to adopt others’ perspectives (i.e., “cognitive empathy”), but decreased accuracy on a cognitive measure of ToM. The authors attributed this apparent discrepancy to anxious participants’ tendency to “over-mentalize” (p. 104) about others’ mental states. Our results support this interpretation.

Some studies on emotion recognition show that socially anxious participants are more accurate than are non-socially-anxious participants at identifying negative emotions such as fear (Arrais et al., 2010; Hirsch & Clark, 2004; Surcinelli, Codispoti, Montebarocci, Rossi, & Baldaro, 2006), whereas other studies do not (Merckelbach, van Hout, van den Hout, & Mersch, 1989; Philippot & Douilliez, 2005). However, our findings indicated that socially anxious participants made more errors than did comparison participants when the eyes depicted in the task had a negative valence. Cognitive models suggest that individuals with SAD exhibit an initial attentional bias towards, and subsequent avoidance of, threatening faces (Mathews & MacLeod, 2005). In our study, the MIE task required that a person answer each question before proceeding to the next item, thus making avoidance of any of the eyes very difficult. However, socially anxious participants may have sacrificed accuracy to answer quickly and avoid looking closely at the negative expressions. Interestingly, these errors did not arise from a negative interpretive bias in the SAD group, nor the absence of a positive interpretive bias. Indeed, participants with SAD were no more likely than were non-socially-anxious participants to choose a more negative answer on the MIE; nor were non-socially-anxious participants more likely to choose a more positively valenced response. Rather, the SAD group made more errors in which they chose an emotion having the same valence as the correct answer. This finding suggests that individuals with SAD possess impairments in decoding and reasoning about
others’ mental states beyond established interpretation biases.

Unsurprisingly, participants with SAD were more often depressed than non-socially-anxious participants. Although ToM deficits appear in depressed patients (Inoue, Tonooka, Yamada, & Kanba, 2004; Inoue, Yamada, & Kanba, 2006), depressive symptoms were unrelated to ToM performance in the present study. Similarly, intelligence did not predict accuracy on the MASC, consistent with prior research (Dziobek et al., 2006). Thus, ToM deficits in SAD individuals on the MASC were attributable to an impairment specific to social information. Conversely, cognitive ability, as measured by the NART and the verbal reasoning subtest of the WAIS, predicted performance on the MIE, inconsistent with research showing a nonassociation between intelligence and accuracy on this task (Baron-Cohen et al., 2001). Unlike the MASC, the MIE requires participants to choose the correct answer from several one-word choices that require a sophisticated vocabulary (e.g., *aghast*). Although participants could request definitions of words, socially anxious people are likely too embarrassed to do so. Consistent with this interpretation, only verbal IQ predicted accuracy.

Although we had few male participants, they made more errors than women did on the MASC, consistent with other studies that have found lower accuracy in men on ToM tasks (Baron-Cohen, 2002; Baron-Cohen & Hammer, 1997). One potential theory for sex differences is Baron-Cohen and Hammer’s (1997) male brain model of autism, which holds that women’s brains are “more ‘social’ and less ‘spatial,’” whereas men’s brains are “more spatial and less social” (p. 198). This nativist explanation asserts that sex differences in the processing of social information exist from birth. However, recent research suggests that the effect of sex on ToM ability might be weaker than previously thought, and that what differences do exist may be partly due to developmental factors, such as parenting style and language development (Charman, Ruffman, & Clements, 2002). When we controlled for sex, SAD continued to predict performance on the MASC and there were no significant interactions between group and sex. Our findings suggest that sex did not drive group differences on ToM tasks.

Consistent with Dziobek et al. (2006), we found that performance on the MIE was uncorrelated with performance on the MASC. The tasks seem to tap different ToM processes. Understanding others requires diverse skills, including accurate facial and emotion recognition and an understanding of gist and of cognitive and affective mental states. The MIE is a static task requiring participants to identify emotions based on very limited information (i.e., a decoding task), whereas the MASC is a dynamic task providing context, facial expression, inflection, content of speech, and body language of the characters (i.e., a reasoning task). If individuals miss some information (e.g., facial expression), they can still answer the question correctly based on other clues (e.g., inflection). These differences may explain why certain variables, such as sex and IQ, were associated with one task but not the other.

Our study has limitations, including the absence of a non-SAD, anxious comparison group. However, analyses eliminating participants who met criteria for other anxiety disorders did not alter the results, thereby suggesting a specific association between SAD and ToM difficulties. Similarly, because individuals in both the SAD group and the non-SAD group met criteria for other Axis I disorders, it is unclear whether other psychopathology contributed to group differences. Although our study suggests an association between SAD and lower performance on ToM tasks, the relation may extend to other anxiety disorders. Future studies might benefit from comparing individuals with SAD to individuals with no history of psychopathology. Furthermore, our results showed significant, but small, group differences on ToM performance. It is unclear how these differences may contribute to the clinical presentation of SAD. Less accurate mind-reading abilities may foster misunderstandings of social cues and subsequent heightened anxiety in individuals with SAD. Finally, we did not record the ethnicity of our participants, and so we cannot rule out the possibility that cultural diversity may have affected their interpretation of the social information presented in the tasks.

Additional research is necessary to clarify the relationship between SAD and ToM. Our findings are a partial replication of Washburn (2012), who found socially anxious individuals (in the absence of depression) performed worse on the MIE than did participants without social anxiety. However, the author did not find any association between the valence of the question and accuracy, nor did he find group differences on the MASC. Hence, it is important not only to replicate our findings, but also to identify the specific weaknesses in ToM that individuals with SAD possess. It would be useful to examine whether socially anxious people show deficits on both affective and cognitive ToM tasks, or if they show similar impairments when on nonvisual media (e.g., written or audio ToM tasks). Similarly, prospective studies that investigate whether ToM impairments are a cause or an effect of the disorder may have important implications for etiology and treatment. If further research also shows that socially anxious individuals have impairments in ToM, we may be...
able to improve these skills through novel treatment techniques such as cognitive remediation therapy.

Conflict of Interest Statement
The authors have no interests to declare.

References


Received: September 9, 2013
Accepted: February 20, 2014
Available online 1 March 2014