When ambiguity hurts: Social standards moderate self-appraisals in generalized social phobia

David A. Moscovitch*, Stefan G. Hofmann

Department of Psychology, Boston University, USA

Received 25 January 2006; received in revised form 17 June 2006; accepted 11 July 2006

Abstract

Thirty-nine individuals with generalized social phobia (social anxiety disorder) and 39 nonclinical controls performed a public speech after receiving cues about social standards. Using a novel video manipulation paradigm, one third of participants received cues indicating that standards for performance were high, one third received cues that standards were low, and the remaining third were given no explicit information about expected standards (i.e., standards were ambiguous). Individuals with social phobia performed objectively worse than controls in all conditions, but rated their performance as being worse only in the high and ambiguous standards conditions. These results suggest that in social phobia, negative self-perception is context-dependent. Implications for the cognitive model and treatment are discussed.

© 2006 Elsevier Ltd. All rights reserved.

Keywords: Social anxiety disorder; Self-appraisals; Self-discrepancy; Self-perception; Information processing

Introduction

Social anxiety is thought to arise when individuals are highly motivated to make a good impression on others in social situations but believe they are incapable of doing so (Gilbert, 2001; Leary, 2001; Leary & Kowalski, 1995; Schlenker & Leary, 1982). Cognitive theories suggest that negative perception of self plays a central role in the maintenance of social phobia (Clark & Wells, 1995; Rapee & Heimberg, 1997). On the basis of early learning experiences, individuals with social phobia develop a number of negative assumptions about themselves (e.g., “I’m stupid,” “I’m unattractive,” etc.; Clark & Wells, 1995) that become reinforced over time by selective information processing errors, which occur both within and between social encounters (see Clark & McManus, 2002; Heinrichs & Hofmann, 2001; Hirsch & Clark, 2004). When faced with social threat, individuals with social phobia shift their attention inward and engage in a process of detailed self-monitoring (Spurr & Stopa, 2002), during which they experience spontaneous, recurrent, and excessively negative self-images that they perceive as being accurate (Hackmann, Clark, & McManus, 2000; Hackmann, Surawy, & Clark, 1998).

*Corresponding author. Department of Psychology, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1.
Tel.: +1 519884567x32549; fax: +5197468631.
E-mail address: dmosco@uwatwater.ca (D.A. Moscovitch).

0005-7967/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.brat.2006.07.008
Following a social encounter, socially phobic individuals appraise their own behavior in a manner that greatly exaggerates their shortcomings and minimizes their performance accomplishments (Alden & Wallace, 1995; Norton & Hope, 2001; Rapee & Lim, 1992; Stopa & Clark, 1993). But what is the proximal cause of these negative self-appraisals? From the perspective of the cognitive model, individuals with social phobia form negative mental self-representations based on how they believe the “implicit audience” views them at any given moment (e.g., Rapee & Heimberg, 1997). This intriguing theory suggests that negative self-appraisals are flexible, dynamic, responses that are dependent on the nature and type of information that is available about potential audience evaluators and their standards for performance. Previous studies have demonstrated that patients with social phobia perceive their self-attributes to fall short of the characteristics they believe others expect them to possess (e.g., Strauman, 1989, 1992; Strauman & Higgins, 1987; Weilage & Hope, 1999). Furthermore, socially anxious or phobic individuals under social threat experience state self-discrepancies that are characterized by an underestimation of their abilities relative to others’ standards (Alden, Bieling, & Waller, 1994; Wallace & Alden, 1991). To our knowledge, however, no studies have directly examined the hypothesis that negative state self-appraisals are activated by the nature and type of information that is available about audience performance standards.

Evidence suggests that socially anxious and phobic individuals are attuned to the standards that others hold for their social behavior and that perceived standards may, in turn, influence self-perception, affect, and performance. Baldwin and Main (2001) found that implicit contextual cues signifying “rejection” may activate cognitive networks that are associated with negative self-judgments in high self-conscious female undergraduates. Others have found that socially anxious individuals might employ the self-presentational strategy of purposeful failure as a way to influence others to lower their performance standards to a level they can more confidently match (Baumgardner & Brownlee, 1987). Finally, Wallace and Alden (1995, 1997) found that individuals with social phobia who received unexpected positive feedback following a social encounter reported increased anxiety about a future interaction because they believed that their interaction partners would subsequently hold higher standards for their performance.

In the present study, our objectives were twofold. First, we wished to examine more closely whether individuals with social phobia believe that audience evaluators hold excessively high standards for social performance. Although patients with social phobia tend to report high levels of socially-prescribed perfectionism on trait self-report questionnaires (e.g., Antony, Purdon, Huta, & Swinson, 1998), previous studies examining social phobic beliefs that are activated during actual social situations have found, surprisingly, that patients’ estimations of others’ standards do not exceed those of nonanxious controls (Alden et al., 1994; Wallace & Alden, 1991). However, studies reporting this null effect did not control for the possibility that socially anxious participants may have assumed that audience members were aware that they were socially anxious. Thus, anxious participants may have rated perceived standards in line with their belief that audience members expected them to perform poorly as a result of their anxiety and, therefore, held low expectations for them. In the present study, we attempted to control for this possible confound by informing participants that audience members “will not know anything about you or the purpose of the experiment.” We expected that under these conditions, individuals with social phobia would provide ratings of audience standards that were significantly higher than those of controls. We also expected that individuals with social phobia would provide lower predicted ratings of their own performance in comparison to controls.

Second, we examined whether self-appraisals and social performance are moderated by the nature and type of information presented to individuals about social standards. To answer this question, we randomly assigned patients with social phobia and healthy controls to one of three conditions that differed according to the nature and availability of information about social standards. Prior to performing a socially threatening task, one third of participants were exposed to cues indicating that standards for performance were high, one third were exposed to cues indicating that standards were low, and the remaining third were given no explicit information about expected standards (i.e., standards were ambiguous). We expected the self-appraisals, performance, and affect of individuals with social phobia to be more negative than controls in the high standards condition relative to the low standards condition. In addition, since previous studies have demonstrated that individuals with social phobia tend to interpret ambiguous social information in a negative manner (e.g., Amir, Foa, & Coles, 1998; Stopa & Clark, 2000), we expected patients’ self-appraisals, affect, and performance in the no standards condition to resemble those of patients in the high standards condition.
Finally, we wished to explore the accuracy of participants’ performance appraisals across conditions by comparing the effects of group and condition on self versus observer ratings. Although previous studies have shown that individuals with social phobia tend to appraise their social performance less accurately than controls (when self ratings are compared to those of objective observers), we did not know how our experimental manipulation might moderate the accuracy of participants’ appraisals and, therefore, we did not predict a specific pattern of results for these analyses.

Method

Participants

The sample consisted of 39 individuals with generalized social phobia and 39 healthy control participants. Most of the clinical participants (N = 29) were recruited through the Center for Anxiety and Related Disorders, a large community outpatient clinic at Boston University. The remaining clinical participants (N = 10) and all control participants were recruited through the community via newspaper, Internet, and flyer advertisements. In exchange for their participation, community participants were paid $40.00, while outpatients received a treatment voucher valuing $40.00.

Trained evaluators assessed all individuals prior to their participation in the study. Outpatients were evaluated with the Anxiety Disorders Interview Schedule for DSM-IV: Lifetime Version (ADIS-IV-L; DiNardo, Brown, & Barlow, 1994), as part of the standard intake procedure of our center. Outpatient recruitment, interview, and interviewer training procedures were identical to those reported in Brown, DiNardo, Lehman, and Campbell (2001).

Non-outpatient clinical participants and all nonclinical participants were pre-screened over the phone and subsequently evaluated with the Mini-Anxiety Disorders Interview Schedule for DSM-IV (Mini-ADIS; Brown, DiNardo, & Barlow, 1994). The Mini-ADIS differs from the ADIS-IV-L only in its omission of the assessment of past diagnoses. The Mini-ADIS rather than the ADIS-IV-L was used to assess non-outpatients because past diagnostic history did not influence the decision to include or exclude participants from the present study.

For every diagnosis, ADIS interviewers assigned a 0–8 clinical severity rating (CSR) to indicate the degree of distress and impairment associated with the disorder (0 = “none” to 8 = “very severely disturbing/disabling”). An ADIS CSR of 4 or higher reflected the presence of a condition that was clinically significant in terms of distress or impairment. The mean CSR rating for the principal diagnosis (generalized social phobia) in the clinical sample was 5.85 (SD = .67, range = 5–7). There were 45 total comorbid diagnoses, with 15 clinical participants receiving one additional diagnosis, 9 receiving two additional diagnoses, and 4 participants receiving three additional diagnoses. The most frequent comorbid diagnoses were major depressive disorder (n = 10), dysthymia (n = 8), generalized anxiety disorder (n = 6), panic disorder with agoraphobia (n = 4), specific phobia (n = 4), and obsessive-compulsive disorder (n = 3). All diagnoses were assigned by upper-level doctoral students and corroborated by senior clinicians.

Participants who met DSM-IV diagnostic criteria for a principal diagnosis of generalized social phobia were eligible to participate in the study if they were also not actively psychotic, manic, suicidal, homicidal, or substance-abusing, were not currently receiving psychotherapeutic treatment for social phobia, and if taking psychotropic medications, had not altered medication type or dose within the last month. Nineteen eligible clinical participants declined participation. The most common reasons for declining participation were too busy to participate (n = 8) and apprehension about some aspect of the study procedure (n = 7).

Control participants were included in the study if they had no current DSM-IV disorders and endorsed no significant symptoms of social phobia (i.e., reported fear levels below a CSR of 3 out of 8 on all of the social situations listed in the ADIS). Although lifetime history of psychopathology was not systematically assessed in control participants, those who reported that they were diagnosed in the past with a mental disorder were included in the study provided that they did not meet current diagnostic criteria for that (or any other) disorder.

Of the overall sample, the majority was male (61.5%) and average age was 30.62 (SD = 10.30, range = 18–63). Participants were predominantly Caucasian (67.9%), with remaining participants identifying...
as Asian (10.3%), African-American (9.0%), Hispanic (2.6%), and Other (e.g., biracial; 10.3%). Median annual income was 25,000, with 31 individuals (39.7%) earning 18,000 or less, 22 individuals (28.2%) earning $20,000–$39,000, 16 individuals (20.5%) earning $40,000–$60,000, and 4 individuals (5.1%) earning above $60,000 per year. Participants completed an average of 16.06 (SD = 2.32; range = 11–23) years of education.

Socially phobic and control participants did not differ in age, t (76) = .50, p = .62, years of education, t (76) = .51, p = .12, or ethnicity, \( \chi^2 (6, N = 78) = 3.11, p = .80 \). Each group was comprised of 24 men and 15 women. Eight men and five women from each group were randomly assigned to each of the three experimental conditions. Participants in the three experimental conditions did not differ according to age, F (2, 75) = .45, p = .64, years of education, F (2, 75) = .86, p = .43, or ethnicity, \( \chi^2 (12, N = 78) = 8.57, p = .74 \). Across the three conditions, participants with social phobia did not differ according to comorbid depression status (including MDD, dysthymia, and depressive disorder, not otherwise specified), \( \chi^2 (2, N = 39) = 3.28, p = .19 \), or in their level of depression symptom severity as measured on the Beck Depression Inventory (BDI; Beck & Beamesderfer, 1974), F (2, 37) = 1.71, p = .17.

**Procedure**

Eligible participants were individually scheduled for an appointment. They provided written consent and completed a self-report questionnaire package. Participants were then led into a laboratory, where they sat in a reclining chair facing a video camera, which stood in front of a one-way mirror. A television set stood to their right. Participants were given questionnaires to hold on their lap and instructed to complete each one at various points during the experiment. The experimenter sat in an adjacent room behind the one-way mirror and communicated with participants via intercom.

Participants were told that they would be required to give an impromptu 10-min videotaped speech on randomly chosen topics. They were informed that the speech would be observed by the experimenter and that the video would later be evaluated by “other people who will not know anything about you or the purpose of the experiment.” After hearing these initial instructions, participants rated their predicted level of performance and perceived audience standards, each on a 0–10 scale.

Participants were then randomly assigned to one of three experimental conditions: high standards, low standards, and no standards. In each condition, participants viewed a video clip of an individual speaking for 2-min about “tourist attractions in Boston.” The individual featured in the video clip was a volunteer in our laboratory whose speech was planned and recorded prior to the experiment, although participants were not privy to this information until the conclusion of the study. Participants received a scripted set of oral instructions over the intercom prior to watching the video clip. All participants viewed the same video clip; however, the oral instructions received with the video differed across the three experimental conditions. The script for the high standards condition contained the following information:

> In our experience, people prefer to watch another person perform a task before they have to do it themselves. So, before you give your speech, I would like you to get a sense of what the task will be like by showing you an example of a speech on video. As you watch the short video clip, keep in mind that not all speech performances are alike. A lot of different people have participated in this experiment before you, from both our Center and the community at large, and they have shown a wide range of anxiety and performance abilities. Truthfully, most people have performed quite a bit better than the individual in the video that you will now watch. Nevertheless, this video clip should still give you a good sense of what the task will be like.

In the low standards condition, participants received instructions that were identical to those in the high standards participants, except they were told that, “... most people have not performed quite as well as the individual in the video that you will now watch.” In the no standards condition, participants viewed the video without receiving any information about performance demands or expectations. Participants were told that they were watching the video in order to assess their level of concentration, and instructions directed their attention to superficial aspects of the video. The no standards script contained the following information:

> Before you give your speech, we would like to assess your level of concentration. To do this, we will show you a short video clip. In the clip, you will see a person talking about the things he likes to do around...
Boston. Your task is to count the total number of times the person in the video says the word ‘Boston.’
Please pay very careful attention to how many times the word ‘Boston’ is repeated. You may use the
bottom of the page in your booklet to help you keep track. After the video clip is finished, please record
your final tally at the bottom of page 3 of the booklet in front of you and place a circle around it.

Participants were then given detailed speech instructions, which were modeled after Beidel, Turner, Jacob,
and Cooley (1989). They were asked to speak for a maximum of 10 min, but were allowed to terminate their
speech at any time by raising a “stop” sign in the air. Participants were told that the speech topics were
chosen randomly (however, the topics for all participants were capital punishment, abortion, and cloning),
and were asked to hold the paper containing the topics face down (i.e., not to look at the topics) until they began
their speech.

Following their speech, participants rated their own performance on a 0–10 scale and their level of anxiety
during the speech on a 0–100 scale. They were then debriefed, compensated for their participation, and
dismissed. Two independent raters, blind to the purpose of the study and the condition and diagnostic status
of participants, watched and rated the videotaped speeches.

Clinical interviews

Anxiety Disorders Interview Schedule for DSM-IV: Lifetime Version (ADIS-IV-L; DiNardo et al., 1994).
The ADIS-IV-L is a semi-structured interview that evaluates DSM-IV anxiety, mood, somatoform, and
substance use disorders, and screens for the presence of other conditions. Although inter-rater reliability was
not tested in the present sample, Brown et al. (2001) reported high ADIS-IV-L reliability for diagnosing social
phobia (Kappa = .77).

Mini Anxiety Disorders Interview Schedule for DSM-IV (Mini-ADIS; Brown et al., 1994). The Mini-ADIS
does not assess past diagnoses, but is otherwise identical to the ADIS-IV-L in content and reliability of current
diagnoses. The Mini-ADIS has been used in several treatment outcome studies (e.g., Brown & Barlow, 1995)
and was used here to screen non-outpatient clinical participants and controls.

Self-report measures

Social Phobia and Anxiety Inventory (SPAI; Turner, Beidel, Dancu, & Stanley, 1989). The SPAI assesses
cognitive, somatic, and behavioral dimensions of social anxiety. It has been psychometrically well-validated in
both clinical (e.g., Beidel, Borden, Turner, & Jacob, 1989; Beidel, Turner, & Stanley, 1989; Turner, Stanley,
& Beidel, 1989) and non-clinical samples (e.g., Osman, Barrios, Haupt, & King, 1996). The SPAI consists of the
social phobia and agoraphobia subscales, each of which is scored separately. In the present study, internal
consistency was high for both the social phobia (z = .99) and agoraphobia (z = .94) scales.

Personal Report of Confidence as a Speaker Scale (PRCS; Gilkinson, 1942; Paul, 1966). The PCRS consists
of 30 Yes/No items that measure public speaking confidence, with possible scores ranging from 0 (no fear) to
30 (extreme fear). The scale is internally consistent (Klorman, Weerts, Hastings, Melamed, & Lang, 1974) and
valid (Lombardo, 1988).

Beck Depression Inventory (BDI; Beck & Beamesderfer, 1974). The BDI is a 21-item self-report inventory
that measures symptoms of depression. The BDI has good internal consistency (z = .89) and test–retest
reliability (ICC = .91), as well as adequate convergent and divergent validity in patients with social phobia
(Coles, Gibb, & Heimberg, 2001). BDI scores in the present study demonstrated excellent internal consistency
(z = .94).

Demographic Questionnaire. Relevant demographic information was collected via a short, author-
constructed measure.

Speech task

Of all social situations, public speaking is the most prevalent fear among individuals with social phobia
(e.g., Mannuzza et al., 1995; Stein, Torgrud, & Walker, 2000) and also the most common fear in the general
population (e.g., Pollard & Henderson, 1988; Stein et al., 2000). Therefore, compared to other social tasks, the speech task typically produces a high level of discomfort in all individuals with social phobia. Moreover, an impromptu speech represents an ideal laboratory situation to study factors related to social anxiety because this situation shows a high degree of external validity (i.e., no role play is necessary) and it can be more easily controlled and experimentally manipulated than other common socially challenging situations, such as conversations or social interactions with other people.

**Manipulation video**

All participants viewed a video clip of an individual giving an impromptu 2-min speech on the neutral topic of “tourist attractions in Boston.” Two independent observers viewed the video clip and rated the social performance of the volunteer featured in the video. Their averaged ratings yielded an overall performance score of 7.50 (0 = very poor; 10 = very good).

**Manipulation checks**

After the manipulation was introduced, participants rated perceived level and clarity of expected performance standards, both on 0–10 scales. In addition, to ensure that the effect of the manipulation was not confounded by individuals’ opinions about the topic presented in the video during the manipulation, participants rated their level of agreement (0–10) with the videotaped confederate who spoke about tourist attractions in Boston.

**Ratings of audience standards**

In order to examine group differences in perceived audience standards, participants were asked, “What level of performance do you think the people watching and evaluating your speech on video expect from you in the upcoming speech?” Prior to making this 0–10 rating (0 = very poor; 10 = very good), participants were informed that the people evaluating their speech will know nothing about them or the purpose of the experiment. Ratings were made at baseline before participants were randomized to one of the three conditions. This measure of others’ standards was adapted from previous research examining abilities and standards in social anxiety (e.g., Alden & Wallace, 1991; Wallace & Alden, 1991, 1995).

**Participants’ ratings of their own performance**

Before assignment to condition, participants rated how well they expected to perform on a 0–10 scale (0 = very poor; 10 = very good). Using the same scale, participants rated how well they thought they actually performed immediately after they finished their speech.

**Observer ratings of participants’ performance**

Speech performance was videotaped and assessed for quality by two trained, independent raters, blind to the purpose of the study and the diagnostic status and experimental condition of each participant. The main variable of interest was overall performance, which was rated by observers on a 0–10 scale, enabling a direct comparison of estimates of performance made by observers versus those made by participants themselves. This single-item measure of performance has been used in previous research with this population (e.g., Norton & Hope, 2001; Woody, 1996) and was used here to capture the “rough and dirty” (or “felt sense”; Clark & Wells, 1995) estimates that individuals with social phobia are likely to make about their own and others’ standards and performance before and after they encounter social situations.

Raters were student volunteers (one master’s student and one upper level undergraduate student) who were trained over several weeks by viewing numerous practice speeches that were videotaped during the pilot phases of the study. Once they began their independent ratings of the study tapes, they met with one another
periodically (five times in total) to review their ratings, discuss rating discrepancies, and if necessary, watch rated speeches together to calibrate their rating methods.

Subjective anxiety during the speech

Immediately after completion of the speech, participants reported their current anxiety level on a 0–100 scale, corresponding to Subjective Units of Distress (SUDS) ratings used in many previous studies.

Results

Trait measures

Table 1 displays descriptive statistics and results of 2 (group) by 3 (condition) ANOVAs examining each self-report questionnaire separately for clinical versus control participants. Results showed significant group effects across all measures (values are indicated in the table), but no significant effects of condition, (all $F's < 2.47$, all $p's > .09$, all $\eta^2_p < .07$) and no significant interaction effects (all $F's < 2.57$, all $p's > .08$, all $\eta^2_p < .07$).

Manipulation checks

A 2 (group) by 3 (condition) univariate ANOVA examining perceived standards after exposure to the manipulation revealed a significant effect of condition, $F(2, 72) = 4.98$, $p = .01$, $\eta^2_p = .13$. Follow-up Bonferroni comparisons indicated that participants rated standards in the high standards condition as being significantly higher ($M = 6.46$, $SD = .41$) than the low ($M = 4.75$, $SD = .42$; $p = .01$) and no ($M = 4.97$, $SD = .42$; $p = .04$) standards conditions. No other significant effects were found (all $F's < 2.31$, all $p's > .13$, all $\eta^2_p < .03$).

A 2 (group) by 3 (condition) univariate ANOVA examining perceived clarity of performance standards after exposure to the manipulation revealed a significant effect of condition, $F(2, 72) = 3.39$, $p = .04$, $\eta^2_p = .09$. Post-hoc comparisons using Bonferroni’s correction were non-significant. In order to determine the nature of the omnibus effect, we also conducted comparisons using (the less conservative) Fisher’s LSD, which indicated that participants rated standards in the no standards condition as being significantly less clear ($M = 3.69$, $SD = 3.10$) than the high ($M = 5.77$, $SD = 3.02$; $p = .02$) and low ($M = 5.42$, $SD = 3.18$; $p = .05$). No other significant effects were found (all $F's < 1.87$, all $p's > .16$, all $\eta^2_p < .05$).

Finally, a 2 (group) by 3 (condition) univariate ANOVA was conducted to examine participants’ agreement with the opinion of the speaker on the video. As expected, this analysis revealed no significant effects (all $F's < 1.46$, all $p's > .24$, all $\eta^2_p < .04$).

Reliability of observer ratings of performance

For the two observers’ ratings of participants’ performance, the one-way random effects model intraclass correlation coefficient (ICC) was adequate, ICC = .76, $p = .00$. All rating discrepancies were resolved by

<table>
<thead>
<tr>
<th>Measure</th>
<th>Patients ($M (SD)$)</th>
<th>Range</th>
<th>Controls ($M (SD)$)</th>
<th>Range</th>
<th>$F$</th>
<th>df</th>
</tr>
</thead>
<tbody>
<tr>
<td>BDI</td>
<td>16.16 (9.52)</td>
<td>0.00–40.00</td>
<td>1.41 (2.89)</td>
<td>0.00–16.00</td>
<td>88.30*</td>
<td>1.71</td>
</tr>
<tr>
<td>PRCS</td>
<td>22.21 (4.82)</td>
<td>11.00–30.00</td>
<td>5.23 (2.89)</td>
<td>2.00–13.00</td>
<td>386.46*</td>
<td>1.72</td>
</tr>
<tr>
<td>SPAI social</td>
<td>125.27 (22.11)</td>
<td>80.07–172.35</td>
<td>20.12 (19.02)</td>
<td>0.00–79.22</td>
<td>494.40*</td>
<td>1.70</td>
</tr>
<tr>
<td>SPAI ag</td>
<td>23.85 (13.40)</td>
<td>2.00–54.00</td>
<td>4.95 (4.88)</td>
<td>0.00–22.00</td>
<td>64.99*</td>
<td>1.72</td>
</tr>
<tr>
<td>SPAI total</td>
<td>101.49 (18.97)</td>
<td>60.07–136.05</td>
<td>15.18 (16.45)</td>
<td>0.00–70.22</td>
<td>456.67*</td>
<td>1.70</td>
</tr>
</tbody>
</table>

Note: BDI = Beck Depression Inventory; PRCS = Personal Report of Confidence as a Speaker Scale; SPAI = Social Phobia Anxiety Inventory; * $p < .001$. 

Table 1
Descriptive statistics and ANOVA group effects for self-report questionnaires completed by study participants
computing the mean of the two raters’ ratings. Thus, we also calculated the average measures ICC, which applies the Spearman-Brown correction to estimate the reliability of this mean score measure. Results indicated that the reliability of this mean score measure was high, ICC = .86, \( p = .00 \).

**Primary analyses**

**Participants’ ratings of audience standards and predicted performance**

Specific, a priori hypotheses were tested using planned one-way between-group contrast analyses. The use of planned contrasts, which is consistent with the recommendations of the American Psychological Association Task force on Statistical Inference (Wilkinson, 1999), allow for enhanced conceptual clarity and increased statistical power compared to omnibus tests (Rosenthal, Rosnow, & Rubin, 2000). Analyses confirmed our hypothesis that individuals with social phobia would provide ratings of audience standards that were significantly higher than those of controls, \( F(1, 74) = 4.49, p = .04, \eta_p^2 = .06 \). Analyses also confirmed our hypothesis that individuals with social phobia would provide ratings of their own performance that were significantly lower than those of controls, \( F(1, 76) = 14.92, p < .001, \eta_p^2 = .16 \). Means and standard errors are presented in Fig. 1.

**Participants’ retrospective ratings of their own performance**

Planned contrasts were conducted to test our hypothesis that individuals with social phobia would appraise their performance more negatively than controls in the high and no standards conditions relative to the low standards condition. Indeed, as illustrated in Fig. 2, between-group contrasts were significant within the high standards condition, \( F(1, 24) = 10.14, p < .01, \eta_p^2 = .30 \), and within the no standards condition, \( F(1, 24) = 37.73, p < .001, \eta_p^2 = .61 \), but not within the low standards condition, \( F(1, 24) = .05, p = .83, \eta_p^2 = .00 \). Means and standard deviations are presented in Table 2.
Observer ratings of participants’ performance

Planned contrasts were conducted to test our hypothesis that observers would rate the overall performance of individuals with social phobia significantly lower than that of controls in the high and no standards conditions relative to the low standards condition. Contrary to prediction, between-group contrasts were significant in each of the three conditions: the high standards condition, $F(1, 24) = 5.03, p = .03, \eta^2_p = .17$, the low standards condition, $F(1, 24) = 6.94, p = .02, \eta^2_p = .22$, and the no standards condition, $F(1, 24) = 11.17, p < .01, \eta^2_p = .32$. As indicated in Table 2, individuals with social phobia were judged by independent observers to perform poorly relative to controls across all conditions.

Subjective anxiety during the speech

Planned contrasts were conducted to test our hypothesis that individuals with social phobia would be significantly more anxious than controls in the high and no standards conditions relative to the low standards condition. Data were missing (due to answers that were left blank) from 8 control participants and 6 patients. Contrary to prediction, between-group contrasts were significant in each of the three conditions: the high standards condition, $F(1, 21) = 5.62, p = .03, \eta^2_p = .21$, the low standards condition, $F(1, 17) = 24.52, p < .001, \eta^2_p = .59$, and the no standards condition, $F(1, 20) = 28.47, p < .001, \eta^2_p = .59$. As indicated in Table 2, individuals with social phobia reported significantly higher levels of anxiety than controls across all conditions.

Exploratory analyses

Congruity of observer and participant ratings of performance

In order to explore the accuracy of participants’ performance appraisals across conditions, a 2 (rater: self and observer) by 2 (group) by 3 (condition) repeated-measures ANOVA was conducted. This analysis

![Fig. 2. Retrospective self-appraisals of speech performance in the three experimental conditions for participants with generalized social phobia (patients) and controls.](image)
compared the effects of group and condition on participant versus observer performance ratings. Results revealed significant main effects of group, $F(1, 72) = 37.84, p = .00, \eta^2_p = .34$, and rater, $F(1, 72) = 43.65, p = .00, \eta^2_p = .34$, and significant two-way interaction effects of rater by condition, $F(2, 71) = 3.50, p = .04, \eta^2_p = .09$, and group by condition, $F(2, 72) = 3.50, p = .04, \eta^2_p = .09$. There were no other significant main effects or two-way interactions (all $F$'s  0.51, all $p$'s  0.60, all $\eta^2_p$  0.01).

Most importantly, there was a significant three-way rater by group by condition interaction, $F(2, 71) = 4.39, p = .02, \eta^2_p = .11$, indicating that group differences in the extent to which self and observer ratings diverged depended on condition. Follow-up 2 (group) by 2 (rater) repeated-measures ANOVAs examining self versus observer ratings for each condition separately showed that the group by rater interaction was not significant in the high standards condition, $F(1, 24) = 1.14, p = .30, \eta^2_p = .05$. The other conditions also failed to reach significance, although trends toward significance were observed in both the low standards, $F(1, 24) = 3.92, p = .06, \eta^2_p = .14$, and no standards conditions, $F(1, 24) = 3.67, p = .07, \eta^2_p = .13$. Thus, while both groups equally underestimated their performance in the high standards condition, only individuals with social phobia underestimated their performance in the no standards condition and only controls underestimated their performance in the low standards condition.

Discussion

The present study advances our understanding of information processing in social phobia. First, as hypothesized, we demonstrated that in anticipation of a social event, individuals with social phobia believe that others hold excessively high standards for their performance. A likely reason for our novel finding was that participants in the present study were told that audience evaluators would know nothing about them. Previous studies (e.g., Alden et al., 1994; Wallace & Alden, 1991) may have found no differences between patient and control ratings of audience standards because participants with social phobia in those studies may have assumed that audience members were aware of their clinical status and, therefore, expected them to perform poorly. As a result, patients in previous studies may have rated others’ standards lower than they would in real life interactions with strangers.

Second, our results indicate that information about social standards moderates self-appraisals in individuals with social phobia. In the absence of explicit information about expected performance standards (i.e., when standards were ambiguous), individuals with social phobia rated their performance as having been poor—as poor as when performance standards were unambiguously high. On the other hand, when patients were exposed to cues that led them to believe that expected standards were low, their self-appraisals were indistinguishable from those of controls. Thus, patients’ perception of self as “deficient” may become activated by contexts in which others’ standards are seen as being either high or unclear.

Third, previous studies have consistently found that individuals with social phobia tend to minimize their performance accomplishments by exaggerating the extent to which they underperform (Alden & Wallace, 1995; Norton & Hope, 2001; Rapee & Lim, 1992; Stopa & Clark, 1993). In the present study, performance ratings by independent observers supported the notion that patients with social phobia exhibit poor social performance that is reliably judged as being deficient on a number of behavioral dimensions. In contrast to previous studies, however, both patients and controls significantly underestimated their performance relative
to the ratings of objective observers, with neither group exaggerating this underestimation on the whole. Rather, whether and how much each group underestimated performance depended on the condition to which they were exposed, with patients showing a trend toward underestimating their performance relative to controls only when information about standards was ambiguous. Thus, individuals with social phobia may minimize the quality of their performance only under social conditions in which performance expectations are poorly defined and open to interpretation. On the other hand, when specific, unambiguous social cues are available indicating that performance expectations are low, individuals with social phobia appear capable of appraising their performance more positively and more accurately.

These findings are especially important because in the majority of real life social situations, cues about expected performance standards are typically absent. Our findings suggest that when confronted with a social situation that contains no explicit informational cues about expected standards for performance, individuals with social phobia both “set the bar” too high in anticipation of their performance and retrospectively view themselves in a distorted manner that confirms their fears of failure. Interestingly, these effects appear to be unrelated to participants’ state anxiety levels.

Our findings are consistent with previous studies indicating that socially anxious and phobic individuals tend to offer neutral (Constans, Penn, Ihen, & Hope, 1999; Hirsch & Matthews, 2000) or negatively biased (Amir et al., 1998; Stopa & Clark, 2000) interpretations of ambiguous social scenarios, while nonanxious controls tend to generate explanations that are significantly more positive. Indeed, healthy controls in the present study judged their performance most positively and accurately in the ambiguous standards condition. This adaptive, positive bias for controls may become disrupted by unambiguous information about social standards, as in the other two conditions.

Our results also have significant implications for improving cognitive-behavioral therapy (CBT) outcomes for patients with social phobia. Given that exposure-based extinction learning in anxiety disorders is context-dependent (e.g., Bouton, 2004; Craske & Mystkowski, in press; Powers, Smits, Leyro, & Otto, in press), successful CBT outcomes typically require repeated extinction trials (exposures) across multiple feared contexts, particularly those in which patients will be most vulnerable to relapse at the conclusion of therapy (e.g., Bouton, 2002). In this vein, findings from the present study imply that patients should be exposed during treatment to multiple contexts in which social standards are ambiguous or poorly defined. Social interactions such as conversations, meeting strangers, or dating encounters would be prime exposure targets, as they are typically less structured than public speaking situations and, thus, inherently contain performance expectations that are highly ambiguous.

In addition to conducting exposure within these contexts, intervention techniques should also be added to help patients with social phobia “lower the bar” for their desired performance to a realistic level by setting appropriate goals prior to exposure, a skill that socially anxious individuals might lack (Hiemisch, Ehlers, & Westermann, 2002). Furthermore, exercises in which patients are instructed to intentionally violate (or fail to achieve) perceived performance standards may help them learn—contrary to their beliefs—that the consequences of social mishaps are not catastrophic (e.g., Foa, Franklin, Perry, & Herbert, 1996; Hofmann, 2004). Finally, treatment strategies should be employed to help individuals with social phobia enhance their overall social performance, whether performance difficulties are caused by patients’ use of self-protective safety behaviors or by genuine skills deficits.

Several limitations of this study should be noted. First, because a high percentage of patients in our study had comorbid depression and dysthymia (as is characteristic of individuals with social phobia), it is unclear whether our results would generalize to nondepressed patients with social phobia. It should be emphasized that patients across the three conditions had statistically equal levels of depression, defined both categorically and dimensionally. However, since the two groups in our study differed in levels of depression as a function of group membership rather than chance, we did not control for depression via analyses of covariance (Miller & Chapman, 2001). Thus, future research should include a clinical control group—preferably a depressed, non-socially phobic one—to enable within-clinical-group comparisons and clarify whether the effects observed in the present study are specific to social phobia. In addition, replicating our findings in both naturalistic studies and investigations of social standards during other types of social interactions such as conversations would greatly strengthen their external validity and generalizability.
Furthermore, measurement decisions may have influenced our surprising finding that the experimental manipulation did not moderate levels of anxiety. We operationalized anxiety in the present study as participants’ subjective levels of distress immediately after the speech, and it is possible that results would have differed if we instead chose to measure maximum levels of arousal during the speech itself. Finally, an issue of concern in almost any study of this nature involves the subjective experience of the participants. Although several manipulation checks were used, it was virtually impossible to know for sure that participants experienced each of the conditions in the way we have assumed they did, especially if the manipulation led to implicit priming effects of which the participants were not consciously aware. Thus, the precise mechanisms that were responsible for this effect remain unclear and require clarification in future studies.

Despite its limitations, this study enhances our understanding of the psychopathology of social phobia. Our findings clearly challenge the notion that social phobia is characterized by automatic, global distortions in self-perception, and provide support for the view that negative self-evaluation depends on the nature and availability of information about the social standard within a particular context. Our results suggest that when the social standard is poorly defined and open to interpretation, individuals with social phobia retrospectively appraise their performance in a negative manner that exaggerates their shortcomings. However, in the presence of specific, unambiguous cues indicating that the social standard is low, this negative bias is absent and patient self-appraisals are both more positive and more accurate. Future research should continue to investigate the notion that biased information processing in social phobia may vary depending on the cues individuals encounter in various social contexts and on the expectancies and beliefs that are primed by those cues.

Acknowledgements

David A. Moscovitch is now at the Department of Psychology, University of Waterloo. This manuscript was adapted from the first author’s doctoral dissertation. The research was partially supported by a Sigma Xi Grants-in-Aid-of-Research Award and a Boston University Psychology Department Felicia Sorembe Lambrose Dissertation Award.

We thank David H. Barlow, Tibor P. Palfai, Michael W. Otto, and Stefan Schulz for their insightful comments on earlier versions of this manuscript. We thank Michael K. Suvak for his statistical guidance, and Tina In-Albon and Alicia E. Meuret for their assistance with the psychophysiological equipment. In addition, we thank our lab volunteers Jared Matas, Micah Brosbe, and Nick Carde for their invaluable contributions to this project. Finally, we gratefully acknowledge those individuals who kindly agreed to participate in the study.

References
