Psychometric Evaluation of the Fear of Positive Evaluation Scale in Patients With Social Anxiety Disorder

Justin W. Weeks[1], **Richard G. Heimberg**[2], **Thomas L. Rodebaugh**[3], **Philippe R. Goldin**[4], and **James J. Gross**[4]

[1]Center for Evaluation and Treatment of Anxiety, Department of Psychology, Porter Hall 200, Ohio University, Athens, OH, 45701, USA

[2]Adult Anxiety Clinic of Temple, Temple University, Philadelphia, PA, 19122, USA

[3]Washington University, Department of Psychology, St. Louis, MO, 63130, USA

[4]Department of Psychology, Stanford University, Department of Psychology, Stanford, CA, 94305, USA

Abstract

The Fear of Positive Evaluation Scale (FPES; Weeks, Heimberg, & Rodebaugh, 2008) was designed to assess *fear of positive evaluation*, a proposed cognitive component of social anxiety. Although previous findings on the psychometric properties of the FPES have been highly encouraging, only one previous study has examined the psychometric profile of the FPES in a sample of patients with social anxiety disorder (Fergus et al., 2009). The primary purpose of the present study was to conduct a large multi-site examination of the psychometric profile of the FPES among patients with a principal diagnosis of social anxiety disorder (*n* = 226; generalized subtype: 97.8%). Responses of non-anxious control participants (*n* = 42) were also examined. The factorial validity, internal consistency, test-retest reliability, construct validity, and treatment sensitivity of the FPES were strongly supported by our findings. Furthermore, an FPES cutoff score was identified for distinguishing levels of fear of positive evaluation characteristic of patients with social anxiety disorder from those characteristic of the control group. Results provide additional support for the psychometric properties of the FPES in clinical samples.

Keywords

Cognitive-behavioral models of social anxiety disorder (SAD) (Clark & Wells, 1995; Rapee & Heimberg, 1997) have long held that social anxiety is, in part, a response to perceived negative evaluation by others. The construct of fear of negative evaluation (FNE) consists of feelings of apprehension about others' negative evaluations, distress over these evaluations, and the expectation that others will evaluate one negatively (Watson & Friend, 1969).

Corresponding author: Justin W. Weeks, Telephone: (740) 593-1094; Fax: (740) 597-3299; weeksj@ohio.edu.

Empirical support for the proposition that social anxiety is a response in part to FNE is extensive (e.g., Coles, Turk, Heimberg, & Fresco, 2001; Horley, Williams, Gonsalvez, & Gordon, 2004; Mansell & Clark, 1999).

A recently updated and extended cognitive-behavioral model of SAD (Heimberg, Brozovich, & Rapee, 2010) has highlighted *fear of positive evaluation* (FPE) as an additional cognitive component of SAD. FPE is distinct from, albeit strongly related to, FNE (Weeks, Heimberg, & Rodebaugh, 2008). FPE pertains to the sense of dread associated with being evaluated *favorably* and *publicly*, which begs a direct social comparison of the self to others and therefore causes a person to feel conspicuous and in the spotlight (see Heimberg et al., 2010; Weeks, Heimberg, & Rodebaugh, 2008, for more detailed overviews).

The Fear of Positive Evaluation Scale (FPES) is a self-report questionnaire that was designed to assess this construct (Weeks, Heimberg, & Rodebaugh, 2008). Several studies have evaluated the psychometric properties of the FPES (Fergus et al., 2009; Weeks, Heimberg, & Rodebaugh, 2008; Weeks, Heimberg, Rodebaugh, & Norton, 2008), and findings to date are highly encouraging. The FPES has demonstrated good internal consistency (all αs > .80) in both undergraduate (e.g., Weeks, Heimberg, & Rodebaugh, 2008; Weeks, Heimberg, Rodebaugh, & Norton, 2008) and clinical (Fergus et al., 2009) samples and strong 5-week test-retest reliability (intraclass correlation coefficient = .70) in an undergraduate sample (Weeks, Heimberg, Rodebaugh, & Norton, 2008).

FPES scores relate strongly to FNE and social anxiety in both undergraduate (e.g., Weeks, Heimberg, & Rodebaugh, 2008; Weeks, Heimberg, Rodebaugh, & Norton, 2008) and clinical (Fergus et al., 2009) samples, supporting its convergent validity. In addition, FPES scores relate strongly to tendencies to *disqualify positive social outcomes* (a long-recognized cognitive tendency in SAD; see Heimberg & Becker, 2002) in both undergraduate and clinical samples (Weeks, 2010). Furthermore, FPES scores relate more strongly to measures of social anxiety than to measures of depression and generalized anxiety symptoms/worry (Fergus et al., 2009; Weeks, Heimberg, & Rodebaugh, 2008; Weeks, Heimberg, Rodebaugh, & Norton, 2008) or to measures of anxiety sensitivity, perfectionism, and self-compassion (Weeks, Heimberg, Rodebaugh, & Norton, 2008), supporting the scale's discriminant validity.

Importantly, findings to date suggest that FPE is distinct from FNE. FPES scores account for unique variance in social anxiety above and beyond that accounted for by FNE in both undergraduate (Weeks, Heimberg, & Rodebaugh, 2008; Weeks, Heimberg, Rodebaugh, & Norton, 2008) and clinical (Fergus et al., 2009) samples. When measured across multiple weeks, FPE and FNE maintain their distinctiveness, with no evidence that one construct prospectively causes the other over the short term (Rodebaugh, Weeks, Gordon, Langer, & Heimberg, in press). Furthermore, FPE (but not FNE) relates positively to discomfort associated with, and negatively to perceived accuracy of, positive social feedback (Weeks, Heimberg, Rodebaugh, & Norton, 2008).

Patients with SAD obtain higher scores on the FPES than patients with other anxiety disorders (Fergus et al., 2009). Moreover, the FPES is sensitive to the effects of cognitive-

behavioral therapy for SAD (Fergus et al., 2009). These findings not only support the construct validity of the FPES but also highlight its potential clinical utility with these patients. Importantly, however, FPES scores of patients with SAD have yet to be compared to those of non-anxious controls – this essential test is also necessary to establish potential cutoff scores for identifying individuals with clinically severe levels of FPE within the general population. In addition, given that FPE is proposed to be a core cognitive component of social anxiety and that social anxiety has been previously reported to relate negatively to quality of life ratings (Hambrick, Turk, Heimberg, Schneier, & Liebowitz, 2003), it was expected that FPES scores would relate positively to disability ratings and negatively to quality of life ratings. These findings would also further inform the continued evaluation of the validity of the FPE construct.

Previous support for the psychometric profile of the FPES notwithstanding, only Fergus et al. (2009) have examined the psychometric characteristics of the FPES in a modest sample of patients with SAD ($n = 51$), and, thus, further study of the psychometric properties of the FPES in patients with SAD remains essential. To this end, the primary goal of the present study was to conduct a multi-site examination of the psychometric profile of the FPES in a large ($n = 226$) sample of patients with a principal diagnosis of SAD: The present study is the first to examine the psychometric properties of the FPES in a sample of patients with principal SAD (as opposed to patients with SAD, whether principal or secondary). In addition, the present study examined a more ethnically diverse sample (see **Participants** section) than that which has been examined previously; combined with the fact that a much larger clinical sample was examined in the present study (more than 4 times greater than that examined by Fergus et al.), the psychometric properties of the FPES as reported in the present study can reasonably be expected to be more stable and generalizable than those reported previously.

Overview of Study Hypotheses

More specifically, the following hypotheses were addressed in a sample of patients with principal SAD: (a) The factor structure of the FPES will be unifactorial, as previously reported in undergraduate (Weeks, Heimberg, & Rodebaugh, 2008; Weeks, Jakatdar, & Heimberg, 2010) and clinical (Fergus et al., 2009) samples, and invariant across site; (b) The distribution of FPES scores will be unimodal symmetric, as previously reported in undergraduate (e.g., Weeks, Heimberg, & Rodebaugh, 2008) and clinical (Fergus et al., 2009) samples; (c) The internal consistency and test-retest reliability of the FPES will be strong; (d) Patients with SAD will obtain greater FPES scores than non-anxious control (NAC) participants; (e) FPES scores will correlate positively with measures of social anxiety and (f) will correlate more strongly with measures of social anxiety than with measures of depression or worry; (g) FPES scores will relate positively to disability and negatively to quality of life ratings; and (h) FPES scores will be sensitive to treatment effects. In addition, we attempted to identify (i) a cutoff score for the FPES to distinguish between individuals with SAD and NACs, as determined by receiver operating characteristic (ROC) analyses. The present study is the first to address the research questions (d), (g), and (i); the first to examine the treatment sensitivity of the FPES utilizing controlled treatment data; and the first to examine the test-retest reliability of the FPES in a clinical sample.

Method

Participants

The majority of participants in the present study were patients ($n = 226$) who were evaluated at the Adult Anxiety Clinic of Temple University (AACT; $n = 99$) or Stanford University ($n = 127$). All patients at both sites were evaluated prior to treatment. All patients received a principal diagnosis of SAD based on the criteria specified in the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed. [*DSM–IV*]; American Psychiatric Association, 1994) (see the **Semistructured Diagnostic Interviews** section below for a detailed description of the interview process). The majority of patients across sites (97.8%) met the criteria for the generalized subtype of social anxiety disorder.[1] Patients with principal diagnoses other than SAD or with comorbid schizophrenia, clinically significant depression that was associated with prominent risk of self-harm, substance abuse or dependence in the past 6 months, or an organic mental disorder were excluded from the study. In addition, 52.7% met criteria for one or more comorbid diagnoses (with the majority of comorbid disorders being: generalized anxiety disorder: $n = 56$ [24.8% of overall patient sample]; major depressive disorder: $n = 21$ [9.3%]; panic disorder: $n = 10$ [4.4%]; specific phobia: $n = 10$ [4.4%]; and dysthymic disorder: $n = 9$ [4.0%]). Patients from both sites were excluded for current psychotherapy at the time of intake and history of neurological disorders. Patients at Stanford only were additionally excluded for current pharmacotherapy or history of head trauma. Eighteen patients (19.2%) at the AACT were receiving pharmacotherapy at time of intake.

Data were also provided by non-anxious control (NAC; $n = 42$) participants who were recruited from the greater community surrounding Stanford University and compensated for their participation. NAC participants were selected so as to reflect the composition of the Stanford patient sample on age, race, and sex. NAC participants also participated in a semistructured diagnostic interview. None of the NAC participants met criteria for any current Axis I diagnosis. Although NAC participants completed a number of self-report measures (see Table 2), only their FPES and Brief Fear of Negative Evaluation-Straightforward (BFNE-S; see **Measures** section for details) scores were analyzed in the present study.

Patients from the two clinics were equivalent in terms of sex ratio, χ^2 (1, $N = 226$) = 1.34, $p = .25$, and race, χ^2 (1, $N = 226$) = 2.88, $p = .09$. However, they differed in age, F (1, 226) = 7.07, $p = .008$. The AACT sample ($M = 30.16$; $SD = 10.72$) had a lower mean age than the Stanford sample ($M = 33.63$; $SD = 8.82$). However, age was not associated with FPES scores across the clinical samples ($r = .007$, $p = .92$). Furthermore, given that patients at Stanford were excluded for current pharmacotherapy at time of intake, whereas 19.2% of the patients at the AACT were receiving pharmacotherapy at time of intake, it is worth noting

---

[1]Note that an *a priori* decision was made to include all patients from our clinical sample in our analyses, including those diagnosed with the non-generalized subtype: (a) so that our clinical sample be more representative of the wide range of symptoms that clinicians and researchers are likely to encounter when dealing with SAD patients; and (b) to prevent range restriction within our clinical sample. However, a follow-up analysis was performed in response to a comment from an anonymous reviewer, in which we excluded all patients who met criteria for the non-generalized subtype and re-ran all of our analyses. The results from these analyses were substantively identical to those reported for our full clinical sample and are available upon request.

that FPES scores did not differ significantly as a function of current medication status within the AACT subsample, $F (1, 97) = 0.03$, $p = .85$, nor did FPES scores differ within the overall clinical sample according to medication status, $F (1, 224) = 0.31$, $p = .58$. Because the clinical samples differed significantly on only one demographic variable (i.e., age) and that variable was not related to FPES scores, they were pooled into a larger group for the primary analyses.

Given that NAC participants were recruited to reflect the composition of patients at Stanford but not the AACT, it is important to note that the AACT clinical sample was equivalent to the NAC sample in terms of sex ratio, $\chi^2 (1, N = 141) = 0.95$, $p = .33$; age, $F (1, 139) = 2.12$, $p = .15$; and race, $\chi^2 (1, N = 141) = 1.45$, $p = .23$. The pooled clinical sample did not differ from the NAC sample on sex ratio, $\chi^2 (1, N = 268) = 0.30$, $p = .58$; race, $\chi^2 (1, N = 268) = 0.29$, $p = .59$; or age, $F (1, 268) = 0.25$, $p = .62$. See Table 1 for specific details on the demographic characteristics of SAD patients and NAC participants.

Measures

All participants completed the FPES. In addition, patients completed a series of measures during their pre-treatment assessment that allowed for the evaluation of the construct validity of the FPES.

Fear of Positive Evaluation Scale (FPES; Weeks, Heimberg, & Rodebaugh, 2008)—The 10-item FPES uses a 10-point ordered-category rating scale, ranging from 0 (*not at all true*) to 9 (*very true*). Two reverse-scored items are included (for the purpose of potentially detecting response biases) but are not utilized in calculating the total score (see Weeks, Heimberg, & Rodebaugh, 2008, for additional details on this issue). See Introduction for a review of previous findings on the psychometric properties of the FPES.

Semistructured Diagnostic Interview

Anxiety Disorders Interview Schedule for DSM-IV-Lifetime Version (ADIS-IV-L; DiNardo, Brown, & Barlow, 1994)—The ADIS-IV-L provides probes and questions that assist in assigning *DSM-IV* diagnoses for a subset of psychiatric disorders, which include anxiety disorders, depressive disorders, and substance use disorders. In addition, Clinician's Severity Ratings (CSRs) for diagnoses are made on a 9-point ordered-category rating scale ranging from 0 to 8, with scores of 4 and above indicating clinical severity. The ADIS-IV-L was administered to all patients at both sites, as well as to all NACs (at Stanford). Training criteria outlined by Brown, DiNardo, Lehman, and Campbell (2001) were satisfied by all interviewers in the present study. Brown et al. (2001) reported a kappa of .77 for a principal diagnosis of social anxiety disorder in a sample of 362 anxiety disorder patients who received two independent ADIS-IV-L interviews. Furthermore, 66 of the patients from the AACT clinical sample (66.7%) were also interviewed by a second assessor who administered the social anxiety disorder module of the ADIS-IV–L, and 100% agreement with the original principal diagnosis of SAD was obtained ($\kappa = 1.0$). Moreover, there was 100% agreement between the CSRs for this subsample to within 1 point of accuracy across independent assessors. However, interrater agreement of ADIS-IV-L diagnoses was not assessed in the Stanford patient sample.

Measures for the Assessment of Social Anxiety (Primary convergent measures)

Convergent validity would be demonstrated if the FPES were to be significantly correlated with the following measures of social anxiety:

Liebowitz Social Anxiety Scale (LSAS; Liebowitz, 1987; Fresco et al., 2001)— The 24-item LSAS can be administered as a clinician-administered or self-report measure and was designed to assess fear and avoidance in social interaction (11 items) and performance (13 items) situations. Items are rated on separate 4-point ordered-category scales of fear and avoidance, which range from 0 (*none* and *never*, respectively) to 3 (*severe* and *usually*, respectively). The LSAS yields a total score, as well as separate scores for fear and avoidance of social interaction and performance situations. Because no hypotheses of the present study related to the distinction between fear and avoidance, only the total score was used. The clinician-administered LSAS demonstrates good internal consistency (α = . 96; Heimberg et al., 1999) and correlates significantly with scores on other measures of social anxiety (Heimberg, Mueller, Holt, Hope, & Liebowitz, 1992). The self-report version of the Liebowitz Social Anxiety Scale (LSAS-SR) (Fresco et al., 2001; Rytwinski et al., 2009) was derived from the clinician administered LSAS (Liebowitz, 1987) and consists of the same questions and rating scale as the clinician-administered version. The LSAS-SR shows good measurement characteristics (Fresco et al., 2001; Rytwinski et al., 2009). Specifically, there is ample evidence for the test-retest reliability, internal consistency, convergent and discriminant validity, and construct validity of the LSAS-SR. The clinician-administered LSAS was completed by all patients at the AACT in the present study and exhibited excellent (α = .96) internal consistency in this subsample. The self-report LSAS was completed by all patients at Stanford and exhibited excellent (α = .97) internal consistency in this subsample. Given that the psychometric properties of the clinician-administered and self-report versions of the LSAS have been established to be essentially identical (e.g., see Fresco et al., 2001; Rytwinski et al., 2009), responses to the LSAS were examined as a single measure across sites.[2]

Social Interaction Anxiety Scale (SIAS; Mattick & Clarke, 1998) and Social Phobia Scale (SPS; Mattick & Clarke, 1998)—The SIAS is a measure of anxiety in dyads and groups, and the SPS is a measure of anxiety associated with being observed by others. Both scales consist of 20 items, which are scored on a 5-point ordered-category rating scale ranging from 0 (*not at all characteristic or true of me*) to 4 (*extremely characteristic or true of me*). The SIAS and SPS have demonstrated strong internal consistency in both clinical and undergraduate samples (all αs > .84 [SIAS] and .86 [SPS]; Heimberg et al.,1992; Mattick & Clarke, 1998), and strong test-retest correlations have been reported in a sample of patients with social anxiety disorder at 4 and 12 weeks (all *r*s > .91, all *p*s < .001; Mattick & Clarke, 1998).

---

[2]It is worth noting that, although responses to the clinician-administered and self-report versions of the LSAS were examined as a single measure across sites, results were substantively identical when examining these as separate measures within each clinical sample. However, we opted to focus our report on the combined scores across sites given that the psychometric properties of these distinct administration formats have been demonstrated to be essentially identical, and as a means of increasing the stability of our reported effects.

Rodebaugh, Woods, and Heimberg (2007) have suggested utilizing only the 17 straightforward (-S) SIAS items to calculate the total score, reporting that the straightforward items are more valid indicators of social interaction anxiety than the reverse-scored items in both undergraduate and clinical samples and that the reverse-scored items may primarily assess extraversion. This 17-item score, hereafter referred to as the SIAS-Straightforward (SIAS-S) score, has demonstrated excellent internal consistency ($\alpha = .93$) and factorial validity in undergraduate samples and has demonstrated strong construct validity in both undergraduate and clinical samples (Rodebaugh et al., 2007). The 20-item SIAS was administered; however, only the SIAS-S items were utilized here. The SIAS was completed by all patients in the present study, and the SPS was completed by all patients at the AACT. Internal consistency of both the SIAS-S ($\alpha = .89$) and SPS ($\alpha = .92$) was excellent in the clinical sample in the present study.

Measure of Fear of Negative Evaluation (Secondary convergent measure)

Brief Fear of Negative Evaluation Scale-Straightforward Items (BFNE-S: Rodebaugh et al., 2004; Weeks et al., 2005)—The BFNE (Leary, 1983) is a 12-item self-report measure of fear and distress related to negative evaluation from others. Rodebaugh et al. (2004) and Weeks et al. (2005) have suggested utilizing only the 8 straightforward (-S) BFNE items to calculate the total score, given that the reverse-scored items of the BFNE: (a) demonstrate poorer psychometric properties in comparison to the straightforward BFNE items and (b) exhibit response biases associated with education level (Weeks et al., 2005). The BFNE-S has demonstrated excellent internal consistency (all α's > .92), factorial validity, and construct validity in undergraduate (Rodebaugh et al., 2004) and clinical (Weeks et al., 2005) samples. The 12-item BFNE was administered to all patients in the present study; however, only the BFNE-S items were utilized in the present analyses. The BFNE exhibited excellent internal consistency ($\alpha = .91$) in the present clinical sample. Because FPE and FNE are both proposed cognitive components of SAD, it was expected that FPES scores would relate positively to BFNE-S scores in the present study.

Measures for the Assessment of Discriminant Validity

Discriminant validity would be demonstrated if higher correlations were exhibited between the FPES and the convergent measures than between the FPES and the following measures:

Beck Depression Inventory-II (BDI-II; Beck, Steer, & Brown, 1996)—The BDI-II is a 21-item self-report measure of depression. BDI-II items assess depressive symptoms and attitudes that can be rated from 0 to 3 in terms of intensity. Excellent internal consistency has been reported for the BDI-II in both outpatient ($a = .92$) and undergraduate ($a = .93$) samples (Beck et al., 1996), and convergent validity has been demonstrated by positive correlations between scores on the BDI-II and scores on measures of hopelessness and suicidal ideation. Furthermore, scores on the BDI-II are more strongly correlated with clinician-rated depression than with clinician-rated anxiety (Beck et al., 1996). The BDI-II was completed by all patients in the present study. The BDI-II exhibited excellent internal consistency ($\alpha = .91$) in the present SAD sample.

Penn State Worry Questionnaire (PSWQ; Meyer, Miller, Metzger, & Borkovec, 1990)—The PSWQ is a 16-item questionnaire designed to assess trait worry, the core feature of generalized anxiety disorder (GAD), regardless of worry content. Items are rated on a 5-point scale, from 1 (*Not at all typical*) to 5 (*Very typical*). The PSWQ is internally consistent (αs = .86 to .93) and has demonstrated acceptable test-retest reliability (*r*s = .74 to .93) (Molina & Borkovec, 1994; Turk, Heimberg, & Mennin, 2004). In addition, individuals with GAD scored higher on the PSWQ than did individuals with obsessive-compulsive disorder, SAD, and panic disorder (Brown et al., 1992). The PSWQ was administered to all Stanford participants in this study. The PSWQ exhibited excellent internal consistency (α = .95) in the present SAD sample.

Measures of Disability and Quality of Life (Construct Validity)

Two additional measures were administered to patients at the AACT. These measures were included in the present study to further evaluate the construct validity of the FPES in SAD patients, by examining relationships between scores on the FPES and measures of impairment and quality of life.

Liebowitz Self-Rated Disability Scale (LSRDS; Schneier et al., 1994)—The LSRDS comprises 22 items that assess current (past 2 weeks) impairment/disability, as well as the worst impairment ever experienced. LSRDS items are rated on a 4-point scale, from 0 (*Problem does not limit me at all*) to 3 (*Problem limits me severely*). Schneier et al. (1994) reported that the LSRDS was highly internally consistent (α = .92) for both the current and worst ratings. Only current impairment was assessed in the present study. The LSRDS was administered only to patients at the AACT. Ratings for current impairment demonstrated acceptable internal consistency (α = .73) in our sample. Given that FPE is conceptualized as an important cognitive component of SAD, it was expected that FPES scores would relate positively to self-rated disability.

Quality of Life Inventory (QOLI; Frisch, Cornell, Villanueva, & Retzlaff, 1992)—The QOLI assesses the degree to which an individual is satisfied with 16 areas of his or her life. Health, standard of living, friendships, relationship with family, and community are a few examples. These areas are rated on: (a) a 0–2 scale of importance to the individual's life and (b) a scale of −3 to 3 of how satisfied the individual is in each area. The total score is derived by multiplying the ratings for importance and satisfaction for each domain, then averaging across the 16 domains; lower QOLI scores indicate lower life satisfaction. The total score has been shown to be internally consistent (α = .98) and has demonstrated adequate test-retest reliability (*r*s = .80 to .91) (Frisch et al., 1992). QOLI scores were also positively correlated with scores on a clinician-administered life satisfaction interview, peer ratings of life satisfaction, and five self-report measures assessing life satisfaction and subjective well-being. The QOLI was administered to all AACT patients, and demonstrated excellent internal consistency in our sample (α = .87). Given that social anxiety has been previously reported to relate negatively to quality of life ratings (Hambrick et al., 2003), it was expected that FPES scores would relate negatively to QOLI scores.

Procedure

Patients completed the ADIS-IV-L, the LSAS (either clinician-administered [AACT] or self-report [Stanford]), and the battery of self-report questionnaires at intake. The present study utilized Bonferroni corrections for multiple comparisons based on the *False Discovery Rate* (Benjamini & Hochberg, 1995).[3] Non-anxious control participants completed the FPES and BFNE after the ADIS-IV-L and were compensated for their time. Of the 268 participants who were administered the FPES, only a single response was missing for one item (for one patient), for which mean imputation was applied for analyses using the scale scores (but not confirmatory factor analyses, which used items).

Confirmatory factor analysis (CFA) was performed using the structural equation modeling software program Mplus (Muthén & Muthén, 1998–2010). In determining factor structure, global model fit was evaluated using the: (a) comparative fit index (CFI, Bentler, 1990), (b) Tucker-Lewis incremental fit index (TLI; Tucker & Lewis, 1973), and (c) root mean square error of approximation (RMSEA, Steiger & Lind, 1980). Hu and Bentler (1999) have suggested cutoffs for fit indices for evaluating the results of confirmatory factor analyses. Specifically, a cutoff of .95 or above on either the Tucker–Lewis Index (TLI; Tucker & Lewis, 1973) or the comparative fit index (CFI; Bentler, 1990) combined with either a root-mean-square error of approximation (RMSEA) "close to .06" (Hu & Bentler, 1999, p. 1) or a standardized root-mean-square residual (SRMR) "close to .08" (Hu & Bentler, 1999, p. 1) provides a good combination of fit indices to conclude that a structural model has adequate or better fit to the data. Note that the SRMR is not available when fitting models to categorical data.

CFAs using categorical data were conducted using the robust weighted-least squares estimator, referred to as WLSMV in Mplus, with missing data estimation. This estimator is appropriate for categorical variables; the BFNE items were treated as categorical due to their limited 5-point response scale. FPES items were treated as continuous because of their 10-point response scale, the entire range of which was utilized in the sample. Tests of invariance across site were conducted via multiple-group factor analysis. These analyses were conducted to determine whether the factor structure and factor loadings of the FPES and BFNE-S items (and the thresholds of the BFNE-S items) were invariant across sites, thus justifying a joint factor analysis across sites and tests of validity that combined data

---

[3]In response to a comment from an anonymous reviewer, the present study utilized corrections for multiple comparisons based on the *False Discovery Rate* (FDR; Benjamini & Hochberg, 1995) in contrast to classical procedures which control the family-wise error rate. Although a thorough review of the FDR approach to multiple significance testing is beyond the scope of the present paper (see Benjamini & Hochberg, 1995, for details), for the purposes of the present study, the FDR approach is summarized by the following formula: P $i/m_{(x)} * q^*$, in which "P" = corrected *p*-value (to be solved for, per family of hypotheses involving multiple comparisons); "i" = the total number of families of hypotheses involving multiple *p-values* to be tested in the present study (fixed for overall study); "m" = the number of hypotheses within a given family (x) to be tested (tailored to each family of hypotheses); and "q^*" = overall alpha correction level for the present study (i.e., α = .05 [fixed]).

For the present study, there were 3 sets of analyses which involved testing *families of p-values* (see **Overview of Study Hypotheses** section). Specifically, the FDR correction procedure was applied to: the overall set of correlations between the FPES, BFNE-S, and all other study measures (for which $m_{(e)} = 15$; see Table 3); hypothesis e (for which $m_{(e)} = 9$ [a comprehensive series of comparisons of correlated correlation coefficients was conducted among the 2 fear of evaluation-related measures and the 3 social anxiety-related measures; see **Convergent validity** section for an overview of all comparisons]); and f (for which $m_{(f)} = 4$ [a series of comparisons of correlated correlation coefficients was conducted among the FPES, the 2 discriminant measures [BDI-II and PSWQ], and the 2 social anxiety measures which were administered to the full sample [LSAS and SIAS-S] [2*2 = 4]). Accordingly, P $3/m_{(x)}* .05^*$ for the overall study, yielding the following alpha correction values (tailored for each hypothesis): overall set of correlations (Table 3) = .01; e = .0167; f = .0375.

across sites. We also conducted the same analyses for the FPES items alone because the primary concern for the current analyses was whether the FPES items in particular showed an invariant, well-fitting factor structure. We conducted these tests using maximum likelihood estimation and the Satorra-Bentler chi-square (referred to as MLM in Mplus). Chi-square difference tests in MLM and WLSMV were handled in accordance to the methods described in the Mplus manual (Muthén & Muthén, 1998–2010), because neither MLM nor WLSMV yields chi-squares that can be simply subtracted to produce a chi-square suitable for difference testing. Because the MLM estimator does not handle missing data, one participant was excluded from that analysis because of a missing data point.

As part of an independent study (Goldin et al., 2011), a subsample of the patients at Stanford was randomly selected for inclusion in a waitlist control group ($n = 27$) and completed the FPES again after a period of 4.5 months; responses were compared across time points to evaluate the test-retest reliability of the FPES. In addition, 29 Stanford patients were randomly selected to receive individual CBT (Hope, Heimberg, Juster, & Turk, 2000) for SAD (without first being waitlisted) (see Goldin et al., 2011); FPES responses of the CBT and waitlist groups at pre- and post-treatment were examined to evaluate the treatment sensitivity of the FPES.[4]

Results

Preliminary Analyses

Means and *SD*s for patients' and non-anxious controls' responses to all study measures are displayed in Table 2. There were no significant gender differences on FPES total scores in either the clinical (F [1, 224] = 0.12, $p = .73$) or NAC (F [1, 40] = 0.36, $p = .55$) samples.

Preliminary analyses were conducted to examine potential site differences in social anxiety symptom severity. The sites differed with regard to LSAS scores, F (1, 204) = 22.64, $p < .001$. Specifically, the AACT sample ($M = 72.39$; $SD = 21.04$) exhibited lower LSAS scores in comparison to the Stanford sample ($M = 85.45$; $SD = 18.43$). Importantly, the mean LSAS scores across sites were well above the recommended cutoff of 60 for identifying individuals with generalized SAD (see Mennin et al., 2002; Rytwinski et al., 2009).

Furthermore, given the focus of the present study on cognitive components of social anxiety (i.e., fears of positive and negative evaluation), preliminary analyses were conducted to examine whether FPES or BFNE-S scores related significantly to the presence of secondary mood disorders within our clinical sample. A multivariate analysis of variance (MANOVA) was tested, with presence versus absence of secondary mood disorders entered as a fixed factor, and FPES and BFNE-S scores entered as dependent variables. The multivariate effect for secondary mood disorder group was significant, F (2, 223) = 5.55, $p = .004$. Follow-up between-subjects comparisons revealed that patients with a secondary mood disorder obtained significantly greater BFNE-S scores ($M = 35.03$, $SD = 3.51$) than patients without

---

[4]Seventy-five patients at Stanford were initially randomly selected for either inclusion in the waitlist control group (initial $n = 37$; 5 dropped out, and 5 participants had not yet completed the Time 2 FPES at the time of our analyses; thus, obtained $n = 27$) or to receive CBT (initial $n = 38$; 6 dropped out, and 3 participants had not yet completed the Time 2 FPES at the time of our analyses; thus, obtained $n = 29$).

a secondary mood disorder ($M = 31.85$, $SD = 0.41$), $F = 10.83$, $p < .001$, and this effect was of medium size, Cohen's $d = 0.67$. In contrast, patients with and without secondary mood disorders did not differ in their FPES scores, $F = 1.85$, $p = .18$.

Confirmatory Factor Analysis

One of our initial tests concerned whether the BFNE-S and FPES items showed an invariant factor structure (defined as invariant factor loadings for each item and invariant thresholds for each categorical item) across site for the clinical groups. An initial attempt at analysis revealed that a total of nine participants (four from the AACT and five from Stanford) had to be deleted from the analysis because some response options were not used by participants at each site. A test of invariance across this somewhat reduced dataset revealed that the factor structure could be constrained across the two sites, showing reasonable fit for the constrained model (CFI = .93, TLI = .93, RMSEA = .07). Further, removing constraints did not improve model fit ($\chi^2[38, N = 217] = 46.90$, $p = .152$). In the constrained model, each item loaded at least moderately well on its factor (FPE standardized loadings: .34–.85; BFNE standardized loadings: .65–.90), with all loadings statistically significant at the $p < .001$ level. The latent variables of FPE and FNE correlated moderately and positively ($r = .36$, $p = .003$).

To address hypothesis (a) more specifically, as well as to address concerns that the sample had to be reduced to test invariance involving the BFNE items, we next tested the FPES items alone. Because these items were treated as continuous in maximum likelihood estimation, there were no issues with lack of overlap in response options across group (however, one participant was excluded due to a missing data point, which is not permitted when using the Satorra-Bentler chi-square, as we did here). A test of invariance demonstrated that the factor structure was invariant across sites, with acceptable to good model fit for the constrained model (CFI = .96, TLI = .95, SRMR = .09, RMSEA = .06). Removing constraints did not improve fit ($\chi^2[8, N = 225] = 1.20$, $p = .997$). Standardized loadings ranged from .35 to .86 and were all statistically significant at the $p < .001$ level.

Psychometric Properties of the FPES

Distribution of FPES scores—The distribution of FPES scores was unimodal symmetric for patients (skewness = −0.43, $SE = .16$; kurtosis = −0.29, $SE = .32$) and for NAC participants (skewness = 0.71, $SE = .37$; kurtosis = −0.28, $SE = .72$).

Internal consistency and test-retest reliability of FPES scores—The FPES exhibited good internal consistency in the clinical (ɑ = .85), non-anxious control (ɑ = .83), and total (ɑ = .89) samples. Moreover, the FPES demonstrated excellent test-retest reliability among the wait-list participants ($n = 27$) at Stanford (see *Procedure* section) over 4.5 months ($r = .80$).

Group differences on FPES scores—Patients scored significantly higher on the FPES than did NAC participants, $F (1, 266) = 120.52$, $p < .001$, see Table 2. The effect size was large, $d = 2.02$ (Cohen, 1988).

Convergent validity—A Bonferroni correction (see Benjamini & Hochberg, 1995) was applied to the overall set of correlations between the FPES, BFNE-S, and all other study measures ($p = .01$; see Footnote 2). Recall that FPE and FNE related moderately ($r = .36$) at the latent level.[5] Bivariate correlations of scores on the FPES (and BFNE-S, for comparison) and all other study measures are displayed in Table 3. Consistent with expectation, FPES scores were significantly and positively correlated with each of the social anxiety measures.

We also examined whether FPES scores related most strongly to any of the measures of social anxiety or FNE. Significance tests were conducted by using a Fisher's *r*-to-*z* transformation refined by Meng, Rosenthal, and Rubin (1992). An overall Bonferroni correction (see Benjamini & Hochberg, 1995) was applied ($p = .0167$; see Footnote 2). The FPES related equivalently to each of the three social anxiety measures (LSAS, SIAS-S, and SPS) (all zs < 0.99, all ps >.32) and more strongly to all three social anxiety measures than to the measure of FNE (BFNE-S scores) (all zs > 2.44, all ps < .015).

Lastly, we examined whether any of the measures of social anxiety (LSAS, SIAS-S, or SPS) related more strongly to the FPES or BFNE-S. Interestingly, the LSAS related marginally more strongly (albeit not significantly so following Bonferroni correction) to the FPES than to the BFNE-S ($z = 2.06$, $p = .0394$); there was no difference for the SIAS-S or SPS (both zs < 1.50, both ps > .13).

Discriminant validity—FPES scores were significantly and positively correlated with both discriminant measures (see Table 3). Significance tests (Meng et al., 1992) were conducted to determine whether FPES scores related more strongly to the measures of social anxiety which were administered to the full sample (LSAS and SIAS-S) than to measures of depression or worry. An overall Bonferroni correction (see Benjamini & Hochberg, 1995) was applied ($p = .0375$; see Footnote 2). FPES scores were more strongly related to both social anxiety measures than to the BDI-II (both zs > 2.10, both ps < .036).

FPES scores related equivalently to the PSWQ and the LSAS and SIAS-S (both ps > .14). It seemed reasonable to assume that the nature of the obtained relationships between FPES scores, worry, and social anxiety could be influenced by the presence of comorbid GAD in our clinical sample (24.8%), which may have resulted in a higher overall base rate of worry symptoms. Thus, follow-up analyses were performed in which patients with comorbid GAD were excluded. Within this subsample, FPES scores related significantly more strongly to scores on both available social anxiety measures (LSAS and SIAS-S) than to PSWQ scores (both zs > 2.13, both ps < .034). In addition, it is worth noting that the obtained relationship between BFNE-S and PSWQ scores was larger than the relationship between FPES and PSWQ scores. Taken together, the present findings provide strong additional support for the discriminant validity of the FPES in a clinical sample.

Relationships between FPES scores, disability, and quality of life—Consistent with hypothesis, FPES (and BFNE-S) scores were significantly and positively correlated

---

[5]It is important to note that the correlation between two latent factors corrects for error variance and is therefore a superior statistical estimate in comparison to the correlation between constructs based on scale scores.

with disability ratings (LSRDS) (see Table 3; recall that an overall Bonferroni correction of $p = .01$ was applied [see Footnote 2]). However, contrary to expectation, neither the FPES nor the BFNE-S (nor any of the social anxiety measures) related significantly to quality of life ratings (QOLI scores) (both $ps > .011$). Follow-up analyses were performed, for which a single extreme outlying case (i.e., a case who scored 3 interquartile ranges above the mean) on QOLI total raw scores was removed. Prior to removing this outlier, the distribution of QOLI scores was positively skewed and leptokurtic (skewness = 1.74, $SE = .25$; kurtosis = 9.56, $SE = .49$). Upon removing this outlier, QOLI scores were normally distributed (skewness = −0.006, $SE = .25$; kurtosis = 0.86, $SE = .49$), and consistent with hypothesis, the FPES (as well as the BFNE-S) related significantly and negatively to QOLI scores (see Table 3).

Unique variance in social anxiety accounted for by FPES scores—Hierarchical regression analyses were conducted to examine whether FPES scores accounted for variance in social anxiety scores beyond that already accounted for by fear of negative evaluation (BFNE-S scores). Given differences in sample size across responses obtained for the three social anxiety measures (see *Measures* section), three separate regression equations were tested. Specifically, in each regression, the relevant social anxiety score (either the LSAS, SIAS-S, or SPS) served as the criterion variable; BFNE-S scores were entered in the first step, and FPES scores were entered in the second step. A Bonferroni correction was applied $(.05/3 = .017)$, which controlled for the number of regression equations tested.[6]

After the first step, BFNE-S scores were significantly and positively related to all of the social anxiety measures, all R^2s > .07, all Fs > 14.81, all $ps < .001$, as expected. In the second step, FPES scores were also significantly and positively related to, and significantly improved the prediction of, all social anxiety scores, all R^2 s > .08, all F s > 11.26, all ps < .001. Thus, FPES scores accounted for significant variance in social anxiety beyond that already accounted for by FNE (see Table 4).

Sensitivity to treatment effects—We examined change in FPES scores from Time 1 to Time 2 between patients who initially received CBT ($n = 29$) and those who were initially waitlisted ($n = 27$). Mean change in FPES scores following CBT reflected significantly greater improvement ($M = -17.41$, $SD = 16.36$) than mean change following the waiting period ($M = 1.41$, $SD = 10.10$), $F (1, 54) = 26.35$, $p < .001$. Moreover, this effect was large, $d = 1.38$ (CI = 1.24, 1.52). We also calculated change in BFNE-S change scores from Time 1 to Time 2 between these groups, for comparison. Mean change in BFNE-S scores following CBT reflected significantly greater improvement ($M = -9.22$, $SD = 6.23$) than mean change following the waiting period ($M = -1.25$, $SD = 4.10$), $F (1, 54) = 31.63$, $p < .001$. Moreover, this effect was large, $d = 1.51$ (CI = 1.13, 1.88).

---

[6]Note that a correction for multiple comparisons based on the FDR (Benjamini & Hochberg, 1995) was not utilized for testing this particular hypothesis (and therefore, this hypothesis was not included in the total number of families of hypotheses involving multiple *p-values* to be tested in the present study [*i*]), in that this would have yielded a correction that was greater than .05 ($p = [4/3]*.05 = .067$), due to the small number ($m = 3$) of relationships compared. Thus, a classical Bonferroni correction was utilized for testing this specific hypothesis.

Receiver operating characteristic (ROC) analyses—Last, we were interested in identifying a cutoff score for the FPES to distinguish between individuals with SAD and non-anxious controls. To determine the extent to which FPES scores can accurately identify SAD, we conducted ROC analyses, which utilize the association between sensitivity and specificity to derive an *area under the curve* (AUC), indicating the extent to which a measure distinguishes between positive cases (i.e., patients) and negative cases (i.e., NACs) in a given sample, irrespective of base rate. A value of .50 of the AUC indicates chance-level prediction, whereas 1.0 indicates perfect classification.

The ROC curve is displayed in Figure 1. FPES scores demonstrated excellent differentiation between the patient and NAC subsamples (AUC = .91, 95% CI = .87–.94). Potential cutoff scores were evaluated through consideration of both sensitivity (i.e., accurately identifying true positives) and specificity (i.e., accurately identifying true negatives). In the present sample, the FPES cut score that provided the optimal balance was 22, correctly classifying 198/226 patients and 35/42 NACs (sensitivity: 87.6%; specificity: 83.3%).

Discussion

The present study was the first to examine the psychometric properties of the FPES in a large multi-site sample of patients with a principal diagnosis of SAD. Various analyses were conducted to replicate and extend findings involving the factor structure, reliability, and validity of the FPES in a clinical sample. First, confirmatory factor analyses replicated the two-factor fear of evaluation model (with separate factors for FPE and FNE) which was originally reported by Weeks and colleagues (Weeks, Heimberg, & Rodebaugh, 2008; see also Weeks et al., 2010) and subsequently demonstrated in a clinical sample (Fergus et al., 2009). A test of invariance revealed that this two-factor structure could be constrained across the two sites, showing reasonable fit for the constrained model. Furthermore, removing constraints did not improve model fit. FPE and FNE correlated modestly at the latent level in the present sample (see a more detailed discussion of this finding below). More importantly for the purposes of the present study, CFAs replicated the unitary factor structure of the FPES in a large sample of patients with SAD. Second, the distributions of FPES scores were unimodally symmetric for both patients and non-anxious controls. Third, the FPES exhibited good internal consistency in the clinical and non-anxious control subsamples and excellent internal consistency in the overall sample; test-retest reliability was excellent for a clinical subsample over a period of 4.5 months. Fourth, patients obtained significantly higher scores on the FPES than did non-anxious control participants, a large effect. Fifth, FPES scores related positively to all measures of social anxiety.

Sixth, although FPES scores related moderately to measures of depressive symptoms and worry in the overall clinical sample, FPES scores related significantly more strongly to all social anxiety measures than to a measure of depressive symptoms. Furthermore, upon excluding patients with comorbid GAD to account for elevated baserates of worry within our patient sample, FPES scores related significantly more strongly to scores on the social anxiety measures than to scores on the measure of worry. Thus, the FPES demonstrated strong overall discriminant validity with regard to depressive symptoms and worry. Indeed, the discriminant validity of the FPES (particularly with regard to depressive symptoms) was

stronger than that for the BFNE-S (e.g., see **Preliminary Analyses,** and Table 2 for discriminant correlations). Seventh, consistent with expectation, FPES scores related positively to disability and negatively to ratings of quality of life. Eighth, FPES scores accounted for unique variance in social anxiety scores beyond that already accounted for by a measure of FNE. Ninth, FPES scores were strongly sensitive to the effects of CBT for social anxiety disorder. Finally, an FPES cutoff score was identified (22) which demonstrated excellent differentiation between the patient and NAC subsamples, and this cutoff score exhibited good sensitivity and specificity. These data in their totality suggest that the FPES provides a highly reliable and valid index of FPE.

The present study was the first to examine relationships between FPES scores and disability and quality of life ratings in an SAD sample. Consistent with expectation, FPES scores related positively to disability. This finding provides strong additional support for the concept that FPE is a symptom which either co-occurs with or has some causal relationship to social anxiety and provides additional support for the construct validity of the FPES as a clinical assessment for use with SAD populations.

Also consistent with expectation, FPES scores related negatively to quality of life ratings. This is consistent with previous findings that outpatients with SAD report marked deficits in quality of life, with this relationship holding even upon accounting for disability (Hambrick et al., 2003). Furthermore, the obtained negative relationship between FPE and quality of life is consistent with accumulating findings highlighting SAD as being characterized by overall *positivity impairment* (e.g., see Kashdan & Weeks, 2010; Kashdan, Weeks, & Savostyanova, 2011) and provides additional support for the validity of the FPE construct.

Our obtained results also supplement an accumulating body of findings providing support for the distinction between FPE and FNE. The FPES related more strongly to all three social anxiety measures than to FNE in our clinical sample (and in fact, one of the social anxiety measures related marginally more strongly to the FPES than to the BFNE-S). Furthermore, as has been reported previously in undergraduate (Weeks, Heimberg, & Rodebaugh, 2008; Weeks, Heimberg, Rodebaugh, & Norton, 2008) and one clinical (Fergus et al., 2009) sample, FPE accounted for unique variance in social anxiety measures above and beyond FNE.

The FPES appears to be sensitive to treatment-induced changes associated with CBT. The *controlled* effect size for the FPES (comparing CBT to a waitlist control group) was large ($d = 1.38$) – in fact, the 95% CI for this effect (CI = 1.24, 1.52) fell above the lower bound of the 95% CI for BFNE-S-related change due to controlled treatment effects (CI = 1.13, 1.88). These findings thus replicate and extend those reported by Fergus et al. (2009), in that the present study is the first to examine the treatment sensitivity of the FPES utilizing controlled treatment data.

Last, the present study is the first to identify an FPES cutoff score for distinguishing between individuals with SAD and non-anxious individuals. FPES scores demonstrated excellent differentiation between our patient and the NAC subsamples, and a cutoff score of 22 balanced strong sensitivity and specificity. This cutoff score may prove useful to

researchers interested in examining differences between groups with and without clinically severe levels of FPE. Furthermore, given strong support for the psychometric profile of the FPES in samples of SAD patients provided by both previous (Fergus et al., 2009) and present findings, future examinations of the potential utility of the FPES as a diagnostic and/or screening instrument for identifying SAD and as a treatment outcome measure in this population (Fergus et al., 2009) appear warranted. Our replication and extension of the effects found by Fergus et al. in a larger, more ethnically diverse sample suggests that such effects are generalizable to a wide range of clinical samples.

An unexpected finding from the present study was the relatively modest relationship obtained between FPE and FNE at the latent level ($r = .36$), in contrast to findings from several previous studies, which have shown this relationship to be consistently large (rs range from .49– .74) (Weeks, Heimberg, & Rodebaugh, 2008; Weeks et al., 2010), including one clinical sample ($r = .62$) (Fergus et al., 2009). It is unclear why the relationship between FPE and FNE was relatively smaller in this specific clinical sample. Previous findings suggest that both FPE and FNE have dimensional latent structures (Weeks, Norton, & Heimberg, 2009), suggesting that the modest relationship between these constructs in the present sample could not be attributable to a relatively small percentage of patients who feared only one type of social evaluation. It also bears noting that the relationship between BFNE-S scores and scores obtained on one of the social anxiety measures used in the present study (the LSAS) was similarly low ($r = .27$), indicating that a lower-than-expected relationship with the BFNE-S was not specific to the FPES in the present sample. One plausible and parsimonious explanation concerns the fact that the current sample consists entirely of treatment seeking individuals with a principal diagnosis of SAD, of whom the vast majority had the more severe generalized form. It may be that the correlation between the latent traits might be underestimated because very few participants in the present sample had low scores on either measure; this feature of the sample might also explain the acceptable but not excellent fit of the hypothesized two-factor *fear of evaluation* structure. Nonetheless, the preponderance of evidence from previous studies (Fergus et al., 2009; Weeks, Heimberg, & Rodebaugh, 2008; Weeks et al., 2010) suggests that FPE and FNE are strongly related (albeit distinct) cognitive components of social anxiety that show a strong two-factor structure, even across multiple time points (Rodebaugh et al., in press).

Our findings provide further support for the concept that it may be *any* type of evaluation that is feared by persons with social anxiety disorder. Thus, SAD may be better characterized by a fear of evaluation *in general* than by a fear of negative evaluation, as has been traditionally emphasized (e.g., Clark & Wells, 1995; Rapee & Heimberg, 1997). We have argued previously (e.g., see Heimberg et al., 2010; Weeks, Heimberg, & Rodebaugh, 2008; Weeks, Heimberg, Rodebaugh, & Norton, 2008) that FPE functions in part to: (a) reduce social conflict with others; (b) protect one's social status (by reducing competition with socially dominant others); and (c) reduce overall conspicuousness. It has also been argued previously that positive evaluation may be feared because it may predict later negative evaluation (e.g., see Wallace & Alden, 1997). In that the present and previous findings provide strong support for the concept that FPE is a cognitive component of SAD, the modification of existing, or the development of novel, clinician-administered measures

of social anxiety symptoms to focus more systematically on reactions to positive social feedback is warranted. This is an important area for future research.

Several limitations to the present study exist. First, only patients with social anxiety disorder (primarily of the generalized subtype) were included in the patient group in the present study. Moreover, the present study did not examine the relationships between responses to the FPES and other measures in the non-anxious control samples. Therefore, the utility of the FPES in samples of varying diagnoses, larger samples of patients with SAD of the non-generalized subtype, and non-anxious persons from the community are topics for future research. On a related note, the direct examination of FPE levels in specific anxiety disorder populations other than SAD merits attention in future studies, given that this has not yet been examined to date. Second, although diagnostic interrater agreement for the AACT patient sample was demonstrated to be excellent in the present study (see **Semistructured Diagnostic Interview** section), interrater agreement of ADIS-IV-L diagnoses was not assessed in the Stanford patient sample. This marks a limitation to our findings with regard to internal validity. Third, only post-treatment FPES scores were examined to evaluate the sensitivity of the FPES to individual CBT in the present study: Future studies should examine the sensitivity of the FPES to different treatments for SAD, including group cognitive-behavioral therapy and pharmacological treatments. Fourth, it is necessary to extend the validation of the FPES through the inclusion of multi-method data sources. Although the FPES relates as expected with measures of social anxiety, examination of the relationship of FPES responses to behavioral and physiological measures are essential next steps in the evaluation of its construct validity. In conclusion, the present findings further support the FPES as a psychometrically sound index of fear of positive evaluation. Furthermore, our findings provide strong additional support for the concept that FPE is an important cognitive component of SAD. Along with the initial findings by Fergus et al. (2009), the current results suggest that the FPES has promise as a means of investigating fear of positive evaluation in the context of clinically severe social anxiety.

Acknowledgments

References

American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4. Washington, DC: Author; 1994.

Beck, AT.; Steer, RA.; Brown, GK. Beck Depression Inventory Manual. 2. San Antonio, TX: Psychological Corporation; 1996.

Bentler PM. Fit indexes, Lagrange multipliers, constraint changes and incomplete data in structural models. Multivariate Behavioral Research. 1990; 25:163–172.10.1207/s15327906mbr2502_3

Brown TA, Antony MM, Barlow DH. Psychometric properties of the Penn State Worry Questionnaire in a clinical anxiety disorders sample. Behaviour Research and Therapy. 1992; 30:33–37.10.1016/0005-7967(92)90093-V [PubMed: 1540110]

Brown TA, DiNardo PA, Lehman CL, Campbell LA. Reliability of *DSM-IV* anxiety and mood disorders: Implications for the classification of emotional disorders. Journal of Abnormal Psychology. 2001; 110:49–58.10.1037/0021-843X.110.1.49 [PubMed: 11261399]

Clark, DM.; Wells, A. A cognitive model of social phobia. In: Heimberg, RG.; Liebowitz, MR.; Hope, DA.; Schneier, FR., editors. Social phobia: Diagnosis, assessment, and treatment. New York: Guilford Press; 1995. p. 69-93.

Coles ME, Turk CL, Heimberg RG, Fresco DM. Effects of varying levels of anxiety within social situations: Relationship to memory perspective and attributions in social phobia. Behaviour Research and Therapy. 2001; 39:651–665.10.1016/S0005-7967(00)00035-8 [PubMed: 11400710]

Cohen, J. Statistical power analysis for the behavioral sciences. 2. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.

DiNardo, PA.; Brown, TA.; Barlow, DH. Anxiety Disorders Interview Schedule for DSM-IV-Lifetime Version (ADIS-IV-L). New York: Oxford University Press; 1994.

Fergus TA, Valentiner DP, McGrath PB, Stephenson K, Gier S, Jecius S. The Fear of Positive Evaluation Scale: Psychometric properties in a clinical sample. Journal of Anxiety Disorders. 2009; 23:1177–1183.10.1016/j.janxdis.2009.07.024 [PubMed: 19700259]

Fresco DM, Coles ME, Heimberg RG, Liebowitz MR, Hami S, Stein MB, Goetz D. The Liebowitz Social Anxiety Scale: A comparison of the psychometric properties of self-report and clinician-administered formats. Psychological Medicine. 2001; 31:1025–1035.10.1017/S0033291701004056 [PubMed: 11513370]

Frisch, MB. Manual and treatment guide for the Quality of Life Inventory. Minneapolis, MN: National Computer Systems, Inc; 1994.

Frisch MB, Cornell J, Villanueva M, Retzlaff PJ. Clinical validation of the Quality of Life Inventory: A measure of life satisfaction for use in treatment planning and outcome assessment. Psychological Assessment. 1992; 4:92–101.10.1037/1040-3590.4.1.92

Gilbert P. Evolution and social anxiety: The role of attraction, social competition, and social hierarchies. The Psychiatric Clinics of North America. 2001; 24:723–751.10.1016/S0193-953X(05)70260-4 [PubMed: 11723630]

Goldin PR, Werner K, Ziv M, Jazaieri H, Luce K, Hill K, Heimberg R, Gross JJ. Psychological mechanisms underlying individual cognitive-behavioral therapy for social anxiety disorder. 2011 Manuscript submitted for publication.

Hambrick JP, Turk CL, Heimberg RG, Schneier FR, Liebowitz MR. The experience of disability and quality of life in social anxiety disorder. Depression and Anxiety. 2003; 18:46–50.10.1002/da.10110 [PubMed: 12900952]

Heimberg, RG.; Becker, RE. Cognitive-behavioral group therapy for social phobia: Basic mechanisms and clinical strategies. New York: Guilford Press; 2002.

Heimberg, RG.; Brozovich, FA.; Rapee, RM. A cognitive-behavioral model of social anxiety disorder: Update and extension. In: Hofmann, SG.; DiBartolo, PM., editors. Social anxiety: Clinical, developmental, and social perspectives. New York: Elsevier; 2010.

Heimberg RG, Horner KJ, Juster HR, Safren SA, Brown EJ, Schneier FR, Liebowitz MR. Psychometric properties of the Liebowitz Social Anxiety Scale. Psychological Medicine. 1999; 29:199–212.10.1017/S0033291798007879 [PubMed: 10077308]

Heimberg RG, Mueller GP, Holt CS, Hope DA, Liebowitz MR. Assessment of anxiety in social interaction and being observed by others: The Social Interaction Anxiety Scale and the Social Phobia Scale. Behavior Therapy. 1992; 23:53–73.10.1016/S0005-7894(05)80308-9

Hope, DA.; Heimberg, RG.; Juster, HR.; Turk, CL. Managing social anxiety: A cognitive-behavioral therapy approach. New York: Oxford University Press; 2000.

Horley K, Williams LM, Gonsalvez C, Gordon E. Face to face: Visual scanpath evidence for abnormal processing of facial expressions in social phobia. Psychiatry Research. 2004; 127:43–53.10.1016/j.psychres.2004.02.016 [PubMed: 15261704]

Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling. 1999; 6:1–55. Retrieved from http://www.informaworld.com/smpp/title~db=all~content=t775653699.

Kashdan, TB.; Weeks, JW. Social anxiety, positive experiences, and positive events. In: Hofmann, SG.; DiBartolo, PM., editors. Social anxiety: Clinical, developmental, and social perspectives. 2. New York: Elsevier; 2010. p. 448-465.

Kashdan TB, Weeks JW, Savostyanova A. Whether, how, and when social anxiety shapes positive experiences and events: A self-regulatory framework and treatment implications. Clinical Psychology Review. 2011; 31:786–799.10.1016/j.cpr.2011.03.012 [PubMed: 21529701]

Leary MR. A brief version of the Fear of Negative Evaluation Scale. Personality and Social Psychology Bulletin. 1983; 9:371–375.10.1177/0146167283093007

Liebowitz MR. Social phobia. Modern problems in pharmacopsychiatry. 1987; 22:141–173.

Mansell W, Clark DM. How do I appear to others? Social anxiety and processing of the observable self. Behaviour Research and Therapy. 1999; 37:419–434.10.1016/S0005-7967(98)00148-X [PubMed: 10228314]

Mattick RP, Clarke JC. Development and validation of measures of social phobia scrutiny fear and social interaction anxiety. Behaviour Research and Therapy. 1998; 36:455–470.10.1016/S0005-7967(97)10031-6 [PubMed: 9670605]

Meng XL, Rosenthal R, Rubin DB. Comparing correlated correlation coefficients. Psychological Bulletin. 1992; 111:172–175.10.1037/0033-2909.111.1.172

Mennin DS, Fresco DM, Heimberg RG, Schneier FR, Davies SO, Liebowitz MR. Screening for social anxiety disorder in the clinical setting: Using the Liebowitz Social Anxiety Scale. Journal of Anxiety Disorders. 2002; 16:661–673.10.1016/S0887-6185(02)00134-2 [PubMed: 12405524]

Meyer TJ, Miller ML, Metzger RL, Borkovec TD. Development and validation of the Penn State Worry Questionnaire. Behaviour Research and Therapy. 1990; 28:487–495.10.1016/0005-7967(90)90135-6 [PubMed: 2076086]

Molina, S.; Borkovec, TD. The Penn State Worry Questionnaire: Psychometric properties and associated characteristics. In: Davey, GCL.; Tallis, F., editors. Worrying: Perspecitves on theory, assessment, and treatment. Chichester, UK: Wiley; 1994. p. 265-283.

Muthén, LK.; Muthén, BO. Mplus user's guide. 6. Los Angeles, CA: Muthén & Muthén; 1998–2010.

Rapee RM, Heimberg RG. A cognitive-behavioral model of anxiety in social phobia. Behaviour Research and Therapy. 1997; 35:741–756.10.1016/S0005-7967(97)00022-3 [PubMed: 9256517]

Rodebaugh TL, Weeks JW, Gordon EA, Langer JK, Heimberg RG. The longitudinal relationship between fear of positive evaluation and fear of negative evaluation. Anxiety, Stress, & Coping. in press. 10.1080/10615806.2011.569709

Rodebaugh TL, Woods CM, Heimberg RG. The reverse of social anxiety is not always the opposite: The reverse-scored items of the Social Interaction Anxiety Scale do not belong. Behavior Therapy. 2007; 38:192–206.10.1016/j.beth.2006.08.001 [PubMed: 17499085]

Rodebaugh TL, Woods CM, Thissen DM, Heimberg RG, Chambless DL, Rapee RM. More information from fewer questions: The factor structure and item properties of the original and Brief Fear of Negative Evaluation Scale. Psychological Assessment. 2004; 16:169–181.10.1037/1040-3590.16.2.169 [PubMed: 15222813]

Rytwinski NK, Fresco DM, Heimberg RG, Coles ME, Liebowtitz MR, Cissell S, Hofmann SG. Screening for social anxiety disorder with the self-report version of the Liebowitz Social Anxiety Scale. Depression and Anxiety. 2009; 26:34–38.10.1002/da.20503 [PubMed: 18781659]

Schneier FR, Heckelman LR, Garfinkel R, Campeas R, Fallon BA, Gitow A, Liebowitz MR. Functional impairment in social phobia. Journal of Clinical Psychiatry. 1994; 55:322–331. [PubMed: 8071299]

Steiger, JH.; Lind, JC. Statistically-based tests for the number of common factors. Paper presented at the annual meeting of the Psychometric Society; Iowa City, IA. 1980 May.

Tucker LR, Lewis C. A reliability coefficient for maximum likelihood factor analysis. Psychometrika. 1973; 38:1–10.10.1007/BF02291170

Turk, CL.; Heimberg, RG.; Mennin, DS. Assessment. In: Heimberg, RG.; Turk, CL.; Mennin, DS., editors. Generalized anxiety disorder: Advances in research and practice. New York: Guilford; 2004. p. 29-50.

Wallace ST, Alden LE. Social phobia and positive social events: The price of success. Journal of Abnormal Psychology. 1997; 106:416–424.10.1037/0021-843X.106.3.416 [PubMed: 9241943]

Watson D, Friend R. Measurement of social–evaluative anxiety. Journal of Consulting and Clinical Psychology. 1969; 33:448–457.10.1037/h0027806 [PubMed: 5810590]

Weeks JW. The Disqualification of Positive Social Outcomes Scale: A novel assessment of a long-recognized cognitive tendency in social anxiety disorder. Journal of Anxiety Disorders. 2010; 24:856–865.10.1016/j.janxdis.2010.06.008 [PubMed: 20637564]

Weeks JW, Heimberg RG, Fresco DM, Hart TA, Turk CL, Schneier FR, Liebowitz MR. Empirical validation and psychometric evaluation of the Brief Fear of Negative Evaluation Scale in patients with social anxiety disorder. Psychological Assessment. 2005; 17:179–190.10.1037/1040-3590.17.2.179 [PubMed: 16029105]

Weeks JW, Heimberg RG, Rodebaugh TL. The Fear of Positive Evaluation Scale: Assessing a proposed cognitive component of social anxiety disorder. Journal of Anxiety Disorders. 2008; 22:44–55.10.1016/j.janxdis.2007.08.002 [PubMed: 17884328]

Weeks JW, Heimberg RG, Rodebaugh TL, Norton PJ. Exploring the relationship between fear of positive evaluation and social anxiety. Journal of Anxiety Disorders. 2008; 22:386–400.10.1016/j.janxdis.2007.04.009 [PubMed: 17531437]

Weeks JW, Jakatdar TA, Heimberg RG. Comparing and contrasting fears of positive and negative evaluation as facets of social anxiety. Journal of Social and Clinical Psychology. 2010; 29:68–94.10.1521/jscp.2010.29.1.68

Weeks JW, Norton PJ, Heimberg RJ. Exploring the latent structure of two cognitive components of social anxiety: Taxometric analyses of fears or negative and positive evaluation. Depression and Anxiety. 2009; 26:E40–E48.10.1002/da.20414 [PubMed: 19067316]

Weeks JW, Rodebaugh TL, Heimberg RG, Norton PJ, Jakatdar TA. "To avoid evaluation, withdraw": Fears of evaluation and depressive cognitions lead to social anxiety and submissive withdrawal. Cognitive Therapy and Research. 2009; 33:375–389.10.1007/s10608-008-9203-0

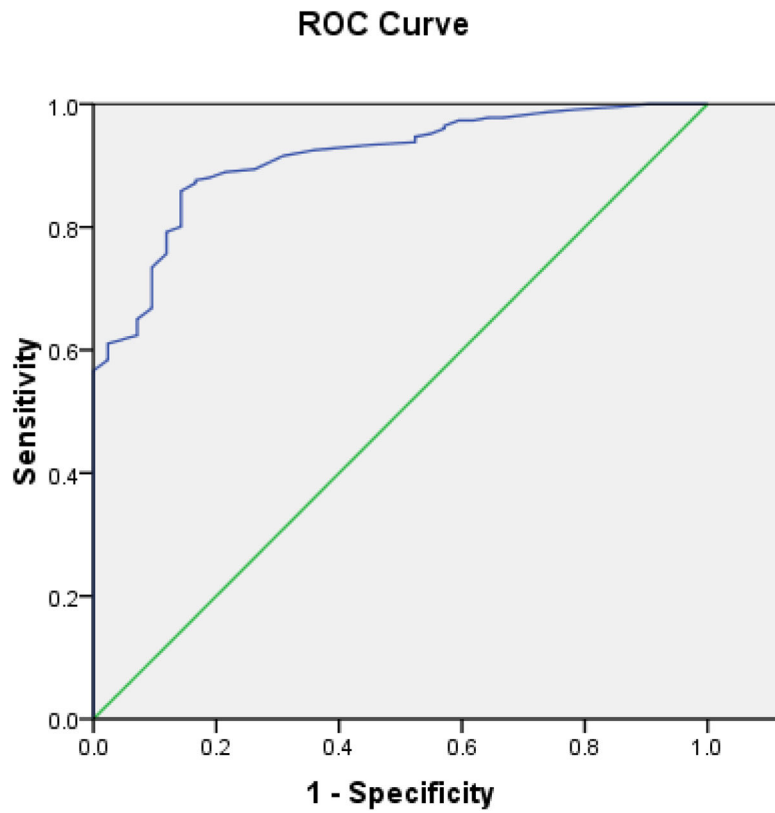## ROC Curve

Figure 1.
Receiver operating characteristic (ROC) curve for Fear of Positive Evaluation Scale scores in the social anxiety disorder patient/non-anxious control samples. Diagonal segments are produced by ties.

Table 1

Descriptive statistics for patients and non-anxious controls on demographic variables.

Variable	Patients (*n* = 226)				Controls (*n* = 42)			
	n	*%*	*M*	*SD*	*n*	*%*	*M*	*SD*
Sex								
Female	108	47.8			22	52.4		
Male	118	52.2			20	47.6		
Race								
Caucasian	129	57.1			22	52.4		
Non-Caucasian	97	42.9			20	47.6		
Age (years)			32.12	9.83			32.93	9.24

Table 2

Means (and standard deviations) on study measures for patients with social anxiety disorder and healthy controls.

Measure	Patients		Controls		t-value, effect size
	M	SD	M	SD	
Fear of Positive Evaluation Scale	39.60	14.92	13.07	10.99	10.54*, 2.02
	n = 226		n = 42		
Brief Fear of Negative Evaluation Scale-Straightforward	32.43	5.57	13.79	4.23	20.59*, 3.77
	n = 215		n = 42		
Liebowitz Social Anxiety Scale [a]	80.59	20.51	16.10	11.37	19.31*, 3.89
	n = 206		n = 41		
Social Interaction Anxiety Scale-Straightforward	44.06	11.22	8.74	7.09	20.06*, 3.76
	n = 223		n = 42		
Social Phobia Scale [b]	33.22	15.52	--	--	--
	n = 99				
Beck Depression Inventory-II	15.69	11.07	1.93	2.97	7.68*, 1.70
	n = 221		n = 40		
Penn State Worry Questionnaire [c]	61.19	11.60	--	--	--
	n = 124				
Liebowitz Self-Rated Disability Scale [b]	12.10	5.80	--	--	--
	n = 91				
Quality of Life Inventory [b]	0.47	2.34	--	--	--
	n = 96				

Notes: Effect size = Cohen's *d*.

*
$p = .001$.

Sample size for patients varies from 206 to 226 (unless noted otherwise, see below) as a result of differences in measure administration across the AACT and Stanford sites;

[a] Liebowitz Social Anxiety Scale was administered as a clinician-administered measure at the AACT, but as a self-report measure at Stanford;

[b] Measure was administered at the AACT only;

[c] Measure was administered at Stanford only;

-- = statistic not computed because scale was not administered to control group.

Table 3

Zero-order correlations between the Fear of Positive Evaluation Scale, the Brief Fear of Negative Evaluation Scale-Straightforward (for comparison), and all other study measures in the overall clinical sample.

Measure	Fear of Positive Evaluation Scale	Brief Fear of Negative Evaluation Scale-Straightforward
Liebowitz Social Anxiety Scale [a]	.43*	.27*
Social Interaction Anxiety Scale-Straightforward	.44*	.54*
Social Phobia Scale [b]	.50*	.55*
Beck Depression Inventory-II	.27*	.42*
Penn State Worry Questionnaire [c]	.29*	.42*
Liebowitz Self-Rated Disability Scale [b]	.42*	.44*
Quality of Life Inventory [b]	−.44*	−.42*

Notes: Sample size for patients varies from 206 to 226 (unless noted otherwise, see below) as a result of differences in measure administration across the AACT and Stanford sites (see Table 2 for measure-specific *n*s);

[a] Liebowitz Social Anxiety Scale was administered as a clinician-administered measure at the AACT, but as a self-report measure at Stanford;

[b] Measure was administered at the AACT only. Furthermore, a single extreme outlier on the Quality of Life Inventory total raw score (with an elevated score, patient sample) was removed for these analyses;

[c] Measure was administered at Stanford only. A family-wise Bonferroni correction of * $p < .01$ was applied (see Footnote 2 for details on correction).

Table 4

Regression weights from hierarchical regression analyses examining the unique variance in social anxiety measures accounted for by measures of fear of evaluation.

Variable	*B*	*SE B*	Beta
Liebowitz Social Anxiety Scale			
Step 1:			
BFNE-S	1.00	.26	.18
Step 2:			
BFNE-S	.67	.25	.18
FPES	.51	.09	.38
Social Interaction Anxiety Scale – Straightforward			
Step 1:			
BFNE-S	1.09	.12	.45
Step 2:			
BFNE-S	.92	.12	.45
FPES	.23	.04	.31
Social Phobia Scale			
Step 1:			
BFNE-S	1.26	.21	.43
Step 2:			
BFNE-S	1.00	.21	.43
FPES	.30	.09	.31

Notes: FPES = Fear of Positive Evaluation Scale; BFNE-S = Brief Fear of Negative Evaluation Scale-Straightforward. All overall adjusted R^2s > .20, all *SE*s < 18.21; all R^2 s > .08, all *F*s > 11.26, all *p*s < .001.