Evaluation of the Clinical Global Impression Scale among individuals with social anxiety disorder

T. I. ZAIDER, R. G. HEIMBERG,[1] D. M. FRESCO, F. R. SCHNEIER
AND M. R. LIEBOWITZ

From the Adult Anxiety Clinic, Department of Psychology, Temple University, Philadelphia, PA; and Anxiety Disorders Clinic, New York State Psychiatric Institute and Columbia University, New York, NY, USA

ABSTRACT

Background. The clinical Global Impression Scale (CGI) is commonly used as a primary outcome measure in studies evaluating the efficacy of treatments for anxiety disorders. The current study evaluated the psychometric properties and predictors of clinicians' ratings on an adapted version of the CGI among individuals with social anxiety disorders.

Method. An independent assessor administered the CGI Severity of Illness and Improvement ratings to 123 patients at baseline and the subset of treated patients again mid- and post-treatment.

Results. Improvement ratings were strongly related to both concurrent Severity of Illness and changes in Severity of Illness ratings from baseline. Additionally, both CGI ratings were positively correlated with both self-report and clinician-administered measures of social anxiety, depression, impairment and quality of life. Measures of social anxiety symptoms accounted for a large portion of the variance in Severity of Illness ratings, with significant additional variance accounted for by measures of impairment and depression. Changes in social anxiety symptoms from baseline accounted for significant variance in Improvement ratings, but no significant additional variance was accounted for by changes in impairment and depressive symptoms.

Conclusions. Our findings support the utility of the CGI as an index of global severity and symptom-specific improvement among individuals with social anxiety disorder.

INTRODUCTION

Global measures of outcome are often used to provide a clinically meaningful summary of an individuals' functioning during and following treatment. Unlike measures designed to identify and assess specific symptoms related to a particular disorder, global ratings attempt to capture a patient's overall well-being at a given point in time. However, the term 'global' remains somewhat elusive and it is unclear which aspects of an individual's functioning actually determine such ratings (Beneke & Rasmus, 1991).

One commonly used measure of global functioning is the Clinical Global Impression Scale

(CGI), a set of ratings made by a clinician in order to assess the overall severity of an individual's symptoms as well as changes in his or her functioning over time. The CGI was first developed for use in psychopharmacology trials as part of the NIMH collaborative study of schizophrenia (Guy, 1976). Since then, it has been used as a standard primary outcome measure in studies investigating the efficacy of pharmacological treatments for psychiatric conditions such as depression (Hellerstein *et al.* 1993), social anxiety disorder (Liebowitz *et al.* 1992; Stein *et al.* 1998), panic disorder (Barlow *et al.* 2000), post-traumatic stress disorder (Brady *et al.* 2000) and binge-eating disorder (Hudson *et al.* 1998). Despite its widespread use as a means of defining treatment response, there has been little evaluation of the validity

[1] Address for correspondence: Dr Richard G. Heimberg, Adult Anxiety Clinic, Department of Psychology, Weiss Hall, Temple University, 1701 North 13th Street, Philadelphia, PA 19122-6085, USA.

and reliability of this measure or its relationship to specific domains of functioning.

One of the few studies that examined the CGI was conducted with a mixed sample of 175 patients with schizophrenia, depression or anxiety disorders who were being treated with various psychotropic medications (Beneke & Rasmus, 1991). Beneke & Rasmus reported a moderate correlation between the Global Improvement item at post-treatment and difference scores in Severity of Illness from pre-treatment to post-treatment. Although the analyses conducted by Beneke & Rasmus (1991) shed some light on the relationship between CGI ratings, they do not tell us anything about how these ratings relate to other, more specific measures of symptom severity and impairment. A study conducted by Leon *et al.* (1993) partly addressed this question by examining the relationship between CGI ratings and symptom-specific measures among patients enrolled in a 16-week psychopharmacology trial for concurrent panic and depressive disorders. Changes in symptom scores accounted for 38–40 % of the variance in CGI Severity of Illness ratings and 26–46 % of the variance in CGI Global Improvement ratings over 8 weeks of treatment. The authors suggested that other factors related to illness, such as appraisal of an individual's distress, discomfort and impairment may account for the remaining variance in CGI ratings.

The purpose of the present study is to evaluate the psychometric properties and determinants of a version of the CGI that was modified for use in studies of individuals with social anxiety disorder (also known as social phobia). The modified version of the CGI has been used as an outcome measure in several studies of pharmacological and psychosocial treatments for social anxiety disorder (Heimberg *et al.* 1998; Liebowitz *et al.* 1999). Unlike previous versions of the CGI, the modified version provides specific instructions to assessors, descriptive anchors for each rating, and a time frame for assessing severity and improvement. Clinicians are asked to make two ratings using a seven-step categorical scale: (1) Severity of Illness within the last week; and (2) Global Improvement from baseline. In doing so, clinicians are instructed to take into account the severity of social anxiety disorder symptoms, as well as impairment and distress resulting from the

disorder, and the presence of related problems (such as depressed mood or substance use).

Modifications to the original Global Improvement rating include the replacement of the terms 'very much improved' and 'much improved' with 'markedly improved' and 'moderately improved', respectively. In addition, descriptive anchors for Global Improvement ratings include explicit reference to corresponding Severity of Illness ratings. For example, 'markedly improved' is defined as 'obviously improved and no more than 3 ('mildly ill') on corresponding severity scale'. Additionally, CGI ratings are assigned following an interview that includes an extensive assessment of the presence and severity of social anxiety symptoms, depressive symptoms and functional impairment.

The CGI, as modified for individuals with social anxiety disorder, is thus explicitly intended to reflect intensity and frequency of core social anxiety symptoms, while taking into account the degree of interference in functioning and any accompanying dysphoria. Individuals with social anxiety disorder often experience depressed mood (Rapee, 1995) and are likely to report disruptions in educational achievement, family relationships, romantic relationships, friendships, and occupational functioning as a result of their symptoms (Schneier *et al.* 1994). These various facets of an individual's overall well-being are expected to contribute to judgements of Severity of Illness and Global Improvement, independently of the presence of social anxiety disorder symptoms. It is possible, for example, that following treatment, an individual remains anxious in certain social situations, but is no longer disabled by the anxiety or the accompanying dysphoric mood.

The independent contribution of depressed mood and impairment to the assessment of social anxiety severity is supported by findings from a study that examined the Clinician's Severity Rating (CSR), a severity rating included in the Anxiety Disorders Interview Schedule (ADIS) (Hope *et al.* 1997). The CSR for social anxiety disorder is a clinician rating of distress and interference associated with social anxiety disorder symptoms. Hope and colleagues (1997) examined the relationship between the CSR for social anxiety disorder and self-report measures of depression, social anxiety and avoidance of social or performance situations. Among 186

individuals with social anxiety disorder, self-reported avoidance behaviour and depression each accounted for significant variance in the CSR (12·2% and 4·6%, respectively), over and above the variance accounted for by self-report measures of social anxiety.

It is of particular interest to examine the properties and determinants of the CGI ratings because they are commonly used to discriminate between responders and non-responders to treatments for social anxiety disorder. For example, a score of 1 (markedly or very much improved) or 2 (moderately or much improved) on the CGI Global Improvement item has been used to define response to treatment in clinical trials evaluating the efficacy of imipramine (Simpson *et al.* 1998), clonazepam (Davidson *et al.* 1993), phenelzine (Liebowitz *et al.* 1992), moclobemide (Schneier *et al.* 1998), paroxetine (Stein *et al.* 1998), and cognitive–behavioural group therapy (Heimberg *et al.* 1998) for social anxiety disorder. Although these studies have demonstrated that the Severity of Illness and Improvement ratings are sensitive to change both within treatment and across treatment conditions, the utility of these ratings can only be fully understood by examining their relationship to external indicators.

We sought to evaluate the modified version of the CGI in a large sample of patients with social anxiety disorder. We examined the distribution properties of the Severity of Illness and Global Improvement items as well as the extent to which these clinician-assigned ratings are related to the patients' subjective reporting of symptom severity, impairment and overall quality of life. This latter question is particularly important since clinical trials often rely on clinicians' impressions of global outcome that may or may not correspond to the patients' subjective experience. Finally, we examined the extent to which depressive symptoms and functional impairment contribute to global ratings of severity and improvement, above and beyond symptoms of social anxiety.

METHOD

Participants

Participants were 123 individuals (52 women, 71 men) between the ages of 18 and 65 (mean = 33·0, s.d. = 10·2) who sought treatment for social anxiety at one of two sites of a collaborative study comparing the efficacy of cognitive–behavioural group therapy, phenelzine, the combination of both, and pill placebo for social anxiety disorder. The two sites from which data were drawn are: the Adult Anxiety Clinic of Temple University (AACT), Philadelphia, Pennsylvania ($N = 65$), and the Anxiety Disorders Clinic of the New York State Psychiatric Institute (NYSPI) ($N = 58$). All participants met DSM-IV criteria for a principal diagnosis of social anxiety disorder as determined by one of the following instruments: the Structured Clinical Interview for DSM-IV – patient edition (SCID-I/P) (First *et al.* 1996) or the Anxiety Disorders Interview Schedule for DSM-IV-Lifetime Version (ADIS-IV-L) (DiNardo *et al.* 1994). Both are semi-structured clinical interviews, designed to assess the presence of current and lifetime DSM-IV anxiety, mood and substance use disorders among individuals ages 18 and above. Individuals were excluded from participation in the present study if they showed evidence of schizophrenia, prominent risk of self-harm, alcohol or substance abuse in the last 6 months, an organic mental disorder, or history of bipolar-I disorder. Individuals seeking treatment at the NYSPI site were also excluded if they met criteria for current major depression. However, at the AACT, individuals who met criteria for current major depression were not excluded, as long as the diagnosis was not principal or co-principal with social anxiety disorder.

The demographic characteristics of participants in this study are shown for the total sample and for each site in Table 1. As shown in Table 1, approximately 59% of the total sample was Caucasian, while the remaining 41% consisted of African-Americans and individuals from other ethnic groups (e.g. Hispanic, Asian, Middle-eastern). Additionally, the majority of the sample reported having at least some college education and having never been married.

Demographic differences between sites were examined using analysis of variance (ANOVA) for age and chi-square analyses for all categorical variables (gender, race, marital status, education). There were no significant differences between sites in age ($F(1, 119) = 0·40$, NS), gender ($\chi^2_{(1)} = 0·31$, NS), or marital status

Table 1. *Demographic characteristics of participants by site*

	Total sample (N=123)		AACT (N=65)		NYSPI (N=58)	
	N	%	N	%	N	%
Age, mean (S.D.)	33·0 (10·2)		32·5 (10·7)		33·7 (9·5)	
range	19–65		19–65		20–61	
Gender						
Female	52	42·3	29	44·6	23	39·7
Male	71	57·7	36	55·4	35	60·3
Race						
Caucasian	71	59·2	49	76·6	22	39·3
African-American	22	18·3	8	12·5	14	25·0
Other	27	22·5	7	10·9	20	35·7
Marital status						
Married	25	20·5	13	20·0	12	21·1
Single, never married	87	71·3	44	67·7	43	75·4
Previously married (separated, divorced or widowed)	10	8·2	8	12·3	2	3·4
Education						
Any high school	26	21·5	4	6·2	22	39·3
Any college	70	57·8	46	70·8	24	42·9
Post-college education	25	20·7	15	23·1	10	17·9

AACT, Adult Anxiety Clinic at Temple University, Philadelphia, Pennsylvania; NYSPI, Anxiety Disorders Clinic of the New York State Psychiatric Institute. Sample sizes vary because of missing data. Percentages may not add to 100% because of rounding errors.

($\chi^2_{(2)} = 3·14$, NS). However, there were ethnic differences between sites. The subsample from NYSPI had fewer Caucasians and more individuals from minority groups than found in the AACT subsample ($\chi^2_{(3)} = 18·16$, $P < 0·001$) (see Table 1). Additionally, there were differences between sites in reported educational attainment with patients from the AACT subsample more likely than those from the NYSPI sample to have had at least some college education ($\chi^2_{(2)} = 19·82$, $P < 0·001$) (see Table 1). Although differences were found between sites on certain demographic variables, these variables were not associated with differences on any of the clinical measures, and therefore were not controlled in subsequent analyses. These site differences suggest that use of data pooled from both sites provides a more diverse sample than available from either site alone.

Procedure

All participants in the present study completed an intake assessment that included a diagnostic interview and completion of self-report measures of social anxiety, depression, and functional impairment. Those who met inclusion criteria were randomly assigned to one

of several treatment conditions in the controlled clinical trial. All participants in the present sample were assessed before treatment began (week 0), at mid-treatment (week 6) and post-treatment (week 12). At each assessment, participants completed self-report inventories and were interviewed by an independent assessor who was blind to the individual's specific treatment condition. Participants were asked to withhold information that would compromise this blind. Independent assessors were masters-level clinicians, psychiatrists and psychiatric nurses who were trained to acceptable levels of reliability. During the independent assessment interview, the interviewer administered the social anxiety disorder section of the ADIS-IV as well as other clinician-rated measures of depression, impairment and social anxiety (see below). The Severity of Illness and Global Improvement ratings of the CGI scale were made by the interviewer immediately following the independent assessment.

Measures

Clinician-administered measures

In addition to the CGI, the assessor completed the following measures.

Liebowitz Social Anxiety Scale (LSAS)
(Liebowitz, 1987)

The LSAS is a 24-item clinician-administered measure designed to assess the extent of fear and avoidance experienced in situations involving social interaction and social performance or scrutiny. Heimberg and colleagues (1999) reported alpha (α) coefficients for all LSAS subscales that ranged from 0·83 to 0·96. Furthermore, the LSAS correlated highly with other self-report and clinician-rated measures of social anxiety and avoidance (ranging from 0·49 to 0·73). In the current study, we used the LSAS Total score, which is the calculated sum of total fear and total avoidance scores.

Hamilton Rating Scale for Depression (HRSD)
(Hamilton, 1967)

The HRSD is a 21-item, clinician-administered measure that assesses symptoms of depression within the previous week. These include recent changes in mood, social activity, appetite, sleep disturbance, energy level, self-esteem, irritability and tension as well as other somatic complaints. Ratings on all items are summed to calculate a total score. The psychometric properties of the HRSD have been investigated with various clinical populations, showing good reliability and validity (see Katz *et al.* 1995).

Disability Profile (DP) (Schneier et al. 1994)

The DP assesses current (past 2 weeks) and worst lifetime impairment in eight domains of functioning, including school, work (outside or inside the home), family, marriage/dating, friendships, hobbies/interests, and daily living activities and chores. The 'current' disability score was used in the present study as an index of clinician-rated impairment. Schneier *et al.* (1994) reported high internal consistency ($\alpha = 0.87$) and moderate to high correlations between the DP and other measures of current disability and social anxiety symptom severity. In addition, scores on the DP discriminated between social anxiety disorder patients and a normal control sample (Schneier *et al.* 1994).

Self-report questionnaires

Social Interaction Anxiety Scale (SIAS)
(Mattick & Clarke, 1998)

The 20-item SIAS assesses the degree of self-reported anxiety associated with situations involving interactions in dyads or groups. Data from several studies support the validity of this measure (Heimberg *et al.* 1992; Brown *et al.* 1997; Mattick & Clarke, 1998). Although the SIAS was administered along with its companion measure, the Social Phobia Scale, which assesses fears of being observed and scrutinized by others (Mattick & Clarke, 1998), for the purposes of brevity, only SIAS scores were used in the present analyses.

Beck Depression Inventory (BDI)
(Beck et al. 1979)

The BDI is a 21-item self-report questionnaire that assesses a broad range of symptoms associated with depression. In a review of research on the psychometric properties of the BDI with clinical and non-clinical samples, Beck *et al.* (1988) concluded that the BDI has strong psychometric properties and high internal consistency. Furthermore, the BDI has has good internal consistency and retest reliability among individuals with social anxiety disorder (Coles *et al.* 2001).

Sheehan Disability Scale (SDS) (Sheehan, 1983)

This three-item questionnaire was used to assess self-reported impairment in occupational functioning, social life/leisure activity, and family life/home responsibilities. The reliability and concurrent validity of this measure has been documented by Leon *et al.* (1997), who reported a coefficient alpha of 0·89. Furthermore, elevated impairment scores were found among 80% of patients with common psychiatric diagnoses, including patients with alcohol and drug dependence, generalized anxiety disorder, major depression, obsessive–compulsive disorder and panic disorder.

Quality of Life Inventory (QOLI)
(Frisch et al. 1992)

The QOLI is a self-report questionnaire that inquires about satisfaction in 16 domains of functioning (e.g. health, relationships, money, work). The participant is asked to rate the importance (0–2) of each domain to his or her life, as well as his or her satisfaction with the domain on a 6-point scale from −3 (very dissatisfied) to +3 (satisfied). The total score is determined by calculating the product of the satisfaction and importance ratings for each area of life, then

averaging all weighted satisfaction scores that have non-zero importance ratings. The psychometric properties of this measure have been evaluated in several different populations, including in-patients, counselling centre clients, undergraduates, and criminal offenders (Frisch *et al.* 1992). High test–retest reliability coefficients were reported, ranging from 0·77 to 0·91 across various samples (Frisch *et al.* 1992). Safren *et al.* (1997) evaluated the QOLI among individuals with social anxiety disorder and found that quality of life scores were inversely associated with various measures of severity of social anxiety, functional impairment and depression, further suggesting its utility in the present sample. This measure was administered at post-treatment (week 12) only.

RESULTS

Site differences on clinical measures

Differences between sites on pre-treatment measures of social anxiety, depression and impairment were examined. Three one-way multivariate analyses of variance (MANOVAs), with site membership as the independent variable, were conducted for pre-treatment scores on measures of social anxiety, depression and impairment, respectively. There were no significant differences between sites on either self-report or clinician-administered measures of social anxiety, impairment and self-reported quality of life.

CGI descriptive statistics and distribution properties

Sites did not differ on CGI Severity of Illness or Improvement ratings at pre-, mid- or post-treatment assessment. Across sites, CGI Severity of Illness ratings ranged from 4 to 7 at pre-treatment (mean = 5·27, s.d. = 0·72), from 2 to 7 at week 6 (mean = 4·65, s.d. = 1·15), and from 1 to 7 at week 12 (mean = 4·21, s.d. = 1·42). CGI Global Improvement ratings ranged from 1 to 5 at both week 6 (mean = 2·79, s.d. = 0·91) and week 12 (mean = 2·32, s.d. = 1·00).

The distribution properties of the CGI were examined by calculating skewness and kurtosis for both ratings across assessment points. Skewness refers to the asymmetry of a distribution, whereas kurtosis refers to the extent to which observations in a distribution cluster around a central point (the 'peakedness' of a distribution). Values > 1 or < -1 generally indicate strong deviation from a normal distribution. CGI Severity of Illness ratings were normally distributed at pre-treatment (skewness = -0.066, kurtosis = -0.463), at week 6 (skewness = -0.238, kurtosis = -0.466), and at week 12 of treatment (skewness = -0.322, kurtosis = -0.565). Global Improvement ratings were normally distributed at week 6 (skewness = -0.056, kurtosis = -0.072) and at week 12 (skewness = 0.432, kurtosis = -0.492). These results suggest that the distribution of CGI ratings did not deviate significantly from normality.

Relationship between CGI ratings

The relationship between the Global Improvement and Severity of Illness ratings was examined at mid-treatment (week 6) and post-treatment (week 12). Zero-order correlations between Global Improvement ratings and concurrent Severity of Illness ratings were high at week 6 ($r = 0.72$) and at week 12 ($r = 0.82$). Because Global Improvement ratings reflect change in functioning from baseline, corresponding changes in Severity of Illness were determined by calculating change in Severity of Illness after adjusting for pre-treatment severity. Linear regression analyses were used to determine residual change in Severity of Illness from pre-treatment to week 6 and from pre-treatment to week 12. Zero-order correlations between Global Improvement ratings and residual change in Severity of Illness were high at week 6 ($r = 0.74$) and week 12 ($r = 0.82$).

Relationship between CGI ratings and measures of social anxiety, depression, impairment and quality of life

Zero-order correlations were calculated in order to determine whether CGI ratings were related to both self-report and clinician-administered measures of social anxiety (SIAS, LSAS), depression (BDI, HRSD), impairment (SDS, DP) and self-reported quality of life (QOLI). These correlations are presented in Table 2. Severity of Illness was significantly correlated with all self-report measures (rs = 0·41 to 0·77) and clinician-administered measures (rs = 0·36 to 0·84) at each assessment point (see Table 2). Zero-order correlations were examined between Global

Table 2.   *Correlations between Clinical Global Impression Severity of Illness* (CGI-SEV), *Clinical Global Impression Improvement* (CGI-IMP), *and self-report and clinician-administered measures*

	CGI-SEV Pre-treatment	CGI-SEV Week 6	CGI-SEV Week 12	CGI-IMP Week 6†	CGI-IMP Week 12‡
Self-report					
SIAS	0·49**	0·71**	0·74**	0·44**	0·63**
BDI	0·48**	0·50**	0·56**	0·25*	0·42**
SDS	0·55**	0·66**	0·77**	0·48**	0·54**
QOLI	−0·45**	—	−0·52**	—	−0·43**
Clinician-administered					
LSAS	0·59**	0·84**	0·83**	0·71**	0·74**
HRSD	0·36**	0·44**	0·60**	0·35*	0·42**
DP	0·71**	0·68**	0·72**	0·48**	0·46**

SIAS, Social Interaction Anxiety Scale; BDI, Beck Depression Inventory; SDS, Sheehan Disability Scale; QOLI; Quality of Life Inventory; LSAS, Liebowitz Social Anxiety Scale; HRSD, Hamilton Rating Scale for Depression; DP, Disability Profile.

† Correlations between Improvement ratings at week 6 and residual change scores in self-report and clinician-administered measures from baseline to week 6.

‡ Correlations between Improvement ratings at week 12 and residual change scores in self-report and clinician-administered measures from baseline to week 12.

* $P < 0.05$; ** $P < 0.01$.

Improvement ratings and residual change in self-report and clinician-administered measures (i.e. change from baseline after adjusting for pre-treatment scores). Global Improvement ratings were significantly correlated with residual changes in self-reported social anxiety and depressive symptoms from baseline to week 6 and with residual changes in self-reported social anxiety, impairment due to social anxiety, and overall quality of life from baseline to week 12. Additionally, Global Improvement ratings were significantly correlated with residual changes in all clinician-administered measures from baseline to week 6 to week 12.

Hierarchical, set-wise regression analyses were conducted to determine to whether functional impairment and dysphoria contribute to the variance in CGI ratings, independently of social anxiety symptoms. Separate analyses were conducted with self-report measures (SIAS, BDI, SDS) and clinician-rated measures (LSAS, HRSD, DP) of social anxiety, depression and impairment, respectively (see Tables 3 and 4). Participants' scores on these measures were entered as predictor variables, with either the Severity of Illness or Global Improvement rating entered as the criterion variable. In analyses predicting Global Improvement, predictor variables were entered in the form of pre-determined residual change scores, representing change from baseline after adjusting for pre-treatment severity. The order in which predictors were entered into the regression equation was determined on an *a priori* basis. Since the presence of social anxiety disorder is typically the primary reason for participation in this study, as well as the focus of the administered treatments, social anxiety symptoms are given primary consideration when Severity and Improvement ratings are assigned. Measures of social anxiety symptoms were thus entered first into the regression equation, followed by simultaneous entry of measures of impairment and depressive symptoms.

As presented in Table 3, at pre-treatment assessment, 27·9% of the variance in the CGI Severity of Illness rating was accounted for by self-reported social anxiety symptoms ($F(1, 112) = 43.36$, $P < 0.005$), with self-reported impairment and depressive symptoms accounting for an additional 14% of the variance ($F(2, 110) = 13.18$, $P < 0.005$). At week 6, self-reported social anxiety symptoms accounted for 46·2% of the variance in Severity of Illness ratings ($F(1, 68) = 58.38$, $P < 0.005$), while self-reported impairment and depression did not account for significant additional variance ($F(2, 66) = 2.40$, NS). At week 12, self-reported social anxiety symptoms accounted for 55% of the variance of Severity of Illness ratings ($F(1, 64) = 79.63$, $P < 0.005$), with significant additional variance accounted for by self-reported impairment and depression ($F(2, 62) = 8.43$, $P < 0.05$).

Similar analyses were conducted to determine whether CGI Severity ratings could be predicted

Table 3. *Hierarchical set-wise regression analysis predicting the severity of illness rating with self-report and clinician-administered measures of social anxiety, depression and impairment*

	Predictors	B	95% CI	Partial r	R^2	F	df
Pre-treatment							
Self-report							
Set 1	SIAS	0·025**	0·017 to 0·032	0·53	0·279	43·36**	1, 112
Set 2	BDI	0·018**	0·003 to 0·032	0·22	—		
	SDS	0·034**	0·012 to 0·057	0·27	0·418	13·18**	2, 110
Clinician-administered							
Set 1	LSAS	0·018**	0·014 to 0·022	0·61	0·376	71·11**	1, 118
Set 2	HRSD	0·019*	0·002 to 0·036	0·20	—	—	—
	DP	0·072**	0·045 to 0·099	0·44	0·540	20·70**	2, 116
Week 6							
Self-report							
Set 1	SIAS	0·046**	0·034 to 0·058	0·68	0·462	58·38**	1, 68
Set 2	BDI	0·002	−0·033 to 0·036	0·01	—	—	—
	SDS	0·06*	0·000 to 0·119	0·24	0·498	2·403	2, 66
Clinician-administered							
Set 1	LSAS	0·034**	0·029 to 0·039	0·85	0·771	192·7**	1, 76
Set 2	HRSD	0·045	−0·028 to 0·037	0·03	—	—	—
	DP	0·060*	0·006 to 0·084	0·26	0·739	3·08*	2, 74
Post-treatment							
Self-report							
Set 1	SIAS	0·061**	0·047 to 0·075	0·74	0·55	79·63**	1, 64
Set 2	BDI	0·013	−0·035 to 0·061	0·07	—	—	—
	SDS	0·089**	0·039 to 0·139	0·41	0·65	8·43**	2, 62
Clinician-administered							
Set 1	LSAS	0·039**	0·033 to 0·046	0·83	0·683	157·52**	1, 73
Set 2	HRSD	0·028	−0·016 to 0·074	0·56	—	—	—
	DP	0·046*	−0·007 to 0·100	0·67	0·710	3·25*	2, 71

SIAS, Social Interaction Anxiety Scale; BDI, Beck Depression Inventory; SDS, Sheehan Disability Scale; LSAS, Liebowitz Social Anxiety Scale; HRSD, Hamilton Rating Scale for Depression; DP, Disability Profile.
* $P < 0.05$; ** $P < 0.005$.

by clinician-administered measures of social anxiety (LSAS), depression (HRSD) and impairment (DP) during the course of treatment (Table 3). The LSAS accounted for 37·6% of the variance in Severity of Illness ratings at pre-treatment assessment $(F(1, 118) = 71·11, P < 0·005)$, with clinician-administered measures of depression and impairment accounting for an additional 16% of the variance $(F(2, 116) = 20·70, P < 0·005)$. A similar pattern was evident at week 6, with clinician-administered measures of depression and impairment adding to the prediction of Severity of Illness ratings independently of social anxiety symptoms $(F(2, 74) = 3·08, P < 0·05)$. At week 12, clinician-assessed social anxiety accounted for 68·3% of the variance in Severity of Illness ratings $(F(1, 76) = 192·7, P < 0·005)$, with significant additional variance accounted for by measures of depressive symptoms and impairment $(F(2, 71) = 3·25, P < 0·05)$.

As shown in Table 4, residual changes in self-reported social anxiety symptoms accounted for 19·5% of the variance in ratings of Global Improvement at week 6 $(F(1, 68) = 16·47, P < 0·005)$, with no additional variance accounted for by self-reported impairment and depressive symptoms $(F(2, 66) = 1·34,$ NS). Similarly, at week 12, self-reported impairment and depressive symptoms did not contribute to the prediction of Global Improvement ratings above and beyond changes in self-reported social anxiety symptoms. Finally, residual change in clinician-administered LSAS scores accounted for significant variance in Global Improvement ratings at mid- and post-treatment assessments (week 6, $F(1, 79) = 94·12, P < 0·005$; week 12, $F(1, 74) = 79·37, P < 0·005$), with no significant additional variance accounted for for by measures of depressive symptoms and impairment at either assessment point.

Table 4. *Hierarchical set-wise regression analysis predicting the Global Improvement rating with changes in self-report and clinician-administered measures of social anxiety, depression and impairment*

	Predictors	B	95% CI	Partial r	R^2	F	df
Week 6							
Self-report							
Set 1	SIAS	0·034**	0·017 to 0·051	0·44	0·195	16·47**	1, 68
Set 2	BDI	−0·002	−0·043 to 0·039	−0·01	—	—	—
	SDS	0·039	−0·014 to 0·091	−0·18	0·227	1·34	2, 66
Clinician-administered							
Set 1	LSAS	0·031**	0·025 to 0·037	0·74	0·544	94·12**	1, 79
Set 2	HRSD	0·025	−0·009 to 0·059	0·16	—	—	—
	DP	0·033	−0·007 to 0·074	0·18	0·575	2·79	2, 77
Post-treatment							
Self-report							
Set 1	SIAS	0·043**	0·029 to 0·057	0·62	0·390	38·42**	1, 60
Set 2	BDI	−0·007	−0·041 to 0·055	0·04	—	—	—
	SDS	−0·035	−0·018 to 0·088	0·17	0·422	1·60	2, 58
Clinician-administered							
Set 1	LSAS	0·032**	0·025 to 0·039	0·72	0·517	79·37**	1, 74
Set 2	HRSD	0·034	−0·009 to 0·077	0·18	—	—	—
	DP	0·038	−0·011 to 0·088	0·18	0·552	2·80	2, 72

SIAS, Social Interaction Anxiety Scale; BDI, Beck Depression Inventory; SDS, Sheehan Disability Scale; LSAS, Liebowitz Social Anxiety Scale; HRSD, Hamilton Rating Scale for Depression; DP, Disability Profile.
* $P<0.05$; ** $P<0.005$.

DISCUSSION

The present study examined properties of a modified version of the CGI, an outcome measure commonly used in clinical trials for the treatment of social anxiety disorder. Overall, our findings support the utility of the Severity of Illness and Global Improvement ratings among individuals with social anxiety disorder. CGI ratings were strongly associated with both self-report and clinician-administered measures of specific symptomatology and impairment.

The descriptive and statistical properties of the Severity of Illness and Global Improvement ratings support the use of parametric statistics at various assessment points during the course of treatment. Furthermore, CGI ratings were strongly related to each other in the present sample. For example, ratings of Global Improvement during treatment were very highly correlated with both concurrent Severity of Illness (i.e. greater improvement was associated with lower patient severity) and with changes in Severity of Illness from baseline. This finding may be due in part to the fact that when assigning Global Improvement ratings, assessors were explicitly instructed to consult and take into account both concurrent and baseline

Severity of Illness ratings. A commonly raised concern about the CGI is that when clinicians assign a rating of global change, their reliance on memory of baseline functioning might compromise the validity of the rating (Leon *et al.* 1993). The strong relationship between Global Improvement ratings at mid- and post-treatment and calculated changes in Severity of Illness ratings supports the assumption that Global Improvement ratings reflect actual changes in functioning from baseline.

The strong relationship between Severity of Illness and Global Improvement ratings raises the question of whether the Global Improvement rating could or should be replaced by a calculation of change in Severity of Illness from baseline to any given point in treatment. In their criticism of the CGI, Beneke & Rasmus (1991) note that single ratings of Global Improvement assigned at a particular point in treatment do not take into account or provide any information regarding baseline differences among patients. One advantage, then, of using changes in Severity of Illness ratings rather than Global Improvement ratings is that it would allow adjustment for differences in baseline severity. As will be discussed, however, differing patterns in the prediction of Severity of Illness and Global

Improvement ratings suggest that these ratings may not be so easily interchangeable.

Severity of Illness and Global Improvement ratings were both highly correlated with self-report and clinician-assessed indices of social anxiety, depressive symptoms and impairment. However, differences emerged when we examined the extent to which variables other than social anxiety symptoms contributed to the variance in Severity of Illness and Global Improvement ratings. Across both self-report and clinician-rated measures, depressive symptoms and functional impairment significantly improved the prediction of Severity of Illness ratings over and above the contribution of social anxiety symptoms. However, changes in depressive symptoms and impairment made no independent contribution to Global Improvement ratings beyond that associated with improvements in social anxiety symptoms.

Severity of Illness ratings therefore appear to be more 'global' than Improvement ratings, reflecting core symptoms of social anxiety disorder as well as dysphoria and the degree of functional capacity in various domains of life. The relative breadth of Severity of Illness ratings in the present sample is consistent with findings from Hope *et al.*'s (1997) study evaluating the ADIS Clinician's Severity Rating (CSR). Like the Severity of Illness rating, the CSR is intended to encompass interference and distress associated with social anxiety disorder. Hope and colleagues reported that among individuals with social anxiety disorder, the CSR was associated with symptoms of social anxiety as well as extensive behavioural avoidance and dysphoria. In contrast, the Global Improvement rating appears to be relatively specific, perhaps because clinicians who rate improvement during or following treatment are likely to focus on the symptoms that are specifically targeted by the intervention, in this case social anxiety.

Overall, the regression models accounted for a significant portion of the variance in the CGI. However, this variance was shared differently in Severity of Illness and Global Improvement ratings, suggesting that each rating reflects different aspects of a patients' clinical outcome. In the light of these findings, it is worth evaluating the role of these ratings as indices of patient functioning and change. Bobes (1998) proposes three types of measurement that should be used to evaluate global improvement in social anxiety: (*a*) objective (i.e. objectively perceived improvement in social anxiety symptoms and avoidance behaviour); (*b*) adaptive (i.e. improvement in the patient's level of functioning and disability relative to their potential); and (*c*) subjective (i.e. subjectively perceived improvement in well-being). Both CGI ratings appear to capture the subjective and objective criteria, although the assessment of improvement in social anxiety is limited to the interview setting. However, measures of functional impairment and disability only contributed significantly to Severity of Illness ratings. To capture changes in impairment during or following treatment (i.e. adaptive measurement), separate measures of this construct may be needed to supplement the CGI Improvement rating. Frisch *et al.* (1992) note that the absence or reduction of symptoms does not necessarily imply improved functioning and satisfaction in various domains of life.

Several limitations of the present study merit consideration. First, the present evaluation of the CGI did not address the issue of reliability, an important psychometric property on which there is little reported data. However, an earlier version of the adapted CGI demonstrated good reliability in a small sample of individuals with social anxiety disorder (Juster *et al.* 1993). The only other known study that examined the inter-rater reliability of the original CGI Severity and Improvement ratings found moderate to low reliability, particularly for the Improvement ratings (Dahlke *et al.* 1992). However, that study examined a sample of inpatients with dementia and used a version of the CGI that offered no instructions to the clinicians assigning the ratings. Therefore, the criteria used to determine what constitutes Severity or Improvement may have been more ambiguous and less consistently applied. Investigators using the CGI as a primary outcome measure should examine the reliability of these ratings among independent assessors. Secondly, the range of scores on the Severity and Improvement ratings were somewhat restricted by the use of a sample whose inclusion into the study required a pre-specified level of severity.

Although the present findings support the overall utility of the CGI for social anxiety disorder, it is important to note that variation in

methodology from study to study may influence determination of these ratings. Differences in the use and content of descriptive anchors, as well as in the prior training of assessors may influence the extent to which particular variables contribute to severity and improvement ratings. For example, in the current study, CGI ratings were assigned in the context of an extensive clinical assessment that included the administration of specific measures of symptoms and functional impairment. These specific measures were administered by a trained assessor, and were used to inform global severity and improvement ratings. Therefore, it is possible that in the absence of these more specific measures, the CGI would operate differently as an outcome measure. Additionally, the fact that clinician-administered measures included in the present analyses were completed by the same person who assigned the CGI ratings may have confounded the relationship between these external validators and the Severity of Illness and Global Improvement ratings. Finally, the generalizability of the current findings may be limited to the version of the CGI modified for patients with social anxiety disorder.

Despite these limitations, our data indicate that the CGI adapted for social anxiety disorder functions well as an index of global severity and specific improvement in a clinical population of treatment-seeking individuals. These ratings were strongly related to one another, as well as to self-report and clinician-rated measures of specific symptomatology and accompanying impairment. Although the present findings support the continued use of the CGI in clinical trials, they also alert us to the importance of supplementing the Global Improvement rating with measures that capture changes in quality of life and impairment in various domains of functioning. Increased clarity as to what variables do and do not contribute to global outcome ratings will better inform the construction of clinically meaningful treatment response criteria.

REFERENCES

Barlow, D. H., Gorman, J. M., Shear, M. K. & Woods, S. W. (2000). Cognitive-behavioral therapy, imipramine, or their combination for panic disorder: a randomized controlled trial. *Journal of the American Medical Association* **283**, 2529–2536.

Beck, A. T., Rush, A. J., Shaw, B. F. & Emery, G. (1979). *Cognitive Therapy of Depression*. Guilford Press: New York.

Beck, A. T., Steer, R. A. & Garbin, M. G. (1988). Psychometric properties of the Beck Depression Inventory: twenty-five years of evaluation. *Clinical Psychology Review* **8**, 77–100.

Beneke, M. & Rasmus, W. (1991). 'Clinical Global Impressions' (ECDEU): some critical comments. *Pharmacopsychiatry* **25**, 171–176.

Bobes, J. (1998). How is recovery from social anxiety disorder defined? *Journal of Clinical Psychiatry* **59**, 12–19.

Brady, K., Pearlstein, T., Asnis, G. M., Baker, D., Rothbaum, B., Sikes, C. R. & Farfel, G. M. (2000). Efficacy and safety of sertraline treatment of posttraumatic stress disorder. *Journal of the American Medical Association* **283**, 1837–1844.

Brown, E. J., Turovsky, J., Heimberg, R. G., Juster, H. R., Brown, T. A. & Barlow, D. H. (1997). Validation of the Social Interaction Anxiety Scale and the Social Phobia Scale across the anxiety disorders. *Psychological Assessment* **9**, 21–27.

Coles, M. E., Gibb, B. E. & Heimberg, R. G. (2001). A psychometric evaluation of the Beck Depression Inventory in adults with social anxiety disorder. *Depression and Anxiety* **14**, 145–148.

Dahlke, F., Lohaus, A. & Gutzmann, H. (1992). Reliability and clinical concepts underlying global judgments in dementia: implications for clinical research. *Psychopharmacological Bulletin* **28**, 425–432.

Davidson, J. R., Potts, N., Richichi, E., Krishnan, R., Ford, S. M., Smith, R. & Wilson, W. H. (1993). Treatment of social phobia with clonezepam and placebo. *Journal of Clinical Psychopharmacology* **13**, 423–428.

DiNardo, P. A., Brown, T. A. & Barlow, D. H. (1994). *Anxiety Disorders Interview Schedule for DSM-IV: Lifetime Version (ADIS-IV-L)*. Graywind: Albany, NY.

First, M., Spitzer, R., Gibbon, M. & Williams, J. (1996). *Structured Clinical Interview for DSM-IV Axis I Disorders – Patient Edition (SCID-I/P, Version 2.0)*. Biometrics Research Department, New York State Psychiatric Institute: New York.

Frisch, M. B., Cornell, J., Villanueva, M. & Retzlaff, P. J. (1992). Clinical validation of the Quality of Life Inventory: a measure of life satisfaction for use in treatment planning and outcome assessment. *Psychological Assessment* **4**, 92–101.

Guy, W. (ed.). (1976). *ECDEU Assessment for Psychopharmacology, Revised Edition*. NIMH Publication: Rockville, MD.

Hamilton, M. (1967). Development of a rating scale for primary depressive illness. *British Journal of Social and Clinical Psychology* **6**, 278–296.

Heimberg, R. G., Mueller, G. P., Holt, C. S., Hope, D. A. & Liebowitz, M. R. (1992). Assessment of anxiety in social interaction and being observed by others: the Social Interaction Anxiety Scale and the Social Phobia Scale. *Behavior Therapy* **23**, 53–73.

Heimberg, R. G., Liebowitz, M. R., Hope, D. A., Schneier, F. R., Holt, C. S., Welkowitz, L., Juster, H. R., Campeas, R., Bruch, M. A., Cloitre, M., Fallon, B. & Klein, D. F. (1998). Cognitive-behavioral group therapy versus phenelzine in social phobia: 12-week outcome. *Archives of General Psychiatry* **55**, 1133–1141.

Heimberg, R. G., Horner, K. J., Juster, H. R., Safren, S. A., Brown, E. J., Schneier, F. R. & Liebowitz, M. R. (1999). Psychometric properties of the Liebowitz Social Anxiety Scale. *Psychological Medicine* **29**, 199–212.

Hellerstein, D. J., Yanowitch, P., Rosenthal, J., Samstag, L. W., Maurer, M., Kasch, K., Burrows, L., Poster, M., Cantillon, M. & Winston, A. (1993). A randomized double-blind study of fluoxetine versus placebo in the treatment of dysthymia. *American Journal of Psychiatry* **150**, 1169–1175.

Hope, D. A., Laguna, L. B., Heimberg, R. G. & Barlow, D. (1997). Relationship between ADIS Clinician's Severity Rating and self-report measures among social phobics. *Depression and Anxiety* **4**, 120–125.

Hudson, J. I., McElroy, S. L., Raymond, N. C., Crow, S., Keck, P. E., Carter, W. P., Mitchell, J. E., Strakowski, S. M., Pope, H. G., Coleman, B. S. & Jonas, J. M. (1998). Fluvoxamine in the treatment of binge-eating disorder: a multicenter placebo-controlled, double-blind trial. *American Journal of Psychiatry* **155**, 1756–1762.

Juster, H. R., Heimberg, R. G. & Mattia, J. I. (1993). Judgment and prediction of response to cognitive-behavioral group therapy for social phobia. Poster presented at the annual meeting of the Association for Advancement of Behavior Therapy, Atlanta, GA.

Katz, R., Shaw, B. F., Vallis, T. M. & Kaiser, A. S. (1995). The assessment of severity and symptom patterns in depression. In *Handbook of Depression, 2nd edn* (ed. E. E. Beckham and W. R. Leber), pp. 61–85. Guilford: New York.

Leon, A. C., Shear, K., Klerman, G. L., Portera, L., Rosenbaum, J. F. & Goldenberg, I. (1993). A comparison of symptom determinants of patient and clinician global ratings in patients with panic disorder and depression. *Journal of Clinical Psychopharmacology* **13**, 327–331.

Leon, A. C., Olfson, M., Portera, L., Farber, L. & Sheehan, D. V. (1997). Assessing psychiatric impairment in primary care with the Sheehan Disability Scale. *International Journal of Psychiatry in Medicine* **27**, 93–105.

Liebowitz, M. R. (1987). Social phobia. *Modern Problems in Pharmocopsychiatry* **22**, 141–173.

Liebowitz, M. R., Schneier, F. R., Campeas, R., Hollander, E., Hatterer, J., Fyer, A., Gorman, J. M., Papp, L., Davies, S., Gully, R. & Klein, D. F. (1992). Phenelzine vs atenolol in social phobia: a placebo-controlled comparison. *Archives of General Psychiatry* **49**, 290–300.

Liebowitz, M. R., Heimberg R. G., Schneier, F. R., Hope, D. A., Davies, S., Holt, C. S., Goetz, D., Juster, H. R., Lin, S. H., Bruch, M. A., Marshall, R. D. & Klein, D. F. (1999). Cognitive-behavioral group therapy versus phenelzine in social phobia: long term outcome. *Depression and Anxiety* **10**, 89–98.

Mattick, R. P. & Clarke, J. C. (1998). Development and validation of measures of social phobia scrutiny fear and social interaction anxiety. *Behaviour Research and Therapy* **36**, 455–470.

Rapee, R. M. (1995). Descriptive psychopathology. In *Social Phobia: Diagnosis, Assessment and Treatment* (ed. R. G. Heimberg, M. R. Liebowitz, D. A. Hope and F. R. Schneier), pp. 41–66. Guilford Press: New York.

Safren, S. A., Heimberg, R. G., Brown, E. J. & Holle, C. (1997). Quality of life in social phobia. *Depression and Anxiety* **4**, 126–133.

Schneier, F. R., Heckelman, L. R., Garfinkel, R., Campeas, A. R., Fallon, B. A., Gitow, A., Street, L., Del Bene, D. & Liebowitz, M. R. (1994). Functional impairment in social phobia. *Journal of Clinical Psychiatry* **55**, 322–331.

Schneier, F. R., Goetz, D., Campeas, R., Fallon, B., Marshall, R. & Liebowitz, M. R. (1998). Placebo-controlled trial of moclobemide in social phobia. *British Journal of Psychiatry* **172**, 70–77.

Sheehan, D. V. (1983). *The Anxiety Disease*. Charles Scribner and Sons: New York.

Simpson, H. B., Schneier, F. R., Campeas, R. B., Marshall, R. D., Fallon, B. A., Davies, S., Klein, D. F. & Liebowitz, M. R. (1998). Imipramine in the treatment of social phobia. *Journal of Clinical Psychopharmacology* **18**, 132–135.

Stein, M. B., Liebowitz, M. R., Lydiard, R. B., Pitts, C. D., Bushnell, W. & Gergel, I. (1998). Paroxetine treatment of generalized social phobia (social anxiety disorder): a randomized controlled trial. *Journal of the American Medical Association* **280**, 708–713.