1-1-1996

The Yale-Brown Obsessive Compulsive Scale: Interview Versus Self-report

Gail Steketee
Boston University

Randy Frost
Smith College, rfrost@smith.edu

Karen Bogart
Smith College

Pergamon

S0005-7967(96)00036-8

Behav. Res. Ther. Vol. 34, No. 8, pp. 675–684, 1996

0005-7967/96 $15.00 + 0.00

THE YALE–BROWN OBSESSIVE COMPULSIVE SCALE: INTERVIEW VERSUS SELF-REPORT

GAIL STEKETEE[1]*, RANDY FROST[2] and KAREN BOGART[2]

[1]School of Social Work, Boston University, 264 Bay State Road, Boston, MA 02215, U.S.A. and
[2]Smith College, Northampton, MA 01063, U.S.A.

(Received 26 June 1995; Revised 16 April 1996)

Summary—Several studies have demonstrated the reliability and validity of the Yale–Brown Obsessive Compulsive Scale (YBOCS) conducted by trained interviewers. The present study examined several aspects of a self-report YBOCS version relative to the usual interview format in two non-clinical samples (*n*s = 46 and 70) and in a clinical OCD sample ($n = 36$) and a clinical non-OCD group ($n = 10$). The self-rated instrument showed excellent internal consistency and test–retest reliability, performing somewhat better than the interview. There was good agreement between symptom checklist categories across the two versions, though clinical subjects reported more symptoms on the self-report form than on the interview. Some order effects were evident for non-clinical subjects only: those who received the self-report first scored lower on both self-report and interview than those who received the interview first. No order effects were observed in the clinical sample. The self-report version showed strong convergent validity with the interview, and discriminated well between OCD and non-OCD patients. Although more study is needed, particularly on clinical samples, these findings suggest that the self-report YBOCS may be a time-saving and less costly substitute for the interview format in assessing OCD symptoms. Copyright © 1996 Elsevier Science Ltd

INTRODUCTION

The Yale–Brown Obsessive Compulsive Scale (YBOCS: Goodman, Price, Rasmussen, Mazure, Fleischmann, Hill, Heninger & Charney, 1989a; Goodman, Price, Rasmussen, Mazure, Delgado, Heninger & Charney, 1989b) offers considerable advantages over other measures of obsessive compulsive (OC) symptoms because scale ratings are not based on the patient's particular type of OC symptoms (e.g. obsessions about contamination and washing rituals, obsessions about harm and checking rituals). Instead, after inquiring about what types of obsessions and compulsions the patient experiences using a standard checklist, interviewers ask patients to identify their main symptoms (obsessions and compulsions) and respond to a series of questions that comprise the YBOC Scale. The scale is divided into the Obsessions subscale and the Compulsions subscale, with additional supplementary items that are not included in subscale scores or total scores. For each subscale, five aspects of obsessive and compulsive pathology are each rated on a scale ranging from 0 (no symptoms) to 4 (extreme symptoms): time spent, degree of interference, distress, resistance (greater resistance is assigned lower scores), and perceived control over the symptom. Subscale scores are summed to yield a YBOCS Total score.

Several psychometric studies of the YBOCS on clinical populations have yielded generally positive results (Goodman *et al.*, 1989a, b; Kim, Dysken & Kuskowski, 1990; Kim, Dysken & Kuskowski, 1992; Woody, Steketee & Chambless, 1995). Somewhat fewer data are available for non-clinical samples (Frost, Steketee, Krause & Trepanier, 1995; Rosenfeld, Dar, Anderson, Kobak & Greist, 1992; Warren, Zgourides & Monto, 1993). Internal consistency (Cronbach's alpha) ranged from 0.69 (Woody *et al.*, 1995) to 0.91 (Goodman *et al.*, 1989a) in clinical patients and was 0.88 in a non-clinical sample (Frost *et al.*, 1995). Subscale alphas for Obsessions and Compulsions subscales were 0.77 and 0.51 for clinical patients (Woody *et al.*, 1995) and 0.78 and 0.84 respectively in a college student sample (Frost *et al.*, 1994). Interrater reliability for the total score was high (0.89–0.93) (Goodman *et al.*, 1989a; Woody *et al.*, 1995). However, test–retest reliability was more variable across studies depending on methodology: intraclass correlations

*Author for correspondence.

ranged from 0.61 for a mean 48-day interval using different raters (Woody *et al.*, 1995) to 0.97 for a 1-week interval by the same rater (Kim *et al.*, 1992).

Evidence for convergent validity was good overall, with same modality measures (e.g. interviewer-rated) showing somewhat higher correlations than non-matched formats (e.g. self-report, behavioral measure) (Frost *et al.*, 1995; Goodman *et al.*, 1989b; Kim *et al.*, 1990; 1992). Woody *et al.* (1995) observed much higher convergent validity after treatment than at pretest, probably because of restriction of range on pretest measures in a treatment-seeking OCD population. Findings for discriminant validity were somewhat more problematic, especially using measures of depression (Goodman *et al.*, 1989b; Woody *et al.*, 1995).

Only one study examined the construct validity of the YBOCS using factor analysis (Danyko, McKay & Neziroglu, 1995) for 83 patients with OCD. As predicted, these authors found distinct, relatively uncorrelated factors for obsessions and compulsions, with only the resistance item from the Obsession subscale failing to load significantly on either factor. Interestingly, Woody *et al.* (1995) also noted some problems with low item-total correlations for the resistance items for both subscales and poor test–retest reliability for resistance to obsessions. They proposed omitting these items from the total scale. Woody *et al.* (1995) also examined the usefulness of including supplemental item No. 12, avoidance. This item performed well with regard to internal consistency and convergent validity with other measures of OC fear and avoidance. They proposed including it in the Total score.

In summary, data gathered from several independent sources indicate that the YBOCS Total score has good reliability and convergent but not discriminant validity, and is sensitive to change following treatment. Addition of avoidance and omission of resistance items might improve the utility of the instrument. Further, although the YBOCS addresses the problem of comprehensive assessment of OCD symptoms that are highly variable in their presentation, in its present interview format it is still a time-consuming, and therefore costly, measure to employ. Studies of non-clinical samples tend to use self-report instruments, despite their probable difficulties with sensitivity and specificity in detecting subclinical and diagnosable OCD symptoms. At present, the YBOCS interview represents the best effort to accurately assess OCD severity, but a self-report version with good reliability and validity would address the difficulty of cost. Such a scale has recently been developed and has shown preliminary promise (Baer, Brown-Beasley, Sorce & Henriques, 1993; Rosenfeld *et al.*, 1992; Warren *et al.*, 1993).

A computer-administered version was compared with a clinician-administered YBOCS in 47 outpatients (31 with OCD and 16 with other anxiety disorders) and 23 nonpatients (Rosenfeld *et al.*, 1992). Results indicated generally good agreement: for 76% of *S*s, Total scores for the two instruments were within three points (out of a maximum of 40). Correlations of subscales and Total score for the two versions in the OCD group ranged from 0.86 to 0.88. For anxiety patients, correlations between versions were in the 0.75 range, and for nonpatients, correlations were poor (0.01–0.03), due to severe restriction of range. As expected, OCD patients scored higher than other groups. Using a cutoff score of 16 or more (a criterion commonly used in medication trials), the two forms of administration identified the same *S*s as falling above or below the cut off, except for two anxiety patients, and with the same two exceptions, effectively discriminated OCD patients from the other samples. Subjects reported being comfortable using the computer version and there was no evidence for a preference for the clinician-rated version.

Two reports provide additional evidence supporting a self-report YBOCS. Baer *et al.* (1993) administered a telephone talking-computer version, an in-person telephone interview, and a self-report questionnaire version of the YBOCS to 18 patients with OCD. All versions were given in counterbalanced order when the self-report and computer versions were compared with the interview version, agreement between scale totals was very high ($rs = 0.97$ and 0.99), and mean Total scores were nearly identical. The time taken to complete scales indicated that *S*s using the computer took longest and the interviewer shortest. However, the very brief times reported (3–5 min) indicate that these data could not refer to both the checklist and the YBOCS scale completion since this usually requires at least 20 min. Interestingly, most patients preferred the live interview and none preferred the computer.

Warren *et al.* (1993) administered a self-report YBOCS to 180 undergraduates and 50 medical patients. Internal reliability was excellent for subscales and Total score (0.88–0.91) and mean scores

were reported to provide baseline data for a normal sample (3.7 for YBOCS Total, indicating a non-obsessive sample). No comparisons with the interview version were made.

The present study was conducted to provide preliminary information from both clinical and non-clinical samples about the psychometric properties of a non-computerized self-report YBOCS compared to the interviewer-rated measure. Burns, Formea, Keortge and Sternberger (1995) have noted the advantages of using nonclinical samples to study clinical problems. Evidence of reliability and validity for the self-report version would support its use as the more cost-effective and efficient instrument for descriptive and clinical research.

METHOD

Subjects

Psychometric studies of the YBOCS were conducted on four separate samples of *S*s. Sample 1 included 47 undergraduate women from a small liberal arts college who were solicited from an introductory psychology course and received class credit for participation. One *S* who was receiving treatment for OCD was not included in the sample, since this was intended to represent a non-clinical sample. Age range for the 46 remaining *S*s was 17–21 yr. Most *S*s were Caucasian, with 12 (26%) from minority groups: 5 Asian, 3 Latino, 3 African American, and 1 Native American. This sample received either the interview or the self-report YBOCS versions in order to examine test–retest reliability.

Sample 2 contained 73 undergraduate women from a small liberal arts college who were solicited from an introductory psychology course and received class credit for participation. Three of the *S*s failed to complete both parts of the study, so only the remaining 70 were used for data analysis. The age range was 17–22 yr. Unfortunately, data on ethnicity were not collected for this sample. Thirty-eight *S*s received the self-report version first, followed by the interview, and the remaining 32 received the reverse order. The order of administration was random. Analyses for this sample were conducted to examine internal consistency and concurrent validity.

Sample 3 consisted of 36 individuals (16 men and 20 women) who sought treatment for obsessive compulsive disorder. Their age ranged from 17 to 63, with a mean of 32.8 yr. All met DSM III-R and DSM IV criteria for OCD, determined by either the Structured Clinical Interview for DSM III-R (kappa for OCD diagnosis was 1.0) or by a structured interview with a highly experienced clinician/researcher of OCD and other anxiety disorders. All *S*s reported that OCD was their primary complaint. Data from the pretreatment assessment are presented. This sample received both versions of the YBOCS in counterbalanced order.

Sample 4 consisted of 10 *S*s (mean age 44.5 yr, range 22–72; 70% female) who met criteria for another Axis I or Axis II disorder for which they sought treatment. Diagnoses were confirmed by a structured clinical interview. This group served as a comparison clinical sample. Primary diagnoses for this group were obsessive compulsive personality disorder ($n = 1$), social phobia (1), trichotillomania (2), generalized anxiety disorder (1), borderline personality disorder (1), panic disorder (2), panic disorder with agoraphobia (1), and post-traumatic stress disorder (1). This sample received both versions of the YBOCS in counterbalanced order, permitting study of internal consistency, concurrent validity, and discriminant validity.

Measures

YBOCS interview. In accordance with methods devised by the YBOCS originators, the interviewer began with a detailed checklist to inquire about obsessions (aggression/harming, contamination, sexual, hoarding/saving, religious, symmetry/exactness, somatic, and miscellaneous) and compulsions (washing/cleaning, checking, repeating, counting, ordering/arranging, hoarding/collecting, and miscellaneous). From this 64-item checklist, the interviewer asked the *S* to select three main obsessions and compulsions that most distressed the *S*. Semi-structured questions about the nature of these symptoms then focused on five areas each for the obsessions and for the compulsions: time spent, interference, distress, resistance, and control. Items were scored on a 0 (none) to 4 (extreme) scale yielding Obsession and Compulsion subscale scores (range 0–20) and a Total score (range 0–40). Subjects did not select their own scores for each YBOCS

item, but rather interviewers asked questions as outlined in the YBOCS manual in order to rate the severity of *Ss*' reported symptoms. This interview required approximately 25 min for non-clinical and 30 min clinical *Ss*.

Self-report YBOCS. The self-report paper-and-pencil form of the YBOCS developed by Baer *et al.* (1993) is very similar to the interview. Subjects are first asked to note the presence of 58 obsessions and compulsions from the same general categories as in the interview. They are then asked to circle the three main obsessions and three main compulsions. The second part of the report asks *Ss* to focus on the main obsessions and main compulsions and to answer five questions for each type as noted above: time spent, interference, distress, resistance, and control. Consistent with the interview format, *Ss* rate each item on a 0 (none) to 4 (extreme) scale. Subjects reported that this measure required approximately 10 min to complete.

Procedure

For the nonclinical *Ss*, both forms of the YBOCS were administered by four undergraduate psychology students trained by the first author who has extensive clinical experience using this instrument. Students had formal training in research and clinical aspects of obsessive compulsive disorder. Training consisted of three phases: (1) a 3-hr meeting at which the content of each YBOCS checklist item was discussed and ratings of all items of the scales were explained; (2) supervised practice interviews; and (3) pilot interviews that were audiotaped and reviewed by the trainer who gave feedback to interviewers. Each experimenter administered the YBOCS interview or self-report to approximately 4 *Ss* each week. To examine interrater reliability, a subset of 14 audiotapes of YBOCS interviews, approximately evenly distributed across interviewers were also scored by the first author who was not informed of *Ss*' recorded scores. The interrater reliability coefficients were all highly significant ($P < 0.001$): Intraclass correlations (ICC) were 0.72 for Obsessions, 0.64 for Compulsions, and 0.80 for the Total score, arguing for the validity of the undergraduate-rated interview data.

For clinical *Ss*, the YBOCS interview was administered by trained clinical interviewers with at least a masters degree in a clinical mental health discipline. All but one were experienced in treating anxiety disorders and OCD in particular and were trained by the first author using a format similar to that described above. In a previous study from which 10 of the current clinical *Ss* were drawn, interrater reliability was very high for YBOCS Total score (ICC = 0.93).

To examine test–retest reliability of the YBOCS, the 46 non-clinical *Ss* from Sample 1 were randomly assigned to receive either the interview twice ($n = 20$) or the self-report twice ($n = 26$), both administered approximately 1 week apart.

To study the relationship between YBOCS forms, non-clinical (Sample 2) and clinical *Ss* (Sample 3) were randomly assigned to receive either the interview or the self-report version first, followed by the other measure about 1 week later to minimize recall of previous responses on the other version. Thirty-two of the non-clinical *Ss* were interviewed first and 38 completed the self-report form first. For the 36 clinical *Ss*, 21 received the self-report YBOCS first; the mean time between occasions was 4.1 days, range 1–34. For clinical non-OCD *Ss* (Sample 4) the same procedures were employed, with half the sample receiving the interview first, and the other half receiving the questionnaire first. There was an interval of approximately 1 week between administrations.

RESULTS

Internal consistency

For the non-clinical *Ss*, we examined internal consistency with Sample 2 rather than Sample 1 because of the larger sample size. As evident from Table 1, the internal consistency for the self-report version was excellent. Alphas were 0.84 or higher (Table 1). Figures for the interview version for this sample were also good, although not quite as high as for the self-report scale, ranging from 0.78 to 0.88. Findings for clinical OCD *Ss* (Sample 3) were somewhat less satisfactory for both YBOCS versions, particularly for the Obsessions subscales which were below 0.60. Alphas

Table 1. Internal consistency (Cronbach's alpha) for non-clinical and clinical samples for self-report and interview YBOCS

	YBOCS self-report	YBOCS interview
Sample 2 (non-clinical; $n = 70$)		
Obessions	0.85	0.80
Complusions	0.84	0.78
Total	0.89	0.88
9-item total*	0.90	0.89
Sample 3 (OCD; $n = 36$)		
Obsessions	0.55	0.56
Complusions	0.71	0.61
Total	0.78	0.74
9-item total	0.77	0.74

*The 9-item total excludes the two resistance items and includes avoidance (item No. 12), as proposed by Woody *et al.* (1995).

for the total scales on both versions were lower than 0.80. The interview did not perform quite as well as the self-report version for the Compulsions subscale.

Consistent with findings from recent research on clinical *S*s using the YBOCS interview format (Woody *et al.*, 1995), the resistance items for obsessions and compulsions collected by interview showed low item-remainder correlations (0.21 for Obsessions and 0.28 for Compulsions). To investigate whether alphas would improve if these items were omitted and the avoidance item (No. 12) was added as Woody *et al.* proposed, we recalculated internal consistency on the revised 9-item scale for both versions. For non-clinical *S*s, the revision increased alpha by only 0.01 for both the self-report and interview versions (Table 1), perhaps because alphas for both were already quite high. Alphas for the revised scale for clinical *S*s were 0.77 for self-report and 0.74 for the interview. Thus, for both the self-report and interview scales, the revised scoring method improved little upon the 10-item standard scoring.

Test–retest reliability

Pearson correlations were used to examine the strength of relationship of scores for individual items and for subscales and Total scores at first and second administration for non-clinical *S*s in Sample 1. As evident from Table 2, self-report items were significantly and at least moderately correlated, with *r*s ranging from 0.40 to 0.83. Subscale and Total scores were all highly related indicating excellent test–retest reliability. Correlations for the interview items were generally lower than for the self-report version, and several were not significant. It is noteworthy that the interview resistance item for both obsessions and compulsions had particularly low correlations. Subscale and Total scores were also lower for the interview, especially for obsessions. However, the sample of 20 *S*s for the interview correlations is lower than desirable.

Mean scores for Sample 1 are given in Table 3. As expected, scores for both versions were in the non-clinical range on both occasions. The self-report scores from both occasions were slightly

Table 2. Test–retest Pearson correlation coefficients for individual items comparing the first and second administration 1 week apart of the self-report and interview YBOCS for nonclinical *S*s (Sample 1)

	Self-report YBOCS ($n = 26$)	Interview YBOCS ($n = 20$)
Obsessions subscale		
Time spent	0.89**	0.45
Interference	0.65**	0.27
Distress	0.66**	0.26
Resistance	0.83**	0.11
Control	0.40*	0.76**
Obsessions total	0.87**	0.55*
Compulsions subscale		
Time spent	0.72**	0.71**
Interference	0.46*	0.32
Distress	0.70**	0.56*
Resistance	0.56**	0.16
Control	0.88**	0.72**
Complusions total	0.82***	0.77**
Total score	0.88***	0.79**

*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$.

Table 3. Means and standard deviations (SD) for interviews and self-report YBOCS for nonclinical *Ss* (Sample 1)

	Time 1			Time 2		
	Mean	SD	Range	Mean	SD	Range
Self-report scores ($n = 25$)						
Obsessions	5.0	3.1	0–12	4.5	3.6	0–13
Complusions	3.5	3.8	0–13	3.6	3.7	0–10
Total	8.4	5.9	0–22	8.0	6.6	0–21
Interview scores ($n = 19$)						
Obsessions	4.5	2.7	1–10	3.7	2.2	0–7
Compulsions	3.5	2.9	0–9	3.4	2.9	0–8
Total	8.0	5.3	1–19	7.2	4.5	0–14

but not significantly higher overall than means for the interview scores. Two-way ANOVAs (time × version) indicated no main effect for version and no interaction of version by time for subscales or Total YBOCS scores. However, the main effect for time was significant for obsessions only [$F(1,42) = 4.33$, $P < 0.05$]; the YBOCS Obsessions subscore was lower at Time 2 than at Time 1 for both versions, but there were no time effects for Compulsions or Total score.

Test–retest reliability of YBOCS checklist

We also tested whether the self-report and interview checklist were reliable from first to second administration for non-clinical Sample 1. We elected to examine categories of obsessions and compulsions, rather than individual checklist items. When we compared the number of items checked within categories using *t*-tests, using Bonferroni corrections, there were no significant differences. There were also no differences in the total number of items checked for either self-report or interview. According to these indices, *Ss*' responses on the checklist appear to be stable over time.

Content validity of the YBOCS checklist

For Sample 2, we examined the self-report and interview checklists to determine whether there were noteworthy differences across versions in the content and number of items that *Ss* reported. We again studied categories of obsessions and compulsions, as well as the total number of items. The most frequently checked categories for both the interview and self-report versions for non-clinical *Ss* were miscellaneous obsessions, miscellaneous compulsions and aggressive obsessions. *t*-tests with a Bonferroni correction to control for number of comparisons, indicated that non-clinical *Ss* checked more aggressive obsessions on self-report than interview versions [$t(69) = -3.37$, $P < 0.005$]. No other categories differed significantly, and there were no differences between versions for the total number of items checked.

Among the 28 checklists available for clinical OCD *Ss* (Sample 3) no significant differences emerged between self-report and interview endorsements. Although not significant, the frequencies for all categories were somewhat higher for self-report assessment compared to the interview, but whether *Ss* are over-reporting or reporting more honestly on the self-report version cannot be determined from these data.

Order effects

Means for both YBOCS versions for non-clinical and OCD *Ss* in Samples 2 and 3 are given in Table 4. As expected, clinical *Ss* showed substantially higher scores than non-clinical *Ss*. To examine order effects, as well as differences between versions, we conducted 2 × 2 ANOVAs (order × version). For non-clinical *Ss*, a main effect emerged for order on both subscales and Total score: $Fs(1,68) = 9.49$ for Obsessions, 7.92 for Compulsions, and 8.74 for Total score, all $Ps < 0.006$. Non-clinical *Ss* ($n = 38$) who received the self-report version first, followed by the interview scored lower overall on both versions than the sample ($n = 32$) who received the interview first, followed by the self-report (Table 4). No significant main effect of version or interaction of version-by-order were found. For the clinical OCD sample, no order effect or order-by-version interactions were found, but version effects were evident for subscales and Total scores: $Fs(1,29) = 9.73$ for Obsessions, 10.00 for Compulsions, and 12.58 for Total score, $Ps < 0.004$.

Table 4. Means and standard deviations (SD) for interview and self-report YBOCS for nonclinical and clinical OCD *Ss*

		Time 1		Time 2	
		Mean	SD	Mean	SD
Non-clinical Sample 2					
		Self-report		Interview	
	Obsessions	2.4	2.6	2.4	2.6
(n = 38)	Complusions	2.4	2.9	2.9	3.1
	Total	4.8	4.8	5.4	5.6
		Interview		Self-report	
	Obsessions	4.2	2.9	4.3	3.7
(n = 32)	Compulsions	4.8	3.2	4.9	4.3
	Total	9.0	5.7	9.2	7.2
Clinical OCD Sample 3					
		Self-report		Interview	
	Obsessions	10.8	2.1	11.4	2.6
(n = 16)	Compulsions	10.7	3.6	12.2	2.9
	Total	21.4	5.4	23.6	5.3
		Interview		Self-report	
	Obsessions	12.4	2.9	10.9	2.9
(n = 15)	Compulsions	11.5	2.9	10.6	2.9
	Total	23.9	5.0	21.5	5.2

Clinical *Ss* scored consistently slightly higher on the interview format compared to self-report regardless of order.

Convergent validity

The subtotal and Total scores of the YBOCS interview and self-report were significantly and strongly correlated with one another among non-clinical *Ss* from Sample 2. Pearson correlations were 0.69 for Obsessions, 0.65 for Compulsions and 0.75 for the Total score (Ps < 0.001). Because of order effects reported above for these *Ss*, we also examined correlations among self-report and interview scales separately for those who received the interview first (n = 32) and for those who received the self-report version first (n = 38). Correlations remained essentially unchanged with the sole exception of the Compulsions scale for the latter group; this was somewhat lower (r = 0.54, $P < 0.001$). As Table 5 indicates, many of the individual YBOCS items were significantly related, with the strongest relationships usually, but not always, evident for matching items. Interestingly,

Table 5. Pearson correlation coefficients for the 10 items from the YBOCS interview and self-reports form given to nonclinical *Ss* and clinical OCD *Ss*

	Time	Interference	Distress	Resistance	Control
Non-clinical Sample 2 (n = 70)					
Interview obsessions			Self-report obsessions		
Time spent	**0.58***	0.34*	0.52*	0.19	0.33*
Interference	0.45*	**0.48***	0.47*	0.33*	0.47*
Distress	0.63*	0.40*	**0.62***	0.26	0.41*
Resistance	0.44*	0.44*	0.51*	**0.30**	0.52*
Control	0.42*	0.38*	0.46*	0.22	**0.51***
Interview compulsions			Self-report compulsions		
Time spent	**0.62***	0.37*	0.52*	0.39*	0.46*
Interference	0.34*	**0.48***	0.46*	0.44*	0.63*
Distress	0.53*	0.44*	**0.59***	0.39*	0.51*
Resistance	0.36*	0.40	0.22	**−0.02**	0.11
Control	0.49*	0.45*	0.57*	0.38*	**0.50***
Clinical OCD Sample 3 (n = 36)					
Interview obsessions			Self-report obsessions		
Time spent	**0.62****	0.08	0.44*	0.17	0.21
Interfernce	0.47*	**0.56****	0.28	0.19	0.20
Distress	0.22	0.10	**0.58****	0.11	0.31
Resistance	−0.01	0.19	0.08	**0.41****	0.05*
Control	0.18	0.17	0.27	0.31	**0.50***
Interview compulsions			Self-report compulsions		
Time spent	**0.76****	0.21	0.46*	0.15	0.20
Interference	0.64**	**0.55****	0.51**	0.10	0.27
Distress	0.47*	0.29	**0.55****	0.29	0.18
Resistance	0.21	0.06	0.10	**0.58****	0.42*
Control	0.23	0.09	0.06	0.55**	**0.44***

*$P < 0.01$, **$P < 0.001$.Correlations between corresponding items for self-report and interview versions are in bold type.

Table 6. Means (standard deviations) and *t*-test comparisons between OCD patients and non-OCD patients on self-report and interview YBOCS.

	OCD patients	Non-OCD patients	*t*	*df*	*P*
Self-report					
Obsessions	10.6 (2.7)	6.6 (5.1)	3.33	(44)	0.01
Compulsions	10.7 (3.3)	6.0 (5.0)	3.62	(44)	0.001
Total	21.3 (5.3)	12.6 (9.3)	3.87	(44)	0.001
Interview					
Obsessions	11.9 (2.7)	6.5 (5.3)	4.39	(44)	0.001
Compulsions	11.8 (2.9)	7.8 (5.7)	3.09	(44)	0.01
Total	23.7 (5.0)	14.3 (9.9)	4.16	(44)	0.001

correlations for both resistance to obsessions and resistance to compulsions for the interview and self-report forms were low and nonsignificant. In fact, self-reported resistance to compulsions was more strongly related to all other interview compulsion items than to resistance.

For the clinical sample, Pearson correlations between interview and self-report versions were also very strong for the subscales and Total score: $r = 0.73$ for Obsessions, 0.78 for Compulsions and 0.79 for Total score, Ps < 0.001. Correlations for corresponding individual items (Table 5) were highly significant, indicating excellent convergent validity of the two versions. Resistance items were not problematic in this sample.

A comparison of mean subscale and Total scores on the self-report and interview versions using *t*-tests yielded mixed results for the two samples. Means were not different for non-clinical *S*s. However for the clinical sample, a consistent pattern was evident for lower scores on the self-report version compared to the interview (refer to Table 4): for Obsessions, $t(35) = 4.06$, $P < 0.001$; for Compulsions, $t(30) = 3.34$, $P < 0.01$; for Total score, $t(35) = 4.37$, $P < 0.001$. However, the differences between means were not large and probably not clinically significant.

Criterion-related validity

To study criterion-related validity of the measures, we compared YBOCS scores for OCD patients with a small sample of 10 non-OCD patients who completed both the self-report and interview. As evident from Table 6, the non-OCD sample had significantly lower subscale and Total scores on both measures. However, using Bonferroni correction, the differences on self-reported obsessions and interview compulsions are only marginal; Total scores remain significantly different for both versions. To compare our findings with those of Rosenfeld *et al.* (1992), we also examined the number of *S*s in each group who fell above or below a commonly used clinical cutoff score of 16 for the self-report and interview measure. For 36 OCD patients, one person fell below the cutoff by one point using both versions, and one additional person did so for the self-report version only (score 11). Among non-OCD patients, on the interview, 5 (50%) scored above the cutoff and on the self-report version, 4 (40%) did so. Thus, both versions of the YBOCS gave similar results: very good sensitivity but relatively poor specificity.

DISCUSSION

Internal consistencies of the subscales and Total score for the self-report measure in the present study were good to excellent, and appeared noticeably better than the interview measure across all three samples. The alphas for the self-report Total score (0.79 for the clinical sample and 0.90–0.91 for non-clinical samples) were equivalent or slightly better than findings from our own and others' previous research (Frost *et al.*, 1995; Goodman *et al.*, 1989a; Woody *et al.*, 1995). It is interesting to note that the alphas for our OCD sample were not noticeably different from those observed in Woody *et al.*'s (1995) sample of 51 OCDs, although in our group, internal consistency was higher for compulsions than obsessions in the interview, the reverse of their findings. Like Woody *et al.*, for resistance items, we also found low item-remainder correlations, as well as low test–retest reliability for the interview and low concurrent validity. However, efforts to apply their suggested revision of the YBOCS scoring to a 9-item scale, omitting both resistance items and including avoidance did not improve internal consistency in either the non-clinical or clinical sample. Nonetheless, the resistance items remain somewhat problematic and warrant further research.

Test–retest correlations indicated excellent reliability over time for the self-report and good reliability for the compulsions scale of the interview. The moderate correlation for the Obsessions interview scale may have been due to a restricted range, although this is also evident for the Compulsion scale (Table 3). The absence of absolute differences as reflected in *t*-tests comparing Time 1 and 2 also argue for the stability of both versions. Findings for the self-report form were somewhat better than results for the interview measure for which our sample was somewhat smaller. These results compare favorably to reports from prior research on the interview form in two other sites (Kim *et al.*, 1990, 1992; Woody *et al.*, 1995).

With regard to validity, the self-report version compared well to the interview measure. Thus, interviewers did not identify more OCD symptoms than patients reporting alone. Within the non-clinical group, few differences were observed on the frequency of items selected from the checklist and for those that were noted, no particular pattern of bias was evident. There is somewhat more cause for concern in the clinical sample where there was a tendency for patients to report more symptoms on the self-report version. This potential bias could result from patients checking more items that essentially represent the same symptom viewed from different perspectives, whereas an interviewer might check only one item that best represented the reported symptom. Certainly, this finding does not represent a threat to validity unless it affects the scoring of severity on the scale itself. We did observe slightly lower scores on the self-report scale in the clinical sample only, indicating a possible need for upward adjustment of self-report scores to match interview findings.

Order effects for the YBOCS scale emerged only for non-clinical *S*s. When these *S*s completed the self-report version before the interview, they rated themselves as less symptomatic than they did after having first talked to an interviewer. Conversely, when they completed the interview first, they scored higher than when the interview was conducted second. Unfortunately, our findings cannot address the question of which rating was more accurate for these *S*s. However, the order effect did not appear to significantly influence the high correspondence between self-report and interview Obsessions and Total score, though it did reduce the correlations between Compulsions scores for one group. The absence of order effects in the clinical sample is reassuring, in that more severe symptoms appear to be less influenced by a preceding assessment of OCD symptoms.

Convergent validity of the self-report version with the interview was evident in the relatively high correlations (0.65–0.75 for non-clinical *S*s and 0.73–0.79 for clinical patients). Given the variance introduced by the 1-week gap between administrations, these figures are likely to be underestimates of the true correlations. Our findings for the non-clinical sample are substantially better than those reported by Rosenfeld *et al.* (1992) for their nonpatients, but it is clear from their mean data that they had considerably greater restriction of range in their relatively small nonclinical sample ($n = 23$) consisting of medical students and staff. Findings for the OCD sample are somewhat below those observed for patients taking computerized YBOCS (0.86 to 0.88; Rosenfeld *et al.*, 1992), but indicate good convergent validity nonetheless, though we note a general trend for the self-report version to yield slightly lower scores than the interview.

In our study, the 10 individual items comprising the YBOCS self-report and interview scales were also moderately to strongly correlated in both the clinical and non-clinical samples. The only exceptions are the two resistance items, but only for the non-clinical sample; these items were at least moderately correlated in the OCD sample. We noted that although the range on resistance in the self-report form for non-clinical *S*s was good, it was restricted in the interview version, possibly accounting for these low correlations. The low scores are not surprising in a group whose symptoms are mild and who therefore have little reason to resist obsessions or compulsions.

Although our non-OCD clinical sample was quite small, we nonetheless obtained good discriminant validity for the self-report YBOCS, as well as for the interview. The non-OCD *S*s scored somewhat higher than the non-clinical samples, but well below the OCD patients on both versions. The cutoff score of 16 showed good sensitivity, though relatively poor specificity. All but two OCD *S*s scored in the clinical range according to self-report, and all but one on the interview, but a large percentage of the non-OCD sample also scored in the clinical range on both versions of the YBOCS.

Overall, the self-report version of the YBOCS performed quite well on psychometric tests employed in the present research. Further study on a larger OCD sample is obviously needed before

conclusions can be clearly drawn about the utility of this instrument for assessing severity of OCD symptoms in clinical research and change following treatment. The present findings do suggest that the self-rated measure is useful for detecting OCD in non-clinical samples. We suggest that it may present a useful and potentially more comprehensive alternative to other self-report measures such as the Maudsley Obsessional–Compulsive Inventory (Hodgson & Rachman, 1977) or the Padua Inventory (Sanavio, 1988; van Oppen, de Haan, van Balkom, Spinhoven, Hoogduin & van Dyck, 1995).

Acknowledgements—Preparation of this manuscript was supported by grant No. MH44190 awarded by the National Institute of Mental Health to Gail Steketee and Dianne Chambless. The authors thank interviewers Bryan Fitzler, MA, of the University of Maine; Martha Gillis, MA, of the American University; Leslie Shapiro, MSW, and Hope Worden, MSW, of Boston University; and Tamara Hartl, BA, Rebecca Christian, BA, and Valerie Castro of Smith College whose assistance was instrumental in completing this project.

REFERENCES

Baer, L., Brown-Beasley, M. W., Sorce, J. & Henriques, A. (1993). Computer-assisted telephone administration of a structured interview for obsessive–compulsive disorder. *American Journal of Psychiatry*, *150*, 1737–1738.

Burns, G. L., Formea, G. M., Keortge, S. & Sternberger, L. G. (1995). The utilization of nonpatient samples in the study of obsessive compulsive disorder. *Behaviour Research and Therapy*, *33*, 133–144.

Danyko, S. J., McKay, D. R. & Neziroglu, F. A. (1995). Factor structure of the Yale–Brown Obsessive Compulsive Scale: a two dimensional measure. *Behaviour Research and Therapy*, *33*, 865–869.

Frost, R. O., Steketee, G., Krause, M. S. & Trepanier, K. L. (1995). The relationship of the Yale–Brown Obsessive Compulsive Scale (YBOCS) to other measures of obsessive compulsive symptoms in a nonclinical population. *Journal of Personality Assessment*, *65*, 158–168.

Goodman, W. K., Price, L. H., Rasmussen, S. A., Mazure, C., Fleischmann, R. L., Hill, C. L., Heninger G. R. & Charney, D. S. (1989a). The Yale–Brown Obsessive Compulsive Scale I. Development, use, and reliability. *Archives of General Psychiatry*, *46*, 1006–1011.

Goodman, W. K., Price, L. H., Rasmussen, S. A., Mazure, C., Delgado, P., Heninger, G. R. & Charney, D. S. (1989b). The Yale–Brown Obsessive Compulsive Scale II. Validity. *Archives of General Psychiatry*, *46*, 1012–1016.

Hodgson, R. and Rachman, S. (1977). Obsessive compulsive complaints. *Behaviour Research and Therapy*, *15*, 389–395.

Kim, S. W., Dysken, M. W. & Kuskowski, M. (1990). The Yale–Brown Obsessive–Compulsive Scale: a reliability and validity study. *Psychiatry Research*, *34*, 99–106.

Kim, S. W., Dysken, M. W. & Kuskowski, M. (1992). The Symptom Checklist-90: obsessive–compulsive subscale: a reliability and validity study. *Psychiatry Research*, *41*, 37–44.

Rosenfeld, R., Dar, R., Anderson, D., Kobak, K. A. & Greist, J. H. (1992). A computer-administered version of the Yale–Brown Obsessive Compulsive Scale. *Psychological Assessment 4*, 329–332.

Sanavio, E. (1988). Obsessions and compulsions: the Padua Inventory. *Behaviour Research and Therapy*, *26*, 169–177.

van Oppen, P., de Haan, E., van Balkom, A. J. L. M., Spinhoven, P., Hoogduin, K. & van Dyck, R. (1995). Cognitive therapy and exposure *in vivo* in the treatment of obsessive compulsive disorder. *Behaviour Research and Therapy 33*, 379–390.

Warren, R., Zgourides, G. & Monto, M. (1993). Self-report versions of the Yale–Brown Obsessive–Compulsive Scale: an assessment of a sample of normals. *Psychological Reports*, *73*, 574.

Woody, S. R., Steketee, G. & Chambless, D. L. (1995). Reliability and validity of the Yale–Brown Obsessive Compulsive Scale. *Behaviour Research and Therapy*, *33*, 597–605.