J Psychopathol Behav Assess (2015) 37:432–441
DOI 10.1007/s10862-014-9465-7

Assessment of Obsessive Compulsive Disorder in Young Children: Psychometric Properties of the Children's Yale-Brown Obsessive Compulsive Scale

Nathan E. Cook · Jennifer B. Freeman · Abbe M. Garcia · Jeffrey J. Sapyta · Martin E. Franklin

Published online: 14 October 2014

Abstract The Children's Yale-Brown Obsessive Compulsive Scale (CY-BOCS) is the most commonly used instrument to assess the symptoms and severity of pediatric obsessive compulsive disorder (OCD). However, only one prior study has evaluated the psychometric properties of the CY-BOCS for assessing young children, ages 5 to 8 years. The limited available evidence suggests that psychometric properties are less favorable with younger children. Thus, the present study aimed to re-examine the technical qualities of the CY-BOCS in a sample of 5 to 8 year olds with early-onset OCD. The sample consisted of 127 younger children, enrolled in a multi-site randomized controlled trial comparing the efficacy of family-based cognitive behavioral therapy to relaxation therapy. The CY-BOCS Total score demonstrated adequate internal consistency, although at a lower level than is typically reported in studies of older children. Internal consistency of the Obsessions and Compulsions subscales was poor. The Total and subscale scores demonstrated good temporal stability over 5 weeks. Agreement between clinician and parent versions was poor at baseline but improved substantially throughout the course of the trial. Results also indicated that the CY-BOCS had good convergent and discriminant validity. Further, certain CY-BOCS items appear more reliable indicators of OCD severity in younger children than others. Limitations, implications, and future directions are discussed, including the potential for further developmentally sensitive refinement of the CY-BOCS for this age group.

Keywords Obsessive-compulsive disorder · Psychometrics · Reliability · Construct validity

N. E. Cook
Department of Psychology, University of Rhode Island, Kingston, RI, USA

J. B. Freeman · A. M. Garcia
Department of Psychiatry and Human Behavior, Alpert Medical School of Brown University, Bradley/Hasbro Children's Research Center, Providence, RI, USA

J. J. Sapyta
Department of Psychiatry and Behavioral Sciences, Duke University Medical Center, Durham, NC, USA

M. E. Franklin
Department of Psychiatry, University of Pennsylvania School of Medicine, Philadelphia, PA, USA

N. E. Cook (✉)
Learning and Emotional Assessment Program, Department of Psychiatry, Massachusetts General Hospital, 151 Merrimac Street, Boston, MA 02881, USA
e-mail: necook@mgh.harvard.edu

Present Address:
N. E. Cook (✉)
Department of Psychiatry, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA

The Children's Yale-Brown Obsessive Compulsive Scale (CY-BOCS; Scahill et al. 1997) is the most commonly used instrument to assess the symptoms and severity of pediatric obsessive compulsive disorder (OCD; Merlo et al. 2007). It is utilized for both clinical and research applications (Lewin and Piacentini 2010) and is recommended as part of a diagnostic assessment for OCD by the American Academy of Child and Adolescent Psychiatry (2012). The CY-BOCS was developed by modifying the Yale-Brown Obsessive Compulsive Scale (Y-BOCS; Goodman et al. 1989), a highly regarded measure of adult OCD symptom severity. A recent review reported that the CY-BOCS was one of only two child and adolescent OCD measures with published research evidence regarding its psychometric properties (Southam-Gerow and Chorpita 2007), which are critical to consider when selecting and appraising clinical assessment measures (Haynes, Smith, and Hunsley 2011; Mash and Hunsley 2005).

To date, eight published studies have examined the psychometric properties of the CY-BOCS and, in general, reported favorable technical qualities. For example, findings from reliability analyses (summarized in Table 1) suggest good internal consistency according to recommended guidelines (Hunsley and Mash 2008). Specifically, across all eight studies the mean unweighted alphas were .84 for the CY-BOCS Total score, .81 for the Obsessions subscale, and .80 for the Compulsions subscale.

Although the CY-BOCS has received considerable research attention, including cross-cultural translation (Yucelen et al. 2006), and results support its technical qualities, extant studies have included samples predominated by older children and adolescents (see Table 1). Many studies include some younger children, however, there are generally few in absolute terms, with some studies including just five participants younger than 7-years-old. In fact, only one study has examined the psychometric properties of the CY-BOCS specifically in a sample of younger children, ages 5 to 8 (Freeman et al. 2011). The paucity of research with 5- to 8-year-olds is concerning given developmental factors relevant to the assessment of OCD with children this young. For example, young children may have difficulty reporting their symptoms or level of distress (Freeman et al. 2012). Also, younger children may be more likely to have OCD symptoms overlap with other disorders, as well as present with compulsions only, without stated obsessions (Freeman et al. 2012). Additionally, Garcia et al. (2009) pointed out that differentiating complex tics from rituals in young children is challenging and may require advanced cognitive and verbal abilities on the part of the child to describe internal states.

Further, modifications to the administration of the CY-BOCS may be necessary with younger children. For example, clinicians might complete the Compulsions subscale prior to the Obsessions subscale (Freeman, Flessner, and Garcia 2011; Freeman et al. 2012). Additionally, administration of the CY-BOCS will increasingly rely on parent-report. The lack of psychometric research on the CY-BOCS with samples of younger children and the need to tailor assessment procedures to ensure accurate information pose a challenge for clinical and research applications.

Moreover, the limited available evidence suggests the psychometric qualities of the CY-BOCS are less favorable in younger children. In terms of internal consistency, the CY-BOCS Total score alpha coefficient in a sample of younger children, with a mean age of 7, was α=.72 (Freeman, Flessner, and Garcia 2011). As a comparison, in Storch et al. (2006) sample of older children, with roughly the same number of participants as Freeman and colleagues' study but a mean age of 11 years, the internal consistency estimate for the CY-BOCS Total score was α=.89 (Storch et al. 2006). The difference between these two alpha coefficients is statistically significant ($W_{45,\ 38}$=2.55, p<.05; Feldt and Kim 2006), such that the internal consistency of the CY-BOCS Total score is significantly lower in younger children. Indeed, as shown in Table 1, Freeman et al. (2011) reported the lowest overall alpha estimates for the CY-BOCS Total and subscale scores.

Additionally, the test-retest reliability of the CY-BOCS can be compared across samples of older and younger children. Both Freeman et al. (2011) and Storch et al. (2004) reported test-retest stability estimates for follow up periods of about 6 to 8 weeks. The mean age of the older sample was 10 years (Storch et al. 2004), compared to 7 years for the younger sample (Freeman, Flessner, and Garcia 2011). The test-retest stability estimates for the CY-BOCS Total score were identical across these two studies (ICC=.79) and the stability of the Obsessions subscale was only slightly lower in younger children (ICC=.66 compared to .70). Conversely, the stability of the Compulsions subscale was substantially higher in younger (ICC=.89) compared to older children (ICC=.76).

These psychometric differences, combined with general developmental challenges to the accurate assessment of OCD symptoms in younger children, highlight the need for additional research regarding the technical qualities of the CY-BOCS as well as potential ways to refine the measure. The present study aimed to address this need by examining the psychometric properties of the CY-BOCS in a sample of younger children, a clinical population that has been largely underrepresented in prior research (Garcia et al. 2009). Our study used a larger sample than all but one previous psychometric evaluation of the CY-BOCS and was gathered in the context of a tightly controlled treatment outcome study. The CY-BOCS data were gathered from evaluators who underwent substantial training and were subjected to strict quality control procedures. Thus, this study aimed to provide more stable reliability and construct validity estimates. Specifically, we sought to contribute to the OCD assessment literature by examining the internal consistency, test-retest stability, and construct validity of the CY-BOCS in younger children.

Method

Participants

Participants were recruited as part of the pediatric obsessive compulsive disorder treatment study for young children (POTS Jr.), a multi-site randomized controlled trial comparing the efficacy of family-based cognitive behavioral therapy to relaxation therapy (RT) for treatment of children ages 5 to 8 with early-onset OCD. Freeman et al. (2012) provided a detailed description of the POTS Jr. methodology. Briefly, children between 5 and 8 years of age with a primary diagnosis of OCD were recruited across the three study sites (Brown University, Duke University, and the University of

Table 1 Summary of CY-BOCS reliability data

Study	Sample			CY-BOCS Total			Obsession subscale			Compulsion subscale		
	N	Mean Age	Age Range	α	ICC_{TR}	ICC_{IR}	α	ICC_{TR}	ICC_{IR}	α	ICC_{TR}	ICC_{IR}
Scahill et al. (1997)	65	12.1	8–17	.87	NR	.84	NR	NR	.91	NR	NR	.66
McKay et al. (2003)	233	10.8	5–17	.95	NR	NR	.92	NR	NR	.94	NR	NR
Storch et al. (2004)	61	10.3	4–18	.90	.79[b]	NR	.80	.70[b]	NR	.82	.76[b]	NR
Storch et al. (2005)	82	10.4	5–18	.76	NR	NR	NR	NR	NR	NR	NR	NR
Storch et al. (2006)	53	11.3	8–17	.89	NR	NR	.89	NR	NR	.79	NR	NR
Yucelen et al. (2006)[a]	19	14.0	8–16	.72	NR	.89	.79	NR	.94	.77	NR	.85
Freeman et al. (2011)	42	6.7	4–8	.72	.79[c]	NR	.64	.66[c]	NR	.71	.89[c]	NR
Conelea et al. (2012)	35	15.6	14–17	.85	NR	NR	.79	NR	NR	.76	NR	NR

Studies listed chronologically. α=Coefficient Alpha. ICC_{TR}=Test Retest Intraclass Correlation Coefficient. ICC_{IR}=Interrater Intraclass Correlation Coefficient. NR=Not Reported

[a] Turkish version of the CY-BOCS

[b] Average test-retest follow-up=5.8 weeks

[c] Average test-retest follow-up=8.2 weeks

Pennsylvania). Inclusionary criteria were a CY-BOCS Total score greater than 16, stable symptoms for more than 3 months, appropriateness for outpatient level of care, and parent availability to participate in treatment. Children were excluded if they required additional or alternative treatment (e.g., presented with a primary diagnosis other than OCD, had a Pervasive Developmental Disorder, were acutely suicidal) or if they might not have benefitted from study treatments (i.e., due to an intellectual disability). Additionally, children were excluded if they presented with factors that threatened the internal validity of the study. Such factors included engagement in concurrent psychotherapy, medication for depression, ADHD, OCD, or tics that had not been stable for greater than 8 weeks, had previously failed a trial of CBT for OCD, or met research criteria for PANDAS.

A total of 127 children met inclusionary criteria and were enrolled in POTS Jr. and ranged in age from 4 to 9 years at baseline assessment (M=6.7, SD=1.2). Two children were 4 years old at baseline but turned 5 during the course of treatment. Additionally, one child had recently turned 9 years old at the time of initial diagnostic assessment and was 8 years old at the time of the initial screening interview. The sample was 52.8 % female (n=67) and predominantly identified as Caucasian (90 %, n=114). The modal grade in school was second grade. Specific phobia (18.9 %; n=23), generalized anxiety disorder (18.3 %; n=22), oppositional defiant disorder (13.9 %; n=17), attention deficit hyperactivity disorder (13.1 %; n=16), social phobia (9.8 %, n=12), and separation anxiety disorder (9.8 %, n=12), were the most common comorbid diagnoses, as assessed via structured diagnostic interview (i.e., Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version; K-SADS-PL; Kaufman et al. 1997). Given the POTS Jr. inclusion/exclusion criteria, comorbid diagnoses were secondary, not primary or co-primary.

Procedure

Children underwent a multi-stage assessment procedure to determine eligibility for POTS Jr., including a screening interview (Gate A), systematic diagnostic assessment and team meeting to determine suitability for the study (Gate B), and a baseline assessment, involving the child and parent(s), conducted by an independent evaluator (IE; Gate C). Throughout the course of the study children were re-assessed by IEs at 5, 9, and 14 weeks. Additionally, naturalistic follow-up assessments were completed at 3, 6, and 12 months.

Independent evaluators were doctoral-level child clinical psychologists with previous experience evaluating child anxiety and OCD. Many of the IEs (at all three sites) had worked on prior pediatric OCD clinical trials and were working within the anxiety treatment clinic at the specific study site, but were not part of the POTS Jr trial. That is, they did not attend study specific meetings, were not involved in screening of calls, were not involved in treatment supervision, and therefore, were blind to the treatment condition.

Important developmental adaptations were undertaken for the assessment process (see Freeman et al. 2012). Given the young age of participants it was expected that assessments would represent their first contact with the mental health system. Thus, efforts were taken to explain terminology and make children and families feel comfortable by, for example, offering fun activities and snacks. The assessment process included psychoeducation to help parents accurately identify and label OCD symptoms. Clinician observation was also used to address similar developmental concerns.

Parent-report was relied upon heavily to address concerns regarding the ability of younger children to complete self-report forms, understand interview questions, and articulate answers. More specifically, if a child denied obsessions they could be inferred from parent report if the parent could provide a solid behavioral description(s) of the ways in which anxiety appeared to be driving their child's behavior. This process of inferring obsessions was done in a conservative manner and only with sufficient evidence from the parent report. For example, if a child engaged in frequent hand washing yet denied specific obsessions, contamination obsessions could be inferred if the parent reported that the child was anxious and/or distressed about washing and was also noted to avoid contact with or show concern related to possible contaminants or germs. Incompleteness or "not just right" obsessions could be inferred for a child engaging in frequent hand washing if the child was again noted to show anxiety and/or distress related to washing and reported needing to wash until it "felt right" or "to get a feeling." In contrast to these examples, in a child who was noted to engage in frequent washing but where the parent could not provide clear behavioral descriptors of anxiety and/other avoidance behaviors, obsessions were not inferred and were rated as absent.

Evaluators were trained to take 5- to 8-year old development into account, including consideration of OCD symptoms in the larger diagnostic picture as well as the function and course of symptoms. For example, a parent might report that a child's OCD causes him or her to disobey directions at home. In such as case, IEs established that the noncompliance was directly caused by obsessions or compulsions, as opposed to alternative diagnostic considerations such as Opposition Defiant Disorder, before including this information in their CY-BOCS ratings. Additionally, interview questions were tailored to the cognitive and developmental abilities of younger children. For example, IEs used concrete examples to facilitate accurate self-report information regarding children's symptoms and behavior. Evaluators gauged the accuracy and helpfulness of child reporting given variability in young children's capacity to reliably report symptoms. Further, IEs were attuned to younger children's shorter attention spans and took care to balance the need for further inquiry on a particular topic or symptom with the need to complete the entire interview and maximize the quality of information gathered. For example, if IEs spent too much time evaluating obsessions the child might be fatigued which could prevent the IE from gathering adequate data to evaluate compulsions. Lastly, the severity of compulsions was assessed first, since younger children were more aware of compulsions compared to obsessions and compulsions were easier for parents to observe and report on.

To ensure reliable assessment, IEs underwent substantial training including, joint interviews and videotape reviews. The IEs met an initial reliability standard of .80 (Cohen's K), reliability was re-assessed bi-monthly via random videotape reviews, and IEs were retrained if they fell below the .80 benchmark (Freeman et al. 2012).

Instruments

Children's Yale-Brown Obsessive-Compulsive Scale (CY-BOCS; Scahill et al. 1997) The CY-BOCS is a 10-item clinician-administered instrument for assessing pediatric OCD symptom severity. Clinicians rate the following five items: 1) how much time OCD symptoms take up each day, 2) how much symptoms interfere with daily life (*interference)*, 3) how much *distress* symptoms cause, 4) how much the child is able to resist the symptoms (*resistance*), and 5) how much *control* the child exhibits over symptoms of OCD. Items are rated on a scale from 0 ("None") to 4 ("Extreme"). The items are rated separately for obsessions and compulsions, yielding two, five-item subscale scores: Obsessions and Compulsions (scores on each can range from 0 to 20). These subscales combine to form a Total score (ranging from 0 to 40). Higher scores on the CY-BOCS indicate more severe OCD symptomatology. Several studies have examined the psychometric properties of the CY-BOCS and demonstrated adequate reliability and construct validity, although, as noted above, some findings suggest less favorable technical properties for assessing young children. As previously described, for POTS Jr. developmentally sensitive administration of the CY-BOCS was facilitated by several adaptations. When completing the CY-BOCS, clinicians incorporate information provided by both the child and parent(s).

In addition to the clinician-rated CY-BOCS, a parent-report version (CY-BOCS-PR) was administered. The CY-BOCS-PR is similarly formatted and includes the same items as the CY-BOCS but is filled out by parents via paper and pencil. The parents who filled out the CY-BOCS-PR were also present for the clinician-administered CY-BOCS assessment. Both the CY-BOCS and CY-BOCS-PR were collected first at the baseline IE assessment (Gate C) and then throughout the course of the study during reassessments by IEs at 5, 9, and 14 weeks.

Children's OCD Impairment Scale (COIS; Piacentini et al. 2003). The COIS is a 33-item, parent-report measure of the impact obsessive compulsive symptoms have on various activities (e.g., making new friends, doing homework, etc.). Parents rate each item on a 4-point scale from 0 ("Not at all") to 3 ("Very much") based on how much their child's OCD has interfered with the activities over the past month. The COIS was collected via parent-report at the baseline IE assessment (Gate C).

Child Behavior Checklist-Parent Report Form (CBCL; Achenbach and Rescorla 2001) The CBCL is a parent rating form measuring a variety of behavioral problems in children ages 6 to 18 years. The CBCL 1.5-5-Parent Report form

(Achenbach and Rescorla 2000) was completed for 5-year-old children in this study (n=19). The CBCL is widely used in general child clinical assessment (Chasel 2002) and has a large body of research demonstrating its favorable psychometric qualities. In addition to items assessing children's strengths and adaptive qualities (e.g., school competence), the CBCL includes 113 problem items that parents rate by endorsing either "Not True," Somewhat or Sometimes True," or "Very True or Often True" based on the preceding 6 months. The problem items combine to form various clinical syndrome scales (e.g., Somatic Complaints, Attention Problems, etc.). A standard score (T-score) at or above 70 on any syndrome scale is described as falling in the *Clinical* range, whereas scores between 65 and 69 are classified as *Borderline,* and scores below 65 as *Normal*. The CBCL data were collected via parent-report at the Gate B diagnostic assessment.

Data Analytic Plan

Several forms of reliability were appraised. The internal consistency of the CY-BOCS Total, Obsessions subscale, and Compulsions subscale scores was examined by calculating coefficient alpha for all 127 participants assessed by IEs at baseline (Gate C).[1] Previous studies of the CY-BOCS examined test-retest stability by reporting intraclass correlation coefficient. Thus, to allow for comparison with existing studies on the CY-BOCS, we examined test-retest stability by computing average one-way random effects intraclass correlation coefficients (ICCs) between baseline and 5 week assessments for subjects in the RT condition only (n=64). Temporal stability was only calculated for RT participants in order to control for the expected treatment-related variation in CY-BOCS scores for children in the CBT condition. Indeed, the CBT condition was superior to the relaxation condition in reducing OCD symptom severity as measured by the CY-BOCS (see Freeman et al. 2014). Thus, examining the RT participants only provided a better estimate of the temporal stability of CY-BOCS scores, separate from treatment effects. An exploratory analysis was conducted to examine agreement between the clinician-administered and parent-report versions of the CY-BOCS. Specifically, Pearson product moment correlation coefficients were computed to determine the magnitude of association between these two versions of the CY-BOCS at baseline and weeks 5, 9, and 14. Reliability estimates were characterized according to recently published interpretive guidelines (Haynes, Smith, and Hunsley 2011). For example, Haynes, Smith, and Hunsley (2011) consider internal consistency coefficient alpha between .70 and .79 *Adequate*, values between .80 and .89 *Good*, and values greater than or equal to .90 *Excellent*. Though reliability will vary given the purpose of assessment and population being assessed, such guidelines are offered as a framework of "good enough" criteria to evaluate scores obtained from psychological tests (Haynes, Smith, and Hunsley 2011).

Construct validity of the CY-BOCS was also appraised. Convergent validity was examined by computing Pearson product moment correlation coefficients between the CY-BOCS Total and subscale scores and the COIS. Discriminant validity was examined by computing Pearson product moment correlation coefficients between the CY-BOCS scores and CBCL syndrome scales measuring depressive symptoms, attention problems, and aggressive behavior. Given the large number of analyses conducted and resultant inflation of the family-wise Type I error rate, only correlations significant at p<.01 were interpreted.

Results

Descriptive statistics are summarized in Table 2. The mean CY-BOCS Total score at baseline for the full sample was 25.6 (SD=4.2), suggesting severe symptomatology. No statistically significant differences were found between females (M=25.8, SD=4.7) and males (M=25.3, SD=3.7) on the CY-BOCS Total score, t (125)=−.63, p=0.53. Also, the CY-BOCS Total score was not significantly correlated with child age (r=−0.05, p=0.57). The mean CY-BOCS-PR Total score at baseline (n=85) was 23.1 (SD=6.4), also suggesting severe symptomatology.

Reliability

The internal consistency of the CY-BOCS Total score at baseline assessment was adequate (α=0.73). However, the internal consistency estimates for the Obsessions (α=0.65) and Compulsions (α=0.60) subscales were poor. For the Total score, several items did not appear to contribute to the overall scale reliability as deletion of these items either improved or had no effect on the scale's alpha coefficient. Specifically, removal of the two *resistance* items increased the scale reliability and removal of the *distress from compulsions* item resulted in the same alpha coefficient. Similarly, for both the Obsessions and Compulsions subscale, deletion of the

[1] We are aware of debate in the literature regarding the use of Cronbach's alpha and the alternative use of a structural equation modeling (SEM) based approach such as Raykov's reliability estimate (see Raykov 1997). However, a primary aim of this study was to compare, as directly as possible, the psychometric properties of the CY-BOCS in our sample of younger children with existing research evidence from other samples and all prior studies reported Cronbach's alpha as a measure of internal consistency. For this reason, along with sample size considerations that limited our ability to conduct SEM based analyses, we computed coefficient alpha, which allowed for a meaningful comparison with prior studies.

Table 2 Descriptive statistics

Scale	M	SD	Range
CY-BOCS			
Total	25.55	4.23	16–37
Obsessions	12.17	2.68	0–18
Compulsions	13.39	2.30	9–19
CY-BOCS-PR			
Total	23.13	6.43	10–38
Obsessions	10.99	3.79	0–19
Compulsions	12.31	3.122	3–19
COIS			
Total	23.72	14.62	0–63
CBCL			
Anxious/Depressed	66.52	10.04	50–92
Withdrawn/Depressed	57.53	7.75	50–83
Somatic complaints	60.92	8.06	50–90
Social problems	57.54	7.60	50–85
Thought problems	68.32	7.46	51–90
Attention problems	57.61	8.43	50–96
Rule breaking behavior	54.49	6.25	50–76
Aggressive behavior	60.32	9.07	50–88

The CBCL syndrome scales are in T score units and all other measures are in raw score units. The CBCL data are from the diagnostic screening assessment (Gate B) and the CY-BOCS and COIS data are from the baseline (Gate C) assessment conducted approximately two-weeks later

Table 3 Temporal stability of CY-BOCS items from baseline to week 5

Item	ICC	p
Time spent on obsessions	0.79	<.001
Interference from obsessions	0.75	<.001
Distress from obsessions	0.65	<.001
Resistance to obsessions	0.42	.014
Control over obsessions	0.33	.060
Time spent on compulsions	0.78	<.001
Interference from compulsions	0.74	<.001
Distress from compulsions	0.65	<.001
Resistance to compulsions	0.61	<.001
Control over compulsions	0.34	.056

resistance items improved subscale reliability. Of note, internal consistency estimates improved substantially for the week 5 assessments with alpha coefficients of 0.89, 0.82, and 0.82 for the Total, Obsessions, and Compulsions scores respectively. By week 5 the *resistance* items and *distress from compulsions* appeared more reliable and if deleted, would not have resulted in improved internal consistency for the Total or subscale scores.

The CY-BOCS Total (ICC=0.68, $p<.001$), Obsessions subscale (ICC=0.69, $p<.001$), and Compulsions subscale (ICC=0.64, $p<.001$) had good test-retest stability from baseline assessment to week 5 among kids in the RT condition. We also examined temporal stability at the item level (see Table 3). For both Obsessions and Compulsions, the *time* and *interference* items demonstrated good stability and both *distress* items appeared moderately stable. The *resistance to compulsions* item was also moderately stable although the *resistance to obsessions* item exhibited poor stability. For both Obsessions and Compulsions the *control* items demonstrated poor temporal stability.

The clinician-administered and parent-report versions of the CY-BOCS were only weakly correlated at the baseline assessment (Total score $r=0.51$, Obsessions $r=0.40$, Compulsions $r=0.42$). By the week 5 assessment agreement had improved slightly but was still moderate overall (Total score $r=0.62$, Obsessions $r=0.60$, Compulsions $r=0.59$). At the week 9 assessment agreement again improved and reached adequate levels of reliability (Total score $r=0.74$, Obsessions $r=0.72$, Compulsions $r=0.70$). By week 14 agreement was good (Total score $r=0.86$, Obsessions $r=0.82$, Compulsions $r=0.86$). All correlations were statistically significant ($p<.001$).

To further examine the relation between clinician and parent-versions of the CY-BOCS, item-level correlations were computed (see Table 4). In general, the *time* and *interference* items showed the strongest association between versions, while the *resistance* and *control* items showed no appreciable correlation at baseline and were only weakly correlated at the week 5 assessment. Interestingly, by the week 14 assessment the agreement between clinician and parent rating had improved substantially and, in fact, *control over compulsions* showed the strongest association between versions.

Construct Validity

The CY-BOCS Total and subscale scores were significantly and positively correlated with the COIS (see Table 5). Of note, the CY-BOCS Obsessions subscale ($r=.27$, $p=.004$) demonstrated a lower, although still statistically significant, correlation with the COIS compared to the Total ($r=.40$, $p<.001$) and Compulsions subscale ($r=.42$, $p<.001$). These correlations represent medium to large effects according to Cohen's (1988) interpretive guidelines.

With regard to discriminant validity, the CY-BOCS demonstrated negligible, non-statistically significant correlations with the CBCL syndrome scales measuring depressive symptoms, attention problems, and aggressive behavior (see Table 5). The correlations generally represent small effect sizes (Cohen 1988).

Table 4 Item-level correlations (r) between clinician and parent-report versions of the CY-BOCS

Item	Baseline	Week 5	Week 9	Week 14
Time spent on obsessions	.63**	.57**	.67**	.72**
Interference from obsessions	.53**	.52**	.75**	.66**
Distress from obsessions	.40**	.45**	.45**	70**
Resistance to obsessions	−.01	.29**	.48**	.56**
Control over obsessions	−.02	.46**	.57**	.76**
Time spent on compulsions	.59**	.34**	.70**	.74**
Interference from compulsions	.49**	.62**	.62**	.68**
Distress from compulsions	.42**	.42**	.59**	.64**
Resistance to compulsions	.20*	.23*	.50**	.74**
Control over compulsions	.03	.30**	.62**	.82**

** $p<.01$, one-tailed

* $p<.05$, one-tailed

Discussion

This study examined the psychometric qualities of the CY-BOCS for assessing the severity of OCD in younger children, ages 5 to 8. To our knowledge this is just the second study to focus on the technical properties of the CY-BOCS with children this young. Our study represents the second largest sample to examine the reliability of the CY-BOCS and the largest sample to examine convergent and discriminant validity with any age group. Further, these data were collected in the context of a tightly controlled randomized clinical trial with careful consideration and ongoing quality control of the reliability of assessments and adherence to age-sensitive procedural adaptations.

In the current study the CY-BOCS Total score demonstrated adequate internal consistency. However, consistent with previous research, the internal consistency estimate is lower than is typically reported in studies of older children. Further, in the current study the internal consistency of the Obsessions and Compulsions subscales was poor. In prior research with younger children (Freeman et al. 2011), the internal consistency of the Obsessions subscale was essentially identical to the estimate reported here, however, the current study found lower internal consistency for the Compulsions subscale than has been previously reported. This is possibly due to the item-level reliability findings. Specifically, the *resistance* and *distress* items on the Compulsions subscale negatively affected the internal consistency estimate.

The Total and subscales scores demonstrated good temporal stability over 5 weeks. However, these stability estimates are lower than estimates reported in previous studies, including the one prior study that focused exclusively on a sample of younger children. Moreover, the current follow-up period was shorter than previous studies [Freeman et al. (2011): average follow-up=8.2 weeks; Storch et al. (2004): average follow-up=5.8 weeks]. Moreover, the present study presents the largest sample for test-retest estimates [Freeman et al. (2011): n=22; Storch et al. (2004): n=37]. Thus, one might have expected higher temporal stability estimates in the current sample. It is important to note, however, that the RT condition in the present study included an active treatment and treatment effects likely obfuscate estimates of temporal stability. In particular, the POTS Jr. trial was preceded by a pilot study (Freeman et al. 2008) and included considerable efforts to carefully develop the treatment protocols. Thus, this possibly explains the lower stability values such that the active

Table 5 Correlations between CY-BOCS, COIS, and CBCL Syndrome Scales

	1	2	3	4	5	6	7	8	9	10	11
1. CY-BOCS Total	–										
2. Obsessions Subscale	.87**	–									
3. Compulsions Subscale	.82**	.44**	–								
4. COIS	.40**	.27**	.42**	–							
5. CBCL-Anxious/Depressed	.05	.05	.03	.32**	–						
6.CBCL-Withdrawn/Depressed	.21*	.14	.22*	.36*	.54**	–					
7. CBCL-Somatic Complaints	.07	.08	.04	.20*	.49**	.35**	–				
8. CBCL-Social Problems	.32**	.24**	.30**	.29**	.52**	.58**	.29**	–			
9. CBCL-Thought Problems	.17*	.15	.14	.42**	.47**	.30**	.26**	.52**	–		
10. CBCL-Attention Problems	.11	.07	.13	.29**	.35**	.35**	.13	.60**	.51**	–	
11. CBCL- Rule Breaking Behavior	.10	.02	.16	.23*	.26**	.23*	.14	.58**	.31**	.42**	–
12. CBCL-Aggressive Behavior	.12	.06	.15*	.36**	.40**	.35**	.05	.58**	.38**	.39**	.55**

The CBCL data are from the diagnostic screening assessment (Gate B) and the CY-BOCS and COIS data are from the baseline (Gate C) assessment conducted approximately two-weeks later

** $p<.01$, one-tailed

* $p<.05$, one-tailed

treatment resulted in symptom reduction and hence, lowered temporal stability. Overall it appears likely that our results are an underestimate of the true stability of the CY-BOCS Total and subscale scores in younger children.

To our knowledge, this is the first study to report comparisons between the clinician-administered and parent-report versions of the CY-BOCS in younger children. Previous research has addressed a self-report version, filled out by adolescents with OCD (Conelea et al. 2012). In the present study, at the baseline assessment agreement between clinician and parent version Total and subscale scores was poor (correlations ranging from .40 to .50). Interestingly, agreement improved substantially throughout the study and by week 14 agreement was good (correlations in the .80s). This improved agreement might have been due to overall lowering of CY-BOCS scores during the course of treatment. It might also be the case that IEs diagnostic skills improved over the course of the study. Moreover, as noted above, several carefully considered developmental assessment adaptations were used for this study. Among them was an increased focus on parental psychoeducation regarding the nature of OCD symptomatology. Thus, increased agreement could signify the improvement in parent's ability to identify and report OCD symptoms over time. This highlights the importance of including parental psychoeducation into the assessment process, particularly with younger children who are likely less reliable reporters of their own behaviors and internal states.

In terms of construct validity, the CY-BOCS was significantly and positively correlated with another measure of OCD impairment, the COIS, providing some evidence for convergent validity. However, the association we found (r=.40) is lower than correlations reported in other studies. In previous studies of the CY-BOCS, correlations with other OCD measures have ranged from .46 to .70, including a correlation of .63 with the NIMH Global OCD Rating Scale reported in the only prior study focused on younger children (Freeman, Flessner, and Garcia 2011). Still, the associations represent medium to large effect sizes. Additionally, the Total and Compulsions subscale scores were more strongly correlated with the COIS (large effect sizes) compared to the Obsessions subscale (medium effect size). This might suggest that among younger children the Compulsions subscale provides a better estimate of OCD-related impairment than the Obsessions subscale. This supports the procedural adaptation of assessing the Compulsion subscale first when assessing OCD symptom severity in younger children.

With regard to discriminant validity, the CY-BOCS demonstrated negligible correlations with scales measuring depression, attention problems, and aggressive behavior. Here again, we found lower correlations than other studies. For example, two previous studies with older child samples (Storch et al. 2004, 2005) reported correlations of r=.53 between the CY-BOCS Total score and a scale measuring aggression. In the current study the correlation between the Total score and the CBCL Aggression scale was much lower (r=.12). Further, our correlations are lower even when compared to the prior study focused on younger children. Freeman and colleagues (2011) reported a correlation between the CY-BOCS Total score and an ADHD scale of r=.34, which is considerably higher than the correlation we found (r=.11). Of note, both associations are still quite a bit lower than those reported in samples of older children. For example, Storch et al. (2004, 2005) reported correlations of r=.47 and r=.56 between the Total score and an ADHD scale. Combined with the convergent validity results discussed above, the correlations reported in this study are generally suppressed compared to extant research. This may be due to lower test-retest reliability of the CY-BOCS in the current study. The interpretation of these findings is complicated, however, by the lack of overlap in the measures used across studies.

A particular strength of the current study is the inclusion of item-level reliability analyses. Such results have not been previously reported with younger children. Overall, the *time* and *interference* items appeared strongest, contributing to internal consistency, demonstrating good temporal stability, and exhibiting the strongest agreement between clinician and parent versions at baseline assessment. In contrast, the *resistance* items appeared problematic. Both *resistance to obsessions* and *resistance to compulsions* adversely affected internal consistency and exhibited poor agreement between clinician and parent versions. Further, *resistance to obsessions* demonstrated poor temporal stability. These findings suggest that the *time* and *interference* items are reliable indicators of OCD severity in younger children when rated by both clinicians and parents. Conversely, the *resistance* items appear less trustworthy.

These findings need to be interpreted in light of a few limitations. First, data for the discriminant validity analyses were collected at different time points. The CBCL data were collected at an earlier assessment than the CY-BOCS data. However, the time-frame between data collection points was only 2 weeks and families had not yet engaged in any treatment related activities. Additionally, only one criterion measure was available to conduct convergent validity analyses. This practice is not uncommon among prior studies of the CY-BOCS as three of the existing studies, including the seminal paper on the CY-BOCS (Scahill et al. 1997), used one convergent validity criterion measure. Also, our examination of temporal stability utilized data from an active comparator. Thus, treatment effects may have attenuated and underestimated the stability of CY-BOCS scores. Still, we used an identical methodology to Freeman et al. (2011), allowing for more direct comparison between study findings.

Given the relative lack of research evidence focusing on the assessment of OCD in children this young, additional work is warranted. Future research should examine the factor structure of the CY-BOCS in a sample of younger children. Previous

factor analytic research on the CY-BOCS has focused on older youth samples (McKay et al. 2003; Storch et al. 2005). Thus, no research evidence exists regarding the internal structure of the CY-BOCS in younger children. Findings with older samples report inconsistent findings regarding the factor structure of the CY-BOCS. McKay et al. (2003) did find support for the two standard Obsessions and Compulsions subscales and suggested a Severity factor composed of the *interference* and *distress* items and a Disturbance factor composed of the remaining items, which both displayed better fit. Similarly, Storch et al. (2005) found no support for the standard Obsessions and Compulsions subscales and reported a revised factor solution similar to McKay and colleagues (2003) except that the frequency items switched from the Disturbance factor to the Severity factor. Clarifying the factor structure of the CY-BOCS in younger children would provide important information regarding the meaningfulness and clinical utility of the Obsessions and Compulsions subscale scores.

Moreover, further developmental refinement of the CY-BOCS appears warranted. Authors have commented on the need for continued development of age sensitive OCD assessment measures (Freeman, Flessner, and Garcia 2011). The current findings support these concerns and suggest, for example, potential item-level refinement of the CY-BOCS for use with younger children. In addition, the current data were collected in the context of a randomized controlled trial with assessments conducted by IEs who were highly trained and implemented age-sensitive procedural adaptations. Future research might examine the "field reliability" of the CY-BOCS when administered by clinicians not engaged in a treatment outcome study. Lastly, a shorter measure might be developed for use in effectiveness contexts.

Further psychometric evaluation of the CY-BOCS in younger children with early onset OCD is necessary given developmental considerations and preliminary evidence of less favorable technical qualities, compared to use with older children. Reliable and accurate identification and measurement of OCD symptoms and severity in younger children is critical given evidence of effective treatments for this population, which, in many cases, can lead to clinical remission (Freeman et al. 2008, 2014). It is hoped that further examination and, where needed, developmental refinement of the CY-BOCS will promote positive outcomes for younger children and their families confronting early onset OCD.

Conflict of Interest Nathan E. Cook declares that there is no conflict of interest, Jennifer B. Freeman declares that there is no conflict of interest, Abbe M. Garcia declares that there is no conflict of interest, Jeffrey J. Sapyta declares that there is no conflict of interest, Martin E. Franklin declares that there is no conflict of interest.

Experiment Participants This study was approved by the Rhode Island Hospital institutional review committee. All participants provided informed consent prior to enrolling in the trial.

References

Achenbach, T. M., & Rescorla, L. A. (2000). *Manual for the ASEBA preschool forms & profiles*. Burlington: University of Vermont, Research Center for Children, Youth, and Families.

Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the ASEBA school-age forms & profiles*. Burlington: University of Vermont, Research Center for Children, Youth, and Families.

American Academy of Child and Adolescent Psychiatry. (2012). Practice parameter for the assessment and treatment of children and adolescents with obsessive-compulsive disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 51*, 98–113.

Chasel, M. L. (2002). Child and adolescent psychological assessment: Current clinical practices and the impact of managed care. *Professional Psychology: Research and Practice, 33*, 446–453.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. New Jersey: Lawrence Erlbaum.

Conelea, C. A., Schmidt, E. R., Leonard, R. C., Riemann, B. C., & Cahill, S. (2012). The children's Yale-brown obsessive compulsive scale: Clinician versus self-report format in adolescents in a residential treatment facility. *Journal of Obsessive-Compulsive and Related Disorders, 1*, 69–72.

Feldt, L. S., & Kim, S. (2006). Testing the difference between two alpha coefficients with small samples of subjects and raters. *Educational and Psychological Measurement, 66*, 589–600.

Freeman, J. B., Garcia, A. M., Coyne, L., Ale, C., Przeworski, A., Himle, M., Compton, S., & Leonard, H. L. (2008). Early childhood OCD: Preliminary findings from a family-based cognitive-behavioral approach. *Journal of the American Academy of Child and Adolescent Psychiatry, 47*, 593–602.

Freeman, J., Flessner, C. A., & Garcia, A. (2011). The children's Yale-brown obsessive compulsive scale: reliability and validity for use among 5 to 8 year olds with obsessive-compulsive disorder. *Journal of Abnormal Child Psychology, 39*, 877–883.

Freeman, J., Garcia, A., Benito, K., Conelea, C., Sapyta, J., Khanna, M., March, J., & Franklin, M. (2012). The pediatric obsessive compulsive disorder treatment study for young children (POTS Jr): developmental considerations in the rationale, design, and methods. *Journal of Obsessive-Compulsive and Related Disorders, 1*, 294–300.

Freeman, J., Sapyta, J., Garcia, A., Compton, S., Khanna, M., Flessner, C., & Franklin, M. (2014). Family-based treatment of early childhood obsessive-compulsive disorder: The pediatric obsessive-compulsive disorder treatment study for young children (POTS Jr) – a randomized clinical trial. *JAMA Psychiatry, 71*, 689–698.

Garcia, A. M., Freeman, J. B., Himle, M. B., Berman, N. C., Ogata, A. K., Ng, J., Choate-Summers, M. L., & Leonard, H. (2009). Phenomenology of early childhood onset obsessive compulsive disorder. *Journal of Psychopathology and Behavioral Assessment, 31*, 104–111.

Goodman, W. K., Price, L. H., Rasmussen, S. A., Mazure, C., Fleischmann, R. L., Hill, C. L., Heniger, G. R., & Charney, D. S. (1989). The Yale-brown obsessive compulsive scale I: Development, use, and reliability. *Archives of General Psychiatry, 46*, 1006–1011.

Haynes, S. N., Smith, G. T., & Hunsley, J. D. (2011). *Scientific foundations of clinical assessment*. New York: Taylor & Francis.

Hunsley, J., & Mash, E. J. (2008). *A guide to assessments that work*. New York: Oxford University Press.

Kaufman, J., Birmaher, B., Brent, D., Rao, U., Flynn, C., Moreci, P., Williamson, D., & Ryan, N. (1997). Schedule for affective disorders and schizophrenia for school-age children-present and lifetime version (K-SADS-PL): Initial reliability and validity data. *Journal of the American Academy of Child and Adolescent Psychiatry, 36*, 980–988.

Lewin, A. B., & Piacentini, J. (2010). Evidence-based assessment of child obsessive compulsive disorder: Recommendations for clinical practice and treatment research. *Child & Youth Care Forum, 39*, 73–89.

Mash, E. J., & Hunsley, J. (2005). Evidence-based assessment of child and adolescent disorders: Issues and challenges. *Journal of Clinical Child and Adolescent Psychology, 34*, 362–379.

McKay, D., Piacentini, J., Greisberg, S., Graae, F., Jaffer, M., Miller, J., Neziroglu, F., & Yaryura-Tobias, J. A. (2003). The Children's Yale-brown obsessive-compulsive scale: Item structure in an outpatient setting. *Psychological Assessment, 15*, 578–581.

Merlo, L. J., Storch, E. A., Adkins, J. W., Murphy, T. K., & Geffken, G. R. (2007). Assessment of pediatric obsessive-compulsive disorder. In E. A. Storch, G. R. Geffken, & T. K. Murphy (Eds.), *Handbook of child and adolescent obsessive-compulsive disorder* (pp. 67–108). Mahwah: Lawrence Erlbaum Associates, Inc.

Piacentini, J., Bergman, R. L., Keller, M., & McCracken, J. (2003). Functional impairment in children and adolescents with obsessive-compulsive disorder. *Journal of Child and Adolescent Psychopharmacology, 13*, 61–69.

Raykov, T. (1997). Scale reliability, Cronbach's coefficient alpha, and violations of essential tau-equivalence with fixed congeneric components. *Multivariate Behavioral Research, 32*, 329–353.

Scahill, L., Riddle, M. A., McSwiggin-Hardin, M., Ort, S. I., King, R. A., Goodman, W. K., Cicchetti, D., & Leckmna, J. F. (1997). Children's Yale-brown obsessive compulsive scale: Reliability and validity. *Journal of the American Academy of Child and Adolescent Psychiatry, 36*, 844–852.

Southam-Gerow, M. A., & Chorpita, B. F. (2007). Anxiety in children and adolescents. In E. J. Mash & R. A. Barkley (Eds.), *Assessment of childhood disorders* (pp. 347–397). New York: Guilford Press.

Storch, E. A., Murphy, T. K., Geffken, G. R., Soto, O., Sajid, M., Allen, P., Roberti, J. W., Killiany, E. M., & Goodman, W. K. (2004). Psychometric evaluation of the children's Yale-brown obsessive-compulsive scale. *Psychiatry Research, 129*, 91–98.

Storch, E. A., Murphy, T. K., Geffken, G. R., Bagner, D. M., Soto, O., Sajid, M., Allen, P., Killiany, E. M., & Goodman, W. K. (2005). Factor analytic study of the children's Yale-brown obsessive-compulsive scale. *Journal of Clinical Child and Adolescent Psychology, 34*, 312–319.

Storch, E. A., Murphy, T. K., Adkins, J. W., Lewin, A. B., Geffken, G. R., Johns, N. B., Jann, K. E., & Goodman, W. K. (2006). The children's Yale-brown obsessive-compulsive scale: Psychometric properties of child- and parent-report formats. *Anxiety Disorders, 20*, 1055–1070.

Yucelen, A. G., Rodopman-Arman, A., Topcuoglu, V., Yazgan, M. Y., & Fisek, G. (2006). Interrater reliability and clinical efficacy of children's Yale-brown obsessive-compulsive scale in an outpatient setting. *Comprehensive Psychiatry, 47*, 48–53.