

The OCI-4: An ultra-brief screening scale for obsessive-compulsive disorder

Amitai Abramovitch^{a,*}, Jonathan S. Abramowitz^b, Dean McKay^c

^a Department of Psychology, Texas State University, San Marcos, TX, USA

^b Department of Psychology and Neuroscience, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

^c Department of Psychology, Fordham University, Bronx, NY, 10458, USA

ARTICLE INFO

Keywords:

OCD
Brief screening measure
Assessment
Treatment sensitivity
Anxiety disorders

ABSTRACT

Obsessive-compulsive disorder (OCD) is a prevalent and burdensome condition that is typically assessed using in-depth interviews or lengthy self-report measures. Accordingly, routine screening in busy non-mental health settings is impractical, and OCD is often under- (or mis-) recognized. We evaluated an ultra-brief version of a widely used self-report measure, the Obsessive-Compulsive Inventory-Revised (OCI-R), which may be employed as a routine screener for OCD. A total of 1087 adults diagnosed with OCD, 1306 unselected adults from the community, and 423 adults with anxiety related disorders completed the OCI-R along with measures of anxiety and mood. Analyses were conducted to reduce the number of items and examine evidence for sensitivity and specificity to OCD clinical status, test-retest reliability, sensitivity to treatment, and convergent and discriminant validity. Four items that optimally assess different dimensions of OCD (washing, checking, ordering, obsessing) were identified. Psychometric evaluation revealed good to excellent test-retest reliability, validity, prediction of clinical OCD status, and sensitivity to treatment. A 4-item version of the OCI-R, called the OCI-4, shows promise as an ultra-brief screening tool for identifying likely OCD in settings where in-depth assessment is impractical. Patients with a positive screen may be referred for further evaluation and appropriate treatment.

1. Introduction

With a lifetime prevalence rate of 2–3 % in the general adult population, obsessive-compulsive disorder (OCD) is a highly prevalent mental health condition (Ruscio, Stein, Chiu, & Kessler, 2010). It is also associated with substantial impairment in daily functioning and work productivity (Markarian et al., 2010), as well as high health care costs (e.g., Moritz, 2008). Fortunately, effective treatments are available in the form of cognitive-behavioral therapy and some forms of medication (Grant, 2014). Yet the majority of people with OCD do not receive proper interventions primarily because of the failure to properly recognize this condition in non-specialized settings. For example, although OCD is regularly encountered in non-psychiatric medical settings such as primary care, obstetrics, and dermatology clinics (e.g., Challacombe & Wroe, 2013; Fineberg et al., 2003; Gros, Magruder, & Frueh, 2013), measures of OCD are seldom administered at such sites because of their length, lack of usefulness as a diagnostic screener, requirement of clinician administration, and cumbersome scoring schemes (Abramowitz, 2008). Further, without sound assessment instruments, OCD is commonly missed by primary care providers (Glazier,

Swing, & McGinn, 2015). Accordingly, there have been calls for screening instruments for OCD that can easily be used in medical clinics (Sussman, 2003) as the disorder lags far behind depression and anxiety related disorders when it comes to proper detection, identification, and referral in direct health service settings.

The cardinal features of OCD are persistent distressing intrusive thoughts associated with anxious apprehension (obsessions) and repeated efforts to reduce this distress using avoidance and overt or covert rituals (compulsions; American Psychiatric Association, 2013). Although the theme of each patient's symptoms is idiosyncratic, research indicates four primary (yet overlapping) domains: contamination (associated with washing compulsions), responsibility for harm and mistakes (associated with checking compulsions), unacceptable (taboo) thoughts, and order/symmetry (McKay et al., 2004). This heterogeneity complicates the clinical identification of OCD and makes it difficult to develop brief assessment instruments as has been done for other mental health problems (e.g., Kroenke, Spitzer, Williams, Monahan, & Löwe, 2007; Löwe, Kroenke, & Gräfe, 2005).

Foa, Kozak, Salkovskis, Coles, and Amir (1998) developed the 42-item Obsessive-Compulsive Inventory (OCI) to improve upon

* Corresponding author at: Department of Psychology, Texas State University, 601 University Drive, San Marcos, TX, 78666, USA.

E-mail addresses: abramovitch@txstate.edu (A. Abramovitch), jabramowitz@unc.edu (J.S. Abramowitz), mckay@fordham.edu (D. McKay).

existing measures, and subsequently revised the scale down to the 18 item Obsessive-Compulsive Inventory-Revised (OCI-R; Foa et al., 2002). With its user-friendly format and good reliability and validity, the OCI-R has enjoyed wide use in OCD specialty clinics and research. Yet the measure remains too lengthy for widespread use as a screener in busy medical settings or community clinics. Accordingly, given the movement toward ultra-brief scales (Rammstedt & Beierlein, 2014), we sought to develop a concise version of the OCI-R that could be used to rapidly identify probable cases of OCD for referral to mental health specialists for further assessment and treatment. This article describes the derivation and evaluation of such a screening instrument with the hope of addressing the under-recognition of OCD. This process first involved identifying the individual OCI-R items that best correspond to the four OCD symptom domains. Subsequently, we examined evidence that this scale (a) demonstrates reliability and validity as a measure of OCD symptoms, (b) is predictive of OCD clinical status, and (c) is sensitive to the effects of treatment. We investigated these parameters relative to the full OCI-R, other measures of OCD symptoms, and measures of anxiety and depression.

2. Methods

2.1. Participants

2.1.1. OCD group

The OCD sample included 1040 adults meeting DSM-IV/DSM-5 criteria for primary OCD, presenting for treatment at sites around the United States, including from the Anxiety and Stress Disorders Clinic outpatient OCD program at the University of North Carolina at Chapel Hill (UNC; $n = 183$), various residential, partial hospitalization, and intensive outpatient OCD treatment centers within the Rogers Behavioral Health System network ($n = 804$), and the Mayo Clinic outpatient OCD program in Rochester, MN ($n = 53$).

2.1.2. Anxiety-related disorders (ARD) group

The ARD sample included 423 individuals diagnosed with a primary DSM-IV/DSM-5 anxiety related disorder (ARDs; Asmundson, 2019). This group was comprised of 179 adults from the Rogers Behavioral Health System and 244 participants from the Anxiety and Stress Disorders Clinic at UNC. The ARD sample included primary panic disorder (18.3 %), panic disorder with agoraphobia (12.3 %), social anxiety disorder (23.3 %), trichotillomania (9.7 %), generalized anxiety disorder (21.9 %), specific phobia (5.9 %), post-traumatic stress disorder (2.5 %), and anxiety disorder not otherwise specified (15.1 %). None of these individuals met criteria for a primary, secondary or tertiary diagnosis of OCD.

2.1.3. Non-clinical community (NCC) sample

We also included a group comprised of 1106 students from the UNC site, and 88 participants from the Mayo Clinic site who completed the OCI-R as part of other clinical studies (total group $n = 1194$). To ensure that this sample was heterogeneous, but unlikely to include individuals meeting diagnostic criteria for OCD, individuals were only included in the NCC sample if they scored below the OCI-R clinical cutoff score for a likely diagnosis of OCD (>21 ; Foa et al., 2002). Detailed demographic information for each sample is presented in Table 1.

2.1.4. Additional samples

Two additional samples were utilized specifically for assessing test-retest reliability and treatment sensitivity. For the treatment sensitivity analyses we used data from 47 participants with primary OCD (64 % female, mean age = 33.82, SD = 19.88) who completed the OCI-R and were administered the Yale-Brown Obsessive-Compulsive Scale (Y-BOCS) before and after receiving manualized exposure and response prevention (ERP) treatment delivered through the outpatient clinic at the UNC site.

Table 1
Demographic and clinical characteristics.

| | OCD ($n = 1040$) | ARD ($n = 423$) | NCC ($n = 1194$) |
|------------------------------|-------------------------|-----------------------|--------------------------|
| Variable | Mean (SD); range / %(n) | | |
| Gender | | | |
| Female | 53.2 % (550) | 49.6 % (210) | 30.1 % (358) |
| Male | 46.8 % (483) | 50.4 % (213) | 69.9 % (833) |
| Age (years) | 31.8 (12.4); 18–79 | 34.5 (13.4); 18–70 | 20.1 (3.5); 18–54 |
| Education | | | |
| Did not complete high school | 9.1 % (65) | 4.1 % (15) | – |
| High school diploma | 25.5 % (182) | 17.8 % (65) | 4.5 % (4) ^a |
| Some vocational or college | 27.4 % (196) | 32.1 % (117) | 17.0 % (15) |
| Vocational or college degree | 15.5 % (111) | 23.3 % (85) | 47.7 % (42) |
| Graduate degree | 22.5 % (161) | 22.7 % (83) | 30.7 % (85) |
| Marital Status | | | |
| Married | 58.3 % (501) | 36.6 % (147) | 90.9 % (80) ^b |
| Not married | 30.7 % (264) | 53.5 % (215) | 9.1 % (8) |
| Separated/Divorced | 6.5 % (56) | 9.9 % (56) | – |
| Widowed | 4.5 % (39) | – | – |
| Ethnicity | | | |
| Asian | 2.4 % (32) | 1.2 % (5) | 5.1 % (60) |
| Black/American | 1.4 % (19) | 2.4 % (10) | 10.5 % (124) |
| Hispanic or Latino | 1.3 % (18) | 2.6 % (11) | 5.3 % (62) |
| White American | 70.8 % (948) | 92.1 % (383) | 75.5 % (889) |
| Other/not identified | 1.0 % (13) | 1.7 % (7) | 3.7 % (43) |
| OCD Severity ^a | | | |
| Y-BOCS Total Score | 25.31 (6.47) | – | – |
| Y-BOCS Obsessions | 12.87 (3.42) | – | – |
| Y-BOCS Compulsions | 12.44 (3.72) | – | – |

OCD = obsessive-compulsive disorder; ARD = anxiety related disorders, NCC = none-clinical controls; Y-BOCS = Yale-Brown obsessive-compulsive scale.

^a All OCD participants received a primary diagnosis of OCD, and those with a total score of zero on the OCI-R were not included in any analyses. Two participants had a Y-BOCS total score of zero.

^a Educational status information was only available for 12 % of the NCC sample.

^b Marital status information was only available for 7 % of the NCC sample.

To examine test-retest reliability, 212 students at Vanderbilt University completed a second administration of the OCI-R 12 weeks after the first administration. This group was 58 % female, had a mean age of 18.9 years (SD = 1.2, range = 18–30), and was ethnically diverse (71 % White American, 17 % Black American, 4 % Asian American, and 3 % Hispanic). The 12-week time interval was chosen because it allowed us to maximize retention while also minimizing practice effects and allowing sufficient time for respondents to experience personal events that might influence their responses to the OCI-R items. Participants were contacted via e-mail to complete the second administration. Across samples, the respective institutional review boards approved data collection, and all participants gave an informed consent.

2.2. Measures

2.2.1. OCI-R (Foa et al., 2002)

All study participants completed the OCI-R. This self-report measure, as described earlier, includes 18 items on which respondents rate their level of distress associated with common OCD (and hoarding) symptoms (e.g., *I check things more often than necessary*) using a 5-point Likert scale ranging from 0 (not at all) to 4 (very much). The 18 items correspond to six symptom factors (washing, checking, ordering, obsessing, neutralizing, and hoarding), each comprised of 3 items. Overall, the OCI-R demonstrates good internal consistency across populations and geographic locations (Cronbach's alpha ranged .81–.95; Hajcak, Huppert, Simons, & Foa, 2004; Hon, Siu, Cheng, Wong, & Foa, 2019), yet the hoarding and neutralizing subscales demonstrate somewhat weaker internal consistency.

2.2.2. *Yale-Brown obsessive-compulsive scale (Goodman, price, Rasmussen, mazure, Delgado et al., 1989, Goodman, Price, Rasmussen, Mazure, Fleischmann et al., 1989)*

The Y-BOCS includes 10 items that assess the following five parameters of obsessions (items 1–5) and compulsions (items 6–10): time, interference, distress, resistance, and control. Items are rated from 0 (none) to 4 (severe), yielding a total score (range = 0–40). Originally developed as a clinician administered semi-structured interview, the Y-BOCS demonstrates good to excellent internal consistency, test–retest reliability, and interrater reliability (Goodman, Price, Rasmussen, Mazure, Delgado et al., 1989; Rapp, Bergman, Piacentini, & McGuire, 2016; Storch et al., 2005). A self-report version (Y-BOCS-SR) was later developed (Baer, Brown-Beasley, Sorce, & Henriques, 1993) that performs well on various indices of reliability and validity (Federici et al., 2010; Steketee, Frost, & Bogart, 1996). Data were available from 500 participants with OCD who completed the Y-BOCS: 375 (75 %) completed the self-report version and 125 (25 %) were administered the interview version. Previous research indicates the two versions may be used interchangeably, and no significant difference in mean total scores across the two administration modalities were found in the present study (self-report = 25.44 [6.73]; interview = 24.90 [5.61]; $t(604) = 0.80, p = .42$).

2.2.3. *Dimensional obsessive-compulsive scale (DOCS; Abramowitz et al., 2010)*

The OCD group completed this 20-item self-report measure that assesses OCD symptom severity across the four most empirically supported symptom dimensions: contamination, responsibility for harm and mistakes, symmetry, and unacceptable (taboo) thoughts. Within each dimension (subscale), five items are rated on a 5-point Likert scale (0–4) to assess time occupied by obsessions and compulsions, avoidance behaviors, associated distress, functional interference, and difficulty disregarding the obsessions and refraining from the compulsions over the past month. The DOCS subscales have good to excellent reliability in both clinical and undergraduate samples ($\alpha = .83-.96$), and test–retest reliability analyses indicate adequate stability of test scores (Abramowitz et al., 2010). The measure converges well with other measures of OCD symptoms and discriminates from general measures of depression, anxiety, stress, and social anxiety in patients and students.

2.2.4. *Beck depression inventory (BDI; Beck, steer, & Brown, 1996)*

This 21-item self-report scale assesses the severity of affective, cognitive, motivational, vegetative, and psychomotor components of depression. The BDI has good psychometric properties (Beck et al., 1996). The OCD group completed the BDI in the present study.

2.2.5. *Penn state worry questionnaire (PSWQ; Meyer, Miller, Metzger, & Borkovec, 1990)*

Participants with OCD completed the PSWQ, a 16-item self-report inventory designed to capture the generality, excessiveness, and uncontrollability dimensions of pathological worry without regard to its specific content. Each item is rated on a 1 (not at all typical of me) to 5 (very typical of me) Likert type scale (e.g., “My worries overwhelm me”). The PSWQ possesses good internal consistency and test-retest reliability in clinical samples and is at least moderately correlated with other measures of trait worry (Molina & Borkovec, 1994).

2.3. *Procedure*

All clinical participants presented for services at one of the study sites and were assessed using one of the following instruments administered by a trained assessor: The Mini International Neuropsychiatric Interview, The Structured Clinical Interview for DSM-IV, and the Anxiety Disorder Interview Schedule for DSM-IV. Subsequently, interviewers reviewed these assessment data with an expert clinician (i.e., the site director or senior clinician) who then met with each patient to review

the data. Kappa (interrater agreement) was not computed at all sites, but this study only included participants for whom the expert clinician and interviewer established full diagnostic consensus that OCD (or an ARD) was the primary diagnosis. A subset of participants in the OCD sample ($n = 500$) completed the Y-BOCS-SR or were administered the Y-BOCS interview as part of their assessment. All participants gave informed consent for their responses to interviews and questionnaires to be used for research purposes.

3. *Results*

3.1. *Item selection*

Selecting OCI-R items to retain for the ultra-brief scale involved a series of conceptually and empirically driven decisions to maximize (a) consistency with the most up-to-date structural analyses of OCD symptoms (i.e., the four theme-based domains described earlier) and (b) sensitivity and specificity. Accordingly, we first eliminated the hoarding items (1, 7, and 13) since hoarding is no longer considered a symptom of OCD. Second, we computed corrected item-total correlations for the three items within each of the remaining five OCI-R subscales (see Table 2) and eliminated items with correlations < 0.70 among participants in the OCD group. This resulted in the elimination of items 2, 16, and 6.

Third, to identify the items that best captured each of the four OCD symptom domains, we computed correlations (using the OCD group) between scores on the remaining OCI-R items and the DOCS subscales (Table 3). Selecting items to retain, however, was less straightforward than expected. The two remaining OCI-R checking items, for example, were identically correlated with the DOCS ‘responsibility’ subscale (which assesses checking rituals). However, we retained item 8, since

Table 2
Corrected item-total correlations within each OCI-R subscale by participant group.

| Subscale and items | | OCD | ARD | NCC |
|---------------------|---|-----|-----|-----|
| Checking | | | | |
| 2 | I check things more often than necessary | .66 | .66 | .25 |
| 8 | I repeatedly check doors, windows, drawers, etc. | .81 | .76 | .33 |
| 14 | I repeatedly check gas and water taps and light switches after turning them off | .73 | .66 | .27 |
| Ordering | | | | |
| 3 | I get upset If objects are not arranged properly | .85 | .82 | .62 |
| 9 | I get upset if others change the way I have arranged things | .79 | .81 | .62 |
| 15 | I need things to be arranged in a particular order | .82 | .85 | .66 |
| Neutralizing | | | | |
| 4 | I feel compelled to count while I am doing things | .70 | .56 | .35 |
| 10 | I feel I have to repeat certain numbers | .77 | .63 | .40 |
| 16 | I feel that there are good and bad numbers | .60 | .41 | .28 |
| Washing | | | | |
| 5 | I find it difficult to touch an object when I know it has been touched by strangers or certain people | .74 | .57 | .41 |
| 11 | I sometimes I have to wash or clean myself simply because I feel contaminated | .86 | .65 | .47 |
| 17 | I wash my hands more often and longer than necessary | .81 | .64 | .45 |
| Obsessing | | | | |
| 6 | I find it difficult to control my own thoughts | .66 | .70 | .51 |
| 12 | I am upset by unpleasant thoughts that come into my mind against my will | .77 | .80 | .59 |
| 18 | I frequently get nasty thoughts and have difficulty getting rid of them | .70 | .72 | .55 |

OCI-R: Obsessive-Compulsive Inventory – Revised; OCD = obsessive-compulsive disorder; ARD = anxiety related disorders, NCC = none-clinical controls.

Table 3
Correlations between OCI-R items* and DOCS subscales among OCD patients (n = 808).

| OCI-R Subscale and item | DOCS subscale | | | |
|-------------------------|----------------|------------|---------------|-----------------------|
| | Responsibility | Symmetry | Contamination | Unacceptable thoughts |
| Checking | | | | |
| 8 | .39 | .27 | .15 | .11 |
| 14 | .39 | .30 | .14 | .13 |
| Ordering | | | | |
| 3 | .15 | .60 | .16 | .08 |
| 9 | .15 | .51 | .26 | .07 |
| 15 | .16 | .60 | .17 | .08 |
| Washing | | | | |
| 5 | .08 | .11 | .74 | -.06 |
| 11 | .10 | .03 | .85 | -.05 |
| 17 | .08 | .01 | .82 | -.09 |
| Obsessing | | | | |
| 12 | .33 | .06 | -.00 | .61 |
| 18 | .32 | .09 | -.07 | .61 |
| Neutralizing | | | | |
| 4 | .21 | .34 | .16 | .13 |
| 10 | .20 | .31 | .05 | .20 |

* only items with a total corrected item correlation ≥ 0.7 are presented; $r_s > 0.07$ were significant at $\alpha = .05$; OCD: Obsessive-compulsive disorder; OCI-R: Obsessive-compulsive inventory – Revised; Y-BOCS = Yale-Brown obsessive-compulsive scale; DOCS: Dimensional obsessive-compulsive scale; Items selected for the ultra-brief scale are in bold font.

this item (“I repeatedly check doors, windows, drawers, etc.”) addresses types of checking compulsions that are more prevalent in OCD (Salkovskis, Millar, Gregory, & Wahl, 2017) compared to those described in item 14 (“I repeatedly check gas and water taps and light switches”). Similarly, OCI-R ordering items 3 and 15 were comparably correlated with the DOCS ‘symmetry’ subscale. Yet because item 3 better captured the distress parameter of OCD, we chose to retain this item. The two remaining OCI-R neutralizing items were only weakly associated with each of four DOCS subscales. This was not surprising given that the neutralizing items pertain exclusively to OCD symptoms involving numbers. Moreover, the neutralizing factor appeared to be subsumed by the ordering factor, and thus redundant. Accordingly, we decided not to retain any neutralizing items. We retained item 11 from the OCI-R washing subscale as it correlated most strongly with the DOCS contamination subscale. Finally, although both OCI-R obsessing items were identically correlated with the DOCS ‘unacceptable thoughts’ subscale, we retained item 18 because the wording best captured the experience of obsessional thoughts that are subjectively resisted. Thus, this process resulted in four items from the OCI-R to be included in the ultra-brief scale: items 8 (checking), 3 (ordering), 11 (washing), and 18 (obsessing).

As a confirmatory step in the item selection process we applied item response theory (IRT) using the program jMetrik (Meyer, 2014). In order to select items that would be sensitive to a potential diagnosis of OCD, while also guarding against overidentification of cases, we chose the basic Rasch model (de Ayala, 2009; Wright & Stone, 1979), and selected the items with the highest discriminability for the highest score choice.

Table 4
Norms for the OCI-4 items and total scores across clinical and non-clinical samples.

| OCI-4 items | OCD | | | ARD | | | NCC | | |
|-------------|-------------|------|-------|-------------|------|-------|-------------|------|-------|
| | Mean (SD) | Mdn. | Range | Mean (SD) | Mdn. | Range | Mean (SD) | Mdn. | Range |
| Total Score | 5.84 (3.39) | 5 | 0–16 | 2.85 (3.18) | 2 | 0–16 | 1.81 (1.57) | 2 | 0–8 |
| Washing | 1.62 (1.58) | 1 | 0–4 | 0.38 (0.84) | 0 | 0–4 | 0.39 (0.67) | 0 | 0–4 |
| Checking | 1.20 (1.35) | 1 | 0–4 | 0.56 (1.02) | 0 | 0–4 | 0.25 (0.56) | 0 | 0–4 |
| Ordering | 1.29 (1.30) | 1 | 0–4 | 1.00 (1.17) | 1 | 0–4 | 0.95 (0.91) | 1 | 0–4 |
| Obsessing | 1.72 (1.56) | 1 | 0–4 | 0.92 (1.27) | 0 | 0–4 | 0.21 (0.52) | 0 | 0–4 |

OCD: Obsessive-compulsive disorder; OCI-4: The 4-item Obsessive-Compulsive Inventory; SD: standard deviation; Mdn=median.

That is, for each item, the discriminability was maximized for selection of scores of ‘4’ on each of the items. This analysis indeed confirmed our selection of items 3, 8, 11, and 18 for the ultra-brief scale, which we termed The 4-item Obsessive-Compulsive Inventory (OCI-4).

3.2. Norms and psychometric properties of the OCI-4

3.2.1. Norms

Means and standard deviations for the OCI-4 total score (the sum of OCI-R items 8, 3, 11, and 18), and individual item scores by group are presented in Table 4. A univariate analysis of variance (ANOVA) revealed a significant main effect for group on the total score, $F(2, 1298) = 642.60, p < .001$, Partial $\eta^2 = 0.33$. Post hoc comparisons (Tukey HSD) revealed that the OCD group had a significantly higher mean score than the ARD group, which had a significantly higher score than the NCC group (all $p_s < .001$).

For the individual items, an omnibus MANOVA revealed a significant multivariate effect, Wilk’s Lambda = 0.59, $F(8, 5302) = 82.37, p < .001$. Univariate ANOVAs conducted for each item indicated significant between-group differences (all $p_s < .001$), and results from the Tukey post hoc tests revealed that on each item, the OCD group scored significantly higher than the ARD and the NCC (all $p_s < .001$). The ARD group scored significantly higher than the NCC group on the ‘obsessions’ and ‘checking’ items ($p_s < .001$), but no significant difference was found between those groups on the ‘ordering’ item ($p = 0.722$) nor the ‘washing’ item ($p = 0.928$). Taken together, these findings provide evidence that the OCI-4 total score and the 4 individual items can discriminate individuals with OCD from unselected community controls, and the total score and obsessing and checking items can discriminate between OCD and ARDs.

3.2.2. Reliability

Although internal consistency is often considered a hallmark of reliability, we chose not to examine internal consistency for the OCI-4 for two main reasons. First, we did not assume unidimensionality of the scale - which is a central assumption of any internal consistency tests (McNeish, 2018). Indeed, correlations among the 4 items were not strong (ranging from $r = .05$ to $r = .37$), owing to the fact that these items were chosen because of their ability to assess different OCD symptom dimensions. Moreover, individuals with OCD usually present with a primary theme, so that a pattern of responses with elevated scores on one or two items and low scores on other items would be the rule rather than the exception. Second, classic coefficients for calculating internal consistency (e.g., Cronbach’s α) rely heavily on the number of items. Therefore, such coefficients calculated in scales with small number of items (e.g., <10 items) significantly underestimate reliability and are considered a contraindication (Graham, 2006). Accordingly, we evaluated reliability using test-retest analyses.

Table 5 shows the mean OCI-4 and OCI-R scores for the test-retest sample ($n = 212$ unselected students) at both time points. Paired t -tests indicated no significant changes in mean scores over the 12-week interval (both $p_s = 0.550$). Pearson correlation coefficients between time 1 and time 2 scores are also reported in Table 5. Strong correlations were detected for both measures, which fell in the range of what is

Table 5

Means, standard deviations, and test-retest coefficients for the OCI-4 and OCI-R ($n = 213$).

| Measure | Time 1 <i>M</i> (<i>SD</i>) | Time 2 <i>M</i> (<i>SD</i>) | Paired <i>t</i> | Test-retest <i>r</i> |
|---------|-------------------------------|-------------------------------|-----------------|----------------------|
| OCI-4 | 2.86 (2.93) | 2.76 (2.97) | 0.60 | .67* |
| OCI-R | 14.23 (11.22) | 14.06 (12.44) | 0.60 | .74* |

* $p < .001$; OCI-4, The 4-item Obsessive-Compulsive Inventory; OCI-R: Obsessive-Compulsive Inventory – Revised; *M*: Mean; *SD*: Standard Deviation.

typically considered evidence of adequate stability of test scores.

3.2.3. Convergent and discriminant validity

As expected, the OCI-4 total score was strongly correlated with the OCI-R total score among the OCD, ARD, and NCC groups ($r_s = 0.87, 0.92$, and 0.76 respectively; all $p_s < .001$). Table 6 shows correlations between scores on the OCI-4 and scores on other measures of OCD and non-OCD symptoms among participants in the OCD group. As can be seen, whereas the OCI-4 was strongly associated with the DOCS total score, it was weakly correlated with measures of general anxiety, depression, and worry. Moreover, these correlations were comparable to those with the OCI-R total score. Notably, weaker correlations were found between the OCI-4 (and OCI-R) and the Y-BOCS. This, however, was not surprising considering that the Y-BOCS assumes an idiographic approach to the assessment of OCD, whereas the OCI-R/OCI-4 takes a nomothetic approach. In concert, this provides evidence for good convergent and discriminant validity for the OCI-4 as a measure of OCD symptoms. Moreover, among people diagnosed with OCD, the OCI-4 possesses convergent and discriminant validity very comparable to that of the OCI-R total score.

3.2.4. Diagnostic sensitivity

We conducted receiver operating characteristic (ROC) analyses, that uses the association between sensitivity and specificity to estimate the area under the curve (AUC) to indicate how well scores on the measure distinguish between the OCD and NCC groups and between the OCD and ARD groups. An AUC of 1.0 indicates perfect prediction, whereas a value of 0.50 indicates the level of chance. In distinguishing individuals with OCD from NCC participants, the AUC estimate for the OCI-4 was 0.86 (95 % CI = .84–.88). In distinguishing individuals with OCD from ARD participants, the AUC estimate for the OCI-4 was 0.76 (95 % CI = .73–.79). These results indicate that the OCI-4 distinguishes individuals with OCD from non-treatment-seeking individuals and those with ARDs quite well. Moreover, the OCI-4 compares extremely well with the full scale OCI-R: AUC = .88 (95 % CI = .74–.80) for OCD vs. NC participants, and AUC = 0.77 (95 % CI = .73–.79) for OCD vs. ARD participants. Fig. 1 graphically displays the AUC estimates for the two measures.

3.2.5. Diagnostically relevant cutoff scores

We next determined the OCI-4 total score with the optimal degree of

Table 6

Correlations between the OCI-4, OCI-R, and symptom measures among patients with OCD.

| Measure | <i>n</i> | OCI-4 | OCI-R |
|--------------------|----------|-------|-------|
| OCD symptoms | | | |
| Y-BOCS total score | 500 | .31 | .30 |
| DOCS total score | 808 | .64 | .68 |
| Other symptoms | | | |
| BAI | 76 | .27 | .26 |
| BDI | 266 | .28 | .29 |
| PSWQ | 485 | .26 | .29 |

OCD: Obsessive-compulsive disorder; OCI-4, The 4-item Obsessive-Compulsive Inventory; OCI-R: Obsessive-Compulsive Inventory – Revised; Y-BOCS = Yale-Brown Obsessive-Compulsive Scale; DOCS: Dimensional Obsessive-Compulsive Scale; BDI: Beck Depression Inventory; BAI: Beck Anxiety Inventory; PSWQ: Penn State Worry Questionnaire.

accuracy for correctly classifying individuals diagnosed with OCD relative to nonclinical individuals. The accuracy of each OCI-4 score was evaluated by calculating its sensitivity and specificity. This analysis revealed that a score of 4 provided the best balance of sensitivity and specificity (Youden’s $J_{MAX} = 0.6$) correctly classifying 80 % of the entire sample of OCD and NCC participants. This cutoff correctly classified 75 % of individuals with OCD (i.e., sensitivity = .75) and 84 % of those in the NCC group (specificity = .84). In classifying those with OCD relative to those with ARDs, a similar set of analyses revealed that a score of 4 also provided the best balance of sensitivity and specificity, correctly classifying 74 % of the members of these groups: 75 % of individuals with OCD (i.e., sensitivity = .75) and 71 % of those in the ARD group (specificity = .71). Since positive and negative predictive values are heavily influenced by the base rate of the condition in the sample, we calculated positive and negative likelihood ratios (PLR, NLR). This computation yielded a PLR = 4.55 (CI, 3.99–5.19), and NLR = 0.3 (CI, 0.27–0.33). Furthermore, these analyses yielded a Diagnostic Odds Ratio = 15.15, that translates to the odds of the OCI-4 screening positive in participants with OCD compared to the odds of the tool screening positive in individuals without OCD.

3.2.6. Sensitivity to treatment

Table 7 shows the treatment sensitivity sample’s mean pre- and post-treatment scores on the OCI-4, OCI-R, and Y-BOCS. Whereas the pre-treatment mean Y-BOCS total score fell within the severe range of symptoms, at post-treatment the mean score was within the mild range, indicating that on average, significant improvement occurred for the individuals who received ERP. Repeated measures *t*-tests indicated that all pre- to post-treatment contrasts were significant at the $p < .001$ level. To compare the magnitude of treatment effects as assessed by the different outcome measures, pre- to post-treatment changes were converted to Cohen’s *d* (Cohen, 1988), a standardized measure of effect size (also shown in Table 7). As expected, the effect size derived from the Y-BOCS (which assesses the most prominent OCD symptoms) was the largest of the group of measures. Effect sizes for the OCI-4 ($d = 1.57$), and the OCI-R ($d = 1.89$), were large, providing evidence that the OCI-4 is sensitive to the effects of empirically supported treatment for OCD. Accordingly, the OCI-4 reflects responsiveness to empirically supported intervention for OCD and thus shows potential as a treatment outcome measure.

4. Discussion

The aim of this study was to derive an ultra-brief version of the OCI-R that can be used to efficiently and effectively identify individuals who may be suffering with OCD and track progress with treatment. The resulting measure, the OCI-4, contains one item assessing each of the four robust theme-based dimensions of obsessive-compulsive symptoms: Washing, checking, intrusive thoughts/obsessions, and ordering. Evidence for reliability was obtained through test-retest analyses, which revealed stability over a 12-week period. Evidence for criterion and construct validity was demonstrated by the finding that the OCI-4 differentiated between groups of participants with and without OCD, correlated strongly with another self-report measure of OCD, and weakly with measures of general anxiety, depression, and worry. Moreover, neither reliability nor validity were substantially sacrificed by reducing the OCI-R from 18 to 4 items.

Of note, the OCI-4 was weakly correlated with the Y-BOCS, the gold standard measure of OCD. Yet this was not the result of reducing the length of the OCI-R, but rather, an artifact of the considerable method variance between the OCI-R/OCI-4 and the Y-BOCS. Whereas the format of the OCI-R/OCI-4 is nomothetic and involves a single distress rating for quintessential (representative) OCD symptoms, the Y-BOCS is an idiographic measure that assesses multiple severity parameters of the respondent’s most prominent obsessions and compulsions as identified using an exhaustive checklist of over 50 specific symptoms. Thus, like

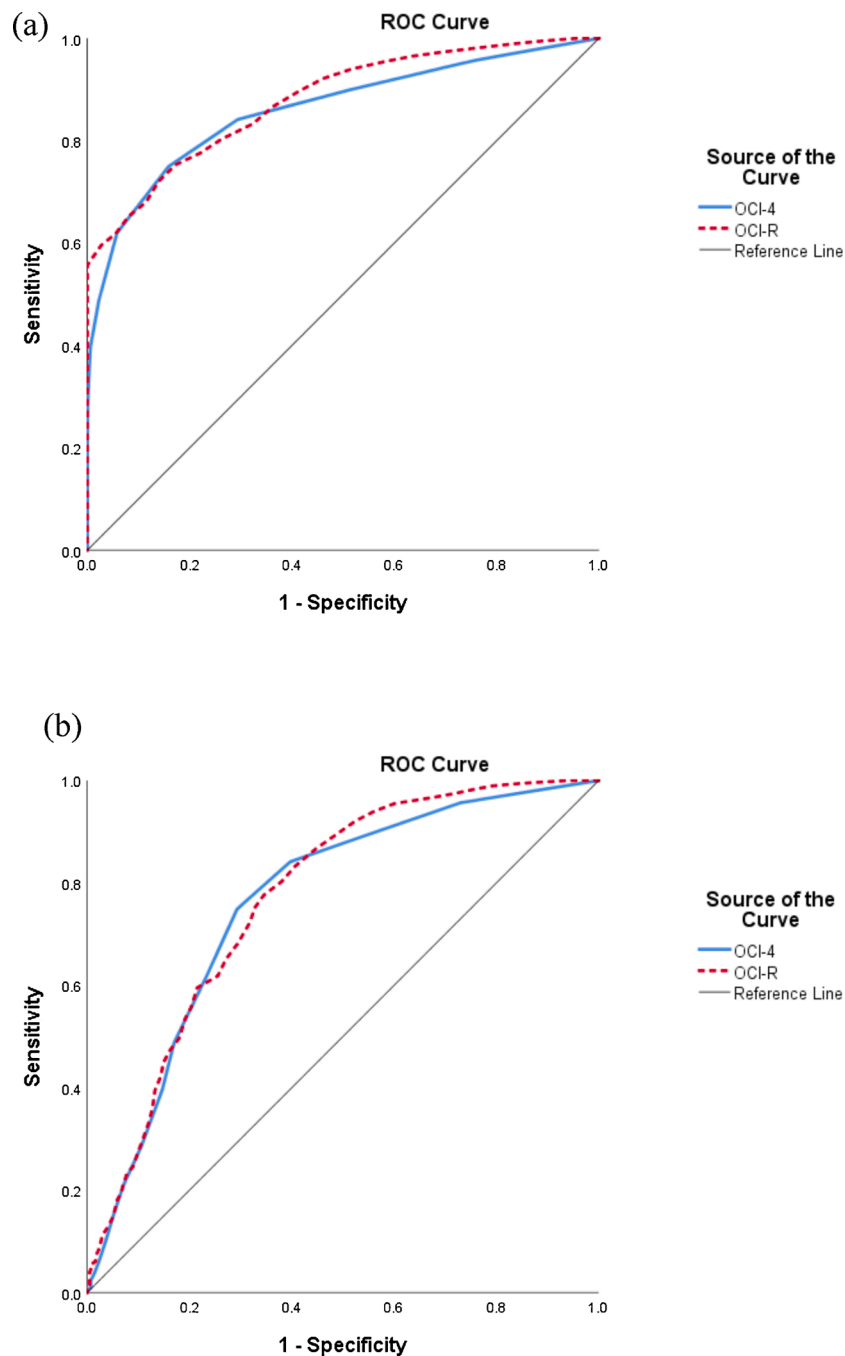


Fig. 1. Receiver operating characteristics (ROC) curves for the (a) OCD-NCC sample and (b) OCD-ARD sample. OCD: obsessive-compulsive disorder; NCC: nonclinical group; ARD: anxiety disorders.

the OCI-R, the OCI-4, to some degree, is likely to confound severity with the types of obsessions and compulsions present. We attempted to limit this problem by including the OCI-R items assessing each OCD symptom domain most broadly on the OCI-4.

Our sensitivity and specificity analyses provide evidence that the OCI-4 is a useful concise screening instrument for busy clinicians and non-psychiatric or non-specialized primary care settings interested in identifying likely cases of OCD. With a high degree of accuracy, the measure differentiates individuals with OCD from those with other ARDs and from unscreened individuals. Indeed, this cut off yielded a PLR, NLR, and Diagnostic OR estimators that are equivalent to those reported for contemporary ultra brief measures including the GAD-7, GAD-2, and PHQ-2 (Arroll et al., 2010, Plummer, Manea, Trepel, &

McMillan, 2016). Indeed the OCI-4's cutoff score yielded a diagnostic OR = 15.15, which translates to a fifteen time greater likelihood of screening positive among individuals with OCD compared to the control groups. The identification of OCD in non-psychiatric health service settings is a public health priority (discussed in Stein et al., 2016). This is due to the social and occupational disability associated with the disorder (Markarian et al., 2010; Ruscio et al., 2010), and the established difficulties in identifying OCD in primary care settings (Glazier et al., 2015; Sussman, 2003). Nevertheless, the OCI-4 is only a screener, and elevated scores are not diagnostic in and of themselves, but rather an indicator of increased likelihood for the presence of OCD to refer for further assessment to establish the presence of a clinical disorder that warrants treatment.

Table 7

Pre- and post-treatment mean scores for 47 patients with obsessive-compulsive disorder treated with exposure and response prevention.

| Measure | Pre-treatment <i>M</i> (<i>SD</i>) | Post-treatment <i>M</i> (<i>SD</i>) | <i>t</i> | Cohen's <i>d</i> |
|---------|--------------------------------------|---------------------------------------|----------|------------------|
| Y-BOCS | 25.66 (4.80) | 10.17 (5.03) | 17.75* | 2.59 |
| OCI-4 | 6.80 (2.90) | 2.72 (2.19) | 10.67* | 1.57 |
| OCI-R | 28.06 (11.70) | 12.15 (9.69) | 12.96* | 1.89 |

Note. Y-BOCS = Yale-Brown Obsessive-Compulsive Scale; OCI-4 = The 4-item Obsessive-Compulsive Inventory ; OCI -R= Obsessive-Compulsive Inventory-Revised.

* $p < .001$. Degrees of freedom for each analysis = 46.

We also found evidence that the OCI-4, like the OCI-R, is sensitive to the effects of treatment for OCD. Thus, the OCI-4 serves the additional purpose of contributing to the assessment and outcomes evaluation mandate that has emerged from third party payers (e.g., Centers for Medicare and Medicaid Services, CMS). To illustrate, CMS has recently required providers to provide assessment for diagnostic purposes, as well as evaluation for the process of treatment and outcome at termination using approved instruments (discussed in Wright et al., 2020). The list of approved measures presently does not include any OCD instruments, despite the aforementioned prevalence and disability associated with the disorder. The availability of an ultra-brief self-report measure that can be easily administered on repeated occasions can quickly fill this gap and increase the likelihood individuals suffering from OCD will be identified and offered appropriate care.

The large multi-sample data set used in this study has both strengths and limitations. The sample was clinically and geographically diverse, increasing the generalizability of our findings; yet the clinical samples were primarily White Americans. Indeed, a critical focus of further study is the assessment of OCD symptoms in non-White American individuals. Items for the OCI-4 were also embedded within the larger pool of items forming the OCI-R, and so it remains to be seen whether the results of the present study generalize to situations in which the OCI-4 is

Appendix A

OCI-4

The following statements refer to experiences that many people have in their everyday lives. Circle the number that best describes **HOW MUCH** that experience has **DISTRESSED or BOTHERED you during the PAST MONTH**. The numbers refer to the following verbal labels:

| | 0 | 1 | 2 | 3 | 4 |
|--|------------|----------|------------|-------|-----------|
| | Not at all | A little | Moderately | A lot | Extremely |
| 1 I get upset if objects are not arranged properly. | 0 | 1 | 2 | 3 | 4 |
| 2 I repeatedly check doors, windows, drawers, etc. | 0 | 1 | 2 | 3 | 4 |
| 3 I sometimes have to wash or clean myself simply because I feel contaminated. | 0 | 1 | 2 | 3 | 4 |
| 4 I frequently get nasty thoughts and have difficulty in getting rid of them. | 0 | 1 | 2 | 3 | 4 |

The 4-Item Obsessive-Compulsive Inventory (OCI-4)

Administration & Scoring

The OCI-4 is an ultra-brief screening scale for OCD, derived from the OCI-R (Foa et al., 2002). It consists of 4 items that a person endorses on a 5-point Likert scale. These items correspond to four of the original OCI-R OCD dimensions:

- 1 Ordering
- 2 Checking
- 3 Washing (contamination)
- 4 Obsessing

The total score is generated by adding the item scores. The possible range of scores is 0–16. The mean total score for persons with OCD is 5.84 ($SD = 3.39$). Recommended total cutoff score is 4, with scores at or above this level indicating the likely presence of OCD.

administered as a stand-alone instrument in non-psychiatric (e.g., primary care) settings. It is, however, noteworthy that Foa and colleagues developed the OCI-R from items embedded within the original OCI (Foa et al., 2002). Additional prospective research on the OCI-4 is warranted. For example, a test-retest study with a clinical sample would provide more powerful evidence of the scale's stability. Further examination of how the OCI-4 and its items correlate with clinical and cognitive phenomena related to OCD will help determine the scale's potential for use in clinical research on OCD. Further, given its utility in a wide range of direct care settings, the OCI-4 will permit additional research on groups that are under-represented in research on OCD, such as members of racial and ethnic minority groups (Williams, Powers, Yun, & Foa, 2010). Finally, assessment of treatment sensitivity was conducted using a rather small sample ($n = 47$).

4.1. Conclusion

The OCI-4 is a 4-item ultra-brief version of the OCI-R that assesses the major dimensions of obsessive-compulsive symptoms and demonstrates good to excellent predictive validity, sensitivity, and specificity, including good discriminability between OCD and ARDs. As such, it can be used to rapidly screen for likely OCD. The measure is also sensitive to the effects of treatment and thus useful for ongoing symptom severity assessment. The complete scale including instructions (derived from the original OCI-R; Foa et al., 2002) can be found in Appendix A.

Declaration of Competing Interest

None.

Acknowledgments

The authors wish to thank Bunmi Olatunji and Bradley Riemann, for their contribution to this study.

References

- Abramowitz, J. S. (2008). Obsessive-compulsive disorder. In J. Hunsley, & E. Mash (Eds.), *A guide to assessment that work* (pp. 275–292). Oxford University Press.
- Abramowitz, J. S., Deacon, B. J., Olatunji, B. O., Wheaton, M. G., Berman, N. C., Losardo, D., et al. (2010). Assessment of obsessive-compulsive symptom dimensions: Development and evaluation of the dimensional obsessive-compulsive scale. *Psychological Assessment*, 22(1), 180–198.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5*. American Psychiatric Association.
- Arroll, B., Goodyear-Smith, F., Crengle, S., Gunn, J., Kerse, N., Fishman, T., et al. (2010). Validation of PHQ-2 and PHQ-9 to screen for major depression in the primary care population. *Annals of Family Medicine*, 8(4), 348–353.
- Asmundson, G. J. G. (2019). A rose by any other name...How should we refer to the collective of conditions characterized by clinically significant anxiety? *Journal of Anxiety Disorders*, 68, Article 102143.
- Baer, L., Brown-Beasley, M. W., Sorce, J., & Henriques, A. I. (1993). Computer-assisted telephone administration of a structured interview for obsessive-compulsive disorder. *The American Journal of Psychiatry*, 150(11), 1737–1738.
- Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Beck depression inventory - Second edition manual*. The Psychological Corporation.
- Challacombe, F. L., & Wroe, A. L. (2013). A hidden problem: Consequences of the misdiagnosis of perinatal obsessive-compulsive disorder. *The British Journal of General Practice: the Journal of the Royal College of General Practitioners*, 63(610), 275–276.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Routledge Academic.
- de Ayala, R. J. (2009). *The theory and practice of item response theory*. Guilford.
- Federici, A., Summerfeldt, L. J., Harrington, J. L., McCabe, R. E., Purdon, C. L., Rowa, K., et al. (2010). Consistency between self-report and clinician-administered versions of the yale-brown obsessive-compulsive scale. *Journal of Anxiety Disorders*, 24(7), 729–733.
- Fineberg, N. A., O'Doherty, C., Rajagopal, S., Reddy, K., Banks, A., & Gale, T. M. (2003). How common is obsessive-compulsive disorder in a dermatology outpatient clinic? *The Journal of Clinical Psychiatry*, 64(2), 152–155.
- Foa, E. B., Huppert, J. D., Leiberg, S., Langner, R., Kichic, R., Hajcak, G., et al. (2002). The Obsessive-compulsive inventory: Development and validation of a short version. *Psychological Assessment*, 14(4), 485–496.
- Foa, E. B., Kozak, M. J., Salkovskis, P. M., Coles, M. E., & Amir, N. (1998). The validation of a new obsessive-compulsive disorder scale: The obsessive-compulsive inventory. *Psychological Assessment*, 10(3), 206–214.
- Glazier, K., Swing, M., & McGinn, L. K. (2015). Half of obsessive-compulsive disorder cases misdiagnosed: Vignette-based survey of primary care physicians. *The Journal of Clinical Psychiatry*, 76(6), e761–767.
- Goodman, W. K., Price, L. H., Rasmussen, S. A., Mazure, C., Delgado, P., Heninger, G. R., et al. (1989). The Yale-Brown obsessive compulsive scale. II. Validity. *Archives of General Psychiatry*, 46(11), 1012–1016.
- Goodman, W. K., Price, L. H., Rasmussen, S. A., Mazure, C., Fleischmann, R. L., Hill, C. L., et al. (1989). The Yale-Brown obsessive compulsive scale. I. Development, use, and reliability. *Archives of General Psychiatry*, 46(11), 1006–1011.
- Graham, J. M. (2006). Congeneric and (Essentially) tau-equivalent estimates of score reliability: What they are and how to use them. *Educational and Psychological Measurement*, 66(6), 930–944.
- Grant, J. E. (2014). Obsessive-Compulsive disorder. *The New England Journal of Medicine*, 371(7), 646–653.
- Gros, D. F., Magruder, K. M., & Frueh, B. C. (2013). Obsessive compulsive disorder in veterans in primary care: Prevalence and impairment. *General Hospital Psychiatry*, 35(1), 71–73.
- Hajcak, G., Huppert, J. D., Simons, R. F., & Foa, E. B. (2004). Psychometric properties of the OCI-R in a college sample. *Behaviour Research and Therapy*, 42(1), 115–123.
- Hon, S. K. H., Siu, B. W. M., Cheng, C. W., Wong, W. C. W., & Foa, E. B. (2019). Validation of the Chinese version of obsessive-compulsive inventory-revised. *East Asian Archives of Psychiatry*, 29(4), 103–111.
- Kroenke, K., Spitzer, R. L., Williams, J. B., Monahan, P. O., & Löwe, B. (2007). Anxiety disorders in primary care: Prevalence, impairment, comorbidity, and detection. *Annals of Internal Medicine*, 146(5), 317–325.
- Löwe, B., Kroenke, K., & Gräfe, K. (2005). Detecting and monitoring depression with a two-item questionnaire (PHQ-2). *Journal of Psychosomatic Research*, 58(2), 163–171.
- Markarian, Y., Larson, M. J., Aldea, M. A., Baldwin, S. A., Good, D., Berkeljon, A., et al. (2010). Multiple pathways to functional impairment in obsessive-compulsive disorder. *Clinical Psychology Review*, 30(1), 78–88.
- McKay, D., Abramowitz, J. S., Calamari, J. E., Kyrios, M., Radomsky, A., Sookman, D., et al. (2004). A critical evaluation of obsessive-compulsive disorder subtypes: Symptoms versus mechanisms. *Clinical Psychology Review*, 24(3), 283–313.
- McNeish, D. (2018). Thanks coefficient alpha, we'll take it from here. *Psychological Methods*, 23(3), 412–433.
- Meyer, J. P. (2014). *Applied measurement with jMetrik*. Routledge.
- Meyer, T. J., Miller, M. L., Metzger, R. L., & Borkovec, T. D. (1990). Development and validation of the Penn state worry questionnaire. *Behaviour Research and Therapy*, 28(6), 487–495.
- Molina, S., & Borkovec, T. D. (1994). The Penn State worry questionnaire: Psychometric properties and associated characteristics. *Worrying: Perspectives on Theory, Assessment and Treatment*, 265–283.
- Moritz, S. (2008). A review on quality of life and depression in obsessive-compulsive disorder. *CNS Spectrums*, 13(9 Suppl 14), 16–22.
- Plummer, F., Manea, L., Trepel, D., & McMillan, D. (2016). Screening for anxiety disorders with the GAD-7 and GAD-2: A systematic review and diagnostic metaanalysis. *General Hospital Psychiatry*, 39, 24–31.
- Rammstedt, B., & Beierlein, C. (2014). Can't we make it any shorter? *Journal of Individual Differences*, 35(4), 212–220.
- Rapp, A. M., Bergman, R. L., Piacentini, J., & McGuire, J. F. (2016). Evidence-based assessment of obsessive-compulsive disorder. *Journal of Central Nervous System Disease*, 8, 13–29.
- Ruscio, A. M., Stein, D. J., Chiu, W. T., & Kessler, R. C. (2010). The epidemiology of obsessive-compulsive disorder in the National Comorbidity Survey Replication. *Molecular Psychiatry*, 15(1), 53–63.
- Salkovskis, P. M., Millar, J., Gregory, J. D., & Wahl, K. (2017). The termination of checking and the role of just right feelings: A study of obsessional checkers compared with anxious and non-clinical controls. *Behavioural and Cognitive Psychotherapy*, 45(2), 139–155.
- Stein, D. J., Kogan, C. S., Atmaca, M., Fineberg, N. A., Fontenelle, L. F., Grant, J. E., et al. (2016). The classification of obsessive-compulsive and related disorders in the ICD-11. *Journal of Affective Disorders*, 190, 663–674.
- Steketee, G., Frost, R., & Bogart, K. (1996). The Yale-Brown obsessive compulsive scale: Interview versus self-report. *Behaviour Research and Therapy*, 34(8), 675–684.
- Storch, E. A., Shapira, N. A., Dimoulas, E., Geffken, G. R., Murphy, T. K., & Goodman, W. K. (2005). Yale-Brown obsessive compulsive scale: The dimensional structure revisited. *Depression and Anxiety*, 22(1), 28–35.
- Sussman, N. (2003). Obsessive-compulsive disorder: A commonly missed diagnosis in primary care. *Primary Psychiatry*, 10(12), 14.
- Williams, M., Powers, M., Yun, Y. G., & Foa, E. (2010). Minority participation in randomized controlled trials for obsessive-compulsive disorder. *Journal of Anxiety Disorders*, 24(2), 171–177.
- Wright, B. D., & Stone, M. H. (1979). *Best test design*. MESA Press.
- Wright, C. V., Goodheart, C., Bard, D., Bobbitt, B. L., Butt, Z., Lysell, K., et al. (2020). Promoting measurement-based care and quality measure development: The APA mental and behavioral health registry initiative. *Psychological Services*, 17, 262–270.