Published in final edited form as:

J Anxiety Disord. 2010 October ; 24(7): 723–728. doi:10.1016/j.janxdis.2010.05.004.

Factor Structure and Measurement Invariance of the Yale-Brown Obsessive Compulsive Scale Across Four Racial/Ethnic Groups

Sarah L. Garnaat, M.A. and Peter J. Norton, Ph.D.

Department of Psychology, University of Houston

Abstract

The Yale-Brown Obsessive-Compulsive Scale (Y-BOCS) is the most commonly used instrument for assessing OCD in clinical trials, but little information is available regarding its appropriateness with patients of diverse racial and ethnic backgrounds. We examined the factor structure and measurement invariance of this widely used measure, across individuals from four racial and ethnic groups utilizing both university and outpatient samples. Results indicated that a two-factor (Obsessions and Compulsions) model fit the undergraduate and outpatient samples and was structurally invariant across racial/ethnic groups. Minimal evidence of non-invariance was observed across racial/ethnic groups, with the exception that items on the Obsessions subscale may, at lower levels, under-estimate obsessive concerns among Black individuals. Overall, the Y-BOCS appears to show invariance across people of different racial and ethnic backgrounds, although caution should be taken in comparing scores obtained from Black patients to current norms given evidence of substantial non-invariance on the Obsessions scale.

Keywords

Yale-Brown Obsessive Compulsive Scale; factor analysis; measurement invariance; psychometric properties

Obsessive Compulsive Disorder (OCD) is an anxiety disorder marked by the recurrence of intrusive thoughts and repeated engagement in behaviors in an attempt to neutralize these thoughts and reduce anxiety. In assessing the severity of obsessive and compulsive behaviors that are the hallmark of OCD, the Yale-Brown Obsessive Compulsive Scale (Y-BOCS) is generally accepted to be the “gold standard” (Antony, Orsillo, & Roemer, 2001). Indeed, in meta-analyses examining the efficacy of pharmacological and behavioral interventions for OCD (Abramowitz, 1997; Norton & Price, 2007), the Y-BOCS was the most frequently utilized primary or secondary measure of treatment outcome. Despite the widespread acceptance and use of the Y-BOCS, however, there are still significant gaps in the psychometric literature establishing the validity of this measure. Two such weaknesses will be addressed herein: establishment of a stable factor structure and cross-cultural validation.

Correspondence concerning this paper should be addressed to Peter J. Norton, Ph.D., Department of Psychology, 126 Heyne Bldg., University of Houston, Houston, TX, 77204-5022, USA, Phone: 713-743-8675, FAX: 713-743-8633, pnorton@uh.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Although many examinations of the factor structure of the Y-BOCS have been undertaken, no structure has been consistently identified across these studies. The scale, as created by Goodman and colleagues (1989a; 1989b) is based on a two-factor structure comprised of an Obsessions factor and a Compulsions factor. This structure was also found by McKay and colleagues (1995) in a clinical sample. Two-factor solutions have also been found in other studies, though the two factors did not mimic the previously found Obsessions and Compulsions factors. Amir and colleagues (1997) identified a factor solution in which the two factors were best described as Disturbance and Symptom Severity. Deacon and Abramowitz (2005) also found a different two factor solution comprised of a Severity factor and a Resistance and Control factor.

This inconsistency in findings across studies may be due in part to limitations of the methodologies employed in these investigations. Some previous investigations (Deacon & Abramowitz, 2005; Fals-Stewart, 1992; Moritz, et al., 2002) have used data reduction techniques, such as Principal Components Analysis (PCA). While PCA is commonly used in the literature, the assumptions made by this technique are generally not suited for use with psychological questionnaire data, such as that of the Y-BOCS (Widaman, 1993). One central reason for this is that PCA ignores measurement error, which is typically considerable in data from self-report psychological assessments (Floyd & Widaman, 1995; Widaman, 1993; Widaman, 2007). Principal Components Analysis also typically makes use of orthogonal rotations which force factors to be uncorrelated. When constructs are believed to be related, such as Obsessions and Compulsions, such an approach is not ideal. Additionally, in PCA, items are regarded as continuous. That is, the ordinal items comprising the scale are treated as though they have interval properties, when this is not the case (e.g. the difference in symptom severity between a response of '0' and '1' is not necessarily the same as between '3' and '4'.) For example, on item 3 (*Distress associated with obsessive thoughts*) it is unclear if the distance between *None* (0) and *Not too disturbing* (1) is identical or similar to the distance between *Very disturbing* (3) and *Near constant and disabling distress* (4). This problem also exists in the application of conventional factor analysis to ordinal data, since this approach also treats variables as continuous. Using continuous data approaches with ordinal data has been repeatedly shown to result in distortion of factor analytic results (Bernstein & Teng, 1989; Jöreskog & Moustaki, 2001). As this method has also been applied in factor analyses of the Y-BOCS (Deacon & Abramowitz, 2005; Fals-Stewart, 1992; McKay, Danyko, Neziroglu, & Yaryura-Tobias, 1995), this has likely also contributed to difficulties in firmly establishing a factor structure. A more appropriate approach to use to investigate the factor structure of the Y-BOCS would be confirmatory factor analysis (CFA) for ordered-categorical variables. This approach treats item responses as ordinal and uses information from the patterns of responses to reconstruct the continuous distributions assumed to be underlying each variable, rather than treating Likert-type items as continuous as is done in the previously discussed approaches. Another limitation in the psychometric validation of the Y-BOCS is in evaluating the appropriateness of this measure for use in cross-cultural comparisons. Measurement invariance and differential item functioning (DIF) studies are used to evaluate whether items function in the same way and are on the same metric across different groups (e.g. race/ethnicity, sex, etc.). If an assessment tool is found to lack measurement invariance across groups, or is found to exhibit DIF across groups, then comparisons between groups may not accurately reflect real group differences and may lead to erroneous conclusions. Similarly, if an individual's score is evaluated against norms from a group where such comparisons are unfounded, this can lead to over-pathologizing or under-identification or -estimation of psychopathology.

Although the presence of OCD has been widely identified across cultural and racial/ethnic groups (e.g. Fontenelle & Hasler, 2008; Fontenelle, Mendlowicz, & Veriani, 2006), there have been few efforts to validate severity measures, such as the Y-BOCS, for use in cross-

cultural contexts. However, what research has been done in this area suggests some of these measures may not function similarly across all cultural and racial/ethnic groups. Measures of OCD symptom severity which have been evaluated for measurement invariance, or differential item functioning, across racial/ethnic groups include the Padua Inventory and the Maudsley Obsessive-Compulsive Inventory (MOCI). Williams and colleagues (2005) investigated DIF of the Padua Inventory across Black and White participants. DIF of the Padua Inventory was identified, resulting in overestimation of scores for Black participants relative to White. A study by Thomas and colleagues (2000) investigating DIF in the MOCI found similar differences in psychometric functioning of this measure between Black and White participants on subscales measuring severity of Cleaning and Checking behaviors. It was found that Black participants were more likely endorse items at a lower level of severity relative to White participants, which could lead to overestimation of the severity of these behaviors. The findings of these studies highlight this importance of validation of OCD measures across racial/ethnic groups, rather than simply assuming no differences exist.

Although these findings suggest non-invariance across groups when assessed using MOCI and Padua Inventory, distinct differences exist between these measures and the Y-BOCS. Most notably, the MOCI and Padua Inventory identify specific beliefs, obsessions, or compulsions (e.g., “I avoid using public telephones because of possible contamination”) whereas the Y-BOCS assesses dimensions (e.g., “Time occupied by obsessive thoughts”) of the overall obsession and compulsions regardless of specific content. As noted by Washington, Norton, and Temple (2008) “measures designed to assess these symptoms may artificially under- or over-pathologize individuals from diverse backgrounds. This may be particularly so with measures of obsessive-compulsive disorder (OCD), as practices such as washing or cleaning, and rule-based behaviors, are likely to be heavily influenced by cultural norms” (pp. 456).

To our knowledge, no such studies have been conducted to evaluate measurement invariance in the Y-BOCS across racial/ethnic groups. As the Y-BOCS is a well-known and frequently used measure of symptom severity of OCD we sought to address the previously noted limitations in the psychometric literature of this measure. First, we undertook evaluation of the factor structure of the Y-BOCS in both community and clinical samples using confirmatory factor analysis for ordered categorical data. Second, we evaluated measurement invariance of the Y-BOCS across four racial/ethnic groups in a community sample. As previous research does not offer a strong reasoning for hypothesizing one factor structure over another, or the presence or absence of measurement invariance across racial/ethnic groups for this measure, we chose to abstain from hypothesizing about the outcome of these analyses.

Method

Sample

Community sample—The community sample was initially comprised of 1036 undergraduates at a large state university who received extra-credit for their participation. Those who self-identified their race/ethnicity as Asian ($n = 216$), Black ($n = 172$), White ($n = 260$), or Hispanic ($n = 183$) and completed the Y-BOCS were selected for inclusion in the present study as these were the only groups with sufficient size to be included in the current analyses. Options for race/ethnicity were presented to participants in multiple-choice format, with options defined by second author based on US national census classifications. Racial and ethnic background data were collected as the parent project was conducted to investigate cross-cultural models of anxiety. Demographic information for the resulting sample is included in Table 1 ($n=831$).

Clinical sample—The clinical sample was comprised of 131 outpatients participating in anxiety disorder treatment trials at the University of Houston Anxiety Disorder Clinic. All participants in the clinical sample met diagnostic criteria for an anxiety or mood disorder, as established by the Anxiety Disorders Interview Schedule for DSM-IV structured interview. Of the sample, 5.3% met criteria for primary or comorbid diagnoses of OCD, and a further 24.4% of the sample reported current obsessive or compulsive symptoms that did not meet criteria for an OCD diagnosis. The Y-BOCS was given as part of a larger battery of measures prior to entry in a treatment condition. Only those who self-identified as one of the racial/ethnic groups in the community sample (i.e. White, Black, Asian, or Hispanic) were included. Demographic information for the clinical sample is presented in Table 1.

Measure

The Y-BOCS used in this study is a self-report measure designed to assess severity of obsessive and compulsive behaviors (Goodman, et al., 1989a; Goodman, et al., 1989b). This scale includes 10 Likert-type items each with five response categories, varying from 0 (no symptoms) to 4 (very severe symptoms). The Y-BOCS has been found to have excellent psychometric properties, including reliability (Goodman, et al., 1989a; Steketee, Frost, & Bogart, 1996) and validity (Goodman, et al., 1989b; Steketee, Frost, & Bogart, 1996) in clinical and non-clinical samples. For the self-report Y-BOCS used in this study, Steketee and colleagues (1999) reported internal consistency (Chronbach's alpha) of .89 and .78 for the total scores of the non-clinical and clinical samples, respectively. A test-retest correlation for the non-clinical sample was reported as .88 and correlations between the self-report and interview versions of the Y-BOCS were .75 for the clinical sample and .79 for the non-clinical sample. This scale is generally regarded as the "gold-standard" assessment of severity of obsessive and compulsive behaviors (Antony, Orsillo, & Roemer, 2001).

Analyses

All analyses were conducted using the Mplus 5.2 software program (Muthén & Muthén, 2007) and items were treated as ordered-categorical.

Confirmatory factor analysis—In order to determine the factor structure of the Y-BOCS, a confirmatory factor analysis was undertaken in the community sample. Each of the factor structures found by McKay and colleagues (1995), Amir and colleagues (1997), and Deacon and Abramowitz (2005), was fit, and the best fitting factor structure was retained. Previously reported factor solutions including a factor determined by less than three items were not included (Kim, Dysken, Pheley, & Hoover, 1994; Moritz, et al., 2002) as these factors are locally unidentified (i.e. two items do not provide sufficient degrees of freedom to independently determine a latent factor.) Additionally, a single-factor structure was fit with all 10 items, as a unifactorial structure has previously been reported in an EFA with the 16-item version of the Y-BOCS (Fals-Stewart, 1992). For each model, latent factors were allowed to covary freely and item residuals were treated as independent unless otherwise specified. After the best-fitting factor structure was determined in the community sample, confirmation of this structure was attempted in the clinical sample. Model fit was evaluated using established recommendations. Recommendations for root mean square error of approximation (RMSEA) suggest that values close to .06 represent good fit (Hu & Bentler, 1999), values less than .08 are indicative of acceptable fit, and values between .08 and .10 represent poor model fit (Brown & Cudeck, 1993). Comparative Fit Index (CFI) is recommended to be > .95 (Hu & Bentler, 1999), and weighted root mean square residual (WRMR) < .90 (Yu, 2002).

Measurement invariance—After an appropriate factor structure was determined, we sought to assess measurement invariance across the four racial/ethnic groups (Asian, Black,

White, Hispanic) to determine whether this measure assessed severity in a comparable way across groups. If this measure was found to be invariant across groups, this would indicate that comparisons of scores across groups would be meaningful. However, if there is non-invariance across groups, such comparisons (e.g. comparing mean differences) would not be appropriate or meaningful.

Measurement invariance in scales with ordinal items has typically been investigated by examining invariance of multiple model parameters, including factor loadings, measurement intercepts, residual variance, and item thresholds, (Millsap & Yun-Tein, 2004). Factor loadings are a measure of the relation between an item and the latent factor. Measurement intercepts represent the intercept for each item, or the value the item would take if the underlying trait was zero and residual variance is the amount of item specific variability not related to the underlying latent factor. Though these three parameters are common in traditional factor analysis, item thresholds are unique to ordinal factor analysis as they represent the boundaries between the response categories of a given item. Testing invariance of threshold evaluates whether these boundaries fall in the same location on the underlying scale across groups (i.e. whether a given response category covers the same levels of severity in one group as in the other).

In the present study all parameters of interest (item thresholds, factor loadings, measurement intercepts, and residual variances) were simultaneously fixed to be equivalent across all four groups. Good fit of this model was interpreted to mean that measurement invariance across the four groups was likely. Poor fit of this model was considered to be indicative of non-invariance within the model. In this case, further pair-wise analyses were conducted, testing measurement invariance between two groups at a time, to determine the “location” of any non-invariance (i.e. the parameter(s) and group(s) where non-invariance is present).

Pairwise comparisons were conducted to further elucidate non-invariance across groups. Significant non-invariance was identified by significant loss of model fit as measured by the Chi-Square Difference of Fit statistic which is included in the DIFFTEST function of Mplus 5.2 (Muthén & Muthén, 2007). Each pair of analyses followed the sequence of steps outlined below. Within each pairwise comparison the constraints imposed in each step were carried through to subsequent steps. First, a baseline model was fit which allowed all parameters to vary freely across groups. Second, invariance of thresholds was tested by constraining the threshold parameters to be equivalent across groups, one item at a time, while allowing all other parameters to vary freely. Third, after any non-invariance of thresholds was identified, the factor structure was imposed and all thresholds found to be invariant were constrained to be equal across groups. The fit of this model served as the basis against which all subsequent models would be evaluated. Fourth, all factor loadings were constrained to be equal across groups and any non-invariant factor loadings were freed in subsequent models. Fifth, measurement intercepts were constrained to be equal across groups and any non-invariant measurement intercepts were freed in subsequent models. Finally, residual variance invariance was tested by constraining residual variances to be equal across groups. After all instances of parameter non-invariance were identified, invariance of factor means, variances, and covariances was investigated to determine whether these values differed between groups.

Results

Confirmatory Factor Analysis

In the community sample, one single-factor model and three two-factor models were fit. These models are represented pictorially in Figure 1 with fit indices in Table 2. Overall fit of all four models tested was poor with most or all fit indices falling outside the generally

accepted ranges for adequate fit for each model. Further investigation of these models indicated that factor loadings for items four and nine were uniformly lower than those of other items across all four models. These items measure resistance to Obsessions and Compulsions, respectively. Considering this pattern of factor loadings and that these items are often regarded as “problematic” (e.g. Kim, et al., 1994; Deacon & Abramowitz, 2005) the decision was made to re-evaluate fit of each of the four models, allowing for residual correlation between items four and nine. The fit of these models is reported in Table 3. The best fitting model was clearly the model reported by McKay and colleagues (1995) with nearly all fit indices falling in an acceptable range. Additionally, a multi-group model was run to assure that this factor structure was appropriate for each of the four racial/ethnic groups and this model fit acceptably (CFI=.99, RMSEA=.07, WRMR=1.44). Finally, the model was run in the clinical sample where this model was also found to exhibit adequate fit (CFI=.99, RMSEA=.09, WRMR=.59).

Measurement Invariance

Due to low rates of endorsement in the highest response categories in the community sample, the two highest response categories were collapsed for all groups, resulting in a scale with four response categories: ‘0’, ‘1’, ‘2’, and ‘3 or 4’. The initial model was fit across all four groups, simultaneously, constraining item thresholds, factor loadings, measurement intercepts, and residual variances. Means and variances were allowed to vary freely for latent factors, as was the covariance between the two latent factors. This model fit the data poorly indicating non-invariance in the model. However, to determine for which groups and parameters this non-invariance occurs, pairwise analyses were undertaken, the fit statistics of which are presented in Table 4. As these analyses require a designation of reference group against which other groups are compared, the White group was chosen as the reference group being that nearly all previous research on the Y-BOCS has been conducted with predominantly White samples (e.g. Goodman, et al., 1989a; Goodman, et al., 1989b).

Asian/White model—After the baseline model was fit in the Asian/White model, thresholds were constrained to be equal across groups, one item at a time. The constraining of thresholds did not result in significant loss of fit, with the exception of the constraint of thresholds of item 10. The third threshold for item 10 was determined to be non-invariant and this parameter was freed in this and subsequent models, resulting in a model whose fit was not significantly different from the baseline model. Next, in subsequent steps factor loadings, measurement intercepts, and residual variances were constrained across groups. None of these constraints resulted in significant loss of model fit, indicating invariance of these parameters. Overall, only one parameter was found to be non-invariant (i.e. the third threshold for item 10) and all others were found to be invariant.

Black/White model—Constraining of thresholds across groups in the Black/White model resulted in significant loss of model fit from the baseline model for many items. For item two, significant loss of model fit occurred when thresholds were constrained, and the first threshold was determined to be non-invariant. This parameter was freed in all remaining models. Constraint of thresholds for items three and four also resulted in significant loss of fit, though for each of these items multiple thresholds were determined to be non-invariant. As a minimum of two thresholds is required to determine the metric of the item and this minimum could not be reached for these items, all thresholds were freed for items three and four. Next, all factor loadings were constrained across groups. Imposition of this constraint did not result in significant loss of model fit, indicating invariance of factor loadings. Measurement intercepts were then constrained across groups and the intercept of item two was determined to be non-invariant. This parameter was subsequently freed. Finally,

residual variances were constrained to be equal across groups resulting in no significant loss of model fit and indicating invariance of residual variances across groups. Overall, substantial non-invariance was found to be present for parameters of items two, three, and four – all items on the Obsessions factor – including multiple thresholds and one intercept. Invariance testing of factor means and variances indicated that these two groups do not differ significantly in average level of Obsessiveness or Compulsiveness or the variability of these factors. However, the relation between the two factors was found to differ between groups, with a stronger relation between these construct in the Black group than in the White group.

Hispanic/White model—After establishing the baseline model thresholds were constrained across groups. All thresholds were determined to be invariant, excepting the first threshold for item 5. This parameter was freed in the present and subsequent models. Factor loadings were constrained to be equal across groups and no significant loss of fit occurred, indicating factor loading invariance. Measurement intercepts were also constrained across groups and these, also, were determined to be invariant, as no significant loss of fit occurred. Finally, residual variances were constrained across groups. The residual variance for item nine was determined to be non-invariant and this parameter was freed. Overall, only one threshold (i.e. first threshold of item five) and one residual variance (i.e. item 9) were found to be non-invariant. All other parameters for this model were found to be invariant across groups.

Comparison of factor means, variances, and covariances

Invariance testing of factor means and covariances in the Asian/White model did not result in significant loss of model fit, indicating there were not differences between the groups in average level of Obsessiveness or Compulsiveness and that the relation between these two factors was the same across groups. However, the variance of the Compulsion factor was found to differ between groups with the variance of the Asian group being less than that of the White group. In the Black/White model, invariance testing of factor means and variances indicated that these two groups do not differ significantly in average level of Obsessiveness or Compulsiveness or the variability of these factors. However, the relation between the two factors was found to differ between groups, with a stronger relation between these constructs in the Black group than in the White group. Like in the Black/White model, no differences between groups were found for factor means or variances for the Hispanic/White model. Again, however, the relation between the two factors was found to differ between groups, with a stronger relation between these constructs in the Hispanic group than in the White group.

Discussion

The present study sought to establish a factor structure for the Y-BOCS using CFA and investigated measurement invariance across four racial/ethnic groups using this factor structure. Factor analyses determined that the original two-factor model proposed by Goodman and colleagues (1989a; 1989b) and empirically supported by McKay et al. (1995) fit the data best. Analyses of measurement invariance were undertaken and minimal non-invariance was detected for Asian and Hispanic groups, relative to the White group, indicating that the Y-BOCS can be appropriately used to make cross-cultural comparisons between these groups. However, substantial non-invariance was found for the Obsessions subscale of the Y-BOCS for the Black group, relative to the White group, suggesting that comparisons between these groups on this subscale are likely not appropriate and may lead to results which are misleading. In particular, examination of thresholds across these two groups indicates that the Y-BOCS may underestimate the interference, distress, and attempts

at resistance due to obsessions among black participants at low to average levels of obsessiveness. How this invariance may affect an individual's total score depends on the individual's level of the underlying trait (see Figure 2). For example, at a level of obsessiveness half a standard deviation above the mean, those in the White group would tend to endorse a score of '2' on question four, measuring interference of obsessions, whereas those in the Black group would tend to endorse a score of '1'. Care should therefore be taken in making clinical decisions with Black patients using the Y-BOCS until more appropriate norms can be established or until such time as measurement invariance can be established in a clinical sample.

The finding of non-invariance of multiple parameters in the Black/White Model is somewhat consistent with the outcome of investigations measurement invariance or DIF in other measures of obsessive and compulsive behaviors. Differential Item Functioning has been found between Black and White groups for both the Padua Inventory (Williams, Turkheimer, Schmidt, & Oltmanns, 2005) and the MOCI (Thomas, Turkheimer, & Oltmanns, 2000). However, in the present study, measurement non-invariance was found only with regard to items measuring Obsessions and was found to underestimate severity of Obsessions at lower levels of the trait, whereas in studies by Williams and colleagues (2005) and Thomas and colleagues (2000) non-invariance was found in compulsions, particularly cleaning and grooming behaviors, and resulted in overestimation of these behaviors. It is unclear why DIF has been identified for compulsions in other measures, whereas non-invariance was found for obsessions for the Y-BOCS and in the opposite direction. One possibility is that this difference is due to the measures themselves. The MOCI and the Padua Inventory inquire about the occurrence of and disturbance caused by specific behaviors (e.g., compulsive washing, beliefs about unlucky numbers), respectively. Thus, these measures are likely to pick up on both cultural differences in socially acceptable behaviors as well as problematic thoughts and behaviors. The Y-BOCS, on the other hand, inquires about the distress and impairment caused by Obsessions and Compulsions in general, rather than about specific behaviors, which may explain the differential findings of non-invariance for this study. Additionally, differences in findings may be due to differences in the samples used (e.g. internet vs. university samples or differences in regional subcultures).

While our findings regarding factor structure were consistent with both the original design of the measure by Goodman and colleagues, as well as the findings by McKay and colleagues, there has been wide variation of factor solutions reported in the literature. The fact that all other previously found factor solutions fit the current data quite poorly may be due to the problems with and variation in methods utilized in previous investigations (e.g. PCA and treating ordinal data as continuous); however given the large number of previously reported factor solutions along with the initial difficulty establishing a factor structure in the current study, further efforts should be made to confirm this structure in other studies. If this cannot be confirmed, this may indicate problems with the construct validity of this measure.

The present study contains a number of other limitations which should be addressed. The first limitation is the relatively low incidence and severity of obsessive and compulsive behaviors due to the use of a community sample. As the clinical sample did not provide sufficient sample sizes, investigation of measurement invariance could not be undertaken in this sample; however future research should seek to confirm the findings of the current study using a clinical sample. Additionally, due to the low incidence of reported OC behaviors in this community sample the top two response categories for each item had to be combined in analyses of measurement invariance. The invariance of the threshold between response categories 3 and 4 should be investigated in future studies. Another limitation of the present study is that, while different racial/ethnic groups were treated as relatively homogeneous, we

did not include assessments for acculturation, ethnic identity, nationality, or other possible dimensions related to culture or racial/ethnic background. This may result in clumping heterogeneous people into a single category. Additionally, the use of a university sample may limit the capacity of this sample to reflect the larger racial/ethnic groups they represent.

Limitations aside, the results of this study highlight the importance of assessing measurement invariance or DIF in OCD-specific measures as well as other measures of anxiety. Studies such as this are important in order to validate the use of these measures in cross-cultural comparisons and we strongly encourage the undertaking of further investigations in this area. In addition to measurement invariance and DIF studies, future research should also seek to develop appropriate norms for groups where cross-cultural comparisons are found to be unsuitable.

Acknowledgments

The project described was supported by Grant Number K01MH073920 from the National Institute of Mental Health awarded to PJN. The content is solely the responsibility of the author and does not necessarily represent the official views of the National Institute of Mental Health or the National Institutes of Health.

References

- Abramowitz JS. Effectiveness of psychological and pharmacological treatment for obsessive-compulsive disorder: A quantitative review. *Journal of Consulting and Clinical Psychology* 1997;65:44–52. [PubMed: 9103733]
- Antony, MM.; Orsillo, SM.; Roemer, L. *Practitioner's Guide to empirically based measures of anxiety*. New York, NY: Kluwer Academic/Plenum Publishing; 2001.
- Amir N, Foa EB, Coles ME. Factor structure of the Yale-Brown Obsessive Compulsive Scale. *Psychological Assessment* 1997;9:312–316.
- Bernstein IH, Teng G. Factoring items and factoring scales are different: Spurious evidence for multidimensionality due to item categorization. *Psychological Bulletin* 1989;105:467–477.
- Brown, MW.; Cudeck, R. Alternative ways of assessing model fit. In: Bollen, KA.; Long, JS., editors. *Testing Structural Equation Models*. Beverly Hills, CA: Sage; 1993. p. 136–162.
- Deacon BJ, Abramowitz JS. The Yale-Brown Obsessive Compulsive Scale: factor analysis, construct validity, and suggestions for refinement. *Journal of Anxiety Disorders* 2005;19:573–585. [PubMed: 15749574]
- Fals-Stewart W. A dimensional analysis of the Yale-Brown Obsessive Compulsive Scale. *Psychological Reports* 1992;70:239–240. [PubMed: 1565726]
- Floyd FJ, Widaman KF. Factor Analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment* 1995;7:286–299.
- Fontenelle LF, Hasler G. The analytical epidemiology of obsessive-compulsive disorder: risk factors and correlates. *Progress in Neuro-Psychopharmacology and Biological Psychiatry* 2008;32:1–15. [PubMed: 17689849]
- Fontenelle LF, Mendlowicz MV, Veriani M. The descriptive epidemiology of obsessive-compulsive disorder. *Progress in Neuro-Psychopharmacology and Biological Psychiatry* 2006;30:327–337. [PubMed: 16412548]
- Goodman WK, Price LH, Rasmussen SA, Mazure C, Fleischmann RL, Hill CL, Heninger GR, Charney DS. The Yale-Brown Obsessive Compulsive Scale: I. Development, use, and reliability. *Archives of General Psychiatry* 1989a;46:1006–1011. [PubMed: 2684084]
- Goodman WK, Price LH, Rasmussen SA, Mazure C, Delgado P, Heninger GR, Charney DS. The Yale-Brown Obsessive Compulsive Scale: II. Validity. *Archives of General Psychiatry* 1989b; 46:1012–1016. [PubMed: 2510699]
- Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling* 1999;6:1–55.

- Jöreskog KG, Moustaki I. Factor analysis of ordinal variables: A comparison of three approaches. *Multivariate Behavioral Research* 2001;36:347–387.
- Kim SW, Dysken MW, Pheley AM, Hoover KM. The Yale-Brown Obsessive Compulsive Scale: measures of internal consistency. *Psychiatry Research* 1994;51:203–211. [PubMed: 8022954]
- McKay D, Danyko S, Neziroglu F, Yaryura-Tobias JA. Factor structure of the Yale-Brown Obsessive Compulsive Scale: a two-dimensional measure. *Behaviour Research and Therapy* 1995;33:865–869. [PubMed: 7677726]
- Millsap R, Yun-Tein J. Assessing factorial invariance in ordered-categorical measures. *Multivariate Behavioral Research* 2004;39:479–515.
- Moritz S, Meier B, Kloss M, Jacobsen D, Wein C, Fricke S, Hand I. Dimensional structure of the Yale-Brown Obsessive Compulsive Scale (Y-BOCS). *Psychiatry Research* 2002;109:193–199. [PubMed: 11927144]
- Muthén, LK.; Muthén, BO. *Mplus Users' Guide*. Los Angeles, CA: Muthén & Muthén; 1997.
- Norton PJ, Price E. A meta-analytic review of adult cognitive-behavioral treatment outcome across the anxiety disorders. *Journal of Nervous and Mental Disorders* 2007;196:521–531.
- Steketee G, Frost R, Bogart K. The Yale-Brown Obsessive Compulsive Scale: Interview versus self-report. *Behaviour Research and Therapy* 1996;34:675–684. [PubMed: 8870295]
- Thomas J, Turkheimer E, Oltmanns TF. Psychometric analysis of racial differences on the Maudsley Obsessional Compulsive Inventory. *Assessment* 2000;7:247–258. [PubMed: 11037392]
- Washington CS, Norton PJ, Temple S. Obsessive-Compulsive symptoms and Obsessive-Compulsive Disorder. *Journal of Nervous and Mental Disorders* 2008;196:456–461.
- Widaman KF. Common factor analysis versus principal component analysis: Differential bias in representing model parameters? *Multivariate Behavioral Research* 1993;28:263–311.
- Widaman, KF. Common factors versus components: Principals and principles, errors and misconceptions. In: Cudeck, R.; MacCallum, RC., editors. *Factor analysis at 100: Historical developments and future directions*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers; 2007. p. 177-203.
- Williams MT, Turkheimer E, Schmidt KM, Oltmanns TF. Ethnic identification biases responses to the Padua Inventory for Obsessive-Compulsive Disorder. *Assessment* 2005;12:174–185. [PubMed: 15914719]
- Yu, CY. Doctoral dissertation. University of California, Los Angeles; Los Angeles, CA: 2002. Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes.

Figure 1.^a. CFA Models

^aFals-Stewart (1992) model originally conducted using a 16-item version of the Y-BOCS. Variances and residual variances are not explicitly represented in the figure and all items were treated as ordered-categorical.

Figure 2.

^a. Threshold plot for Black/White model Obsession items

^aItem thresholds are presented on a standardized scale. Values along the top represent level of the underlying trait for the Obsessions factor.

Table 1

Demographic makeup of community and clinical samples

Demographic Variable	Community Sample (<i>n</i> =831)	Clinical Sample (<i>n</i> =131)
Age (M ± SD)	22.1 ± 4.91	33.4 ± 11.76
Sex (% Female)	76.4	64.6
Race/Ethnicity		
Asian (%)	26.0	6.1
Black (%)	20.7	9.9
White (%)	31.3	66.4
Hispanic (%)	22.0	17.6

Table 2

Fit statistics for CFA models.

Model	χ^2	df	CFI	RMSEA	WRMR
Fals-Stewart	497.98	17	.93	.19	2.51
McKay	352.55	24	.96	.13	1.67
Amir ^a	501.89	17	.93	.19	2.48
Deacon	581.15	17	.93	.20	2.52

^aCorrelation between latent factors fixed to .99 in order to achieve convergence, as initial model estimates of the correlation between latent factors exceeded $r=1.0$.

Table 3

Fit statistics for CFA models with correlated residuals

Model	χ^2	df	CFI	RMSEA	WRMR
Fals-Stewart	406.32	16	.95	.17	2.25
McKay	136.95	24	.99	.08	.99
Amir ^a	411.60	16	.95	.17	2.27
Deacon	419.06	16	.95	.17	2.24

^aCorrelation between latent factors fixed to .99 in order to achieve convergence, as initial model estimates of the correlation between latent factors exceeded $r=1.0$.

Table 4

Chi-Square difference of fit statistic for measurement invariance models^a

Model	Asian/White Model		Black/White Model		Hispanic/White Model	
	χ^2	<i>p</i>	χ^2	<i>p</i>	χ^2	<i>p</i>
Factor Loadings	11.76	.11	7.90	.16	6.73	.46
Measurement Intercept	16.24	.09	14.52	.07	11.77	.30
Residual Variance	21.88	.08	20.45	.06	18.26	.20

^a Chi-Square difference of fit statistics for models of threshold invariance are not included in the table, due to the large number of these models. All fit statistics listed in the above table are for models after any non-invariant parameters were freed.