

PSYCHOMETRIC ANALYSIS OF RACIAL DIFFERENCES ON THE MAUDSLEY OBSESSIONAL COMPULSIVE INVENTORY

Joneis Thomas
Eric Turkheimer
Thomas F. Oltmanns
University of Virginia

Black university students scored significantly higher than White students on the Maudsley Obsessional Compulsive Inventory (MOCI). They tended to endorse more Cleaning and Checking subscale items in the pathological direction. Subsequent analyses examined whether this finding is a reflection of valid group differences in the prevalence of OCD or a psychometric artifact. Structured interviews were conducted to determine the correspondence of MOCI scores with OCD diagnoses. The race difference in endorsement frequency on the MOCI did not extend to OCD diagnoses. The MOCI scores showed modest predictive validity in Whites, but they did not predict interview-based diagnoses in Blacks. Multivariate item response theory was then employed to examine race differences in the Cleaning and Checking subscales. Equivalent item discrimination parameters fit the data for Black and White participants for both subscales. A more restrictive model in which relative item difficulties were also constrained to be equal for Black and White participants did not fit. This interaction between race and item difficulty suggests that the items do not have equivalent psychometric properties in Blacks and Whites.

Keywords: Obsessive-Compulsive Disorder, Maudsley Obsessional Compulsive Inventory, racial differences, test bias, Item Response Theory, assessment, anxiety

Obsessive-Compulsive Disorder (OCD) may be the most disabling form of anxiety disorder. In fact, the Global Burden of Disease Study found that, in a comparison among 107 different forms of disease and injury, OCD is the 10th leading cause of disability worldwide (Lopez & Murray, 1998). It is clearly a global problem and occurs with considerable frequency in both developed and developing countries (Al-Issa & Oudji, 1998). Unfortunately, relatively little is known about the

influence of culture and ethnicity on the expression and course of symptoms in OCD. Most evidence regarding this disorder and its treatment has been collected in developed countries from White participants. Additional data from specific ethnic and minority groups could provide useful information regarding the generalizability of findings from previous studies.

OCD has not been studied extensively in Black people (Bernstein, 1997; Neal-Barnett & Smith, 1997). Only three studies have examined racial differences in the prevalence of OCD, and their results are inconsistent. The most extensive data regarding this issue come from the Epidemiologic Catchment Area (ECA) project, which reported

Correspondence concerning this article and requests for offprints should be addressed to Thomas Oltmanns, Department of Psychology, Gilmer Hall, P.O. Box 400400, University of Virginia, Charlottesville, VA 22904-4400. E-mail: tfo@virginia.edu

that the lifetime prevalence of OCD among Black men and women was significantly lower than that among White men and women (Karno, Golding, Sorenson, & Burnam, 1988). Two smaller studies found differences in the opposite direction. Fabrega, Mezzich, and Ulrich (1988) found that Black adults scored significantly higher than White adults on symptom ratings of obsessions and compulsions. A more recent study of Black and White adolescents also reported a significantly higher risk for OCD among Blacks (Vallen-Basile et al., 1996). Several methodological factors may account for the discrepancies among these studies. Principal among these may be the investigators' choice of assessment instruments and the fact that existing measures of obsessive compulsive disorder have not been validated for use with Black populations.

Efforts to assess clinical symptoms in ethnic minorities are complicated by reliance upon instruments that have not been normed on ethnic minority samples (Butcher, Nezami, & Exner, 1998; McKenry, Everett, Ramseur, & Carter, 1989). In the case of obsessive-compulsive disorder, one of the most widely used self-report measures is the Maudsley Obsessional Compulsive Inventory (MOCI; Hodgson & Rachman, 1977). The MOCI is a 30-item true/false questionnaire that has been used extensively in clinical trials because it is sensitive to treatment effects (Frost, Steketee, Krause, & Trepanier, 1995; Steketee, 1994). It has also been used with college samples to identify nonclinical OCD participants for research purposes (Sternberger & Burns, 1990b). The original psychometric analyses of the MOCI were performed on a sample of 100 British obsessional patients (Rachman & Hodgson, 1980). Sternberger and Burns (1990a) provided the first norms for the MOCI on a nonclinical American sample, in which 87% of the participants were White. Relatively little information is available regarding the validity of the MOCI in different ethnic groups, particularly Blacks.¹

¹Chan (1990) reported a psychometric analysis of the MOCI based on responses from 183 English-speaking undergraduate students in Hong Kong, but his paper does not consider predictive validity in terms of formal diagnoses, and it does not examine measurement invariance.

The present article describes our efforts to examine the implications of a serendipitous finding involving racial differences on the MOCI. In 1991, we began using the MOCI to screen large numbers of Black and White undergraduate students. Our goal was to identify people who might participate in a laboratory study of obsessive-compulsive behavioral problems (Oltmanns & Gibbs, 1995). We noticed that, in comparison to White students, Black students more frequently produced rather high scores on the MOCI. In order to explore the implications of this unexpected result, we employed psychometric procedures to determine whether the apparent racial difference was the result of measurement artifacts or valid group differences in the prevalence of OCD. First, data gathered from structured interviews were examined to determine the correspondence of MOCI scores with OCD diagnoses. Second, a multidimensional item response theory (IRT) analysis was conducted to examine whether the MOCI has equivalent psychometric properties across races.

Psychometric Considerations

An adequate explanation of group differences requires that scores on questionnaires are numerically comparable across groups. The term "measurement invariance" will be used to refer to this condition. Measurement invariance exists when a score in one group has the same meaning, in terms of the underlying construct or latent trait, as the same score in another group. Measurement invariance is important for between-group comparisons for several reasons. If scores across groups are not comparable, then mean differences between groups may be artifactual and misleading (Horn & McArdle, 1992; Reise, Widaman, & Pugh, 1993) and the use of such scores may spuriously make one group appear more pathological than another.

Measurement invariance can be ascertained with item response theory (IRT), which specifies a mathematical model of the relationship between observable test performance and the unobserved (or latent) traits assumed to underlie it (Hamletton & Swaminathan, 1985). Group differences in this mathematical relationship indicate differential responding to items across groups, thereby suggesting measurement variance. Item response theory

models posit that the probability of a response at a given level of ability or latent trait is a function of item discrimination and item difficulty parameters. McDonald (1982) has shown that this model is equivalent to logistic factor analysis of the type used to analyze dichotomously scored items. The discrimination parameter in an IRT model is equivalent to the factor loading divided by the square root of the unique variance in factor analysis terms. The difficulty parameter in an IRT model is equivalent to a negative threshold value divided by the factor loading in factor analysis terms (see McDonald, 1982, for a complete account).

This equivalent relationship indicates that measurement invariance may be present if the factor structure of a scale differs across groups (Benson, 1987) such that factorial invariance (Byrne, Muthen, & Shavelson, 1989) is not obtained. A parallel thus exists between measurement invariance (whether the psychological construct tapped by items is equivalent across groups) and factorial invariance (whether the underlying dimensions represented by items are equivalent across groups). Using structured interviews and an item analysis, we investigated the validity of the MOCI for assessing the severity of OCD symptoms and predicting diagnosis in Blacks and Whites.

Methods

Participants

The participants in this study were undergraduate students at the University of Virginia. Over 3 years, 1,947 students were screened with the MOCI. This group included 113 Black males (6%), 201 Black females (10%), 765 White males (39%), and 868 White females (45%). Within 2 to 8 weeks after they completed the initial questionnaires, 124 of the original participants were interviewed (48 people with high scores on the MOCI and 76 people with low scores on the MOCI). The group of interviewed participants included 22 Black males (18%), 32 Black females (26%), 32 White males (26%), and 38 White females (31%).

Measures

The MOCI (Hodgson & Rachman, 1977) was used to screen participants for obsessive-compulsive

traits and behaviors. The MOCI contains 30 items in a true-false format and yields a total "obsessional-ity" score as well as four subscale scores: Checking, Cleaning, Slowness, and Doubting. Internal consistency and test-retest reliability have been shown to be quite good for the MOCI total score and for the Checking and Cleaning subscales. The scale also has acceptable concurrent validity with other OCD measures (Emmelkamp, Kraaijkamp, & Hout, 1999; Richter, Cox, & Drenfeld, 1994).

Scoring of the MOCI is based on the number of items endorsed in the pathological direction; therefore, a total MOCI score is based on a possible 30 points. The mean for our total sample was 7.12 ($SD = 4.33$). This figure is similar to the norms obtained by Sternberger and Burns (1990b) with a sample of university students ($M = 7.58$, $SD = 4.28$).

Interviews

The selection of participants to be interviewed was based on MOCI scores. We selected a group of high scorers (i.e., MOCI scores of 17 or higher) and a comparison group of low scorers (i.e., MOCI scores of 12 or lower). We chose 17 as the cutoff score for high scorers because it is approximately two standard deviations above the mean for our total sample. The comparison group was matched to the high scoring group on race and gender.

We initially used the Anxiety Disorders Interview Schedule-Revised (ADIS-R; DiNardo & Barlow, 1988) to assess participants for specific anxiety disorders. Diagnoses were made using *Diagnostic and Statistical Manual of Mental Disorders*, revised third edition (*DSM-III-R*; American Psychiatric Association, 1987) criteria. Diagnostic agreement using the ADIS-R has been shown to be quite good for OCD and other forms of anxiety disorder (Chorpita, Brown, & Barlow, 1998). During the last year of the study, our research protocol was modified. The Yale-Brown Obsessive Compulsive Scale (Y-BOCS; Goodman et al., 1989) was used in place of the ADIS-R to assess the severity of OCD symptoms using *DSM-III-R* criteria.² The Y-BOCS has been shown to have good inter-rater agreement and internal consistency

²Eighty-one of the students were interviewed with the ADIS-R; 43 of the students were interviewed using the Y-BOCS.

in many previous studies. The convergent validity of the Y-BOCS with other measures of OCD is also quite good (Amir, Foa, & Coles, 1997; Goodman & Price, 1992; Nakagawa, Marks, Takei, de Araujo, & Ito, 1996).

Procedures

The initial screening of participants with the MOCI was completed as part of a mass testing session of introductory psychology students, which is conducted approximately 3 weeks after the beginning of each semester. Within 2 to 8 weeks of the screening, selected participants were interviewed for obsessive-compulsive thoughts and behaviors by a clinical psychology graduate student or a clinical faculty member. With participants' consent, the interviews were videotaped. At the conclusion of the interview, participants received course credit for their participation and were debriefed. On the basis of the interview data, the experimenter decided whether the person met diagnostic criteria for OCD.

Reliability analysis

In order to assess the reliability of OCD diagnoses, the video-taped interviews were reviewed by the first author (JT) for an independent assessment of OCD. Under these circumstances, the author was not blind to race or gender, but she was blind to each student's MOCI score. In cases of disagreement between the author and the original interviewer, the author's diagnosis was used for data analysis. Inter-rater agreement was acceptable (87% agreement between raters and a Kappa value of .61).

Data Analysis

Factor analysis

Factor analytic methods were used to conduct a multidimensional IRT (Muthen & Christofferson, 1981) analysis of MOCI items. The equivalences between factor analysis and IRT allowed us to use the LISCOMP (Muthen, 1987) program for structural equation modeling of dichotomously scored items. LISCOMP permits maximum likelihood tests of the equality of factor structures (i.e., item parameters) of the MOCI in Blacks and Whites. LISCOMP was applied to the Cleaning and Checking subscales. For the Cleaning subscale

there were 314 Blacks and 1,633 Whites. For the Checking subscale, there were 311 Blacks and 1,627 Whites.³ Although significant race differences were also found on the Slowness subscale, factor analyses were not conducted because of psychometric and conceptual limitations previously shown to be associated with this subscale (Rachman & Hodgson, 1980, p. 159).

The analyses began by fitting a one-factor model (Model 1), without equality constraints between races, to the subscale items. The analyses then proceeded with a series of two-factor, oblique models that became progressively more restrictive.

The first two-factor model was an unconstrained model in which the factor structures were allowed to differ between groups (Model 2). The purpose of this model was to detect if its fit was better than the one-factor model. The second two-factor model was a constrained model in which equivalent factor loadings (or item discrimination parameters) were specified between groups (Model 3). A Promax rotation in SAS was applied to the matrix of factor loadings. The difference in fit between the unconstrained and constrained models tests the null hypothesis of no difference in the factor structure between groups.

The last two-factor model was a more restrictive constrained model in which item discrimination and item difficulty parameters were equated between groups (Model 4). This model uses the mean differences in the k latent traits to account for the observed differences in item endorsement frequency. In other words, this model determines if the race differences in the items can be accounted for by a single mean difference between groups on the latent (OCD) trait. Three factor models were also attempted, but LISCOMP did not converge to a solution despite repeated attempts with different starting values for the parameters.

The chi-square (χ^2) statistic was used to compare the fit of competing models. The χ^2 value from the more restrictive model was subtracted from the χ^2 value from an alternative model. The corresponding difference in degrees of freedom

³Missing data account for the difference in the number of students used for analysis of the Cleaning and Checking subscales.

was also determined. Since the statistical significance of chi-square can be influenced by sample size, the Tucker-Lewis Index (TLI; Tucker & Lewis, 1973), which is relatively independent of sample size (Marsh, Balla, & McDonald, 1988), was also used to provide an additional source of validation of the goodness-of-fit of the models. A statistically significant difference in chi-square and the value at which the TLI changes direction indicate that the less restrictive model fits the data better than the more restrictive model.

Results

Preliminary descriptive analyses

A 2×2 Race by Gender analysis of variance (ANOVA) on the entire sample indicated that the total MOCI score for Blacks was higher than Whites, $F(1, 1943) = 111.55, p < .001$. Table 1 shows that Blacks scored nearly one standard deviation higher than Whites and the total MOCI scores for females were significantly higher than males, $F(1, 1943) = 5.86, p < .05$. In addition, a significant interaction revealed that Black males scored higher than Black females; whereas White males scored lower than White females, $F(1, 1943) = 8.30, p < .01$.

To further understand these results, separate ANOVAs were performed for each MOCI subscale (see Table 1). Blacks scored significantly higher than Whites, $F(1, 1943) = 205.26, p < .001$, on the 10-item Cleaning subscale. In fact, Blacks scored almost one standard deviation higher than Whites. Females scored significantly higher than males, $F(1, 1943) = 13.36, p < .001$. As with the total MOCI score, a significant interaction indicated that Black males scored higher than Black females; whereas White males scored lower than White females, $F(1, 1943) = 5.76, p < .05$. For the 9-item Checking subscale, Blacks scored significantly higher than Whites, $F(1, 1943) = 36.97, p < .001$. There was not a main effect for gender. An interaction indicated that Black males scored higher than Black females; White males scored lower than White females, $F(1, 1943) = 8.75, p < .01$. Table 2 provides examples of several items that were endorsed in the pathological direction more frequently by Blacks than Whites. Cleaning items showed the greatest race differences in endorsement frequency.

Analyses of interview data

We were interested in the correspondence of MOCI scores with diagnostic status for high MOCI scorers (i.e., ≥ 17) as compared with low MOCI scorers (i.e., ≤ 12). Race and MOCI score group are not independent, $\chi^2(1) = 29.32, p < .001$. In the entire sample, more Blacks (9.5%, $n = 25$) than Whites (2.7%, $n = 41$) scored in the clinically significant range. Of the 19 Blacks with high MOCI scores who were interviewed, 16% ($n = 3$) met diagnostic criteria for OCD. Approximately the same proportion was obtained for the 35 Blacks who had MOCI scores in the lower range ($n = 5$). Of the 29 Whites with high MOCI scores who were interviewed, 28% ($n = 8$) met diagnostic criteria for OCD. The diagnostic hit rate among the 41 Whites with MOCI scores in the lower range was 13% ($n = 4$). This difference for Whites was not significant, $\chi^2(1) = 1.82, ns$.

Factor analysis

Our next set of analyses involved a multivariate item response theory analysis to study race differences in the psychometric structure of the Cleaning and Checking subscales. These analyses were conducted on the total sample, using a sequence of hierarchical factor models. The results of model fitting are summarized in Table 3.

For the Cleaning subscale, moving from the unconstrained, one-factor Model 1, $\chi^2(70) = 128.37, p < .05$; TLI = .9085, to the unconstrained, two-factor Model 2 resulted in a significant improvement of fit, $\Delta \chi^2(18) = 63, p < .05$; TLI = .9711. This demonstrates that two factors fit the Cleaning subscale better than one factor.

Since two factors fit the data, the next steps determined if the two-factor structure has equivalent psychometric properties across races. First, item discrimination parameters (or factor loadings) were constrained to equality across groups. The constrained, two-factor Model 3 did not produce a significant decrease in model fit from the unconstrained Model 2, $\Delta \chi^2(16) = 17.28, ns$; TLI = .9758. This nonsignificant difference in χ^2 between the more restrictive Model 3 and the less restrictive Model 2 confirms the null hypothesis that the same factor structures apply to Blacks and Whites.

Table 1
Means and Standard Deviations for MOCI Total Scores and for Cleaning and Checking Subscales by Gender and Race

	Blacks				Whites				Effect Size		
	Males (n = 113)		Females (n = 201)		Males (n = 765)		Females (n = 868)		Gender	Race	Interaction
	M	SD	M	SD	M	SD	M	SD			
Total MOCI	10.0	4.8	9.2	4.6	6.3	4.1	7.0	4.2	0.15	0.66	0.37
Cleaning	3.7	2.2	3.5	2.2	1.7	1.6	2.1	1.9	0.22	0.90	0.30
Checking	2.7	1.9	2.2	1.7	1.7	1.7	1.8	1.7	0.03	0.37	0.38

Note. MOCI = Maudsley Obsessional Compulsive Inventory. The values represent the mean number of MOCI items answered in the pathological direction. Maximum score = 30.

Psychometric Analysis of Racial Differences

Table 2
Examples of Endorsement Frequencies for MOCI Cleaning and Checking Items for the Entire Sample

Item	% Endorsement of Pathological Response		
	Blacks	Whites	Difference
Worry unduly about contamination if I touch an animal.	33% ^a	14% ^b	19%
Check letters over and over again before mailing them.	36% ^c	25% ^d	11%
Use only an average amount of soap.	39% ^c	13% ^e	26%
Excessively concerned about cleanliness.	60% ^c	36% ^b	24%
Tend to check things more than once.	61% ^e	43% ^b	12%
Can use well-kept toilets without any hesitation.	36% ^c	12% ^b	24%
Unduly concerned about germs and diseases.	52% ^c	19% ^b	33%

Note. MOCI = Maudsley Obsessional Compulsive Inventory.
^a*n* = 315, ^b*n* = 1,633, ^c*n* = 314, ^d*n* = 1,632, ^e*n* = 313.

Table 3
Results of MOCI Factor Analysis Model Fitting

Subscale	Model	χ^2	<i>df</i>	Change in fit	TLI
Cleaning	Model 1 Unconstrained 1 Factor	128.4	70		0.91
	Model 2 Unconstrained 2 Factor	65.4	52	<i>p</i> < .05	0.97
	Model 3 Item discrimination equal in races	82.7	68	<i>ns</i>	0.98
	Model 4 Item discrimination and difficulties equal in races	167.2	60	<i>p</i> < .05	0.87
Checking	Model 1 Unconstrained 1 Factor	146.4	40		0.89
	Model 2 Unconstrained 2 Factor	46.5	26	<i>p</i> < .05	0.97
	Model 3 Item discrimination equal in races	65.7	38	<i>ns</i>	0.97
	Model 4 Item discrimination and difficulties equal in races	121.5	44	<i>p</i> < .05	0.92

Note. MOCI = Maudsley Obsessional Compulsive Inventory. TLI = Tucker-Lewis Index.

The rotated factors that emerged suggest a washing factor and a contamination factor (see Table 4). These factors are moderately correlated in Blacks (*r* = .47) and in Whites (*r* = .64).

In the final step, both item discrimination and item difficulty parameters were constrained to equality across groups (Model 4). This resulted in a significant decrease in model fit. A statistically significant result was obtained when Model 4 was compared to Model 3, $\Delta \chi^2(8) = 85.47, p < .05$; TLI = .8665. This shows that the less restrictive Model 3 fits better than the more restrictive Model 4, which suggests that the items do not

have the same relationship to difficulty in Blacks and Whites. Thus, Model 4 was rejected. When item discrimination and item difficulty parameters are equated between groups, the psychometric structure of the Cleaning subscale is not the same for Blacks and Whites. In other words, the race difference in endorsement frequency cannot be accounted for solely by a mean difference between groups on the latent trait; something other than the latent (OCD) trait appears to influence responses.

The results for the Checking subscale show a similar pattern. Moving from the unconstrained, one-factor Model 1, $\chi^2(40) = 146.43, p < .05$; TLI = .8851, to the

Table 4
Parameters of the Rotated, Two-Factor Model for the MOCI Cleaning Subscale With Factor Loadings Constrained Equal

Item	Factor 1	Factor 2	Intercepts	
			Blacks	Whites
1. Avoid using public phones	-.000	.882	-1.426	-1.971
5. Worry about contamination if I touch animals	-.291	.926	-.450	-1.091
9. Concerned if I bump somebody	-.022	.520	-.963	-1.095
13. Use a lot of soap	1.035	-.040	-.256	-1.140
17. Excessive concern about cleanliness	.938	.240	.242	-.373
19. Concerned about using toilets	-.190	.895	-.362	-1.170
21. Unduly concerned about germs and diseases	.320	.784	.025	-.873
24. Concerned about touching money	.300	.568	-.725	-.828
26. Take my time washing	.951	-.138	-.450	-.800
27. Use a lot of antiseptics	.991	.086	-.790	-1.158

Note. MOCI = Maudsley Obsessional Compulsive Inventory. The entries in bold denote items loading on the factor.

unconstrained, two-factor Models 2 resulted in an improvement in model fit, $\Delta \chi^2(14) = 99.86$, $p < .05$; TLI = .9658). This indicates that two factors fit the Checking subscale better than one factor.

Constraining item discrimination parameters to equality across groups (Model 3) fit the data better than the less restrictive unconstrained model, as indicated by the non-significant change in χ^2 , $\Delta \chi^2(12) = 19.42$, ns ; TLI = .9682. This demonstrates that the same factor structures apply to both groups (see Table 5). The resulting factors represent a checking factor and an obsessive

thoughts factor. These factors are moderately correlated in Blacks ($r = .36$) and Whites ($r = .37$).

The more restrictive Model 4 (with both item discrimination and item difficulty parameters equal across groups) resulted in a significant worsening of model fit, $\Delta \chi^2(6) = 55.78$, $p < .05$; TLI = .9237, and was rejected. As with the Cleaning items, the Checking items do not have the same relationship to difficulty across groups. When both item discrimination and item difficulty are equated, the psychometric structure of the Checking subscale is different for Blacks and Whites.

Table 5
Parameters of the Rotated, Two-Factor Model for the MOCI Checking Subscale With Factor Loadings Constrained Equal

Item	Factor 1	Factor 2	Intercepts	
			Blacks	Whites
2. Frequent nasty thoughts	.128	1.129	.770	1.033
6. Frequently check things	1.050	-.452	.648	.620
8. Uncontrollable unpleasant thoughts	.287	1.147	1.051	1.023
14. There are unlucky numbers	.263	.130	1.008	1.283
15. Check letters over and over before mailing	.788	-.127	.395	.681
22. Check things over and over	1.096	-.477	-.262	.191
26. Take my time washing	.203	.451	.414	.809
28. Spend a lot of time checking things	1.006	-.086	1.216	1.527

Note. MOCI = Maudsley Obsessional Compulsive Inventory. The entries in bold denote items loading on the factor.

Discussion

We used the MOCI to study the frequency and severity of obsessive-compulsive symptoms in a sample of university undergraduates. Previous research has shown that nonclinical samples exhibit symptoms of OCD, that the MOCI is a valid instrument for identifying OCD symptoms (Sternberger & Burns, 1990b), and that some people in nonclinical samples meet diagnostic criteria for OCD (Burns, Formea, Keortge, & Sternberger, 1995; Gibbs, 1996). Unfortunately, these studies have not examined the usefulness of the MOCI with ethnic minorities. We found unexpectedly that the mean MOCI score for Black university students was almost one standard deviation higher than that for White students.⁴ The present article explored the validity of the MOCI with Black participants using psychometric analyses. Our report complements other reports which also indicate that college samples can provide a means for assessing the validity and utility of psychological measures in different ethnic groups (Chmielewski, Fernandes, Yee, & Miller, 1995).

Interview data

Although significantly more Blacks than Whites had high MOCI scores, Whites with high MOCI scores were almost two times more likely than Blacks with high MOCI scores to meet diagnostic criteria for OCD. Even more interesting is the finding that low scoring Blacks were just as likely as high scoring Blacks to meet diagnostic criteria for OCD. This result demonstrates a lack of predictive validity of the MOCI in Blacks. In other words, the tendency of Blacks to endorse more MOCI items does not seem to be a reflection of pathological OCD symptoms.

The predictive validity of the MOCI was better for Whites than for Blacks. High scoring Whites (28%) were twice as likely as low scoring Whites (13%) to meet the *DSM-III-R* criteria for OCD. The fact that the MOCI did bear some relationship to these

⁴Although the MOCI does not have widely accepted cutoff scores, the mean MOCI score for blacks was not in the pathological range. The mean score for blacks was lower than the mean MOCI score of Hodgson and Rachman's (1977) original clinical sample ($M = 18.9$, $SD = 4.9$) and within one standard deviation of the mean of our total sample.

diagnostic decisions among Whites is consistent with previous studies concerning the use of the MOCI in student populations (Sternberger & Burns, 1990b). This race difference in predictive validity suggests that the MOCI did not have equivalent relations with external variables (i.e., structured interview) in both groups (Drasgow, 1984).

Factor Analysis

We also examined the race differences on the MOCI using factor analytic methods to conduct an item response theory analysis of Cleaning and Checking subscale items. The Cleaning subscale was more appropriately represented by a two-factor solution (washing and contamination), which is consistent with the results of Sternberger and Burns' (1990a) principal components analysis. In our factor analysis of the Cleaning subscale, Items 9 and 24 do not fit the model as well as the other items (see Table 4). The small group differences in endorsement frequency separate these items from the others (see intercept values in Table 4) and may be responsible for their lack of fit. The Checking subscale was more clearly represented by a checking factor and an obsessive thoughts factor.

For both subscales, models in which item discrimination parameters were constrained to be equal fit the data in Blacks and Whites. A more restrictive model in which item difficulty parameters were also constrained to equality across groups did not fit the data. This suggests that the race difference in endorsement frequency cannot be explained solely by a mean difference in the latent trait. Differences in item difficulty between groups did not allow the same factor structure or the same linear combination of items to fit each group. The interaction between race and item difficulty indicates that the psychometric properties of the MOCI are not equivalent in Blacks and Whites. Therefore, measurement invariance was not obtained (Horn & McArdle, 1992; Reise et al., 1993).

Differences in item difficulty may indicate that items are "easier" for Blacks, meaning that less of the latent trait is required in order for an item to be endorsed. This seems to produce the differential response patterns between groups in which Blacks endorse more items. If items are ordered

differently in terms of endorsement frequency, then they are not measuring the latent trait similarly in both groups (Thissen, Steinberg, & Gerrard, 1986). The observed race difference on the MOCI appears to be the result of the item composition of the inventory; there are many items on which Blacks endorse the pathological response more often than Whites. By comparing item difficulty and discrimination parameters, the IRT analysis revealed that some MOCI items do not function the same for Blacks and Whites. Without evidence of measurement invariance or factor invariance, the mean MOCI differences between Blacks and Whites may be uninterpretable (Horn & McArdle, 1992).

Cultural Considerations

The psychological meaning of measurement invariance in a particular data set cannot be evaluated solely in mathematical terms. A theoretical framework is also required for this enterprise (Poortinga & van der Flier, 1988). Indeed, the choice between measurement invariance and valid group differences can be complicated. Not only is measurement invariance in racial group comparisons important, but conceptual invariance is important as well (Jackson, 1989). Given the sparse literature on OCD in Blacks as well as the absence of a theory of race differences in OCD, attempts to account for our obtained race differences are at best speculative.

It seems clear that race differences on the MOCI are not driven by OCD as a mental disorder. We did not set out to explain in this article why the differential response pattern occurred for the MOCI items. Nevertheless, we will offer some suggestions. One important consideration may be simply minority status. People who are not members of the majority group may approach questionnaires with a different set of expectations that influence their responses to particular questions. Cultural differences in various kinds of attitudes and beliefs, especially those regarding issues such as spirituality, may also be related to this phenomenon. Prior exposure to violence, interpersonal conflict, and other anxiety-provoking experiences represent additional considerations that might

help to elucidate differences in the ways that Blacks and Whites responded to certain items on the MOCI. Systematic exploration of the social meanings attached to various forms of danger, contamination, and unacceptable thoughts may all be relevant to this process (Hatch, Friedman, & Paradis, 1996; Neal-Barnett & Smith, 1997; Williams, Chambless, & Steketee, 1998).

Limitations of the Study

An empirical problem involved in studying racial differences in psychopathology relates to sampling and representativeness. For example, in this study uneducated and older Blacks and Whites were not sampled. Blacks and Whites are diverse, differing between and within groups by situational and experiential factors related to national origin, rural versus urban residence, religious affiliation, social class, family structure, and gender (McKenry et al., 1989; Zuckerman, 1990). Therefore, other background variables may legitimately explain group differences (van de Vijver & Poortinga, 1991).

The representativeness of MOCI items themselves may also be called into question. A larger sampling of various obsessive and compulsive symptoms may yield a clearer picture of the driving force behind the race difference on the MOCI. The next step in this research might be to develop more items covering a wider range of OCD symptoms; to administer them to a diverse, non-college sample; and to conduct another IRT analysis of the items.

The relatively high proportion of our control participants (those with low scores on the MOCI) who were considered to meet diagnostic criteria for OCD after completing a structured diagnostic interview is surprising. This suggests that a broad, inclusive definition of OCD may have been employed resulting in the over-diagnosis of OCD. It is also possible that students over-reported OCD symptoms. In contrast, the relatively low rate of diagnosis among high MOCI scorers not only calls into question the predictive validity of the MOCI but may also suggest that high scorers under-reported symptoms when they were interviewed.

Directions for Future Research

In general, questionnaires measuring anxiety symptoms may not be valid with ethnic minorities unless they have been examined for that purpose. This study has raised questions about the utility of the MOCI for screening Black nonclinical participants. The MOCI may be of limited utility in Blacks if the purpose is to use the inventory to find people who are likely to meet diagnostic criteria for OCD. Future efforts in this area should be aimed at developing a self-report measure that is appropriate for screening obsessive-compulsive symptoms in Black participants.

References

- Al-Hssa, I., & Oudji, S. (1998). Culture and anxiety disorders. In S. S. Kazarian and D. R. Evans (Eds.), *Cultural clinical psychology: Theory, research and practice* (pp. 127-151). New York: Oxford University Press.
- American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed., rev.). Washington, DC: Author.
- Amir, N., Foa, E. B., & Coles, M. E. (1997). Factor structure of the Yale-Brown Obsessive Compulsive Scale. *Psychological Assessment*, 9, 312-316.
- Benson, J. (1987). Detecting item bias in affective scales. *Educational and Psychological Measurement*, 47, 55-67.
- Bernstein, D. M. (1997). Anxiety disorders. In W. Tseng & J. Streltzer (Eds.), *Culture and psychopathology: A guide to clinical assessment* (pp. 46-66). New York: Brunner/Mazel.
- Burns, G. L., Formea, G. M., Keortge, S., & Sternberger, L. G. (1995). The utilization of nonpatient samples in the study of obsessive compulsive disorder. *Behaviour Research and Therapy*, 33, 133-144.
- Butcher, J. N., Nezami, E., & Exner, J. (1998). Psychological assessment of people in diverse cultures. In S. S. Kazarian & D. R. Evans (Eds.), *Cultural clinical psychology: Theory, research, and practice* (pp. 61-105). New York: Oxford University Press.
- Byrne, B. M., Muthen, B., & Shavelson, R. J. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin*, 105(3), 456-466.
- Chan, D. W. (1990). The Maudsley Obsessional-Compulsive Inventory: A psychometric investigation on Chinese normal subjects. *Behaviour Research and Therapy*, 28, 413-420.
- Chmielewski, P. M., Fernandes, L. O. L., Yee, C. M., & Miller, G. A. (1995). Ethnicity and gender in scales of psychosis-proneness and mood disorders. *Journal of Abnormal Psychology*, 104, 464-470.
- Chorpita, B. F., Brown, T. A., & Barlow, D. H. (1998). Diagnostic reliability of the *DSM-III-R* anxiety disorders. *Behavior Modification*, 22, 307-320.
- DiNardo, P. A., & Barlow, D. H. (1988). *Anxiety Disorders Interview Schedule-Revised (ADIS-R)*. Albany, NY: Graywind.
- Dragow, F. (1984). Scrutinizing psychological tests: Measurement equivalence and equivalent relations with external variables are the central issues. *Psychological Bulletin*, 95(1), 134-135.
- Emmelkamp, P. M. G., Kraaijkamp, H. J. M., & Van den Hout, M. A. (1999). Assessment of obsessive-compulsive disorder. *Behavior Modification*, 23, 269-279.
- Fabrega, H., Jr., Mezzich, J., & Ulrich, R. F. (1988). Black-white differences in psychopathology in an urban psychiatric population. *Comprehensive Psychiatry*, 29(3), 285-297.
- Frost, R. O., Steketee, G., Krause, M. S., & Trepanier, K. L. (1995). The relationship of the Yale-Brown Obsessive Compulsive Scale (YBOCS) to other measures of obsessive compulsive symptoms in a nonclinical population. *Journal of Personality Assessment*, 65, 158-168.
- Gibbs, N. A. (1996). Nonclinical populations in research on obsessive-compulsive disorder: A critical review. *Clinical Psychology Review*, 16, 729-773.
- Goodman, W. K., & Price, L. H. (1992). Assessment of severity and change in obsessive compulsive disorder. *Psychiatric Clinics of North America*, 15, 861-869.
- Goodman, W. K., Price, L. H., Rasmussen, S. A., Mazure, C., Fleischmann, R. L., Hill, C., Henenger, G. R., & Charney, D. S. (1989). The Yale-Brown Obsessive Compulsive Scale: I, Development, use, and reliability. *Archives of General Psychiatry*, 46, 1006-1011.
- Hamleton, R. K., & Swaminathan, H. (1985). *Item response theory: Principles and applications*. Boston: Kluwer Academic Publishers.
- Hatch, M. L., Friedman, S., & Paradis, C. M. (1996). Behavioral treatment of obsessive-compulsive disorder in African Americans. *Cognitive & Behavioral Practice*, 3, 303-315.
- Hodgson, R. J., & Rachman, S. (1977). Obsessional-compulsive complaints. *Behaviour Research and Therapy*, 15, 389-395.
- Horn, J. L., & McArdle, J. J. (1992). A practical and theoretical guide to measurement invariance in aging research. *Experimental Aging Research*, 18(3), 117-144.
- Jackson, J. (1989). Race, ethnicity, and psychological theory and research. *Journal of Gerontology*, 44(1), 1-2.
- Karno, M., Golding, J. M., Sorenson, S. B., & Burnam, A. (1988). The epidemiology of obsessive-compulsive disorder in five US communities. *Archives of General Psychiatry*, 45, 1094-1099.
- Lopez, A. D., & Murray, C. C. J. L. (1998). The global burden of disease, 1990-2020. *Nature Medicine*, 4, 1241-1243.
- Marsh, H. W., Balla, J. R., & McDonald, R. P. (1988). Goodness-of-fit indexes in confirmatory factor analysis: The effect of sample size. *Psychological Bulletin*, 103, 391-410.
- McDonald, R. P. (1982). Linear versus nonlinear models in item response theory. *Applied Psychological Measurement*, 6(4), 379-396.
- McKenry, P. C., Everett, J. E., Ramseur, H. P., & Carter, C. J. (1989). Research on black adolescents: A legacy of cultural bias. *Journal of Adolescent Research*, 4(2), 254-264.
- Muthen, B. (1987). *LISCOMP: Analysis of linear structural relations with a comprehensive measurement model* [Computer software]. Morrisville, IN: Scientific Software.

- Muthen, B., & Christofferson, A. (1981). Simultaneous factor analysis of dichotomous variables in several groups. *Psychometrika*, *46*, 407-419.
- Nakagawa, A., Marks, I. M., Takei, N., de Araujo, L. A., & Ito, L. M. (1996). Comparisons among the Yale-Brown Obsessive-Compulsive Scale, compulsion checklist, and other measures of obsessive-compulsive disorder. *British Journal of Psychiatry*, *169*, 108-112.
- Neal-Barnett, A. M., & Smith, J. (1997). African Americans. In S. Friedman (Ed.), *Cultural issues in the treatment of anxiety* (pp. 154-174). New York: Guilford.
- Oltmanns, T. F., & Gibbs, N. (1995). Emotional responsiveness and obsessive-compulsive behavior. *Cognition and Emotion*, *9*, 563-578.
- Poortinga, Y. H., & van der Flier, H. (1988). The meaning of item bias in ability tests. In S. H. Irvine & J. W. Berry (Eds.), *Human abilities in cultural context* (pp. 166-183). New York: Cambridge University Press.
- Rachman, S. J., & Hodgson, R. J. (1980). *Obsessions and compulsions*. Englewood Cliffs, NJ: Prentice-Hall.
- Reise, S. P., Widaman, K. F., & Pugh, R. H. (1993). Confirmatory factor analysis and item response theory: Two approaches for exploring measurement invariance. *Psychological Bulletin*, *114*(3), 552-566.
- Richter, M. A., Cox, B. J., & Dorenfeld, D. M. (1994). A comparison of three assessment instruments for obsessive-compulsive symptoms. *Journal of Behavior Therapy and Experimental Psychiatry*, *25*, 143-147.
- Steketee, G. (1994). Behavioral assessment and treatment planning with obsessive compulsive disorder: A review emphasizing clinical application. *Behavior Therapy*, *25*, 613-633.
- Sternberger, L. G., & Burns, G. L. (1990a). Compulsive Activity Checklist and the Maudsley Obsessional-Compulsive Inventory: Psychometric properties of two measures of Obsessive-Compulsive Disorder. *Behavior Therapy*, *21*, 117-127.
- Sternberger, L. G., & Burns, G. L. (1990b). Maudsley Obsessional-Compulsive Inventory: Obsessions and compulsions in a nonclinical sample. *Behaviour Research and Therapy*, *28*(4), 337-340.
- Thissen, D., Steinberg, L., & Gerrard, M. (1986). Beyond group-mean differences: The concept of item bias. *Psychological Bulletin*, *99*(1), 118-128.
- Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, *38*, 1-10.
- Valleni-Basile, L. A., Garrison, C. Z., Waller, J. L., Addy, C. L., McKeown, R. E., Jackson, K. L., & Cuffe, S. P. (1996). Incidence of obsessive-compulsive disorder in a community sample of young adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry*, *35*, 898-906.
- van de Vijver, F. J. R., & Poortinga, Y. H. (1991). Testing across cultures. In R. K. Hamleton & J. N. Zaal, *Advances in educational and psychological testing: Theory and applications* (pp. 277-308). Boston: Kluwer Academic Publishers.
- Williams, K. E., Chambless, D. L., & Steketee, G. (1998). Behavioral treatment of obsessive-compulsive disorder in African Americans: Clinical issues. *Journal of Behavior Therapy & Experimental Psychiatry*, *29*, 163-170.
- Zuckerman, M. (1990). Some dubious premises in research and theory on racial differences. *American Psychologist*, *45*, 1297-1303.