Shorter communication

Psychometric properties of the OCI-R in a college sample

Greg Hajcak [a,b,*], Jonathan D. Huppert [b], Robert F. Simons [a], Edna B. Foa [b]

[a] *Department of Psychology, University of Delaware, 222 Wolf Hall, Newark, DE 19716, USA*
[b] *University of Pennsylvania, Pennsylvania, PA, USA*

Abstract

Two studies examined the psychometric properties of the Obsessive–Compulsive Inventory—Revised (OCI-R; Psychol. Assessment 14 (2002) 485) in a nonclinical student sample. In Study 1, we investigated the factor structure and internal consistency of the OCI-R using a sample of 395 undergraduate students. At a second testing session 1 month later, 178 students completed the OCI-R. Test–retest reliability was examined using data from 94 students who completed the OCI-R in both sessions. Convergent validity was also assessed with the Maudsley Obsessive–Compulsive Inventory (MOCI). In Study 2, we further investigated the convergent and divergent validity of the OCI-R using a new sample of 221 students who completed a battery of measures of obsessive–compulsive symptoms, worry, and depression. There was a significant order effect for both the OCI-R and the MOCI: means of each measure were significantly lower when presented second. Despite the order effect, statistical analyses indicated that the OCI-R has adequate test–retest reliability for the full scale and subscale scores, solid factor structure, and high internal consistency. Convergent validity with other measures of obsessive–compulsive symptoms was moderate to excellent, and divergent validity was good. The results indicate that the OCI-R is a short, psychometrically sound self-report measure of obsessive–compulsive symptoms.
© 2003 Elsevier Ltd. All rights reserved.

Keywords: OCI; OCD; Scale; Assessment; Nonclinical sample; Obsessive-compulsive

1. Introduction

The Obsessive–Compulsive Inventory—Revised (OCI-R; Foa et al., 2002) is a shortened version of the Obsessive–Compulsive Inventory (OCI; Foa, Kozak, Salkovskis, Coles, & Amir,

* Corresponding author. Tel.: +1-302-831-1041; fax: +1-302-831-3645.
 E-mail address: hajcak@udel.edu (G. Hajcak).

1998) that measures obsessive–compulsive symptoms. In addition to providing a total score, the OCI-R yields six subscales: washing, checking, ordering, obsessing, hoarding, and neutralizing. Compared to the OCI, the OCI-R is significantly shorter, eliminates the frequency scale which appears to be redundant, has subscales of three items each which are summed (making them easy to compare), and retains excellent psychometric properties in a mixed sample of patients with obsessive–compulsive disorder, other anxiety disorders, and non-patients.

Although much research on OCD focuses on clinical populations, it is increasingly common for this research to extend to nonpatient samples to investigate constructs that are important to the theoretical understanding of OCD (for reviews, see Burns, Formea, Keortge, & Sternberger, 1995; Gibbs, 1996). For research on OCD to benefit from studies in nonclinical populations, it is crucial for measures that assess OCD symptoms to have sound psychometric properties within both clinical and nonclinical populations. For example, Simonds, Thorpe, and Elliott (2000) found that the original version of the OCI had excellent psychometrics in a population of college students—comparable to the properties reported in clinical populations by Foa et al. (1998). However, Wu and Watson (2002) investigated the factor structure of the OCI in a sample of college students and reported that a different, five-factor structure than that theorized by Foa et al. (1998) better fit the data, though goodness-of-fit indices were still not excellent according to most standards (e.g., Hu, & Bentler, 1999).

Although some data from nonanxious controls were reported for the OCI-R in the Foa et al. (2002) study, the psychometric performance of the OCI-R was not fully examined in a non-clinical student sample. The present study sought to examine the psychometric properties of the OCI-R in a population of college students.

2. Study 1

2.1. Method

2.1.1. Participants

Two large introductory psychology courses at the University of Delaware participated in a subject pool, completing a number of self-report measures of different psychological constructs. From this pool of subjects, 395 students had usable data for the current study. This sample consisted of 242 females, 146 males, and five students who did not identify their gender. A second sample of 178 students (106 women and 72 men) were administered self-report measures 1 month later. Of these students, 94 (61 women and 33 men) had completed the OCI-R in both testing administrations, and they constituted the final sample for test–retest reliability. The vast majority of all samples were Caucasian freshmen. No further details on demographics were obtained.

2.1.2. Measures
2.1.2.1. *OCI-R.* The Obesssive–Compulsive Inventory—Revised (OCI-R; Foa et al., 2002) is a 18-item self-report measure that assesses the distress associated with obsessions and compulsions and has excellent psychometric properties in anxious patients.

2.1.2.2. PSWQ. The Penn State Worry Questionnaire (PSWQ; Meyer, Miller, Metzger, & Borkovec, 1990) is a self-report measure that assesses dysfunctional attitudes about worry and has excellent psychometric properties in both clinical and nonclinical populations (Brown, Antony, & Barlow, 1992).

2.1.2.3. MOCI. The Maudsley Obsessive–Compulsive Inventory (MOCI; Hodgson & Rachman, 1977) is a true/false self-report questionnaire that assesses rituals and obsessions. It has satisfactory test–retest reliability and internal consistency, although alphas for the subscales have been moderate (Emmelkamp, Kraaijkamp, & van den Hout, 1999).

2.1.3. Procedure

The OCI-R and PSWQ were administered to the 395 subjects in the original subject pool. Four weeks later, 200 students were signed up for a second testing session, which included the OCI-R and MOCI; 178 students (106 women, 72 men) from this second sample had usable data. In all cases, the OCI-R was administered before other self-report measures. Ninety-four students (61 women, 33 men) had complete OCI-R data from both the first and second testing administrations, and this sample was used for test–retest reliability. The vast majority of the students were Caucasian freshmen.

2.2. Results

2.2.1. Factor structure

The mean OCI-R total score was 18.91 (SD = 11.38); the subscale means were as follows: washing ($M = 2.41$; SD = 2.55), checking ($M = 2.95$; SD = 2.64), ordering ($M = 4.48$; SD = 3.16), obsessing ($M = 2.92$; SD = 2.82), hoarding ($M = 4.44$; SD = 2.72), and neutralizing ($M = 1.78$; SD = 2.20). The mean MOCI score was 7.02 (SD = 4.66). To assess the factor structure of the OCI-R, data from the first testing session ($N = 395$) was subjected to a confirmatory factor analysis using Proc Calis in SAS version 8.02. Latent factors were allowed to covary, while errors were not. The six-factor structure of the OCI-R was confirmed, using criteria recommended by Hu and Bentler (1999). The model had a significant chi-square ($\chi^2 120) = 272$, $p < 0.0001$), a goodness-of-fit index (GFI) of 0.93, a comparative fit index (CFI) of 0.94, a root mean square residual (RMR) of 0.05, and a root mean square error of approximation (RMSEA) of 0.06. All of these values indicate a good fit for the model except the chi-square. Standardized parameter estimates are available upon request from the first author. Each item has a moderate to high standardized loading on the predicted factor. This suggests that the OCI-R subscales represent separate coherent subtypes of obsessive–compulsive symptoms within a nonclinical sample. A one-factor model was tested for the 18 items, and all indices suggested poor fit compared to the six-factor model ($\chi^2(135) = 1024$, $p < 0.0001$; GFI = 0.76, CFI = 0.66, RMR = 0.11, RMSEA = 0.13).

2.2.2. Internal consistency

Cronbach's alpha was calculated and used to assess the internal consistency of the OCI-R total scale, as well as each of the subscales. The OCI-R total scale demonstrates excellent internal consistency (0.88). The washing (0.76), checking (0.76), ordering (0.84), and obsessing (0.77)

subscales demonstrated excellent internal consistency. The hoarding (0.68) and neutralizing (0.61) subscales showed moderate to good internal consistency.

2.2.3. Test–retest reliability

To assess test–retest reliability, Pearson's r was calculated. Although the OCI-R total ($r = 0.70$) and subscale (washing: $r = 0.72$; checking: $r = 0.75$; ordering: $r = 0.61$; obsessing: $r = 0.77$; hoarding: $r = 0.58$; mental neutralizing: $r = 0.54$) test–retest reliabilities are somewhat lower than those reported in Foa et al. (2002), all coefficients indicate good to excellent test–test reliability.

2.2.4. Convergent validity

Convergent validity of the OCI-R total scale was determined by correlating the OCI-R with another measure of obsessive–compulsive symptoms, the MOCI. The correlations between the OCI-R total and subscale scores with the MOCI total, washing subscale and checking subscale are presented in Table 1. The MOCI total score correlated 0.56 with the OCI-R, indicating moderate convergent validity.

In terms of subscale correlations, the MOCI Cleaning and Checking subscales correlated best with OCI-R Washing (0.49) and Checking (0.58) subscales, respectively. Importantly, the MOCI cleaning and checking subscales were each better correlated with the conceptually similar sub-scale on the OCI-R, relative to the other subscales.

2.2.5. Divergent validity

We assessed the divergent validity of the OCI-R by correlating the OCI-R total and subscale scores with the PSWQ—a measure of pathological worry. Although there is some overlap between worry and obsessive–compulsive symptoms, the OCI-R is intended to measure a construct that is distinct from pathological worry. These correlations are presented in Table 1. Although OCI-R total scores were significantly correlated with a measure of pathological worry, this correlation was significantly lower than the correlation between the OCI-R and the MOCI ($z = 2.03$, $p < 0.05$).

3. Study 2

In Study 1, the correlation between the OCI-R and MOCI total scores was rather low, in contrast to Simonds et al.'s (2000) finding that the original version of the OCI was highly correlated with the MOCI in a sample of college students. Additionally, Foa et al. (2002) report a high correlation between the OCI-R and the MOCI in 34 nonanxious control subjects. Given the discrepant findings in Study 1, we sought to evaluate the convergent validity of the OCI-R in a new sample using two measures of obsessive–compulsive symptoms: the MOCI and the Padua Inventory—Washington State University Revision (PI-WSUR; Burns, Keortge, Formea, & Sternberger, 1996).

In addition, because Foa et al. (2002) report high correlations between the OCI-R and the Beck Depression Inventory (BDI; Beck et al., 1979), the current study also assessed the divergent validity of the OCI-R using the BDI. Finally, to replicate the findings in Study 1, the PSWQ was also administered in Study 2.

Table 1
Convergent validity for the OCI-R subscale and total scores with the MOCI total, and cleaning and checking subscale scores (Studies 1 and 2), and with the PI-WSUR total, and contamination and checking subscales scores in Study 2, and divergent validity with the BDI and PSWQ

OCI subscale	Convergent validity						Divergent validity		
	MOCI (Study 1/Study 2)			PI-WSUR			BDI Total	PSWQ Total	PSWQ Total
	Total (n = 178/221)	Cleaning (n = 178/221)	Checking (n = 178/221)	Total (n = 221)	Contamination (n = 221)	Checking (n = 221)	(n = 221)	Study 1 (n = 393)	Study 2 (n = 221)
Washing	0.50/0.54	**0.49/0.50**	0.38/0.40	0.67	**0.71**	0.62	0.19*	0.30	0.26
Checking	0.46/0.62	0.21**/0.39	**0.58/0.57**	0.63	0.41	**0.74**	0.30	0.40	0.34
Ordering	0.36/0.45	0.32/0.28	0.37/0.41	0.51	0.44	0.54	0.12 (ns)	0.26	0.32
Obsessing	0.38/0.40	0.19*/0.19*	0.23**/0.40	0.49	0.33	0.47	0.51	0.26	0.35
Hoarding	0.34/0.36	0.16*/0.20*	0.32/0.30	0.38	0.23	0.42	0.31	0.32	0.36
Neutralizing	0.39/0.41	0.24**/0.27	0.31/0.42	0.56	0.33	0.53	0.25	0.23	0.11 (ns)
Total score	**0.56/0.65**	0.37/0.42	0.51/0.58	**0.75**	0.57	0.76	0.39	0.42	0.42

OCI-R, Obsessive-Compulsive Inventory—Revised; BDI, Beck Depression Inventory; PSWQ, Penn State Worry Questionnaire. All p-values <0.001, except where indicated with a *(p-values<0.05), **(p-values<0.01), or ns (=not significant). Bold values indicate correlations with similar scales.

3.1. Method

3.1.1. Participants

A new sample of University of Delaware undergraduates in an introductory psychology course was used for Study 2. All students completed a packet of questionnaires, and received course credit for their participation. To control for potential order effects, the MOCI was given before the OCI-R in Study 2, counterbalancing the order used in Study 1. Usable data from 221 students comprised the final sample for Study 2. The sample consisted of 159 women and 62 men. The vast majority of students identified themselves as first (80.1%), or second year (14.9%) students. In terms of race, 81.4% of the students identified themselves as Caucasian, 5.4% as African American, 5.9% as Hispanic, 5.0% as Asian, and 5.3% as Native American.

3.1.2. Measures

3.1.2.1. PI-WSUR. The Padua Inventory—Washington State University Revision (Burns et al.,1996) is a 39-item self-report measure that was designed to differentiate OCD symptoms from worry. The PI-WSUR yields a total score and subscale scores for: (a) contamination obsessions and washing compulsions; (b) dressing/grooming compulsions; (c) obsessional thoughts of harm to self/others; and (d) obsessional impulses to harm self/other. Burns et al. (1996) report excellent psychometric properties for the PI-WSUR.

3.1.2.2. BDI. The Beck Depression Inventory (BDI; Beck et al., 1979) is a 21-item self-report scale used to assess cognitive and physical symptoms of depression. It has widely been used in psychological research, with good reliability and validity.

3.2. Results

3.2.1. Descriptive statistics

In Study 2, the mean OCI-R total score was 11.95 (SD $= 9.30$); the subscale means were as follows: washing ($M = 1.58$; SD $= 2.06$), checking ($M = 1.32$; SD $= 2.11$), ordering ($M = 3.01$; SD $= 2.71$), obsessing ($M = 1.73$; SD $= 2.31$), hoarding ($M = 3.31$; SD $= 2.33$), and neutralizing ($M = 1.00$; SD $= 1.60$). The mean MOCI score was 8.58 (SD $= 5.11$). The total OCI-R scores were significantly lower in Study 2 than in Study 1 ($F(1, 220) = 51.10$, $p < 0.001$). However, relative to Study 1, the means for the MOCI were significantly higher in Study 2 ($M = 7.02$, SD $= 4.66$ vs. $M = 8.58$, SD $= 5.11$, respectively; $F(1, 177) = 7.71$, $p < 0.01$). In addition, all of the subscale scores for both the OCI-R and the MOCI were significantly lower when given second, relative to the other measure. Thus, the results from Studies 1 and 2 collectively indicate that both the OCI-R and the MOCI are susceptible to significant order effects, when given with other self-report measures of obsessive–compulsive symptoms.

3.2.2. Factor structure and internal consistency

Because the OCI-R means were significantly different between Study 1 and Study 2, we sought to reexamine the factor structure and internal consistency in the data from Study 2 using the same procedures as Study 1. Despite having higher scores, the factor structure remained essentially the same, and the fit was slightly better than in Sample 1 ($\chi^2(120) = 217$, $p < 0.0001$; GFI $= 0.91$, CFI $= 0.94$, RMR $= 0.04$, RMSEA $= 0.06$).

In terms of internal consistency, the OCI-R total scale demonstrates excellent internal consistency (0.88). Again, the washing (0.75), checking (0.81), ordering (0.88), and obsessing (0.81) subscales demonstrated excellent internal consistency. The two exceptions are again the hoarding (0.65) and neutralizing (0.47) subscales.

3.2.3. Convergent validity

Convergent validity of the OCI-R total scale was determined by correlating the OCI-R with the MOCI and PI-WSUR total scores. The correlations between the OCI-R total and subscale scores with the MOCI total, washing subscale and checking subscale are presented in Table 1. The MOCI total score correlated 0.65 with the OCI-R, indicating moderate convergent validity. In terms of subscale correlations, the MOCI cleaning and checking subscales correlated best with OCI-R washing (0.50) and checking (0.57) subscales, respectively, indicating that these subscales were measuring similar constructs.

Correlations between the OCI-R total and subscale scores with the PI-WSUR total and contamination and checking subscale scores are presented in Table 1. The OCI-R appears to have excellent convergent validity with the PI-WSUR for total score (0.75) as well as for washing (0.71) and checking (0.74) subscale scores.

Given the higher correlations with the OCI-R and the PI-WSUR, we examined correlations between the MOCI and the PI-WSUR. The correlation between MOCI total score and PI-WSUR total score was 0.59, the MOCI-cleaning correlated 0.54 with the PI-WSUR contamination subscale, and the checking subscales correlated 0.60 (all p-values described for correlations between MOCI and PI-WSUR were <0.001). For both total and subscale scores, the OCI-R demonstrated excellent convergent validity with the PI-WSUR, and moderate convergent validity with the MOCI; the PI-WSUR also showed only moderate convergent validity with the MOCI.

3.2.4. Divergent validity

The divergent validity of the OCI-R was assessed by correlating the OCI-R with measures of pathological worry and depression. The results of the current study indicate that although the means of the OCI-R total and subscale scores are influenced by the order of presentation, the correlation with the PSWQ remained nearly the same ($r = 0.42$; see Table 1). Again, the correlation between the OCI-R and PSWQ was significantly lower than both the correlation between the OCI-R and MOCI ($z = 3.41$, $p < 0.001$) and between the OCI-R and PI-WSUR ($z = 5.48$, $p < 0.001$). Controlling for PSWQ, the partial correlation between OCI-R and MOCI ($r = 0.59$, $p < 0.001$) and between the OCI-R and PI-WSUR ($r = 0.72$, $p < 0.001$) further support the notion that the OCI-R measures a construct distinct from pathological worry.

In this study, we extended the examination of the divergent validity of the OCI-R by correlating it with a measure of depression. A moderate correlation between the OCI-R and the BDI ($r = 0.39$) was found, however, this correlation was significantly lower than the correlation between the OCI-R and MOCI ($z = 3.79$, $p < 0.001$) and between the OCI-R and the PI-WSUR ($z = 5.86$, $p < 0.001$); controlling for BDI, the partial correlations between OCI-R and MOCI ($r = 0.60$, $p < 0.001$) and between OCI-R and PI-WSUR ($r = 0.71$, $p < 0.001$) indicate that the

OCI-R is measuring a concept distinct from depression. The correlations between the BDI and the OCI-R total and subscale scores are presented in Table 1.

4. Discussion

Overall, the OCI-R appears to have excellent psychometric properties. Foa et al. (2002) demonstrated this within a large clinical sample and within a small sample of nonanxious controls. The present study extends these findings by demonstrating that the OCI-R continues to have good reliability and a solid factor structure in large samples of nonanxious college students. The OCI-R has good convergent and divergent validity with other measures related to anxiety and depression.

Although the test–retest reliabilities reported in the current study were somewhat lower than those reported by Foa et al. (2002), the OCI-R total and subscale scores demonstrated good test–retest reliability across a 4-week period. This discrepancy could be due to the fact that the test–retest reliabilities reported by Foa et al. (2002) were calculated for a 1-week interval between administrations. Convergent validity was moderate with the MOCI, but excellent with the PI-WSUR. Interestingly, the correlation between the MOCI and PI-WSUR was also moderate. One possible explanation for these findings is that both the OCI-R and PI-WSUR are scored on a Likert-scale, whereas the MOCI is scored with a true–false format. The lower correlations between the OCI-R/PI-WSUR and the MOCI could reflect differences in response format, or may be explained by the weaker psychometric properties of the MOCI (Emmelkamp et al., 1999; Sternberger & Burns, 1990).

In addition to excellent convergent validity, the OCI-R was found to assess a construct distinct from depression and pathological worry. Notably, the BDI correlated best with the obsessing subscale of the OCI-R, perhaps because of the relationship between depressive rumination and obsessing.

We found significant differences between our mean OCI-R scores (for both total and subscale scores) in Studies 1 and 2. To investigate whether these differences in means were accompanied by different patterns of responding, we performed an additional factor analysis and re-analyzed the internal consistency of the OCI-R total and subscales in the second study. Despite significant differences in mean scores, the factor structure and internal consistencies were nearly identical in Studies 1 and 2. Importantly, we found a similar order effect for the MOCI: means were significantly lower when the MOCI was given after the OCI-R, relative to when it preceded the OCI-R. Foa et al. (2002) reported similarly reduced means when the items of the OCI-R were carved out of the original version of the OCI compared to when the OCI-R was administered alone, though this effect was not found in clinical patients with OCD. Thus, it is possible that the questionnaire order effect in the present studies is specific to nonclinical populations. This may be due to fatigue, habituation, or to increased understanding of the severity of pathology being questioned when presented with multiple measures. Future studies that employ self-report measures of OCD symptoms in student samples should be wary of questionnaire order effects. Given these order effects, the cutscores reported from the ROC curves in Foa et al. (2002) may be overly conservative if the OCI-R is given after another measure of OCD.

Although the present study assessed the performance of the OCI-R in two large student populations, the vast majority of the students in the present study identified themselves as Caucasian. Additionally, most of the students in the present study were first year students taking introductory psychology. Thus, the subjects were homogenous in terms of age and ethnicity, and future studies should evaluate the OCI-R in more diverse community samples. Furthermore, future studies should examine the relationship between the OCI-R and PI-WSUR subscales, in both clinical and student populations to better understand the structure of OCD symptoms.

In sum, the results of the present study are particularly important considering the theoretical utility of assessing obsessive–compulsive symptoms and related constructs in large student samples. Because the OCI-R is psychometrically sound and is quick to administer and score, it is an ideal measure for use in screening and research in large student populations.

References

Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1979). *Cognitive therapy of depression*. New York: Guilford Press.

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

Burns, G. L., Keortge, S. G., Formea, G. M., & Sternberger, L. G. (1996). Revision of the Padua Inventory of Obsessive Compulsive Disorder Symptoms: Distinctions between worry, obsessions, and compulsions. *Behavior Research and Therapy*, *34*, 163–173.

Brown, T. A., Antony, M. M., & Barlow, D. H. (1992). Psychometric properties of the Penn State Worry Questionnaire in a clinical anxiety disorders sample. *Behavior Research and Therapy*, *30*, 33–37.

Emmelkamp, P. M. G., Kraaijkamp, H. J. M., & van den Hout, M. A. (1999). Assessment of obsessive-compulsive disorder. *Behavior Modification*, *23*, 269–279.

Foa, E. B., Kozak, M. J., Salkovskis, P., Coles, M. E., & Amir, N. (1998). The validation of a new Obsessive-Compulsive Disorder scale: The Obsessive-Compulsive Inventory. *Psychological Assessment*, *10*, 206–214.

Foa, E. B., Huppert, J. D., Leiberg, S., Langner, R., Kichic, R., Hajcak, G., & Salkovskis, P. (2002). The Obsessive-Compulsive Inventory: Development and validation of a short version. *Psychological Assessment*, *14*, 485–496.

Gibbs, N. A. (1996). Nonclinical populations in research on obsessive-compulsive disorder: A critical review. *Clinical Psychology Review*, *16*, 729–773.

Hodgson, R. J., & Rachman, S. (1977). Obsessional-compulsive complaints. *Behavior Research and Therapy*, *15*, 389–395.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1–55.

Meyer, T. J., Miller, M. L., Metzger, R. L., & Borkovec, T. D. (1990). Development and validation of the Penn State Worry Questionnaire. *Behavior Research and Therapy*, *28*, 487–495.

Simonds, L. M., Thorpe, S. J., & Elliott, S. A. (2000). The Obsessive-Compulsive Inventory: Psychometric properties in a nonclinical student sample. *Behavioral and Cognitive Psychotherapy*, *28*, 153–159.

Sternberger, L. G., & Burns, G. L. (1990). Compulsive Activity Checklist and the Maudsley Obsessional-Compulsive Inventory: Psychometric properties of two measures of Obsessive-Compulsive Disorder. *Behavior Therapy*, *21*, 117–127.

Wu, K. D., & Watson, D. (2002). Further investigation of the Obsessive-Compulsive Inventory: psychometric analysis in two nonclinical samples. *Journal of Anxiety Disorders*, *17*, 305–319.