Reliability and validity of the Thai version of the Yale–Brown Obsessive Compulsive Scale – Second Edition in clinical samples

Thanita Hiranyatheb
Ratana Saipanish
Manote Lotrakul
Department of Psychiatry, Faculty of Medicine, Ramathibodi Hospital, Mahidol University, Bangkok, Thailand

Purpose: The Yale–Brown Obsessive Compulsive Scale (Y-BOCS) is regarded as the most acceptable tool for measuring obsessive–compulsive disorder symptom severity. Recently, the Yale–Brown Obsessive Compulsive Scale – Second Edition (Y-BOCS-II) was developed for better measurement. The study reported here aimed to evaluate the psychometric properties of the Thai version of the Yale–Brown Obsessive Compulsive Scale – Second Edition (Y-BOCS-II-T).

Patients and methods: The original version of the Y-BOCS-II was translated into Thai, which involved forward translation, synthesis of the translation, and back translation. Modification and cross-cultural adaptation were completed accordingly. The developed Y-BOCS-II-T, together with the Hamilton Rating Scale for Depression, was administered to 41 patients who had a primary diagnosis of obsessive–compulsive disorder. The patients then completed the Pictorial Thai Quality of Life instrument and Patient Health Questionnaire. Lastly, the Global Assessment of Functioning (GAF) and the Clinical Global Impression – Severity Scale (CGI-S) of all patients were blindly rated by another experienced psychiatrist who was not the interviewer.

Results: The mean total score of the Yale–Brown Obsessive Compulsive Scale – Second Edition – Severity Scale (Y-BOCS-II-SS) and the Yale–Brown Obsessive Compulsive Scale – Second Edition – Symptom Checklist (Y-BOCS-II-SC) were 18.44 (standard deviation = 10.51) and 15.85 (standard deviation = 9.58), respectively. The Y-BOCS-II-T had satisfactory internal consistency (Cronbach’s alpha = 0.94 for the Severity Scale, and Kuder–Richardson Formula 20 = 0.90 for the Symptom Checklist). Inter-rater reliability was excellent for both the Y-BOCS-II-SS and Y-BOCS-II-SC. Factor analysis of Y-BOCS-II-SS items revealed a two-factor component associated with obsession and compulsion. The Y-BOCS-II-SS correlated highly with the CGI-S and GAF (r = 0.75 and −0.76, respectively), but the Y-BOCS-II-SC correlated moderately (r = 0.42 for CGI-S; r = −0.39 for GAF). The Y-BOCS-II-SS and Y-BOCS-II-SC slightly to moderately correlated with the Hamilton Rating Scale for Depression, Patient Health Questionnaire, and Pictorial Thai Quality of Life, which might indicate the comorbidity depression and its effect on quality of life.

Conclusion: The Y-BOCS-II-T is a psychometrically reliable and valid measure for the assessment of both severity and characteristics of obsessive–compulsive symptoms in Thai clinical samples.

Keywords: Thailand, obsessive–compulsive disorder, assessment, obsession, compulsion, depression

Introduction

The Yale–Brown Obsessive Compulsive Scale (Y-BOCS) is accepted as the gold-standard measurement tool for symptom severity of obsessive–compulsive disorder (OCD) and is widely used by clinicians. It consists of two main parts: the Severity Scale and
the Symptom Checklist. However, the instrument has some limitations, such as, the “resistance against obsessions” item had the lowest correlation with the Y-BOCS Severity Scale and weighed lightly on the Obsession Severity subscale, the insensitivity of the Y-BOCS to slight changes in the severity of symptoms, the lack of an avoidance rating scale, and the ambiguity of some items in Symptom Checklist.

The Yale–Brown Obsessive Compulsive Scale – Second Edition (Y-BOCS-II) was developed in 2010 to overcome these limitations: it replaced the “resistance against obsessions” item with the “obsession-free interval”, extended the scoring from five- to six-point, integrated avoidance into the scoring of the Severity Scale items, and modified the Symptom Checklist content and format.

As with the Y-BOCS, the Y-BOCS-II consists of two main components: the Symptom Checklist and the Severity Scale. The former is subdivided into broad categories of symptoms while the latter assesses many aspects of severity. High internal consistencies were found for the Symptom Checklist and Severity Scale total scores ($\alpha=0.91$ and 0.89, respectively), with high inter-rater reliability (intraclass correlation coefficient [ICC] =0.96) and one-week test–retest reliability (ICC =0.85). The Y-BOCS-II strongly correlated with the original Y-BOCS and other clinician ratings of OCD symptom severity, which implied good construct validity. However, it modestly correlated with the Penn State Worry Questionnaire and Inventory of Depressive Symptomatology – Self Report, suggesting limited discriminant validity. Overall, studies have suggested that the Y-BOCS-II is a reliable and valid measure.

In Thailand, there have never been any validated measurements for OCD. It is problematic for Thai psychiatrists and OCD patients to communicate with each other about the severity and characteristics of symptoms in clinical practice. Studies about OCD in Thailand have also been limited because of the lack of standard measuring tools. Hence, a measurement tool for OCD symptoms and severity in our country is needed.

Therefore, the main objective of the study reported here was to evaluate the psychometric properties of the Thai version of the Yale–Brown Obsessive Compulsive Scale – Second Edition (Y-BOCS-II-T), including its reliability, validity, factor structure, and correlation between subscales.

Subjects and methods

Subjects and procedure

Ninety-eight patients diagnosed with OCD by trained psychiatrists were recruited from the outpatient clinic of the Department of Psychiatry, Ramathibodi Hospital, Bangkok. Of these, 21 were excluded due to illiteracy, intellectual disability, severe psychosis, or physical disorders causing impaired mental function. The remaining 77 patients were invited to participate in this study, but 36 were not available. Therefore, the total number of participants in the study was 41. The mean age of participants was 38.41±16.75 years, 61% were male, 61% had received a tertiary education or higher, and 53% were single. The mean duration of participants’ OCD illness was 14.5 years (standard deviation [SD] =13.07).

The diagnosis of OCD was confirmed by the use of the Thai version of the Mini-International Neuropsychiatric Interview, which is a structured diagnostic interview instrument for Diagnostic and Statistical Manual of Mental Disorders, fourth edition (DSM-IV), disorders. A research assistant, who was a clinical psychologist, was trained to administer this instrument.

All participants were given an explanation of the purpose and methods of the study. After giving consent, all participants were interviewed by one of four interviewers: two psychiatrists and two psychiatric nurses. The interview process included assessing participants’ OCD symptoms using the Y-BOCS-II-T and severity of depression using the Thai version of the Hamilton Rating Scale for Depression (HAM-D). The Global Assessment of Functioning (GAF) and Clinical Global Impression – severity scale (CGI-S) were assessed by an experienced psychiatrist (ML) who was not aware of the scores of the other measures. The Patient Health Questionnaire (PHQ-9) and Pictorial Thai Quality of Life (PTQL) self-report measurements were also administered.

All interviewers were trained to use the Y-BOCS-II-T and the Thai version of the HAM-D by attending didactic seminars, observing videos of patients being interviewed by an experienced psychiatrist, and practicing the administration of the measurement tool with each other. Each interviewer had to record at least five of their interviews with patients, and ten recordings were randomly selected to assess inter-rater reliability.

The study was approved by the Ethics Committee of the Faculty of Medicine, Ramathibodi Hospital, Mahidol University, Bangkok.

Measures

Y-BOCS-II

The Y-BOCS-II consists of two main components. In the 67-item Symptom Checklist, the Y-BOCS-II-SC, 29 items include obsessions, 29 items include compulsions, and nine items include avoidance. Ten items in the Severity Scale, the Y-BOCS-II-SS, assess time spent on obses-
sions and compulsions; obsession-free interval; resistant to compulsion; control over obsessions and compulsions; distress associated with obsessions; distress if compulsions prevented; and interference from obsessions and compulsions, which can be scored between 0–5 with upper limit of total score of 50.1

After obtaining permission from the copyright holder, the original version of the Y-BOCS-II was translated into Thai independently by three of the authors, with consensus between the translators achieved through forward translation and synthesis of the translation. The back translation was done by a native English-speaking English-language teacher. Finally, modifications and cross-cultural adaptations were completed accordingly.

GAF
The GAF is a numeric scale (from 0 to 100) used by mental-health clinicians to rate subjectively the social, occupational, and psychological functioning of adults – for example, how well or adaptively a patient is dealing with problems in life. The clinician assesses either the symptom severity or the level of functioning, whichever is the worse of the two.9 The higher score the patient gets, the less symptoms or the higher function they have.

CGI-S
The CGI-S is commonly used to measure the symptom severity of patients with mental disorders. It is a seven-point scale on which 1 = normal, not at all ill; 2 = borderline mentally ill; 3 = mildly ill; 4 = moderately ill; 5 = markedly ill; 6 = severely ill; and 7 = extremely ill. This rating is based on observed and reported symptoms, behavior, and function in the past 7 days. Clearly, symptoms and behavior can fluctuate over a week; the score should reflect the average severity level across the 7 days.10

PTQL
The PTQL is a self-report tool used to measure mental illness in Thais both in a clinical and a community setting. It consists of 25 items divided into six domains: Physical, Cognitive, Affective, Social Function, Economic, and Self-Esteem. All items have sufficient discriminant power. It was found to be statistically significantly different (P = 0.001) between those people with mental disorders and normal people. It also demonstrated a high level of concurrent validity association with the World Health Organization Quality of Life-BREF (Pearson correlation coefficient = 0.92). The reliability coefficient for the alpha coefficients of the PTQL was 0.88.12

HAM-D
The HAM-D is a well-accepted research tool for measuring the severity of depression and response to treatment. It is a semi-structured interview measure widely used by clinicians. The Thai version of the HAM-D has good internal consistency (alpha coefficient = 0.74) and its concurrent validity, as compared with the Global Assessment Scale, is also satisfactory (Spearman’s correlation coefficient = −0.82).8

PHQ-9
The PHQ-9 is a self-report measure, consisting of nine questions based on the DSM-IV criteria for a major depressive episode. It refers to symptoms experienced by patients during the 2 weeks prior to answering the questionnaire. The Thai version of the PHQ-9 has satisfactory internal consistency (Cronbach’s alpha = 0.79) and moderate convergent validity compared with the HAM-D (r = 0.56; P < 0.001).11

Data analysis
The collected data were analyzed using SPSS software (v 17.0; IBM Corporation, Armonk, NY, USA). Descriptive statistics were used to examine the scoring of the Y-BOCS-II-T-SC, Y-BOCS-II-T-SS, PTQL, HAM-D, PHQ-9, GAF, and CGI-S, which were reported as mean ± SD. Internal consistency for the Symptom Checklist was calculated with the Kuder–Richardson Formula 20,18 and for the Severity Scale with Cronbach’s alpha coefficient.14 The acceptable alpha values ranged from 0.70 to 0.95.11 ICCs were calculated for testing the inter-rater reliability. An ICC of 0.60 or above was considered acceptable for adequate reliability.16 Zero-order correlation17 was computed to determine the correlation between Y-BOCS-II-SS and Y-BOCS-II-SC because of the effect of age on the correlation between these two measures, while Pearson’s correlation was used in determining convergent and divergent validity among the other measures.

The factor structure of the Y-BOCS-II-T was first examined with a confirmatory factor analysis to verify two models of two-factor solutions. The first model conformed with the Obsession and Compulsion factors, and the second conformed with the Interference/Severity and Resistance/Control factors.5,18,19 Following Storch et al,5 several statistical methods were used in confirmatory factor analyses, including chi-square, the goodness-of-fit index, the normed fit index, the comparative fit index, the root mean-square residual (RMR), and the root mean-square error of approximation (RMSEA). It was decided that a nonsignificant chi-square value would indicate a good fit between a model and the data, a normed fit index value of above 0.95 would be considered a good
model fit, goodness-of-fit index and comparative fit index values of above 0.90 would be considered an indication of an adequate fit, and an RMSEA of less than 0.06 and RMR of less than 0.08 would be accepted for model fit.\(^20\)

Subsequent exploratory factor analysis (EFA) was undertaken with principal-axis factoring and promax rotation. According to principal-axis factoring, the first factor was extracted and the highest square correlation between the variable and the factor was examined.\(^21\) The remaining factors were also extracted using residual matrices. Promax rotation was used as an oblique method of rotation to allow for any relationship between the factors to be identified.\(^3\) The determination of factors based on the basis of eigenvalues (>1.00) and scree plot. Items with a pattern matrix value 0.4 or greater were considered as loading on a factor.

**Results**

Using the Y-BOCS-II-T, the Y-BOCS-II-SS average total score was found to be 18.44 (SD=10.51). The summation of positively endorsed items in the Y-BOCS-II-SC total scores ranged from 1 to 39 (mean=15.85, SD=9.58). The age of participants correlated significantly with the Y-BOCS-II-SS and Y-BOCS-II-SC total scores (\(r=-0.46\) and \(-0.53\), respectively).

According to the Mini-International Neuropsychiatric Interview, 27 patients had a comorbid diagnosis, with the most prevalent being lifetime psychosis (eight patients), followed by major depressive episode—melancholia (seven patients), generalized anxiety disorder, and anorexia nervosa binge eating/purging (six patients).

The mean scores of HAM-D and PHQ-9 in the participants with OCD were 6.03±6.27 and 7.42±5.97, respectively. The results of the depression scales implied normal to mild depression in the participants. The mean score of PTQL was 34.71±11.10, which demonstrates a moderate quality of life. The mean scores of GAF and CGI-S were 62.24±9.36 and 3.59±0.99, respectively.

**Reliability**

**Internal consistency**

The internal consistency of the Y-BOCS-II-SS was assessed with Cronbach’s alpha. The internal consistencies for the total, Obsession subscale, and Compulsion subscale scores of the Y-BOCS-II-SS were 0.94, 0.92, and 0.90, respectively.

The internal consistency of the Y-BOCS-II-SC was calculated using the Kuder–Richardson Formula 20 formula. The results revealed a high internal consistency for the total, Obsession subscale, and Compulsion subscale scores but a moderate level in the score for the Avoidance subscale (0.90, 0.80, 0.84, and 0.57, respectively).

**Inter-rater reliability**

By using ICCs, the results demonstrated excellent correlation (\(r=0.99\)) between four raters for the Y-BOCS-II-SS. The same result was found in the Y-BOCS-II-SC (\(r=0.99\)).

**Factor structure of the Y-BOCS-II-T**

**Confirmatory factor analysis**

First, the two-factor structure containing the Obsession and Compulsion factors was examined as “Model 1”. Items 1–5 were indicated to load on the Obsession factor and Items 6–10 were indicated to load on the Compulsion factor. Results revealed this two-factor model was a poor fit to meet established guidelines (Table 1); also, the Obsession and Compulsion factors were highly correlated (\(r=0.91\)). Factor loadings for the individual items, presented in Table 2, indicated that the strongest factor loadings were for Items 4 (distress associated with obsessions), 10 (interference from compulsions), 5 (interference from obsessions), 9 (distress if compulsions prevented), 8 (control over compulsions), and 3 (control over obsession).

Further, the two-factor solution that consisted of Interference/Severity and Resistance/Control factors (Model 2) was tested. The Interference/Severity factor consisted of Items 1–3 and 6–8, whereas the Resistance/Control factor consisted of Items 4, 5, 9, and 10. Once more, the two-factor model was repeatedly a poor fit (Table 1). The two factors were very strongly correlated (\(r=0.95\)). Factor loadings for the individual items on the Resistance/Control factor with the strongest factor loadings were, again, for Items 4 (distress associated with obsessions), 10 (interference from compulsions), 5 (interference from obsessions), 9 (distress if compulsions prevented), 8 (control over compulsions), and 3 (control over obsession).

**EFA**

Due to there not being good model fit, as determined by confirmatory factor analysis, EFA was undertaken.

**Table 1 Relative fit of confirmatory factor-analysis models**

<table>
<thead>
<tr>
<th>Model</th>
<th>(\chi^2)</th>
<th>df</th>
<th>(P)</th>
<th>GFI</th>
<th>NFI</th>
<th>CFI</th>
<th>RMR</th>
<th>RMSEA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>62.8</td>
<td>34</td>
<td>&lt;0.001</td>
<td>0.77</td>
<td>0.84</td>
<td>0.92</td>
<td>0.11</td>
<td>0.15</td>
</tr>
<tr>
<td>2</td>
<td>75.0</td>
<td>34</td>
<td>&lt;0.001</td>
<td>0.72</td>
<td>0.81</td>
<td>0.88</td>
<td>0.12</td>
<td>0.17</td>
</tr>
</tbody>
</table>

**Notes:** Model 1 is Obsession (Items 1–5) and Compulsion (Items 6–10); Model 2 is Interference/Severity (Items 1–3, 6–8) and Resistance/Control (Items 4, 5, 9, and 10).

**Abbreviations:** CFI, comparative fit index; \(df\), degrees of freedom; GFI, goodness-of-fit index; NFI, normed fit index; RMR, root mean-square residual; RMSEA, root mean-square error of approximation.
Principal-axis factoring with a promax rotation on the Y-BOCS-II-SS yielded two factors with eigenvalues greater than 1 (6.66 and 1.12). The two factors were classified as Y-BOCS-II Compulsion (Factor 1) and Obsession (Factor 2). These two factors accounted for 71.21% of the variance in the Y-BOCS-II-SS. However, interference from Obsession items demonstrated a dual loading on both the Obsession and the Compulsion factors. The pattern matrix for each item on these two factors is presented in Table 3.

### Correlations between the Y-BOCS-II-SS items

Table 4 shows the correlations between each item and the Y-BOCS-II-SS total score. Most of the items correlated well with the total score except for resistance to compulsions. The mean scores for all items and the values of alpha (if item deleted) are also demonstrated.

### Construct validity

Correlations between study measures are presented in Table 5. The total score of the Y-BOCS-II-SS correlated well with the CGI-S and GAF, which implies strong convergent validity. Divergent validity was examined via correlations between the Y-BOCS-II-SS scores and measures of depression and quality of life. The results revealed low correlations between the total Y-BOCS-II-SS and HAM-D scores, with a higher correlation between the obsession subscale and HAM-D than the compulsion subscale and HAM-D. The correlation between the total Y-BOCS-II-SS and PHQ-9 was moderate, and the correlation between the PHQ-9 and obsession subscale was higher than the PHQ-9 and compulsion subscale. The total Y-BOCS-II-SS score and PTQL were moderately correlated.

With the Y-BOCS-II-SC, the correlation between the Y-BOCS-II-SC and both the CGI-S and GAF was moderate. The Y-BOCS-II-SC moderately correlated with the HAM-D and PHQ-9, and slightly correlated with the PTQL.

### Discussion

The aim of study reported here was to evaluate the psychometric properties of the Y-BOCS-II-T. It demonstrated excellent internal consistencies for the total, Obsession subscale, and Compulsion subscale scores of the Y-BOCS-II-SS, which are similar to those obtained in the original paper. Confirmatory factor analysis with two models of two-factor solutions yielded a poor fit to the data, with several statistical tests not fitting the established guideline. The results of the analyses of these two models showed the same pattern as in past studies. However, the result of the EFA of the Y-BOCS-II-SS revealed two-factor components associated...
with Obsession and Compulsion, except for interference from obsessions that loaded on both factors. This relationship pattern may come from the close relationship between obsession and compulsion in terms of interference with daily living. The correlations between each item and the Y-BOCS-II-SS total score mostly related very well.

Similarly, high internal consistencies were found in the total, Obsession subscale, and Compulsion subscale scores of the Y-BOCS-II-SC. An exception was found with the Avoidance subscale, which was similar in Storch et al’s study, where moderate internal consistency was found. This result may be due to the heterogeneity of item content or domains avoided by participants.

Inter-rater reliability was excellent, as demonstrated by the high ICCs for both the Y-BOCS-II-SS and Y-BOCS-II-SC (r=0.99); this is the same pattern reported in the original Y-BOCS-SS and Y-BOCS-SC research. Convergent and divergent validity were also studied. The Y-BOCS-II-SS total score, not surprisingly, correlated well with the CGI-S, which rates the severity of mental illness. This association was stronger than the correlations between the Y-BOCS-II-SS and other measures of depressive symptoms. Moreover, this pattern of correlation is similar to that found in the development and psychometric evaluation studies of both the Y-BOCS and Y-BOCS-II. The Y-BOCS-II-SS total score had an inverse correlation with the total score of GAF, which demonstrated the higher the score of severity, the lower the score of functioning. This pattern of correlation supported the good divergent validity of the Y-BOCS-II-SS. In addition, the Y-BOCS-II-SS total had low to moderate correlation with depression scales such as the HAM-D and PHQ-9, which might be a reflection of the comorbidity of depression in OCD patients. In addition, the moderate correlation between the total Y-BOCS-II-SS and PTQL might imply the effect of OCD symptom severity on patients’ quality of life.

The Y-BOCS-II-SC total moderately correlated with the Y-BOCS-II-SS total, which demonstrates the incongruence between the severity and number of symptoms. The correlation between the Y-BOCS-II-SC total and both the CGI-S and GAF was moderate, with a slight correlation between the Y-BOCS-II-SC total and PTQL. The results indicate that severity of symptoms has a stronger impact on functioning and quality of life than the number of symptoms.

### Study limitations

This study had some limitations. First, the limited number of OCD participants was a major setback to doing the factor analysis of the Y-BOCS-II-SC, so the findings of this are not reported in this paper. However, Y-BOCS-II-SC could be used for factor analysis in the future. Second, the tools used for testing convergent validity were limited for OCD severity, since, to the best of our knowledge, this is the first study on OCD measurement. We had only the CGI-S and GAF as tools to find the convergent validity for the Y-BOCS-II-SS and none for the Y-BOCS-II-SC. Third, the test–retest reliability and treatment sensitivity of the Y-BOCS-II were not examined in this study. Test–retest reliability is widely used to determine the stability of a scale over time. It needs to be assessed at a separate time, usually 1 to 2 weeks later. Unfortunately, most of the participants’ follow-up appointments at the hospital occurred 1 month or more following their completion of the Y-BOCS-II-T, which is too long a period after which to assess test–retest reliability.
Conclusion
The Y-BOCS-II-T proved to be a reliable and valid measure for the assessment of both severity and characteristics of obsessive–compulsive symptoms, indicating that this could be a very useful measure for clinicians in Thailand to assess OCD symptoms and their severity. Further studies may look at factor analysis of the Y-BOCS-II-SC.

Acknowledgments
This study was supported by a grant from the Faculty of Medicine, Ramathibodi Hospital, Mahidol University, Bangkok, Thailand. The authors are grateful to Rungthip Prasertchai, Wanwisa Ketkaew, and Sudawan Jullagate for their assistance in data collection. They also wish to thank Umaporn Udomsubpayakul for her assistance with statistical methodology.

Author contributions
TH and RS conceived the study, collected and analyzed the data, prepared the manuscript. ML participated in the study design, and assisted in data collection. All authors took part in drafting the article or revising it critically for important intellectual content. All authors read and approved the final version of the manuscript for publication.

Disclosure
The authors report no conflicts of interest in this work.

References