Psychometric Properties of Indonesian Version of The Oxford Happiness Questionnaire

Etti Rahmawati¹, Juliana Irnayanti Saragih², Nuovi Adeline³

¹Faculty of Psychology, University of Sumatera Utara, Indonesia
etty.rahmawati7@gmail.com
²Faculty of Psychology, University of Sumatera Utara, Indonesia
julie_psusu@yahoo.com
³Faculty of Psychology, University of Sumatera Utara, Indonesia
12099na@gmail.com

Abstract — Oxford Happiness Questionnaire (OHQ) is an instrument constructed by Peter Hills and Argyle to measure happiness. OHQ is considered a good instrument, has been translated into several languages and has been used widely in several countries for researchs. The objective of this study is to examine the psychometric properties of the Indonesian version of OHQ. The empirical finding is expected to further validate the Classical Test Theory approach used before. Evaluation of psychometric properties was analyzed using the Item Response Theory (IRT) with Graded Response Model (GRM). A total of 450 people with the age of 18 to 40 have participated in the study. The results have showed that the items in Indonesian version of OHQ have reasonably high discriminating power and are accurate in measuring happiness. As a result, it can be stated that the Indonesian version of OHQ has good psychometric properties and can be used as a reliable measurement tool for evaluating individual happiness in Indonesia for research purpose.

Keywords — Oxford Happiness Questionnaire, Happiness, Psychometric Properties, IRT, GRM.

I. INTRODUCTION

Oxford Happiness Questionnaire constructed by Michael Argyle and Peter Hills has been used widely in the research setting. OHQ has been adapted into several languages. Psychometric properties that have been reported showed that OHQ have an excellent measurement results reliability as reported by Hills and Argyle [1] and Baqer [2], amounted to 0.91. Doğan & Totan [3] on the OHQ which has been translated into Turkish amounted to 0.92, and amounted to 0.84 in Brazilian language version by Dama'sio, Zanon, & Koller [4]. Estimation of reliability coefficient that was reported was estimated using internal consistency approach.

The validity testing that has been conducted reported that the OHQ has unidimensional factor structure [1],[3],[4]. Criteria based validation proved that the OHQ positively correlated with the extraversion personality dimensions and negatively correlated with neuroticism [1], [5]–[7]. In addition, the OHQ score positively correlated with agreeableness, conscientiousness and openness [6].

This research aimed to examine the psychometric characteristics of OHQ which has been translated into Indonesian (OHQ-I). The results of empirical testing is expected to complement the results of the validation using classical test theory approach that have been conducted previously. Psychometrics properties testing conducted using item response theory approach with graded response models (GRM). GRM is a model that is suitable for multilevel response categories such as the Likert scale [8], [9]. In using GRM, items allowed to have a different amount of response categories because it will not affect the parameter estimation as well as the parameters interpretation.

II. METHOD

Four hundred and fifty subjects, aged 18 to 40 (mean = 25.21, SD = 5.66) years, Male = 193, Female = 257) were selected by nonrandom sampling. Unidimensionality assumption testing is very important in the use of IRT. Unidimensionality data testing in this research was conducted by principal component analysis. Goodness of fit test to determine
a model that can describe the actual data were conducted by comparing the value of likelihood ratio test (LRT) in the constrained model (making the value of item discriminating power parameters same for all item) and unconstrained model (let the value of item discriminating power parameters different for each item). If the LRT difference is significant, then the model with more parameters were more capable to explain the data. GRM was used for item parameter estimation.

III. RESULT

A. Exploratory Factor Analysis (EFA)

The analysis showed that the sample used was adequate to conduct EFA, and EFA was appropriate to be used as an analytical tool indicated by the value of KMO above 0.50 and the value of Barlett's test of Sphericity < 0.05. EFA showed that the amount of total variance that can be explained by factor 1 amounted to 32.74 % bigger than the 20% limit (Reckase, 1979; Deng, Wells & Hambleton, 2008) to be stated that unidimensional assumptions are met. Results are presented in Table 1.

<table>
<thead>
<tr>
<th>KMO</th>
<th>Value</th>
<th>Criterion</th>
<th>Information</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.90</td>
<td>&gt; 0.50</td>
<td>Accepted</td>
</tr>
<tr>
<td>Barlett's test of Sphericity</td>
<td>$\chi^2=5169.98$ (p=0.001)</td>
<td>&lt; 0.01</td>
<td>Accepted</td>
</tr>
<tr>
<td>Total Variance Explained</td>
<td>32.74%</td>
<td>&gt; 20%</td>
<td>Unidimensional</td>
</tr>
</tbody>
</table>

B. Goodness of Fit

Based on the goodness of fit results by comparing the value of LRT on the constrained GRM model with the unconstrained GRM model, it can be found that unconstrained GRM was better fit to the data, thus the analysis of item parameters will be done using this model. Goodness of fit results are presented in Table 2.

<table>
<thead>
<tr>
<th>LRT Constrained (M1)</th>
<th>LRT Unconstrained (M2)</th>
<th>M1-M2</th>
<th>p</th>
<th>Fit Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>-19576.4</td>
<td>-19383.5</td>
<td>385.69</td>
<td>&lt;0.001</td>
<td>M2 more informative</td>
</tr>
</tbody>
</table>

C. Item Parameter Estimation

The estimation results showed that the OHQ-I items discriminating power parameter ($\alpha$) amounted 0.532 until 3.608, concluded that all item has discriminating power that considered good because it has a value above 0.50. Based on the location parameter or threshold ($\beta$) it can said to be good if scattered along the continuum of the latent traits, the continuum span that is considered to represent the latent attribute is in the range of -4 to +4. Based on the estimation results, it is known that 89.66% (26 items) has a value in the range of -4 to +4 and 10.34% (3 items) also measured outside these ranges, namely item 8 (-5.161 to 4.225), 14 (-3.033 to 6.576), and 23 (-2.581 to 5.584), although these three-item measured the latent trait in a very broad range of areas, but these three-item has discriminating power that considered good. In addition, the test information value estimation indicated that the OHQ-I was able to measure happiness trait well covering 94% of the width of the range of -4 to +4 happiness latent trait with the total information in the amount of 117.74 from 124.57 on the entire range of latent trait that can be measured. If each items were viewed, it can be seen that the OHQ-I items can function well in almost the entire range of the latent trait. Details of the threshold estimation ($\beta$), item discriminating power, item information, and test item information are presented in Table 3.
Thus, it can be concluded that the OHQ-I items are accurate and have a good discriminating power to measure happiness in almost the entire range of happiness latent trait. Thus, it can be concluded that the OHQ-I can be used as a measuring tool to evaluate individual happiness in Indonesia.

**REFERENCES**


