Rasch analysis of the Neck Bournemouth Questionnaire: Turkish version, validity and reliability study

Abstract

Background/aim: Multidimensional evaluation of patients with chronic neck pain is important for planning the treatment program. The aim of this study was to investigate validity and reliability of the Turkish version of the Neck Bournemouth Questionnaire (NBQ).

Materials and methods: The internal construct validity of the NBQ was examined by the fit of the data to the Rasch measurement model. External validity of the NBQ was evaluated by testing for expected associations of Rasch transformed NBQ score with the corresponding variables through the process of convergent validity. The reliability of the NBQ in terms of both internal consistency and test-retest reliability was assessed by Person Separation Index (PSI) and differential item functioning (DIF) by time effect.

Results: It was determined that the questionnaire has 2 factors. None of the items of Factor 1 (F1) and Factor 2 (F2) showed DIF. The reliability of F1 (Cronbach’s Alpha=0.89, PSI=0.87) and F2 (Cronbach’s Alpha=0.77, PSI=0.87) was good with Cronbach’s Alpha and PSI. There was good correlation between NBQ/F1 and (Neck Disability Index) NDI (r=0.673), (Neck Pain and Disability Scale) NPDS (r=0.709). Also, there was a correlation between NBQ/F2 and Beck Depression Inventory (BDI) (r=0.552), Beck Anxiety Inventory (BAI) (r=0.410).

Conclusion: The Turkish version of the Neck Bournemouth Questionnaire is valid and reliable.

Key Words: Health related quality of life, disability, chronic neck pain
1. Introduction

Neck pain is a health problem that nearly half of all individuals in the world experience at least once in their lifetime [1]. It is known that physical, behavioral and mental health is adversely affected, the level of disability increases and the level of health-related quality of life decreases significantly in patients with neck pain[2, 3]. The appropriate use of outcome measurements is very important to determine the most effective treatment program depending on the evaluation. While the measurement questionnaires designed to assess pain, disability and quality of life for patients with low back pain are relatively high, they are limited for patients with neck pain. For these reasons, the Neck Bournemouth Questionnaire (NBQ) was adapted by Bolton and Humphreys in 2002 from the Bournemouth Questionnaire, which was developed for low back pain [4]. The Turkish version of the Bournemouth Questionnaire which developed for low back pain patients has been shown to be valid and reliable [5].

NBQ consists of 7 questions that question the pain intensity, daily life activities, social activities, anxiety, emotional aspects of depression, kinesiophobia and the ability to control pain. The items in the questionnaire are specific to the patients with neck pain and each question evaluates a different parameter. In this context, Neck Pain and Disability Scale (NPDS) [6], Neck Disability Index (NDI) [7], Beck Depression Inventory (BDI) [8], Beck Anxiety Inventory (BAI) [9], and Tampa Scale for Kinesiophobia Scale (TSK)[10] were used to evaluate the validity of the questionnaire. The fact that the items of the questionnaire are short and clear, provide practicality to the researchers and clinicians during the application. NDI [11] and NPDS [12] are the most frequently used questionnaires for patients with neck pain. The advantages of NBQ over NDI and NPDS are the evaluation of anxiety, depression, and kinesiophobia, as well as NBQ being
shorter, more practical and more sensitive in measuring the time-dependent change of symptoms [13, 14]. There are five versions of this questionnaire (Italian, German, Portuguese, Dutch and French) but there is no Turkish version available [13, 15-18]. Therefore, the aim of this study is to investigate the validity and reliability of the Turkish version of NBQ.

2. Materials and Methods

The study was approved by Hacettepe University Non-invasive Clinical Research Ethics Committee (Ethics Committee Registration No: GO 17/844). Written permission was obtained from Jennifer E Bolton on 18.09.2017 for the Turkish version of the NBQ by translation and cultural adaptation.

One hundred and twenty five patients (18-65 years old) with neck pain (at least for the previous 3 months) due to mechanical or cervical disc herniation were included in the study. The participants signed the consent form to be included in the study. Individuals who were illiterate, with malignant disease, motor weakness due to herniated cervical disc, with loss of function due to disease were excluded from the study. For the test-retest study, the first and second filling of the NBQ was apart by 1-3 days. For the test-retest reliability, 43 individuals were planned to be included, but during this time, the patients who had a change in the severity of their symptoms due to additional treatment were excluded and the study was completed with 40 individuals.

The translation and cultural adaptation of NBQ were carried out according to the guideline established by Beaton et al.

Translation and cultural adaptation steps:
1st step- Translation: The translation of the NBQ was carried out by a physiotherapist and a linguist whose native language is Turkish and can speak English fluently. Two people who created the translation independently created two separate translation texts.

2nd step- Synthesis: The two translators discussed about each version and created a consensus version.

3rd step- Back Translation: The questionnaire was translated back to English by two linguists whose native language was Turkish and could speak Turkish fluently.

4th step- Expert Committee Review: 5 physiotherapists with at least two years of experience in the field of low back-neck health and two native English speakers created an expert committee. The physiotherapists included in the committee have validity and reliability studies in this field. The expert committee evaluated the translations in terms of cultural adaptation and conformity, and formed the pre-final version of the questionnaire.

5th step- Pretesting: The last version of the questionnaire was applied as a pilot to 35 patients and it was determined by the expert committee that the questionnaire was understandable.

For the external validity of NBQ, questionnaires that are known to be valid and reliable in patients with neck pain were used. NDI was developed to assess the level of disability in patients with neck pain. NDI consist of 10 questions in total. The subsections of NDI are designed to assess pain, personal care, lifting, reading, headaches, concentration, work, driving, sleeping, and recreation [7]. NPDS assesses factors such as pain severity, participation in social life, sleep, mood, driving, stiffness in the neck. NPDS consists of 20 questions in total [6]. BDI was prepared to measure the individual's behavior and thoughts specific to depression. In the BAI, the questions assess the tendency to anxiety.
BAI and BDI consist of 21 questions scored between 0-3 [8, 9]. TSK includes injury / re-injury and fear-avoidance parameters in work-related activities. TSK consists of 17 questions [10].

**2.1. Statistical Analysis**

**Validity**

Factor analysis was performed to assess the unidimensionality of the NBQ prior to continue with Rasch modelling. The internal construct validity of the NBQ was examined by fit of the data to the Rasch measurement model [19], while the external validity of the NBQ was assessed by testing for expected associations of Rasch transformed NBQ score with the corresponding variables through the process of convergent validity. The Rasch analysis includes the sequential steps [20] of (i) *Rescoring of NBQ items showing disordered thresholds*, (ii) *After deletion of the misfitting items, analysis for overall model and individual item fit*, (iii) *Examination for differential item functioning (DIF) for gender, age (<44/>44), body mass index (BMI<=26/>26) and duration of pain (DP<=48/>48)* and (iv) *Test for local independency and unidimensionality*. In terms of external validity, the association of Rasch transformed NBQ score with NDI, NPDS, BDI, BAI, TSK was analyzed by Spearman’s correlation coefficient.

**Reliability**

The reliability of the NBQ in terms of both internal consistency and test-retest reliability was examined by Person Separation Index (PSI) and DIF by time effect. PSI [21] which is equivalent to Cronbach's alpha [22] but has the linear transformation from the Rasch model, is a measure of internal consistency. Minimum Cronbach’s alpha values of 0.7 and 0.90 are suggested for group and individual use, respectively. For test-retest reliability of NBQ, DIF was carried out to verify the invariance of item difficulty
hierarchy across the first and the second assessment (DIF by time). Data were analysed using RUMM2020 [23].

3. Results

A total of 125 chronic neck pain (CNP) patients, 102 female and 23 male, were included in the study. The gender-based distribution of the socio-demographic characteristics of all the participants is shown in Table 1. Bartlett’s test was 420,929 and the KMO was 0.83. According to the KMO and Barlett Test results, the number of samples is sufficient for factor analysis and also the sample of our study is suitable for factor analysis. As a result of the factor analysis, two factors with eigenvalues >1 explained 67.2% of total variance. Considering the content of the questionnaire, it was decided to interpret it with two factors. When the rotated factor loads of the items were examined, Items 1, 2, 3, 6, 7 were found to be included in Factor 1 (F1), 4 and 5 were included in Factor 2 (F2).

Interpretability of the NBQ was assessed by the percentage of respondents with incomplete questionnaires and by the percentage of respondents reporting the best or worst score (ceiling and floor effect). While none of the respondents had incomplete data for all items of NBQ, the floor (ceiling) effects were 0 (0), 0 (1.6%) and 4.8% (0.8%) for the total, F1 subscale and F2 subscale, respectively.

3.1. Rasch Analysis on NBQ/F1 (Items 1,2,3,6,7)

Starting with 5 items, two items (Item 3 and 7) displayed disordered thresholds, thus the adjacent categories were collapsed together. Following this, all items were found to fit the model (given a Bonferroni adjustment fit level of 0.01) (Table 1). Overall mean item fit residual was 0.084 (SD 1.158) and mean person fit residual was -0.451 (SD 1.193). Item-trait interaction was non-significant, supporting the invariance of items (chi-square: 5.62 (df=5), p=0.344). When DIF was tested for the variables mentioned above, none of
the items showed DIF. Also, the scale satisfied the requirements of local independence
and unidimensionality.

Insert Table 1 here.

The PSI was good (0.89) indicating the ability of the scale to differentiate between 4
groups of patients and Cronbach’s alpha was 0.87. When the test-retest was examined via
DIF by time, none of the items showed DIF.

When the targeting of the final 5-item NBQ/F1 was evaluated, the scale was well-targeted
to the patients with a mean person score of -0.032 and mean item score of 0 (Figure 1).

Insert Figure 1 here.

3.2. Rasch Analysis on NBQ/F2 (Items 4,5)

Starting with 2 items, Item 4 displayed disordered thresholds, thus necessitating the
classification of adjacent categories together. Following this, both items were found to fit
the model (given a Bonferroni adjustment fit level of 0.025) (Table 2). Overall mean item
fit residual was 0.172 (SD 0.404) and mean person fit residual was -0.626 (SD 0.879).

Item-trait interaction was non-significant, supporting the invariance of items (chi-square:
2.29 (df=2), p=0.318). When DIF was tested for the variables mentioned above, none of
the items showed DIF. Also, the scale satisfied the requirements of local independence
and unidimensionality.

Insert Table 2 here.

The PSI was good (0.77) indicating the ability of the scale to differentiate between 3
groups of patients and Cronbach’s alpha was 0.78. When the test-retest was examined via
DIF by time, none of the items showed DIF.
When the targeting of the final 2-item NBQ/F2 was evaluated, patients on average had lower “disability” levels (mean person score: -0.239) than the average difficulty of the scale items (mean item score: 0) (Figure 2).

3.3. External construct validity

When the correlations of NBQ Rasch transformed score with the NDI, NPDS, BDI, BAI, TSK were examined, there was positive correlation between NBQ/F1 and NDI (r=0.673), NPDS (r=0.709), BDI (r=0.338), BAI (r=0.405), TSK (r=0.330). Also, there was positive correlation between NBQ/F2 and NDI (r=0.359), NPDS (r=0.458), BDI (r=0.552), BAI (r=0.410), TSK (r=0.223) (Table 3).

Insert Table 3 here.

4. Discussion

As a result of our study, it was determined that the Turkish version of the Neck Bournemouth Questionnaire which is suitable for biopsychosocial model developed for patients with CNP is valid and reliable. According to the results of the factor analysis, it was found that the questionnaire was of a two-factor structure. F1 included items related to pain and function, and F2 included items related to anxiety and depression.

Each item of NBQ represents a different field. Each field represented can be affected by various parameters such as cultural characteristics, age, pain duration. Since the content of the questionnaire is so rich, it is more appropriate to determine its validity and reliability by using a modern psychometric approach, Rasch analysis. The Rasch analysis allows the total score to be converted to the linear score.

According to the factor analysis, it was determined that the two factors of NBQ explain the total variance better and when the content is examined, it has been determined that
the questionnaire has two factors. In the original article of the questionnaire published by Bolton et al. in 2002, the questionnaire was interpreted with a single factor structure [4]. In the Italian version of the NBQ, published in 2014 by Geri et al., it was found that the questionnaire had a two-factor structure for the first time. Then, according to Rasch analysis of the Italian version of NBQ, which was published by Geri et al. in 2015, it was determined that NBQ had two factors. According to this study, items 1, 2, 3, 6 and 7 included F1, 4 and 5 in F2. F1 was defined as “pain and function” and F2 was defined as “anxiety and depression” [13, 24]. Our results are consistent with the studies published in the literature in recent years, when the factor loadings of each item in the questionnaire are taken into consideration, items 1, 2, 3, 6 and 7 included F1, 4 and 5 in F2.

Items 3, 4 and 7 displayed disordered thresholds. When we examine the contents of these items, we think that they have important contributions to the NBQ and it is necessary for the protection of the biopsychosocial aspect of the question.

The internal consistency of the Turkish version of the NBQ is quite high (Cronbach’s alpha was 0.87). The PSI value (0.89) is also good. High Cronbach's alpha and PSI indicate that the variables in the study were homogeneous and the questionnaire was reproducible.

In our study, according to personal differences such as gender, age, BMI, duration of pain, no item showed DIF. This shows that the answers to the questions are not affected by these variables. In addition, the absence of time-dependent DIF also indicates that the test-retest reliability is high and the reliability does not change within a certain period of time.

External construct validity is analyzed by valid and reliable scale and questionnaires for the individuals included in the study [25]. NPDS, NDI, BDI, BAI, and TSK were used in
our study and the highest correlation level was between NPDS and NBQ. The reason why the highest relationship is between these two questionnaires is that both of the questionnaires are evaluating pain and function as well as depression. However, although the content of the two questionnaires seems similar, NBQ is richer in content in terms of anxiety and kinesiophobia [14]. The stronger relation of NDI and NPDS with F1 of the NBQ indicates that these questionnaires assess pain and function rather than anxiety and depression. In addition, we think that NBQ has a low-intermediate relationship with TSK, BAI, BDI and this is due to the fact that these questionnaires and scales are not specific to neck painful individuals.

According to Deyo et al., the ideal questionnaire is a short, practical questionnaire that minimizes the burden of data collection and analysis [26]. Based on our results, we think that NBQ is an ideal questionnaire because it is short, practical and contains understandable questions. It is known that the severity of symptoms associated with neck pain changes over time. For this reason, NBQ interrogation according to time period is very important in terms of reflecting the clinical changes.

As a result of our study, we also think that NBQ contains all the parameters needed to evaluate the quality of life and it can give an idea about the quality of life of patients with CNP.

As a result, Rasch analysis showed that the Turkish version of NBQ is valid and reliable in patients with CNP. The NBQ is practical, comprehensible, suitable for biopsychosocial model, sensitive to time-dependent changes as well as a questionnaire that provides objective results for the planning and maintenance of clinical trials.
References


10.1097/BRS.0b013e31817144e1


10.1177/0748175611400289


p. 44-9.


<table>
<thead>
<tr>
<th>Items</th>
<th>Location</th>
<th>SE</th>
<th>Individual</th>
<th>Chi-Square</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item</td>
<td></td>
<td></td>
<td>Item Fit</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Residual</td>
<td></td>
<td>Test Statistics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 1</td>
<td>-0.463</td>
<td>0.053</td>
<td>-0.333</td>
<td>0.322</td>
<td>0.571</td>
</tr>
<tr>
<td>Item 2</td>
<td>0.293</td>
<td>0.052</td>
<td>-0.505</td>
<td>0.674</td>
<td>0.412</td>
</tr>
<tr>
<td>Item 3</td>
<td>0.628</td>
<td>0.055</td>
<td>-1.246</td>
<td>3.475</td>
<td>0.062</td>
</tr>
<tr>
<td>Item 6</td>
<td>-0.132</td>
<td>0.050</td>
<td>0.841</td>
<td>0.025</td>
<td>0.874</td>
</tr>
<tr>
<td>Item 7</td>
<td>-0.326</td>
<td>0.052</td>
<td>1.664</td>
<td>1.128</td>
<td>0.288</td>
</tr>
</tbody>
</table>

SE: Standart Error
### Table 2. Fit of NBQ/F2 items to Rasch model

<table>
<thead>
<tr>
<th>Items</th>
<th>Location</th>
<th>SE</th>
<th>Individual Item Fit</th>
<th>Chi-Square Statistics</th>
<th>Test p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item 4</td>
<td>-0.011</td>
<td>0.059</td>
<td>0.457</td>
<td>1.397</td>
<td>0.237</td>
</tr>
<tr>
<td>Item 5</td>
<td>0.011</td>
<td>0.054</td>
<td>-0.114</td>
<td>0.893</td>
<td>0.345</td>
</tr>
</tbody>
</table>

SE: Standard Error
Table 3. Results of external validity

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>NBQ/F1</th>
<th>NBQ/F2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>r</td>
<td>p</td>
</tr>
<tr>
<td>NDI</td>
<td>0.673</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>NPDS</td>
<td>0.709</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>BDI</td>
<td>0.338</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>BAI</td>
<td>0.405</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>TSK</td>
<td>0.330</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

F1: Factor 1  
F2: Factor 2  
NBQ: Neck Bournemouth Questionnaire  
NDI: Neck Disability Index  
NPDS: Neck Pain and Disability Scale  
BDI: Beck Depression Inventory  
BAI: Beck Anxiety Inventory  
TSK: Tampa Scale for Kinesiophobia
Figure 1. Targetting of NBQ/F1 to patients

Figure 2. Targetting of NBQ/F2 to patients. Abbreviations: SD, Standart Deviation