Cross Cultural Adaptation, Validity, Reliability and paperResponsiveness of the Arabic Version of Northwick Park Neck Pain Questionnaire

Section A -Research

Cross Cultural Adaptation, Validity, Reliability and Responsiveness of the Arabic Version of Northwick Park Neck Pain Questionnaire

Yassmina T. El Farouk[1], Amal H. Mohammed[2], Rania R. Mohamed[3]

Abstract

Background: Neck pain can affect the quality of life; it has been found that those who work are more likely to develop neck pain. Self-administered questionnaires are becoming more often used as an assessment tool in clinical practice.

Purpose:this study was done to translate and subjected to determine the Northwick Park Neck Pain questionnaire face validity, content validity, construct validity, responsiveness, feasibility; internal consistency reliability, and test-retest reliability in neck pain patients.

Methodology: Twenty experts (2 panels) with PhD, and 135 patients with chronic neck pain participated in this study. Forward Arabic translation of the Northwick Park Neck Pain questionnaire, Making the first version of the Arabic translation; Blind back-translation of the questionnaire's initial translation, face and content validity of the pre-final Arabic questionnaire were examined and measured by scale content validity. Expert proportion of clearance, Index of content validity (CVI), Expert proportion of relevance, missed item index, Time needed to answer the questions were measured. Cronbach's coefficient alpha to measure internal consistency reliability, Intraclass Correlation Coefficient (ICC) to examine test re-test reliability and find the degree of association between NPQ and neck disability index (NDI) questionnaire, Responsiveness for Arabic version of Northwick Park neck pain questionnaire was assessed.

Results: The study revealed that the scale Index of clarity was 97.77%, Expert proportion of clearance was 97.53%, the scale CVI (S-CVI) was 100%, expert proportion of relevance was 100%, The correlation between total score of NPQ and total score of NDI questionnaire was strong positive significant correlation ($r = 0.929$, $p = 0.001$). Also, the correlation between total score of NPQ and all items of NDI was moderate to strong positive correlation ($r = 0.651$ to 0.795, $p < 0.001$) , Cronbach's alpha for the Arabic version of NPQ was 0.898 , ICC for total score was 0.961, with 95% CI (0.946-0.973) , the questionnaire items were filled by 100% in all sheets and it needed less than 3 minutes to be answered in about 87.4% of all sheets. There were no missed items, The Arabic version of NPQ questionnaire showed moderate responsiveness as the Standardized Response Mean (SRM) of total score was 0.62.

Conclusion: The translated Arabic version of Northwick Park Neck Pain Questionnaire has adequate validity, feasibility, reliability sufficient to measure symptom severity and functional status in neck pain patients**.**Also, it is found that the questionnaire has a moderate degree of responsiveness.

Keywords: Validity – Reliability – Feasibility- Responsiveness- NQP.

1. Demonstrator of Physical Therapy, Basic Science Department, Faculty of Physical Therapy, October 6 University, Cairo, Egypt**.**
2. Professor of Physical Therapy, Basic Science Department, Faculty of Physical Therapy, Cairo University, Cairo, Egypt.
3. Assistant professor of Physical Therapy, Basic Science Department, Faculty of Physical Therapy, Cairo university, Cairo, Egypt.
Corresponding author: Yassmina T. El Farouk   **Email: yasmina.tarek.pt@o6u.edu.eg**

INTRODUCTION:

Neck pain, a complex condition, is a significant issue in contemporary life. Across all age

1132

Eur. Chem. Bull. 2023,12(4), 1132-1143

Cross Cultural Adaptation, Validity, Reliability and paperResponsiveness of the Arabic Version of Northwick Park Neck Pain Questionnaire

Section A -Research

categories, women seem to experience neck pain more frequently than men do**(1).**The fourth significant disease that accounts for a person's year of incapacity is neck pain**(2).**Both the frequency of neck pain and the potentially hazardous rise in texting on mobile devices, particularly among young people, have risen in recent years **(3)**. It can make it difficult for a person to do things like sit still for long periods of time, sleep poorly, and negatively affect their quality of life. It also has a connection to missed work. **(4) (5)**.

Some of the English neck measures that have been developed included the Neck Disability Index (NDI) **(6).**the Pain Disability Index (PDI) **(7).**the Northwick Park Neck Pain Questionnaire (NPQ) **(8)**, the Copenhagen Neck Functional Disability Scale (CNFDS) **(9)**. the Neck Pain and Disability Scale (NPDS) **(10).**the Neck Bournemouth Questionnaire (NBQ)**(11),** and the Functional Rating Index (FRI)**(12).**

Validity and reliability are the two most crucial and fundamental factors to consider when assessing any measurement tool or instrument for use in qualitative research. Validity and Reliability promote transparency while lowering the chance of researcher bias introduction. **(13).**They are important concepts for modern research since they increase the accuracy of assessments and evaluations of research work. **(14)**.

A questionnaire's validity is determined by how well it captures the study concept, or alternatively, the variable it is intended to assess. A questionnaire's content validity, construct validity, criterion validity, and face validity are all evaluated as part of the validity assessment.**(15).**

The degree of consistency or stability with which a questionnaire measures the topic that it is intended to measure is referred to as the questionnaire's reliability or precision. A decrease in random error is related to an increase in a questionnaire's reliability. Test-retest estimate, alternative form reliability, split-half reliability, internal consistency reliability, and inter-observer reliability are all included in the evaluation of a questionnaire's reliability.**(15).**

One type of validity is responsiveness, which is defined as the capacity to recognize significant clinical changes. Numerous criteria are used to assess an instrument's responsiveness. The clinical change that was observed and the background noise are related in a variety of ways, per responsiveness indices. The average change score between the first assessment and the follow-up is divided by the change scores' standard deviation to determine the standardized response mean. (SRM). Another often used responsiveness metric is effect size. (Average change divided by standard deviation of initial scores). It is thought that the SRM is a more reliable predictor of responsiveness than effect size, which is more directly tied to the extent of clinical improvement because it is unaffected by sample size. **(16).**

MATERIALS AND METHODS:

1. Design of the Study
This study was a prospective observational study.

2. participants and sample size
This study was carried out in Armed Forces clinics Of Remaya and the Outpatient Clinic of the Faculty of Physical Therapy, October 6 University to test validity,reliability and responsiveness of Arabic version of NQP in patients with neck pain. The study was approved by the faculty of physical therapy ethical committee review board (No:P.T.REC/012/003836). This study followed the recommendations of **Borsa et al. (2012) (17)** and **Sousa and Rojjanasrirat, (2011) (18)** for evaluating validity,reliability and responsiveness.

All of the study's 135 patients with neck pain had previously been examined by an orthopedic doctor and had experienced persistent neck pain for more than three months, their pain level not less than 4 by visual analogue scale (VAS) , patients were of both sexeswith Arabic reading and writing competence. Patients were excluded if they had a history of neck malignancies, traumas, or acute infections Patients with rheumatological disease, fibromyalgia, or other chronic pain conditions and Patients who have had neck surgery or fractures.Two expert panels; each consisted of ten experts were participated in the study with at least 10 years of experience or a doctoral degree, they worked primarily with the Arabic people and has English and Arabic fluency.

3. Study stages:
The translation of the Northwickpark neck pain questionnaire into Arabic language was approved by the author then the process proposed by **Sousa and Rojjanasrirat, (2011) (18)** was used to translate and adapt the Northwick Park neck pain questionnaire into Arabic.

3.1- Forward translation: Two forward-translated versions of an English-language survey into Arabic were created (Arl and Ar2). Forward translation was done by two translators, both of whose native tongues are Arabic but who came from different backgrounds. One of the translators was proficient in Arabic and had knowledge of medical jargon and the tool's architecture. The second translator was skilled at interpreting cultural and linguistic nuances of Arabic.

3.2- Development of the preliminary initial translated Arabic version:
a) The researchers compared and combined the two versions (Ar l and Ar 2).
b) Correcting linguistic inconsistencies and errors in Arabic

1134

Eur. Chem. Bull. 2023,12(4), 1132-1143

Cross Cultural Adaptation, Validity, Reliability and paperResponsiveness of the Arabic Version of Northwick Park Neck Pain Questionnaire

Section A -Research

C)The initial version translation into Arabic then emerged.

3.3- Blind back-translation of the questionnaire's:

a- Two back-translated versions of the questionnaire were produced. (En1 and En2)

b- Back translation was done by two translators, however their backgrounds were different.

C-One translator had extensive knowledge of English health terms and the tool's build, while the other was familiar with the linguistic and cultural quirks of the English language.

3.4- Comparison of (En1 and En2):

• The researchers compared the back-translated versions of En1 and En2, as well as En1 and En2 with the integrated form of the questionnaire, in terms of instructions, items, response format, language, sentence structure, meaning, and relevance. They found no differences between the two, hence the initial translated Arabic translation of the questionnaire was accepted as the pre-final Arabic translation.

3.5- The pre-final version evaluation for face and content validity

Ten experts from the initial expert panel were asked to assess the clarity of each component of the instrument (face validity)**.**In order to evaluate the clarity of the instruction, the items, and the response words, dichotomous questions (clear/unclear) were utilized,If the answer was unclear, they were required to give suggestions for improving its clarity. For The content equivalence

(content-related validity) of each item on the pre-final Arabic version of the questionnaire was then rated by the second expert panel using the following scale: a) Not relevant; b) Unable to determine relevance; c) Relevant but requires modest adjustment; 4: Very concise and relevant, and offers suggestions to further its relevance (a and b considered not relevant, c and d considered relevant), To make the items clearer and more relevant, modifications were made, as recommended by the two expert panels .Then the pre-final version was designated as the final Arabic Version after passing expert face and content validity testing.

5.6- complete the psychometric evaluation of the pre-final form

The final Arabic version of the questionnaire was conducted on 135 patients with chronic neck pain, Patients filled out both Arabic versions of Northwick Park neck pain questionnaire and neck disability index to measure construct validity of the questionnaire. The assessment of administration duration and the number of missing answers per question were used to determine feasibility. The patients completed it again after one week for the Re-test process. Responsiveness of the Arabic version of Northwick Park neck pain questionnairewhich is the measurement's capacity to identify change over time was assessed. **(19).** It is normally evaluated using the Standardized Response Mean (SRM) and the Effect Size (ES). **(20).**
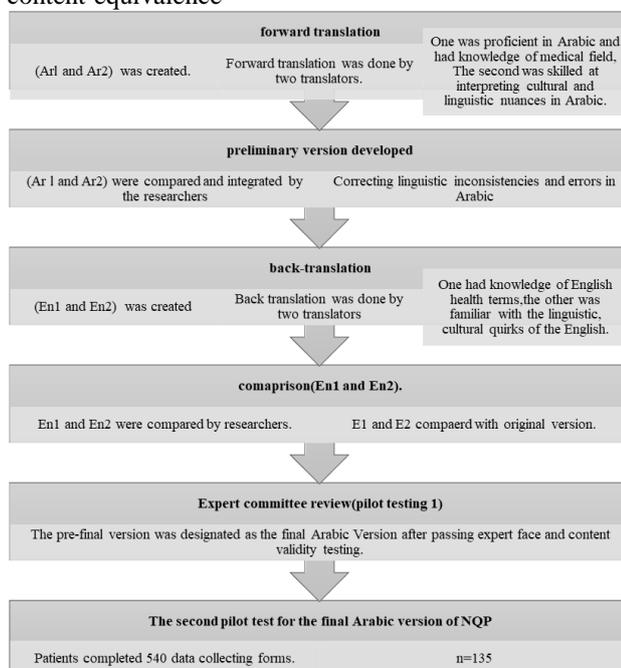

Figure (1). Flowchart of the translation process of the current study

STATISTICAL ANALYSIS:

Descriptive statistical analysis on the sample was performed using means and standard deviations for numerical data and using frequency

and percentage for categorical data. Clarity index and expert proportion of the clearance were used for face validity. Index of content validity (CVI), scale content validity indices (S-CVI) and expert

1135

Eur. Chem. Bull. 2023,12(4), 1132-1143

Cross Cultural Adaptation, Validity, Reliability and paperResponsiveness of the Arabic Version of Northwick Park Neck Pain Questionnaire

Section A -Research

proportion of relevance were used to test the content validity. Construct validity was investigated through the correlation between NPQ questionnaire and NDI questionnaire using Pearson correlation coefficients. Cronbach's alpha was used to measure the internal consistency reliability. Test–retest reliability was measured using intraclass correlation coefficient (ICC). Feasibility measured by calculation of missed item index and average time needed to answer the questionnaire. Responsiveness was assessed by using Standardized Response Mean (SRM) and Effect Size (ES) and responsiveness to change was assessed by paired t test. The level of significance for all statistical tests was set at $p < 0.05$. Statistical analysis was conducted through the statistical package for social studies (SPSS) version 25 for windows (IBM SPSS, Chicago, IL, USA).

RESULTS:

Subject characteristics

135 subjects with chronic neck pain participated in this study, their characteristics as shown in table 1.

Table 1. *General characteristics of the subjects.*

	Mean ±SD	Minimum	Maximum
Age (years)	36.56 ± 12.01	19	65
Weight(kg)	79.96 ± 13.73	45	118
Height (cm)	169.28 ± 9.55	149	193
BMI (kg/m²)	27.77 ± 3.38	19.14	36.2
Duration of illness (year)	2.53 ± 1.58	0.33	5
	N	**%**	
Sex distribution			
Females	100	74	
Males	35	26	
Diagnosis			
Disc lesion	83	61.5	
Cervical spondylosis	50	37	
Cervical rib syndrome	2	1.5	

Face validity

In order to test the face validity of the Arabic version of the NPQ questionnaire, 10 experts participated. There were 8 experts from the basic sciences department holding Ph.D. and two orthopedists. Their mean ± SD experience years of the expert panel for face validity was 14.8 ± 3.67 years with minimum of 11 years and maximum of 22 years.

The mean scale index of clarity of Arabic version of NPQ was 97.77% which is excellent. The index of clarity of Arabic version of NPQ ranged from 80% to 100%. The mean expert proportion of clearance of the Arabic version of NPQ was 97.53% which is excellent. The expert proportion of clearance ranged from 88.88% to 100%.

Content validity:

10 experts participated to examine the content validity of Arabic version of NPQ questionnaire. There were 8 experts from basic sciences department hooding Ph.D., one orthopedist with Ph.D. and one orthopedist with 17 years' experience. Their mean ± SD experience years of the expert panel for content validity was 14.9 ± 2.64 years with minimum of 11 years and maximum of 19 years.

The Arabic version of NPQ demonstrated excellent content validity, the scale CVI (S-CVI) was 100%. The mean expert proportion of relevance of Arabic version of NPQ was 100% which is excellent.

Construct validity

The correlation between total score of NPQ and total score of NDI questionnaire was strong positive significant correlation (r = 0.929, p = 0.001). Also, the correlation between total score of NPQ and all items of NDI was moderate to strong positive correlation (r = 0.651 to 0.795, p < 0.001). The correlation between sections of NPQ and the corresponding items of NDI were moderate to strong positive significant correlation (p < 0.001 (Table 2).

1136

Eur. Chem. Bull. 2023,12(4), 1132-1143

Table 2. *Correlations between NPQ questionnaire and NDI questionnaire:*

NPQ questionnaire	NDI										
	Pain Intensity	**Personal Care**	**Lifting**	**Work**	**Headaches**	**Concentration**	**Sleeping**	**Driving**	**Reading**	**Recreation**	**Total**
Pain Intensity	.892**	.511**	.386**	.593**	.527**	.439**	.483**	.642**	.497**	.615**	.697**
Pain And sleeping	.684**	.662**	.443**	.628**	.593**	.551**	.587**	.626**	.676**	.634**	.757**
Pins, Needles or Numbness In Arms At Night	.588**	.643**	.558**	.602**	.550**	.559**	.652**	.488**	.646**	.608**	.738**
Duration Of Symptoms	.625**	.493**	.387**	.617**	.686**	.513**	.635**	.602**	.564**	.662**	.732**
Carrying	.406**	.502**	.791**	.467**	.336**	.337**	.556**	.321**	.520**	.401**	.600**
Reading And Watching Tv	.544**	.589**	.454**	.754**	.581**	.610**	.584**	.504**	.480**	.641**	.711**
Working/Housework, Etc	.570**	.739**	.622**	.672**	.572**	.621**	.720**	.512**	.586**	.658**	.793**
Social Activities	.661**	.724**	.551**	.660**	.559**	.612**	.720**	.587**	.675**	.694**	.806**
Driving	.298**	.246**	0.043	.301**	.284**	.210*	.246**	.901**	.214*	.269**	.321**
Total Score	.795**	.775**	.651**	.795**	.706**	.674**	.779**	.762**	.730**	.783**	.929**

r value: Pearson correlation coefficient   p value: Probability value   **: Significant at $p < 0.001$ *: Significant at $p < 0.01$

Cross Cultural Adaptation, Validity, Reliability and paperResponsiveness of the Arabic Version of Northwick Park Neck Pain Questionnaire

Section A -Research

RELIABILITY

- Internal consistency of the Arabic version of NPQ:

Cronbach's alpha for the Arabic version of NPQ was 0.902 that means it had excellent internal consistency.

- Test-retest reliability of the Arabic version of NPQ:

The Arabic version of NPQ questionnaire showed high test-retest reliability in all sections; ICC for total score was 0.961, with 95% CI 0.946-0.973). (Table 3)

Table 3. *Test-retest reliability of Arabic version of NPQ:*

NPQ		ICC	(95% CI)		P value
			Lower bound	Upper bound	
Section1	**Pain Intensity**	0.813	0.737	0.867	0.001
Section2	**Pain and sleeping**	0.901	0.861	0.930	0.001
Section3	**Pins, Needles or Numbness in Arms at Night**	0.861	0.805	0.901	0.001
Section4	**Duration of Symptoms**	0.878	0.828	0.913	0.001
Section5	**Carrying**	0.877	0.827	0.913	0.001
Section6	**Reading and Watching TV**	0.839	0.774	0.885	0.001
Section 7	**Working/Housework, Etc**	0.876	0.825	0.911	0.001
Section8	**Social Activities**	0.938	0.913	0.956	0.001
Section9	**Driving**	0.895	0.839	0.932	0.001
	Total score	0.961	0.946	0.973	0.001

ICC, Inter class correlation coefficient value; CI, Confidence Interval; P value, Probability value

Floor and ceiling effect

A total of <10% respondents scoring "4" indicated that an item does not show significant ceiling effects. The respondents scoring "0" were from 13.3 to 30.4%. The response distributions for each item showed that all response categories were used for all items with no significant floor and ceiling effects. (Table 4).

Table 4. *Item response distribution:*

	Response category (%)				
	0	**1**	**2**	**3**	**4**
1	17.8	20.0	40.0	15.6	6.7
2	20.0	48.9	14.1	12.6	4.4
3	22.2	45.2	17.0	9.6	5.9
4	20.0	27.4	28.1	17.8	6.7
5	13.3	40.0	22.2	19.3	5.2
6	13.3	37.0	31.1	16.3	2.2
7	20.0	48.9	19.3	9.6	2.2
8	30.4	36.3	20.0	11.1	2.2
9	20.7	28.1	9.6	1.5	3.7

Feasibility:

The questionnaire needed an average2.82 ± 1.27 min to be answered with maximum of 7 min and minimum of 1 min. There were no missing items. The frequency and percentage of timetaken to answer the questions in minutes showed in table 5.

1138

Table 5. *Frequency distribution of time needed to fill the questioner in minutes:*

Time (min)	Frequency	Percent
1	61	45.2
2	29	21.5
3	28	20.7
4	12	8.9
5	1	.7
6	3	2.2
7	1	.7
Total	135	100

Responsivenessof NPQ:

Responsiveness statistics were assessed for both scores by using the standardized response mean (SRM) and the effect size (ES). The SRM is defined as the mean change in the scores between baseline and follow-up, and this change is divided by the standard deviation (SD) of the individual changes in the scores. The higher the SRM, the greater the responsiveness is **(20).**

Responsiveness was done to include the reaction of patients on items of questionnaire, so the changes between the first and second responses between the two measurements was calculated.

The Arabic version of the NPQ questionnaire was responsive to change as there was a significant difference in total score between 1st and 2nd measurements.

The Arabic version of NPQ questionnaire showed moderate responsiveness as the Standardized Response Mean (SRM) of total score was 0.62. (Table 6).

Table 6. *Responsiveness of NPQ:*

	Responsiveness of NPQ								
	Ist measurement		2nd measurement		MD	t-value	p-value	Effect size	SRM
Sections	mean	SD	mean	SD					
1	1.73	1.13	1.23	1.04	0.5	6.78	0.001	0.44	0.58
2	1.33	1.07	1.17	1.04	0.16	2.85	0.01	0.15	0.25
3	1.32	1.10	1.15	1.04	0.17	2.64	0.01	0.15	0.17
4	1.64	1.18	1.41	1.09	0.23	3.44	0.001	0.19	0.30
5	1.63	1.10	1.39	1.03	0.24	3.91	0.001	0.22	0.34
6	1.57	0.99	1.45	0.95	0.12	1.90	0.06	0.12	0.16
7	1.25	0.96	1.24	0.97	0.01	0.13	0.89	0.01	0.01
8	1.19	1.06	1.12	1.04	0.07	1.53	0.13	0.07	0.13
9	1.04	1.05	0.76	0.97	0.28	4.13	0.001	0.28	0.35
Total score	35.64	21.17	30.67	20.34	4.97	7.21	0.001	0.23	0.62

SD, Standard deviation; MD, Mean difference; t- value, Paired t test; P value, Probability value; SRM, standardized response mean.

DISCUSSION

This study was designed to translate and test face validity, content validity, feasibility, internal consistency reliability, test-retest reliability and responsiveness of Arabic language version of Northwick Park neck pain questionnaire.

1139

Cross Cultural Adaptation, Validity, Reliability and paperResponsiveness of the Arabic Version of Northwick Park Neck Pain Questionnaire

Section A -Research

The limitation of everyday activities brought on by neck pain plays a significant role in the planning of treatment. The evaluation should be based mostly on the patient's subjective experience of pain and disability because, regardless of the cause of the pain, it is difficult to make an objective assessment. It is more accurate to use a multidisciplinary approach to assess the patient's functioning in everyday life for diagnostic purposes rather than concentrating on the examination findings. For this goal, particular questionnaires have been created for the evaluation of patients and the quantitative measurement of treatment outcomes. A scale's acceptability, ease of use, high reliability, validity, and responsiveness to clinical changes are just a few of the criteria that can help determine whether it is the right scale for a certain situation. **(Kose et al., (2007) (21).** In contrast to questionnaires evaluating general health, localized pain and function questionnaires are thought to have superior validity because they are focused on a single body region. **Yilmaz et al., (2022) (22)**

Validity of the Translated Arabic Version of NQP: Face validity
The results demonstrated that the translated NQP has excellent face validity as it proved to have simple, clear, and understandable words.
According to expert's opinion, the index of clarity of all nine sections was 97.77% and the mean value of expert proportion of clarity was 97.53%.
The NPQ is simple for patients to complete, simple to score, and offers an objective measure to assess outcomes in patients with acute or chronic neck pain, according to **(Leak et al., 1994) (8).**

Content validity
Also, the study showed that the degree to which the content of the questionnaire is an adequate reflection of the construct to be measured, wasexcellent content validity as CVI (S-CVI) was 100% and the mean value of proportion of relevance was 100%.
It is proved that the items of the questionnaire are directly relevant to neck pain in patients with chronic neck pain and its associated functional disabilityassessment.
The current study's findings supported by **Polit and Beck's (2006)(23)** assertion that a scale would be considered to have excellent content validity if it contained items with item indexes of content validity (I-CVI) that met the following requirements: I-CVI of 1.00 for three to five experts, and a minimum I-CVI of.78 for six to ten experts, as well as S-CVI

of.90 or higher. If the initial evaluation indicates the need for significant item improvements, the recommended standards may require two rounds of expert review.
The S-CV1/Ave of 0.90 or above is the minimum acceptable index, and items that do not attain the minimum acceptable indices are revised and re-evaluated, according to **Waltz et al. (2005) (24).Chiu et al. (2001) (25),**found that the questionnaire has good content validity.

Construct validity
The correlation between total score of NPQ and total score of NDI questionnaire was strong positive significant correlation (r = 0.929, p = 0.001). Also, the correlation between total score of NPQ and all items of NDI was moderate to strong positive correlation (r = 0.651 to 0.795, p < 0.001). The correlation between sections of NPQ and the corresponding items of NDI were moderate to strong positive significant correlation (p < 0.001.
This was almost in agreement with **Hoving et al., (2003) (26)** assessment of the validity of the NDI and NPQ as outcome measures in whiplash injury, which found that the correlations between the two were r= 0.88.
Yilmaz et al., (2022) (22) foundthatNPQ with the NDI Spearman's correlation coefficient was calculated to be 0.648, which approximately agreed with the current study, it was concluded that there is a good correlation between the two scales.
According to **Chiu et al., (2001) (25)**, the questionnaire had good construct validity (Spearman correlation coefficient of 0.59 when the score was correlated with that of a generic 42-item Chinese health questionnaire).

Feasibility of the Translated Arabic Version of NPQ
The Arabic version of **NPQ** had high feasibility because its items were filled by 100% in all sheets and it needed less than 3 minutes to be answered in about 87.4% of all sheets.The current study's findings were consistent with those of **Van et al. (2015) (27),** who claimed that a missing rate on the item level was acceptable if no single item had a missing rate higher than 10% and that a completion time was acceptable if 95% of the sheets were finished in less than 15 minutes. This result was almost in agreement with the Spanish version of the NPQ as 90% of the patients completed the questionnaire and they took less than 10 minutes to complete it. **(González et al.,2001) (28).**

Cross Cultural Adaptation, Validity, Reliability and paperResponsiveness of the Arabic Version of Northwick Park Neck Pain Questionnaire

Section A -Research

Internal Consistency and Test-Retest Reliability of the Translated Arabic Version of NPQ:

After one week from the test, Cronbach's alpha for the Arabic version of NPQ was 0.898 (0.868- 0.931) that means it had excellent internal consistency.Therefore, **George and Mallery (2003) (29)** define good internal consistency as a value between 0.7 and 0.9.

The Arabic version of NPQ questionnaire showed high test-retest reliability in all sections; ICC for total score was 0.961, with 95% CI 0.946-0.973.
ICC for section I was 0.813, with 95% CI 0.737-0.867. ICC for section 2 was 0.901, with 95% CI 0.861-0.930. ICC for section 3 was 0.861, with 95% CI 0.805-0.901. ICC for section 4 was 0.878, with 95% CI 0.828-0.913. ICC for section 5 was 0.877, with 95% CI 0.827-0.913. ICC for section 6 was 0.839, with 95% CI 0. 0.774-0.885. ICC for section 7 was 0.876, with 95% CI 0.825-0.911. ICC for section 8 was 0.938, with 95% CI 0.913-0.956. ICC for section 9 was 0.895, with 95% CI 0.839-0.932.

This came in agreement with **Lee et al., (2010) (30)**as Internal consistency and the intraclass correlation coefficient (ICC) were used to assess reliability, as well as Cronbach's alpha. The translated versions of the NPQ had good test-retest reliability, ICC=.83. The value of Cronbach's alpha for NPQ was discovered to be.87.
Also, it came in agreement with **Aguirre et al., (2013) (31)**due to its strong internal consistency (Cronbach Coefficient Alpha of 0.86 and good test-retest reliability, ICC of 0,89).
According to **Chiu et al., (2001) (25)**, They claimed that the questionnaire showed very good internal consistency and test-retest reliability (intraclass correlation coefficient, 0.95; Cronbach's alpha, 0.87); these findings were consistent with those of the current study.
The internal consistency of the English version of the Northwick Park Neck Pain Questionnaire was almost in agreement with the results of the current study, with a Cronbach's alpha of 0.79 for the entire scale. **(Sim et al., 2006) (32).** In the validity and reliability (French version), The NPQ's ICC score was 0.84. **(Wlodyka-Demaille et al., (2002) (33)**Additionally, the ICC value for NPQ (Brazilian version) was discovered to be 0.96. **(Almeida et al. (2022) (34).**
According to a study done by**Kose et al., (2007) (21)**study examined Turkish individuals with neck pain using four disability scales (NDI, NPDS, NPQ,

and CNFDS) for comparison. They discovered that for the NDI, NPQ, CDS, and NPDS, respectively, the ICC scores were 0.86, 0.85, 0.84, and 0.81, and that the Cronbach's alpha coefficient varied from moderate to high on all four scales, with coefficients ranging from 0.80 (for the NPQ and CNFDS) to 0.94. (For NPDS). They suggested that for Turkish-speaking patients with neck pain, all scales had adequate reliability and validity.
Yilmaz et al., (2022) (22)in their study showed that The NPQ had a good internal consistency (Cronbach alfa = 0.704) and excellent test-retest reliability (Intraclass correlation coefficient = 0.995) which came in agreement with our study.

Responsiveness
Further study should be done to calculate responsiveness after physical therapy treatment.
The Arabic version of NPQ questionnaire showed moderate responsiveness as the Standardized Response Mean (SRM) of total score was 0.62. The previous results of responsiveness came in agreement with Argentinean version of NPQ as the Sensitivity to change was good (r=0,661). (**Aguirre et al., 2013) (31)**

The English version of the Northwick Park Neck Pain Questionnaire's sensitivity to change was almost in agreement with the results of the current study, with a standardized response mean of 0.71.(**Sim et al., (2006) (32).**
Northwick park neck pain questionnaire is a simple questionnaire that is easy for patients to complete, suitable , quick and easy tool for the therapist to assess neck pain and its resultant disability and gives us a percentage of the patient's level of function, it takes less than 3 minutes to complete and all items of the questionnaire is clear and easy to understand , it also provide a good follow up tool, thus NPQ is the right questionnaire for assessing pain and functional disability that is directly related to neck pain.
The final version is used as a basis for the subsequent study, which will be carried out to look at the complete psychometric characteristics of the Arabic version of the NPQ.

LIMITATIONS:
Due to the lack of treatment, there was a moderate change between the first and second measurements on calculating responsiveness to change, so further studies should be done after treatment.

Cross Cultural Adaptation, Validity, Reliability and paperResponsiveness of the Arabic Version of Northwick Park Neck Pain Questionnaire

Section A -Research

CONCLUSION:

The Arabic translation of the Northwick Park Neck Pain Questionnaire has sufficient face validity, content validity, construct validity, feasibility, internal consistency reliability, and test-retest reliability to assess the severity of symptoms and functional status in neck pain patients. The questionnaire is also discovered to be moderately responsive.

REFERENCES:

1. **Safiri, S., Kolahi, A. A., Hoy, D., Buchbinder, R., Mansournia, M. A., Bettampadi, D., Ashrafi-Asgarabad, A., Almasi-Hashiani, A., Smith, E., Sepidarkish, M., Cross, M., Qorbani, M., Moradi-Lakeh, M., Woolf, A. D., March, L., Collins, G., & Ferreira, M. L. (2020).** Global, regional, and national burden of neck pain in the general population, 1990-2017: systematic analysis of the Global Burden of Disease Study 2017. *BMJ (Clinical research ed.)*, *368*, m791. https://doi.org/10.1136/bmj.m791

2. **Al-Najjar, H. M. M., Mohammed, A. H., &Mosaad, D. M. (2020).** Effect of ice massage with integrated neuromuscular inhibition technique on pain and function in subjects with mechanical neck pain: randomized controlled trial. *Bulletin of Faculty of Physical Therapy*, *25*, 1-7.

3. **Smith E, Hoy DG, Cross M et al (2014).** The global burden of other musculoskeletal disorders: estimates from the Global Burden of Disease 2010 study. Ann Rheum Dis 73:1462–1469. **https://doi.org/10.1136/annrheumdis-2013-204680**

4. **Hogg-Johnson, S., van der Velde, G., Carroll, L. J., Holm, L. W., Cassidy, J. D., Guzman, J., ... &Peloso, P. (2008).** The burden and determinants of neck pain in the general population: results of the Bone and Joint Decade 2000–2010 Task Force on Neck Pain and Its Associated Disorders. *European Spine Journal*, *17*, 39-51.

5. **de Cássia Pereira Fernandes, R., da Silva Pataro, S. M., De Carvalho, R. B., &Burdorf, A. (2016).** The concurrence of musculoskeletal pain and associated work-related factors: a cross sectional study. *BMC Public Health*, *16*(1), 1-9.

6. **Vernon, H. (2008).** The Neck Disability Index: state-of-the-art, 1991-2008. *Journal of manipulative and physiological therapeutics*, *31*(7), 491-502.

7. **Tait, R. C., Chibnall, J. T., & Krause, S. (1990).** The pain disability index: psychometric properties. *Pain*, *40*(2), 171-182.

8. **Leak, A. M., Cooper, J., Dyer, S., Williams, K. A., Turner-Stokes, L., & Frank, A. O. (1994).** The Northwick Park Neck Pain Questionnaire, devised to measure neck pain and disability. *British journal of rheumatology*,*33*(5),469–474. https://doi.org/10.1093/rheumatology/33.5.469

9. **Jordan, A., Manniche, C., Mosdal, C., &Hindsberger, C. (1998).** The Copenhagen Neck Functional Disability Scale: a study of reliability and validity. *Journal of manipulative and physiological therapeutics*, *21*(8), 520–527.

10. **Wheeler, A. H., Goolkasian, P., Baird, A. C., & Darden, B. V. (1999).** Development of the Neck Pain and Disability Scale: item analysis, face, and criterion-related validity. *Spine*, *24*(13), 1290.

11. **Bolton, J. E., & Humphreys, B. K. (2002).** The Bournemouth Questionnaire: a short-form comprehensive outcome measure. II. Psychometric properties in neck pain patients. *Journal of manipulative and physiological therapeutics*, *25*(3), 141-148.,

12. **Feise, R. J., & Menke, J. M. (2001).** Functional rating index: a new valid and reliable instrument to measure the magnitude of clinical change in spinal conditions. *Spine*, *26*(1), 78-87

13. **Singh AS. (2014).** Conducting case study research in non-profit organisations. Qualitative Market Research: An International Journal

14. **Tavakol M, Dennick R. (2011).** Making sense of Cronbach's alpha. *Int J Med Educ*; 2: 53

15. **Galanis, P. (2013).** Validity and reliability of questionnaires in epidemiological studies. *Archives of Hellenic Medicine/ArheiaEllenikesIatrikes*, *30*(1).

16. **Joy C, Robert S, Allan D, Nicolas B and James H (2000):** Responsiveness of the short form-36, disability of the arm, shoulder, and hand questionnaire, patient-rated wrist evaluation, and physical impairment measurements in evaluating recovery after a distal radius fracture Journal of Hand Surgery. 25, 2, 330-340.

17. **Borsa, J. C., Damásio, B. F., & Bandeira, D. R. (2012).** Cross-cultural adaptation and validation of psychological instruments: Some

1142

Cross Cultural Adaptation, Validity, Reliability and paperResponsiveness of the Arabic Version of Northwick Park Neck Pain Questionnaire

Section A -Research

considerations. *Paidéia (Ribeirão Preto)*, *22*, 423-432.

18. **Sousa, V. D., &Rojjanasrirat, W. (2011).** Translation, adaptation and validation of instruments or scales for use in cross- cultural health care research: a clear and user- friendly guideline. *Journal of evaluation in clinical practice*, *17*(2), 268-274.

19. **Gummesson I, Atroshi C and Ekdah I (2003):** The Disabilities of the Arm, Shoulder and Hand (DASH) outcome questionnaire: Longitudinal construct validity and measuring self-rated health change after surgery, BMC Musculoskeletal Disorder, 4, 11.

20. **Jeon S, Lee J, Chung M , Baek G , Oh J, and Gong H (2011):** Responsiveness of the Korean version of the disabilities of the arm, shoulder and hand questionnaire (K-DASH) after carpal tunnel release. Clinics in Orthopedic Surgery, 3, 147-151. doi:10.4055/cios.2011.3.2.147.

21. **Kose, G., Hepguler, S., Atamaz, F., & Oder, G. (2007).** A comparison of four disability scales for Turkish patients with neck pain. *Journal of rehabilitation medicine*, *39*(5), 358–362. https://doi.org/10.2340/16501977-0060

22. **Yilmaz, E., Aydin, T., & Pasin, Ö. (2022).** Cross-cultural translation, validity and reliability of the Turkish version of the Northwick Park Neck Pain Questionnaire.

23. **Polit, D. F., & Beck, C. T. (2006).** The content validity index: are you sure you know what's being reported? Critique and recommendations. *Research in nursing & health*, *29*(5), 489-497.

24. **Waltz C, Strickland O, Lenz E 2005.**Measurement in Nursing and Health Research, Springer Publishing Company; New York; 3re edition.

25. **Chiu, T. T., Lam, T. H., & Hedley, A. J. (2001).** Subjective health measure used on Chinese patients with neck pain in Hong Kong. *Spine*, *26*(17), 1884–1889. https://doi.org/10.1097/00007632-200109010-00013

26. **Hoving, J. L., O'Leary, E. F., Niere, K. R., Green, S., & Buchbinder, R. (2003).** Validity of the neck disability index, Northwick Park neck pain questionnaire, and problem elicitation technique for measuring disability associated with whiplash-associated disorders. *Pain*, *102*(3), 273-281.

27. **van Veen, M. J., Birnie, E., Poeran, J., Torij, H. W., Steegers, E. A., &Bonsel, G. J. (2015).** Feasibility and reliability of a newly developed antenatal risk score card in routine care. *Midwifery*, *31*(1), 147-154.

28. **González, T., Balsa, A., Sáinz de Murieta, J., Zamorano, E., González, I., & Martin-Mola, E. (2001).** Spanish version of the Northwick Park Neck Pain Questionnaire: reliability and validity. *Clinical and experimental rheumatology*, *19*(1), 41–46

29. **George, D., & Mallery, M. (2003).** Using SPSS for Windows step by step: a simple guide and reference

30. **Lee, K. W., Seo, H. D., Jung, K. S., Kim, S. H., & Chung, Y. J. (2010).** Reliability and validity of Korean version Northwick Park Neck Pain Questionnaire in neck pain patients. *Physical Therapy Korea*, *17*(3), 68-76.

31. **Aguirre, M. V., Rodríguez, M. G., Clarett, M., Iribarne, J. I., Martínez, M., Battistotti, R., ... & Orsini, E. (2013)**. Cultural adaptation and Argentine validation of the Northwick Park Neck Pain Questionnaire in the hospitals of the Autonomous City of Buenos Aires. *Revista de la Facultad de CienciasMedicas (Cordoba, Argentina)*, *70*(2), 76-82.

32. **Sim, J., Jordan, K., Lewis, M., Hill, J., Hay, E. M., &Dziedzic, K. (2006).** Sensitivity to change and internal consistency of the Northwick Park Neck Pain Questionnaire and derivation of a minimal clinically important difference. *The Clinical journal of pain*, *22*(9), 820-826.

33. **Wlodyka-Demaille, S., Poiraudeau, S., Catanzariti, J. F., Rannou, F., Fermanian, J., & Revel, M. (2002).** French translation and validation of 3 functional disability scales for neck pain. *Archives of physical medicine and rehabilitation*, *83*(3), 376-382.

34. **Almeida, M., Dibai-Filho, A. V., Guirro, R., Guirro, E., Fonseca, M., & Avila, M. A. (2022).** Psychometric properties of the Brazilian short version of the Northwick Park Neck Pain Questionnaire. *Clinical rehabilitation*, 2692155221086198. Advance online publication. https://doi.org/10.1177/02692155221086198.