

Psychometric validation of the Spanish neck disability index and development of a shortened 5-item version for chronic neck pain

DOI: <https://doi.org/10.5114/pq/218497>

Claudio Chamorro Lange^{ID}, Hernán Andrés de la Barra Ortiz^{ID}

Exercise and Rehabilitation Sciences Institute, School of Physical Therapy, Faculty of Rehabilitation Sciences, Universidad Andres Bello, Santiago, Chile

Abstract

Introduction. Chronic neck pain (CNP) is a prevalent condition that compromises quality of life and functional performance. The neck disability index (NDI) is widely used to assess neck-related disability, yet updated psychometric evidence for Spanish-speaking populations with CNP is limited. This study aimed to validate the Spanish NDI and evaluate the psychometric performance of a shortened clinical form in adults with CNP.

Methods. A cross-sectional study was conducted in adults with CNP referred to physical therapy ($n = 123$). Participants completed the Spanish NDI. Construct validity was assessed through exploratory factor analysis (EFA) using Spearman correlations and principal components extraction assuming a one-factor solution, with a 0.40 loading criterion. Sampling adequacy was verified with the Kaiser–Meyer–Olkin index. Internal consistency was examined with Cronbach's alpha. Confirmatory factor analysis (CFA), using a robust maximum likelihood estimator, evaluated the model fit of the best-performing item set. Model fit indices included the Comparative Fit Index (CFI), Tucker–Lewis Index (TLI), Root Mean Square Error of Approximation (RMSEA), and Standardized Root Mean Square Residual (SRMR).

Results. A one-factor structure emerged in the EFA, with items S1 (pain), S2 (personal care), S3 (lifting), S4 (reading), and S10 (recreation) showing the strongest loadings. Removing low-loading items improved internal consistency ($\alpha = 0.71$ vs. 0.64 for the full scale). The CFA supported the reduced 5-item model, demonstrating satisfactory fit (CFI = 0.99, TLI = 0.97, RMSEA = 0.04, SRMR = 0.04), supporting the structural validity of the instrument for assessing disability in individuals with CNP.

Conclusions. The Spanish NDI demonstrated adequate construct validity in adults with CNP. The shortened 5-item version showed improved internal consistency and strong model fit, offering a clinically efficient alternative for rapid disability screening. Further validation in different populations and clinical contexts is recommended.

Key words: neck disability index, chronic neck pain, psychometrics, validation study, factor analysis, disability

Introduction

Chronic neck pain (CNP) is a highly prevalent musculoskeletal condition that contributes substantially to disability, reduced quality of life, and long-term healthcare utilisation worldwide [1–3]. Recent global meta-analytic evidence shows that neck pain affects about one quarter of adults, with prevalence estimates ranging from 23% in the Americas to nearly 28% in Europe, highlighting its considerable impact across regions. Neck pain also ranks among the leading causes of musculoskeletal disability globally, second only to low back pain in most epidemiological surveys [2, 3]. This substantial burden reinforces the need for valid and culturally adapted tools to assess disability in individuals with CNP accurately. Given its persistent course and functional repercussions, accurate assessment of neck-related disability is essential for guiding clinical decision-making, monitoring treatment outcomes, and advancing research in rehabilitation sciences [4–7].

Patient-reported outcome measures (PROMs) play a central role in capturing patients' perceptions of functional limitations and disability [7, 8]. Among these instruments, the neck disability index (NDI) is the most widely used and recommended PROM for evaluating neck-specific disability due to its clinical relevance, ease of application, and strong measurement properties [8–10]. Originally developed in English, the NDI has been translated into several languages, including Spanish

and Portuguese, to support its use across diverse populations [11–14]. Notably, an updated Portuguese validation in individuals with CNP demonstrated robust psychometric performance and highlighted the importance of reassessing translated versions using contemporary analytical methods [14, 15].

Although a Spanish version of the NDI exists, researchers have not re-examined its structural validity in individuals with CNP using modern psychometric techniques. Earlier Spanish adaptations relied primarily on classical test theory and did not include factor-analytic approaches that are now essential for evaluating dimensionality, item performance, and overall construct validity [11]. In parallel, evidence from other languages – including Portuguese, English, and Chinese – has identified item redundancies and supported the development of abbreviated versions that retain validity while improving feasibility in clinical practice [10–15]. Despite these advances, no validated shortened Spanish version of the NDI is currently available.

Ensuring the psychometric quality of measurement instruments requires examining both reliability and validity [16, 17]. Internal consistency, typically assessed using Cronbach's alpha, is a key indicator of reliability and reflects the degree to which items coherently measure the same underlying construct, supporting the instrument's stability across contexts [18, 19]. Construct validity can be evaluated through exploratory and confirmatory factor analysis (EFA and CFA), which

Correspondence address: Hernán Andrés de la Barra Ortiz, Exercise and Rehabilitation Sciences Institute, School of Physical Therapy, Faculty of Rehabilitation Sciences, Universidad Andres Bello, Avenida Fernández Concha 700, Las Condes, Santiago, Chile, Postal code: 7591538, e-mail: hdelabarra@unab.cl, handresdelabarra@yahoo.es; <https://orcid.org/0000-0002-3927-1743>

Received: 08.12.2025

Accepted: 24.02.2026

Citation: Lange CC, de la Barra Ortiz HA. Psychometric validation of the Spanish neck disability index and development of a shortened 5-item version for chronic neck pain. *Physiother Quart.* 2026;34(4):16–22; doi: <https://doi.org/10.5114/pq/218497>.

assess a measure's internal structure and determine whether item groupings align with theoretical expectations [20]. These approaches help identify low-performing or redundant items and improve the instrument's overall measurement precision.

Given the high prevalence of CNP and the clinical need for efficient, valid tools to assess disability, an updated psychometric evaluation of the Spanish NDI is warranted. Therefore, this study aimed to evaluate the construct validity and internal consistency of the Spanish version of the NDI in adults with CNP and to examine the psychometric performance of a shortened clinical version.

Participants and methods

Design

A cross-sectional observational study was conducted to validate the Spanish version of the NDI. The study followed the Consensus-based Standards for the Selection of Health Measurement Instruments (COSMIN) guidelines and current recommendations for evaluating the structural validity and internal consistency of patient-reported outcome measures [21, 22].

Participants

Participants were recruited from three private clinics in the eastern sector of the Santiago Metropolitan Region, Chile, between April 2023 and December 2024. Adults presenting with CNP were consecutively invited to participate. The study was conducted in accordance with the principles of the Declaration of Helsinki [21]. The study procedures and objectives were explained to all participants, and written informed consent was obtained before participation, including consent for the use of NDI-derived data for research purposes.

The sample size calculation followed COSMIN recommendations, which specify that studies assessing structural validity should include at least 100 participants to ensure stable factor solutions and adequate parameter estimation [22, 23]. COSMIN recommendations suggest a minimum of approximately 5–7 participants per item, with at least 100 participants, for factor analysis procedures, whereas other psychometric guidelines have proposed a more conservative ratio of approximately 10 participants per item to obtain stable and reliable factor solutions [15, 24]. Given that the NDI comprises 10 items, a minimum sample of 100 participants met the COSMIN threshold, while an ideal range of 100–200 participants aligned with recommended item-to-sample ratios. A convenience sampling strategy was employed, prioritising participant availability and accessibility.

Selection criteria

The inclusion criteria required participants to be adults (≥ 18 years), of either sex, Spanish speakers, and to present with clinically defined CNP, characterised by pain or discomfort localised between the superior nuchal line and the T1 spinous process or the shoulder girdle, persisting for at least three months [25]. Participants were also required to have an NDI score of ≥ 5 and to report resting neck pain of ≥ 3 on the Numeric Pain Rating Scale (NPRS) [14, 26].

Exclusion criteria included: musculoskeletal injuries of the neck or shoulders within the past three months; neurological signs such as paraesthesia, sensory loss, weakness, or discoloration in the neck, arms, forearms, or hands; autoimmune disorders (e.g., dermatomyositis, systemic lupus erythemato-

sis, hepatic porphyria, carcinoid cutaneous syndrome, pella-gra); a history of cancer or tumours within the past five years; other diagnosed chronic pain conditions; or incomplete responses to the NDI [26].

Instrument: NDI

The validated Spanish version of the NDI was administered to all participants. This version has demonstrated excellent internal consistency (Cronbach's $\alpha = 0.937$) [11] and retains the original structure of the English instrument: 10 sections rated on a 0–5 Likert scale, where higher scores indicate greater disability [12]. The sections include S1 pain intensity, S2 personal care, S3 lifting, S4 reading, S5 headaches, S6 concentration, S7 work, S8 driving, S9 sleeping, and S10 recreation. Participants completed the NDI in a self-administered format, without examiner interference, in a private, well-lit room to ensure comfort and standardisation. All participants provided informed consent before completing the questionnaire.

Statistical analysis

Statistical analyses were conducted using Stata 16.0 (StataCorp®) [27]. Prior to descriptive analyses, quantitative demographic variables were assessed for normality using the Shapiro–Wilk test and for homoscedasticity using Levene's test to determine the suitability of applying parametric statistics. Normally distributed variables are reported as means and standard deviations, whereas categorical variables are presented as frequencies and percentages. The psychometric evaluation followed the updated COSMIN guidelines for assessing structural validity and internal consistency [21, 22].

Construct validity was first examined through exploratory factor analysis (EFA). Given the ordinal nature of the NDI items, the inter-item correlation matrix was computed using Spearman's Rho. Sampling adequacy was evaluated using the Kaiser–Meyer–Olkin (KMO) index, with values ≥ 0.70 considered acceptable [14], and Bartlett's test of sphericity was used to verify the suitability of the correlation matrix for factor extraction.

Principal component extraction was applied assuming a one-factor solution, in line with the theoretical unidimensional structure of the NDI and with the aim of identifying items contributing most strongly to the underlying construct. Factor loadings, communalities, and uniqueness values were derived from the unrotated solution. Because a unidimensional structure was specified a priori, rotation (varimax rotation) was used only for item-level inspection and did not influence factor retention or interpretation. Items with factor loadings < 0.40 were considered low-performing and excluded from subsequent analyses. A scree plot was generated to visually inspect factor retention. Internal consistency of the retained items was assessed using Cronbach's α .

A unidimensional structural equation model (SEM) was estimated using a robust maximum likelihood estimator to confirm the factor structure identified in the EFA [20]. Model fit was evaluated using the chi-square test, the root mean square error of approximation (RMSEA, 90% CI), the comparative fit index (CFI), the Tucker–Lewis index (TLI), and the standardised root mean square residual (SRMR). Acceptable fit was defined as CFI and TLI ≥ 0.90 and RMSEA and SRMR ≤ 0.08 [14, 28]. The coefficient of determination (CD) was reported to quantify the proportion of variance explained by the model. A path diagram with standardised factor loadings was generated to illustrate the final SEM.

Results

Table 1 summarises the characteristics of the study participants. All individuals had been referred to physical therapy due to CNP. A total of 123 participants were enrolled, of whom 32.5% were men ($n = 40$) and 67.5% were women ($n = 83$), with a mean age of 34.9 years (± 7.3). Regarding occupational demands, 77.2% ($n = 95$) reported predominantly static work, characterised by prolonged sitting and limited movement, whereas 22.8% ($n = 28$) indicated performing dynamic tasks requiring greater physical activity. Participants also reported their daily sitting duration: 43.1% sat for 4–8 hours per day, and 35.8% sat for more than 8 hours.

After confirming eligibility criteria for CNP, participants underwent an assessment of resting pain and completed the Spanish version of the NDI. The mean resting NPRS score was 3.93 (± 1.4), and the mean NDI score was 29.8 (± 5.3). No statistically significant differences were observed between men and women in either pain intensity or disability scores.

Table 1. Demographic and clinical characteristics of the participants

Participants [<i>n</i> (%)]	
Total	123 (100)
Men	40 (32.5)
Women	83 (67.5)
Mean age (mean $\pm$ SD) ^{N,h}	
Total	34.9 $\pm$ 7.3
Men	34.7 $\pm$ 7.2
Women	34.9 $\pm$ 7.3
Posture at work [<i>n</i> (%)]	
Static	95 (77.2): women = 62, men = 33
dynamic	28 (22.8): women = 21, men = 7
Hours sitting [<i>n</i> (%)]	
0–4 hours	26 (21.1): women = 19, men = 7
4–8 hours	53 (43.1): women = 34, men = 19
> 8 hours	44 (35.8): women = 30, men = 14
NPRS score (mean $\pm$ SD) ^{N,h}	
Total	3.93 $\pm$ 1.4
Men	3.85 $\pm$ 1.2
Women	3.92 $\pm$ 1.4
NDI score (mean $\pm$ SD) ^{N,h}	
Total	29.8 $\pm$ 5.3
Men	29.7 $\pm$ 5.3
Women	29.8 $\pm$ 5.2

NDI – neck disability index, NPRS – numeric pain rating scale
^N variable met normality assumptions (Shapiro–Wilk test)
^h met homoscedasticity assumptions (Levene’s test)
No statistically significant differences were observed between genders (t -test, $p > 0.05$).

Exploratory factor analysis (EFA)

Table 2 presents the inter-item correlation matrix used for the exploratory factor analysis (EFA). The matrix was computed using Spearman’s rho to account for the ordinal nature of the NDI items. Overall, the correlations were positive, with the strongest associations observed between S1–S2, S1–S3, S1–S4, S1–S10, S2–S3, S3–S4, S3–S10, and S4–S10. These patterns suggest that pain intensity (S1), personal care (S2), lifting (S3), and reading (S4) share substantial common variance with other sections, supporting the presence of an underlying unidimensional structure.

Table 2. Spearman correlation matrix of the NDI sections

Items	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
S1	1.00									
S2	0.50	1.00								
S3	0.34	0.34	1.00							
S4	0.35	0.18	0.24	1.00						
S5	0.06	0.07	0.04	0.01	1.00					
S6	0.06	0.01	0.22	0.18	0.25	1.00				
S7	0.12	0.01	0.01	0.20	0.02	0.11	1.00			
S8	0.10	0.23	0.37	0.19	0.09	0.10	0.04	1.00		
S9	0.26	0.20	0.15	0.06	0.09	0.11	0.06	0.10	1.00	
S10	0.40	0.33	0.40	0.34	0.20	0.09	0.05	0.18	0.10	1.00

Bold values indicate the highest inter-item correlations. The table presents the inter-item correlations among all 10 sections of the NDI. Sections with the strongest associations include S1 (pain intensity), S2 (personal care), S3 (lifting), S4 (reading), and S10 (recreation).
S5 – headaches, S6 – concentration, S7 – work, S8 – driving, S9 – sleeping, S10 – recreation

Table 3 presents the results of the exploratory factor analysis (EFA) applied to the NDI correlation matrix. The likelihood ratio (LR) test (χ^2) yielded a p -value < 0.05 , indicating that the hypothesis of item independence can be rejected. This demonstrates that the correlation matrix contains sufficient shared variance among items to justify factor analysis. Additionally, the Kaiser–Meyer–Olkin (KMO) statistic was 0.70 (0.697), supporting the adequacy of the sample for factor extraction.

The exploratory factor analysis revealed a one-factor solution, supported by an eigenvalue greater than 1.0 and by the inflection point observed in the scree plot (Figure 1). Factor loadings derived from the single-factor solution are presented in Table 3. This factor represents the latent construct of neck disability. The NDI sections with the strongest factor loadings were S1 (pain intensity), S2 (personal care), S3 (lifting), S4

Table 3. Exploratory factor analysis (EFA) of the Spanish version of the neck disability index

Variable (NDI items)	Factor loadings (factor 1)	Uniqueness	Communalities
S1	0.65	0.58	0.43
S2	0.58	0.67	0.34
S3	0.60	0.64	0.36
S4	0.48	0.77	0.23
S5	0.14 [†]	0.98	0.02
S6	0.23 [†]	0.95	0.05
S7	0.13 [†]	0.98	0.02
S8	0.38 [†]	0.84	0.14
S9	0.28 [†]	0.92	0.08
S10	0.59	0.66	0.34

S1 – pain intensity, S2 – personal care, S3 – lifting, S4 – reading, S5 – headaches, S6 – concentration, S7 – work, S8 – driving, S9 – sleeping, S10 – recreation
The likelihood ratio (LR) test indicated that the correlation matrix differed significantly from the identity matrix (χ^2 , $p < 0.05$). Factor loadings were derived from a single-factor solution. Items with factor loadings < 0.40 were considered low-performing[†].

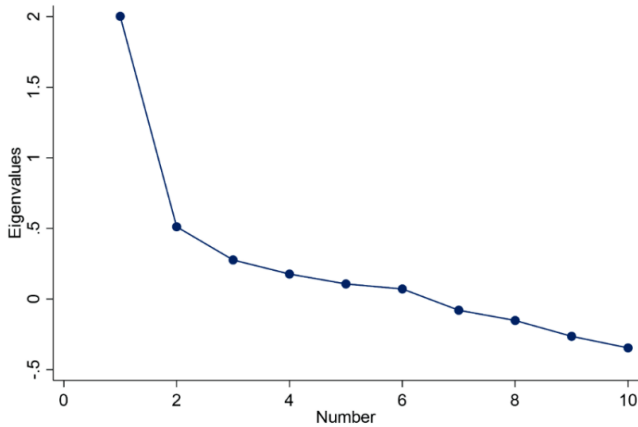


Figure 1. Scree plot of eigenvalues from exploratory factor analysis

(reading), and S10 (recreation), indicating that these items contributed most to the underlying dimension. Items S5 (headaches), S6 (concentration), S7 (work), S8 (driving), and S9 (sleeping) demonstrated the lowest loadings and contributed minimally to the latent construct.

Internal consistency was examined using Cronbach’s alpha. When all 10 items were included, reliability was modest ($\alpha = 0.64$). After excluding items with factor loadings below the predefined threshold (S5–S9), internal consistency improved to $\alpha = 0.71$, indicating acceptable reliability for the reduced five-item version.

Confirmatory factor analysis (CFA)

A structural equation model (SEM) was constructed using neck disability as the latent variable and the NDI sections that demonstrated the strongest factor loadings in the EFA – S1, S2, S3, S4, and S10 –as observed indicators [14, 20]. The purpose of this model was to confirm the factorial structure identified in the exploratory phase and to evaluate the extent to which these items represent the underlying construct.

Table 4 presents the results of the confirmatory factor analysis and the corresponding model fit indices. Standardised factor loadings ranged from 0.46 to 0.74, indicating moderate-to-strong associations with the latent construct. The highest values were observed for S1 (pain intensity), S2 (per-

Table 4. Confirmatory factor analysis of the 5-item Spanish version of the neck disability index (NDI)

Standardised factor loadings between the latent variable and NDI sections		
Section	Standardised factor loadings (95% CI)	R ²
S1	0.74 (0.60–0.88)	0.56
S2	0.62 (0.47–0.76)	0.38
S3	0.53 (0.37–0.70)	0.28
S4	0.46 (0.29–0.63)	0.21
S10	0.56 (0.40–0.73)	0.32

S1 – pain intensity, S2 – personal care, S3 – lifting, S4 – reading, S10 – recreation
All factor loadings are standardised.
Model fit indices: χ^2 ($p = 0.26$), RMSEA = 0.04 (90% CI 0.00–0.14), CFI = 0.99, TLI = 0.97, SRMR = 0.04, CD = 0.75
All indices indicate good model fit according to recommended thresholds (CFI/TLI ≥ 0.90 , RMSEA/SRMR ≤ 0.08).
CFI – comparative fit index, RMSEA – root mean square error of approximation, SRMR – standardised root mean square residual, TLI – Tucker–Lewis index

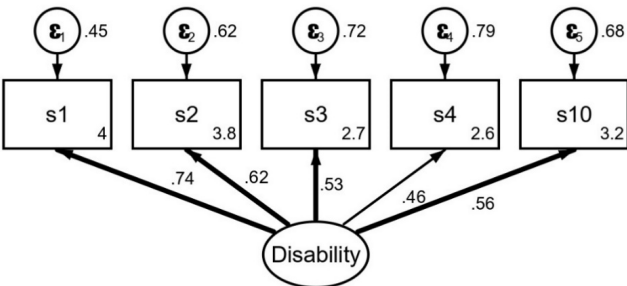


Figure 2. Path diagram of the neck disability index showing standardised factor loadings for the five retained items: S1 – pain intensity, S2 – personal care, S3 – lifting, S4 – reading, and S10 – recreation

sonal care), and S3 (lifting). These sections also exhibited the largest R^2 values, indicating that they explained the greatest proportion of variance in the latent construct.

The goodness-of-fit indices supported an adequate model fit, with a non-significant chi-square test ($p > 0.05$), CFI and TLI values above 0.90, and RMSEA and SRMR values below 0.08, consistent with recommended thresholds for acceptable structural validity. Figure 2 displays the final SEM path diagram, illustrating the standardised relationships between the selected NDI items and the latent variable neck disability.

Discussion

Construct validity is essential when evaluating instruments such as the NDI because it determines whether the scale accurately measures disability in individuals with neck pain. In contrast to content or criterion validity, construct validity specifically examines whether the questionnaire items reflect the latent construct – disability – rather than unrelated dimensions [16, 17, 19, 20]. Factor analysis, a key method for assessing this property, reduces complex data patterns by identifying observable variables that meaningfully correspond to underlying factors [14, 20].

In this study, the construct validity and internal consistency of the Spanish version of the NDI were evaluated in patients with CNP. The analysis revealed that, after removing items with insufficient factor loadings, the most robust factor structure comprised sections S1 (pain intensity), S2 (personal care), S3 (lifting), S4 (reading), and S10 (recreation). Model fit indices confirmed that this reduced structure provided a better representation of the latent construct, with improved internal consistency. Among these, S1, S2, and S10 emerged as the items with the strongest influence on the model, highlighting their central role in capturing disability within this population and supporting the suitability of a shortened version of the NDI.

Unidimensional nature of the NDI

Disability is a multifaceted construct that is difficult to quantify directly in clinical practice. As a PROM, it reflects an individual’s perceived functional capacity or health-related quality of life at a given moment, typically assessed through concise, self-administered questionnaires [8, 9]. These instruments offer clinicians and researchers essential insights into the patient’s experience, complementing traditional objective assessments and supporting shared decision-making in conditions such as neck pain, where symptom burden and functional impact are substantial [8, 9, 29, 30].

The NDI, adapted from the Oswestry Disability Index, is among the most widely used and validated PROMs for assessing disability in individuals with neck pain. It has been translated into more than 20 languages and is recommended

as a core instrument for evaluating pain intensity and functional limitations in cervical disorders, alongside tools such as the VAS and the NPRS [12, 13, 31, 32]. Its widespread use, documented reliability, and clinical relevance support its application across diverse clinical and research contexts [2, 13, 32].

In this study, factor analysis confirmed a unidimensional structure for the Spanish version of the NDI, indicating that a single latent factor adequately represents the construct of cervical spine disability. This finding reinforces the conceptual coherence of the instrument. Although the factor loadings varied across items, the pattern of correlations suggests that each section contributes unique and non-redundant information to the construct. Uniformly high inter-item correlations would suggest redundancy; however, the observed variability supports the multidimensional expression of disability within a unified latent domain.

It is also important to recognise that neck pain encompasses a broad range of clinical presentations, varying in duration, severity, and associated neurological features. These differences may influence the relevance and performance of specific NDI items across patient subgroups. While the unidimensionality observed in this study reinforces the NDI as a cohesive measure of cervical spine disability, ongoing evaluation in diverse clinical populations remains necessary to ensure that individual items maintain clinical and psychometric relevance across varying contexts.

The short version of the NDI

The sections S1 (pain intensity), S2 (personal care), S3 (lifting), S4 (reading), and S10 (recreation) demonstrated the strongest factor loadings and relevant inter-item correlations in individuals with CNP, likely reflecting their contribution to the underlying construct of disability in this population. Pain intensity (S1) is a central feature of CNP. Although participants reported moderate pain levels (NPRS mean 3.93 ± 1.4), even persistent moderate pain can meaningfully disrupt daily functioning. Psychosocial factors such as catastrophising, anxiety, and depression – well-established contributors to CNP – further amplify disability perception and symptom chronicity [2, 33].

Personal care (S2) requires adequate cervical spine mobility, and limitations in tasks such as dressing or grooming often emerge when pain restricts neck movement. As described in the International Classification of Functioning, Disability and Health (ICF), these daily activities constitute essential components of personal functioning and are frequently compromised in chronic musculoskeletal conditions [34–36]. Similarly, lifting (S3) involves cervical and shoulder loading, which can exacerbate symptoms, particularly in individuals with concomitant myofascial pain syndrome characterised by hypersensitive trigger points in cervical and periscapular musculature [36, 37].

Reading (S4) requires sustained static postures, increasing cervical muscle tension and fatigue. In this sample, most participants reported sitting for 4–8 hours per day, a behaviour associated with increased muscle activation and higher pain reports during static tasks [37, 38]. Prolonged device use, especially with suboptimal ergonomics, may further contribute to symptoms associated with neck pain related to digital device use, often referred to as “tech neck” [39, 40].

Recreation (S10) encompasses social and leisure activities and often reflects the broader impact of chronic pain beyond work-related functioning. Restrictions in this domain are strongly associated with psychosocial distress and higher perceived disability, making this item particularly relevant for capturing the lived experience of individuals with CNP.

In summary, these findings provide preliminary support for the five-item version; however, its psychometric performance should be further examined through independent validation and assessment of additional measurement properties.

NDI validation

Previous studies have proposed reducing the NDI to a shorter, five-item version using both conceptual criteria and Rasch analysis [41]. Conceptually, sections S1 (pain intensity), S5 (headache), and S9 (sleep) have been considered for exclusion because they do not directly align with disability as defined by the ICF [34, 41]. Within the ICF framework, disability refers to limitations in activity and participation, meaning that symptoms such as pain intensity, headache, or sleep disturbance may not fully operationalise this construct. Rasch analyses have also reported suboptimal item performance for S3 (lifting) and S4 (reading), with lifting being particularly influenced by occupational demands, which limits its applicability across individuals with cervical spine pain. Several Rasch-based validations similarly identified low responsiveness in the headache item (S5), as headache is not a predominant clinical feature in most cervical spine pain presentations [12, 13, 42]. It is important to note that most shortened NDI models have been developed in samples with non-specific neck pain rather than in populations with CNP. Consequently, variations in item retention largely reflect differences in study populations as well as methodological choices. Rasch analysis estimates person ability and item difficulty to evaluate item fit, differential item functioning, and measurement precision. In contrast, factor analysis captures the underlying dimensional structure by modelling inter-item correlations, thus offering a complementary perspective on construct validity [20].

The selection of Rasch analysis, factor analysis, or a combined approach should therefore align with the study objectives [12]. When the primary goal is to clarify latent dimensionality and the items demonstrate stable performance, factor analysis provides a robust and widely accepted framework. Conversely, when the objective is to explore item-level behaviour in greater depth – such as assessing redundancy, variability, or clinical relevance – additional psychometric approaches may complement factor analysis in future studies. Importantly, conceptual decisions regarding item exclusion must be supported by empirical evidence to avoid removing items that may be clinically meaningful in specific populations. Restricting item selection solely to theoretical criteria may overlook dimensions of disability that are highly relevant for individuals with CNP. Studies specifically conducted in CNP populations highlight the importance of items such as S2 (personal care) and S10 (recreation), which consistently represent essential components of disability [12]. For example, a Brazilian cohort study identified a five-item structure including personal care, concentration, work, driving, and recreation with strong factorial validity, reinforcing the relevance of S2 and S10 for accurately characterising functional limitations in individuals with CNP [14, 15].

Limitations

This study has several limitations that should be acknowledged. First, the sample was drawn from a single Spanish-speaking population, which may limit the generalisability of the findings to other cultural or linguistic contexts. Second, the analyses focused exclusively on structural construct validity and internal consistency; therefore, other important measurement properties – such as criterion validity, test–retest reliability, and responsiveness – were not assessed and should

be addressed in future studies. Third, although the findings provide robust support for the unidimensional structure of the instrument, they apply specifically to individuals with CNP and may not be extrapolated to other cervical spine pain subgroups, such as those with acute or non-specific presentations. Furthermore, as the EFA and CFA in this study were conducted within the same sample, independent validation of the five-item structure is warranted. Future research involving more diverse populations and complementary psychometric evaluations would help strengthen the external validity and broader applicability of both the full and abbreviated versions of the NDI.

Conclusions

This study validates both the Spanish version of the NDI and a shortened five-item version specifically tailored for individuals with CNP. The findings support a unifactorial structure that accurately reflects the underlying construct of cervical spine disability, and the exclusion of items with low factor loadings improved the internal consistency of the instrument. The abbreviated version, comprising sections S1, S2, S3, S4, and S10, demonstrated adequate construct validity and enhanced reliability, indicating that it represents a psychometrically robust alternative to the full questionnaire.

Future studies should apply and validate this abbreviated version in other populations to assess its concurrent validity with the full NDI and with other shortened forms proposed in previous research. As this study was conducted in a Spanish-speaking cohort, additional cross-cultural evaluations are needed to determine its generalisability across different settings and languages.

Overall, the validated abbreviated version offers an efficient and reliable tool for assessing disability in individuals with CNP, with strong applicability in Spanish-speaking contexts and the potential for broader use pending further cross-cultural validation.

Ethical approval

The research related to human use complied with all the relevant national regulations and institutional policies, followed the tenets of the Declaration of Helsinki, and was approved by the Ethics Committee of the Metropolitan Health Service East (Santiago, Chile) (approval No.: 20200234, October 26, 2022).

Informed consent

Informed consent was obtained from all individuals included in this study. All participants provided informed consent prior to enrolment.

Disclosure statement

No author has any financial interest or has received any financial benefit from this research.

Conflict of interest

The authors state no conflict of interest.

Funding

This research received no external funding.

References

- [1] Fandim JV, Nitzsche R, Michaleff ZA, Pena Costa LO, Saragiotto B. The contemporary management of neck pain in adults. *Pain Manag.* 2021;11(1):75–87; doi: 10.2217/pmt-2020-0046.

- [2] Kazeminasab S, Nejadghaderi SA, Amiri P, Pourfathi H, Araj-Khodaei M, Sullman MJM, Kolahi A-A, Safiri S. Neck pain: global epidemiology, trends and risk factors. *BMC Musculoskelet Disord.* 2022;23(1):26; doi: 10.1186/s12891-021-04957-4.
- [3] Cohen SP. Epidemiology, diagnosis, and treatment of neck pain. *Mayo Clin Proc.* 2015;90(2):284–99; doi: 10.1016/j.mayocp.2014.09.008.
- [4] Safiri S, Kolahi A-A, Hoy D, Buchbinder R, Mansournia MA, Bettampadi D, Ashrafi-Asgarabad A, Almasi-Hashiani A, Smith E, Sepidarkish M, Cross M, Qorbani M, Moradi-Lakeh M, Woolf AD, March L, Collins G, Ferreira ML. 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.* 2020; 368:m791; doi: 10.1136/bmj.m791.
- [5] Goode AP, Freburger J, Carey T. Prevalence, practice patterns, and evidence for chronic neck pain. *Arthritis Care Res.* 2010;62(11):1594–601; doi: 10.1002/acr.20270.
- [6] Hogg-Johnson S, van der Velde G, Carroll LJ, Holm LW, Cassidy JD, Guzman J, Côté P, Haldeman S, Ammendolia C, Carragee E, Hurwitz E, Nordin M, Peloso P. 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. *J Manipulative Physiol Ther.* 2009;32(2 Suppl):46–60; doi: 10.1016/j.jmpt.2008.11.010.
- [7] Mercieca-Bebber R, King MT, Calvert MJ, Stockler MR, Friedlander M. The importance of patient-reported outcomes in clinical trials and strategies for future optimization. *Patient Relat Outcome Meas.* 2018;9:353–67; doi: 10.2147/PROM.S156279.
- [8] Churrua K, Pomare C, Ellis LA, Long JC, Henderson SB, Murphy LED, Leahy CJ, Braithwaite J. Patient-reported outcome measures (PROMs): a review of generic and condition-specific measures and a discussion of trends and issues. *Health Expect.* 2021;24(4):1015–24; doi: 10.1111/hex.13254.
- [9] Ramirez MM, Shepherd MH, Melnick SJ, Hanebuth C, Bazemore C, Couce L, Hendren S, Horn ME. Patient-reported outcome measures in physical therapy practice for neck pain: an overview of reviews. *J Patient Rep Outcomes.* 2023;7(1):97; doi: 10.1186/s41687-023-00637-0.
- [10] Howell ER. The association between neck pain, the Neck Disability Index and cervical ranges of motion: a narrative review. *J Can Chiropr Assoc.* 2011;55(3):211–21.
- [11] Andrade Ortega JA, Delgado Martínez AD, Ruiz RA. Validation of the Spanish version of the neck disability index. *Spine.* 2010;35(4):E114–8; doi: 10.1097/brs.0b013e3181afea5d.
- [12] MacDermid JC, Walton DM, Avery S, Blanchard A, Etruw E, McAlpine C, Goldsmith CH. Measurement properties of the neck disability index: a systematic review. *J Orthop Sports Phys Ther.* 2009;39(5):400–17; doi: 10.2519/jospt.2009.2930.
- [13] En MCC, Clair DA, Edmondston SJ. Validity of the Neck Disability Index and Neck Pain and Disability Scale for measuring disability associated with chronic, non-traumatic neck pain. *Man Ther.* 2009;14(4):433–8; doi: 10.1016/j.math.2008.07.005.
- [14] Barreto FS, Avila MA, Pinheiro JS, Almeida MQG, Ferreira C de SB, Fidelis-de-Paula-Gomes CA, Dibai-Filho AV. Less is more: five-item Neck Disability Index to assess chronic neck pain patients in Brazil. *Spine.* 2021; 46(12):688–93; doi: 10.1097/BRS.0000000000003872.
- [15] Pontes-Silva A, Avila MA, Fidelis-de-Paula-Gomes CA, Dibai-Filho AV. The Short-Form Neck Disability index has

- adequate measurement properties in chronic neck pain patients. *Eur Spine J*. 2021;30(12):3593–9; doi: 10.1007/s00586-021-07019-4.
- [16] Ahmed I, Ishtiaq S. Reliability and validity: importance in medical research. *J Pak Med Assoc*. 2021;71(10):2401–6; doi: 10.47391/JPMA.06-861.
- [17] Souza AC de, Alexandre NMC, Guirardello E de B. Psychometric properties in instruments evaluation of reliability and validity. *Epidemiol Serv Saude*. 2017;26(3):649–59; doi: 10.5123/S1679-49742017000300022.
- [18] Tavakol M, Dennick R. Making sense of Cronbach's alpha. *Int J Med Educ*. 2011;2:53–5; doi: 10.5116/ijme.4dfb.8dfd.
- [19] Streiner DL. Starting at the beginning: an introduction to coefficient alpha and internal consistency. *J Pers Assess*. 2003;80(1):99–103; doi: 10.1207/S15327752JPA8001_18.
- [20] Marsh HW, Guo J, Dicke T, Parker PD, Craven RG. Confirmatory factor analysis (CFA), exploratory structural equation modeling (ESEM), and set-ESEM: optimal balance between goodness of fit and parsimony. *Multivariate Behav Res*. 2020;55(1):102–19; doi: 10.1080/00273171.2019.1602503.
- [21] Bibbins-Domingo K, Brubaker L, Curfman G. The 2024 revision to the declaration of Helsinki: modern ethics for medical research: modern ethics for medical research. *JAMA*. 2024; doi: 10.1001/jama.2024.22530.
- [22] Mokkink LB, Prinsen CAC, Bouter LM, Vet HCW de, Terwee CB. The COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) and how to select an outcome measurement instrument. *Braz J Phys Ther*. 2016;20(2):105–13; doi: 10.1590/bjpt-rbf.2014.0143.
- [23] Gagnier JJ, de Arruda GT, Terwee CB, Mokkink LB; Consensus group. COSMIN reporting guideline for studies on measurement properties of patient-reported outcome measures: version 2.0. *Qual Life Res*. 2025;34(7):1901–11; doi: 10.1007/s11136-025-03950-x.
- [24] Boateng GO, Neilands TB, Frongillo EA, Melgar-Quinonez HR, Young SL. Best practices for developing and validating scales for health, social, and behavioral research: a primer. *Front Public Health*. 2018;6:149. doi:10.3389/fpubh.2018.00149.
- [25] Misailidou V, Malliou P, Beneka A, Karagiannidis A, Godolias G. Assessment of patients with neck pain: a review of definitions, selection criteria, and measurement tools. *J Chiropr Med*. 2010;9(2):49–59; doi: 10.1016/j.jcm.2010.03.002.
- [26] de la Barra Ortiz HA, Avila MA, Miranda LG, Liebano RE. Effect of high-intensity laser therapy in patients with non-specific chronic neck pain: study protocol for a randomized controlled trial. *Trials*. 2023;24(1):563; doi: 10.1186/s13063-023-07599-0.
- [27] Boston RC, Sumner AE. STATA: a statistical analysis system for examining biomedical data. *Adv Exp Med Biol*. 2003;537:353–69; doi: 10.1007/978-1-4419-9019-8_23.
- [28] Tabachnick BG, Fidell LS. *Using Multivariate Statistics*. 7th ed. Upper Saddle River: Pearson; 2018.
- [29] Bobos P, MacDermid JC, Walton DM, Gross A, Santaguida PL. Patient-reported outcome measures used for neck disorders: an overview of systematic reviews. *J Orthop Sports Phys Ther*. 2018;48(10):775–88; doi: 10.2519/jospt.2018.8131.
- [30] Young K, Steinhaus M, Gang C, Vaishnav A, Jivanelli B, Lovecchio F, Qureshi S, McAnany S, Kim HJ, Iyer S. The use of patient-Reported Outcomes Measurement Information System in spine: a systematic review. *Int J Spine Surg*. 2021;15(1):186–94; doi: 10.14444/8024.
- [31] Vernon H. The Neck Disability Index: State-of-the-art, 1991–2008. *J Manipulative Physiol Ther*. 2008;31(7):491–502; doi: 10.1016/j.jmpt.2008.08.006.
- [32] Modarresi S, Lukacs MJ, Ghodrati M, Salim S, MacDermid JC, Walton DM; CATWAD Consortium Group. A systematic review and synthesis of psychometric properties of the Numeric Pain Rating Scale and the Visual Analog Scale for use in people with neck pain. *Clin J Pain*. 2021;38(2):132–48; doi: 10.1097/AJP.0000000000000999.
- [33] Ortego G, Villafañe JH, Doménech-García V, Berjano P, Bertozzi L, Herrero P. Is there a relationship between psychological stress or anxiety and chronic nonspecific neck-arm pain in adults? A systematic review and meta-analysis. *J Psychosom Res*. 2016;90:70–81; doi: 10.1016/j.jpsychores.2016.09.006.
- [34] Vargus-Adams JN, Majnemer A. International Classification of Functioning, Disability and Health (ICF) as a framework for change: revolutionizing rehabilitation: Revolutionizing rehabilitation. *J Child Neurol*. 2014;29(8):1030–5; doi: 10.1177/0883073814533595.
- [35] Alford VM, Ewen S, Webb GR, McGinley J, Brookes A, Remedios LJ. The use of the International Classification of Functioning, Disability and Health to understand the health and functioning experiences of people with chronic conditions from the person perspective: a systematic review. *Disabil Rehabil*. 2015;37(8):655–66; doi: 10.3109/09638288.2014.935875.
- [36] Ezzati K, Ravarian B, Saberi A, Salari A, Reyhanian Z, Khakpour M, Chabok SY. Prevalence of cervical myofascial pain syndrome and its correlation with the severity of pain and disability in patients with chronic non-specific neck pain. *Arch Bone Jt Surg*. 2021;9(2):230–4; doi: 10.22038/abjs.2020.48697.2415.
- [37] Christensen SWM, Palsson TS, Krebs HJ, Graven-Nielsen T, Hirata RP. Prolonged slumped sitting causes neck pain and increased axioscapular muscle activity during a computer task in healthy participants – a randomized crossover study. *Appl Ergon*. 2023;110:104020; doi: 10.1016/j.apergo.2023.104020.
- [38] Opara M, Kozinc Ž. Which muscles exhibit increased stiffness in people with chronic neck pain? A systematic review with meta-analysis. *Front Sports Act Living*. 2023;5:1172514; doi: 10.3389/fspor.2023.1172514.
- [39] Jung K-S, Jung J-H, In T-S, Cho H-Y. Effects of prolonged sitting with slumped posture on trunk muscular fatigue in adolescents with and without chronic lower back pain. *Medicina*. 2020;57(1):3; doi: 10.3390/medicina57010003.
- [40] Tsantili A-R, Chrysikos D, Troupis T. Text Neck Syndrome: disentangling a new epidemic. *Acta Med Acad*. 2022;51(2):123–7; doi: 10.5644/ama2006-124.380.
- [41] Walton DM, MacDermid JC. A brief 5-item version of the Neck Disability Index shows good psychometric properties. *Health Qual Life Outcomes*. 2013;11(1):108; doi: 10.1186/1477-7525-11-108.
- [42] van der Velde G, Beaton D, Hogg-Johnston S, Hurwitz E, Tennant A. Rasch analysis provides new insights into the measurement properties of the neck disability index. *Arthritis Rheum*. 2009;61(4):544–51; doi: 10.1002/art.24399.