Clinimetric properties of the Nepali version of the Pain Catastrophizing Scale in individuals with chronic pain

Saurab Sharma1,2
Pascal Thibault3
J Haxby Abbott2
Mark P Jensen4

1Department of Physiotherapy, Kathmandu University School of Medical Sciences, Dghukhel, Nepal; 2Centre for Musculoskeletal Outcomes Research, Dunedin School of Medicine, University of Otago, Dunedin, Otago, New Zealand; 3Department of Psychology, McGill University, Montreal, Québec, Canada; 4Department of Rehabilitation Medicine, University of Washington, Seattle, USA

Background: Pain catastrophizing is an exaggerated negative cognitive response related to pain. It is commonly assessed using the Pain Catastrophizing Scale (PCS). Translation and validation of the scale in a new language would facilitate cross-cultural comparisons of the role that pain catastrophizing plays in patient function.

Purpose: The aim of this study was to translate and culturally adapt the PCS into Nepali (Nepal version of PCS [PCS-NP]) and evaluate its clinimetric properties.

Methods: We translated, cross-culturally adapted, and performed an exploratory factor analysis (EFA) of the PCS-NP in a sample of adults with chronic pain (N=143). We then confirmed the resulting factor model in a separate sample (N=272) and compared this model with 1-, 2-, and 3-factor models previously identified using confirmatory factor analyses (CFAs). We also computed internal consistencies, test–retest reliabilities, standard error of measurement (SEM), minimal detectable change (MDC), and limits of agreement with 95% confidence interval (LOA95%) of the PCS-NP scales. Concurrent validity with measures of depression, anxiety, and pain intensity was assessed by computing Pearson’s correlation coefficients.

Results: The PCS-NP was comprehensible and culturally acceptable. We extracted a two-factor solution using EFA and confirmed this model using CFAs in the second sample. Adequate fit was also found for a one-factor model and different two- and three-factor models based on prior studies. The PCS-NP scores evidenced excellent reliability and temporal stability, and demonstrated validity via moderate-to-strong associations with measures of depression, anxiety, and pain intensity. The SEM and MDC for the PCS-NP total score were 2.52 and 7.86, respectively (range of PCS scores 0–52). LOA95% was between −15.17 and +16.02 for the total PCS-NP scores.

Conclusion: The PCS-NP is a valid and reliable instrument to assess pain catastrophizing in Nepalese individuals with chronic pain.

Keywords: measurement error, outcome measure, Bland–Altman plot, reliability, measurement properties, pain assessment

Introduction

Pain catastrophizing, which has been defined as an exaggerated negative cognitive response related to pain, has been shown to be associated with a broad array of important quality of life domains such as pain intensity, pain interference, and psychological function in individuals with chronic pain from different cultures. Catastrophizing is responsive to various pain treatments and has been found to mediate the effects of many of these. Thus, catastrophizing remains an important psychosocial domain to assess in patients with chronic pain. The most common measure of pain catastrophizing is the 13-item Pain Catastrophizing Scale (PCS).
The PCS has three subscales assessing pain-related rumination, magnification, and helplessness.\(^1,^3\) This three-factor structure has been confirmed in other English-speaking samples,\(^9,^10\) although research has shown that two-factor solutions show better fit in some samples.\(^11,^12\) Non-English translations have also been shown to yield both two and three factors.\(^13,^14\) However, given the strong associations among the factors, researchers tend to use the total PCS score more often than the subscale scores.\(^4,^5,^7\)

The original English version of the PCS and its translated versions are psychometrically robust. The total score of PCS has repeatedly been shown to be reliable as evidenced by good-to-excellent internal consistency\(^2,^3,^10,^14,^17,^20,^21\) and excellent test–retest reliability.\(^2,^14,^17,^20,^22\) Standard error of measurement (SEM) and minimal detectable change (MDC) of the PCS range from 1.6 to 4.6\(^14,^15,^21,^22\) and 8.8 to 12.8,\(^14,^15,^17,^21\) respectively. Concurrent validity of the PCS is supported via moderate-to-strong correlations (ie, \(r \geq 0.30\)) with measures of depression,\(^2,^14,^16,^17\) anxiety,\(^1,^2,^14,^17\) and pain intensity.\(^1,^2,^10,^11,^14,^16,^17\)

The PCS has been translated into a number of different languages (https://eprovide.mapi-trust.org/instruments/pain-catastrophizing-scale#languages). The availability of translated measures is critical to facilitate research examining the role that language and culture may play in how catastrophizing is related to patient function. Thus, the aim of this study was to translate the PCS into another language – this time, Nepali (Nepali version of PCS [PCS-NP]) – and assess its clinimetric properties. Consistent with the extant research, we hypothesized that the PCS-NP would demonstrate either a two-factor structure or a three-factor structure\(^1,^12\) and that the factors would be strongly associated with one another, supporting the use of the total scale in this population. We further hypothesized that the PCS-NP scales would demonstrate 1) good-to-excellent internal consistencies (ie, Cronbach’s \(\alpha \geq 0.70\),\(^2,^14,^17,^20,^21,^23,^24\) 2) excellent 2-week test–retest reliability (ie, intraclass correlation coefficient [ICC] \(\geq 0.75\)),\(^14,^15,^17,^20,^22,^24\) 3) SEM between 1.6 and 4.6,\(^14,^15,^21,^22\) 4) MDC between 8.8 and 12.8,\(^14,^15,^17,^21\) and 5) concurrent validity via moderate-to-strong correlations (ie, \(r \geq 0.30\)) with measures of depression, anxiety, and pain intensity.\(^2,^14,^16,^17\)

**Methods**

We first translated the PCS to Nepali and performed cognitive debriefing of the instructions and items to ensure that the items were culturally appropriate in 30 individuals with musculoskeletal pain. Next, we evaluated the clinimetrics of the PCS-NP in two independent samples. Ethical approval of the research was obtained from the Institutional Review Committee of the Kathmandu University School of Medical Sciences, Dhulikhel, Nepal (reference number: 105/14), and the study complies with ethical guidelines of the Declaration of Helsinki. Written informed consent was obtained from all the research participants before data collection. Those participants who could not read or provide a written signature provided a verbal consent, and a witness signed on their behalf.

**Translation procedure**

Translation of PCS into Nepali was performed using standard guidelines.\(^2^3\) First, we obtained permission from the developer (Professor Michael Sullivan) to translate the PCS into Nepali. Then, two forward translators translated the PCS into Nepali. The two forward translated versions were synthesized to a single first translated (FT) version. The FT version was then back-translated into English by three independent native English speakers. An expert committee meeting was held, which comprised language experts, translators (who were available for a face-to-face meeting), and research methodologist. Any discrepancies were discussed, and Nepali version of the PCS that is simple and comprehensible was finalized during the meeting. Translators remotely located were contacted via e-mails and Skype, and any important suggestions were noted. The final Nepali version was back-translated into English, which was sent to the last author (MPJ) for review.

The resultant version was tested in a sample of 30 individuals with either acute musculoskeletal pain or persistent musculoskeletal pain for cognitive debriefing. Minor changes were made in sentence structure, and simpler words replaced difficult words so that most participants could understand the final PCS-NP. The final Nepali version, its back-translation, and translation history were sent to the developer, who reviewed and approved it. The Nepali version is licensed and distributed by Mapi Research Trust, and researchers intending to use PCS-NP should contact Mapi Research Trust for its use (https://eprovide.mapi-trust.org/instruments/pain-catastrophizing-scale).

The translation history is summarized in Figure 1.

**Clinimetric evaluation**

Participants

Two independent adult (age \(\geq 18\) years) samples of Nepali-speaking individuals with chronic pain (pain lasting for >3 months) were recruited. The first sample included individuals with chronic pain (N=144) recruited from a tertiary care hospital (n=44; consecutive patients) and a community (n=100; by door-to-door survey) in rural Nepal with pain intensity \(\geq 4\) out of 10 on a 0–10 numerical pain rating scale (NPRS) (0= “no pain” and 10= “maximum pain”). The sec-
The original English version of PCS

FT 1 (translated by a nonmedical native Nepali translator)  

FT 2 (translated by a Nepali physiotherapist)

FT synthesis
Discrepancies between the two forward translations resolved after discussion facilitated by the lead author (SS)

BT 1a
BT 2a
BT 3a

Expert committee review
- Consisted of all the translators, researchers, methodologists, and original measure developer
- Remotely located experts were contacted via e-mails or Skype
- Reviewed all the translation version of the PCS and reports
- Prefinal PCS-NP was produced after consensus

Pretesting
- Prefinal version tested on 30 individuals with musculoskeletal pain
- Every participant was probed to inquire if he/she correctly identified the real meaning of the original English version
- Asked if any changes in the questionnaire will make it easier for them to understand

Creation of the final version of Nepali PCS-NP
- Changes on the grammar and sentence structure were made using feedback by participants during the pretesting

Approval of the PCS-NP by the developer
- PCS-NP was back-translated to English
- All the written reports of the translation history was sent to the developer for the review as a single Excel document
- The translation history was reviewed by the developer and was approved
- PCS-NP was sent to www.mapi-trust.org

Figure 1 Translation history of PCS-NP.

Note: All the back translators were performed independently by native English speakers blind to the original English version of the PCS.

Abbreviations: BT, backward translation; FT, forward translation; PCS, Pain Catastrophizing Scale; PCS-NP, Nepali version of PCS.

The second sample, also consisted of individuals with chronic pain (N=275) who were required to report ≥3 out of 10 pain on an NPRS, was recruited mostly from the community (n=252). The reminder (n=23) of the sample was recruited from the same hospital as that used to recruit the first sample. All the participants in both samples provided data via interview.

Measures

Participants in the first sample were administered the PCS-NP once, and participants in the second sample were administered the PCS-NP twice (2 weeks apart), with 244 (90%) of original sample providing retest data. A summary of the measures used in the study is provided in Table 1 and described in more detail later.

Beck Anxiety Inventory (BAI)

The 21-item BAI was used to assess perceived anxiety in the first sample. Each item is scored on a 4-point Likert scale ranging from 0 to 3 where 0 = “not at all” and 3 = “severely, I could barely stand it”. Total score can range from 0 to 63, with higher scores indicating more anxiety. The Nepali version of BAI – the version used in this study – has been shown to...
improvement.32,33 We classified participants with a GRoC and "stable" groups, considering 1-point change as significant < 4 indicate improvement, and scores 4 indicate worsening.

The Nepali version of BDI is a reliable score can range from 0 to 63, with higher scores indicating symptoms is rated on the 4-point scale, ranging from 0 to 3. Total options designed for that item. Severity of depression symptoms in the second sample on a 7-point Likert scale ranging from 1 to 7.31–33 Score 4 represents “no change”, scores >4 indicate improvement, and scores <4 indicate worsening. The GRoC was used to categorize the sample to “improved” and “stable” groups, considering 1-point change as significant improvement.32,33 We classified participants with a GRoC score of 4 as stable and participants with GRoC scores of 5–7 as improved. Using this classification, participants in the stable group were used for computing the test–retest stability, SEM, MDC, and limits of agreement statistics.34

Other measures
Demographics information collected included age, sex, religion, ethnicity, education, and occupation. Additionally, total duration of chronic pain was recorded in months and pain site was assessed using a pain diagram.

Table 1 Nepali versions of PROMs used in the study

<table>
<thead>
<tr>
<th>Name</th>
<th>Study</th>
<th>Item</th>
<th>Scale</th>
<th>Construct assessed</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCS-NP</td>
<td>EFA, CFA</td>
<td>13</td>
<td>0–4, ordinal</td>
<td>Pain catastrophizing</td>
</tr>
<tr>
<td>BDI-NP31</td>
<td>EFA</td>
<td>21</td>
<td>0–3, ordinal</td>
<td>Depression</td>
</tr>
<tr>
<td>BAI-NP31</td>
<td>EFA</td>
<td>21</td>
<td>0–3, ordinal</td>
<td>Anxiety</td>
</tr>
<tr>
<td>GRoC-NP31</td>
<td>CFA</td>
<td>7</td>
<td>1–7, ordinal</td>
<td>Global rating of change</td>
</tr>
<tr>
<td>PROMIS</td>
<td>CFA</td>
<td>1</td>
<td>1–5, ordinal</td>
<td>Average pain intensity</td>
</tr>
</tbody>
</table>

Abbreviations: CFA, confirmatory factor analysis; EFA, exploratory factor analysis; BDI-NP, Nepali version of Beck Anxiety Inventory; BDI-NP, Nepali version of Beck Depression Inventory; GRoC-NP, Nepali version of Global Rating of Change; PCS-NP, Nepali version of Pain Catastrophizing Scale; PROMs, Patient-Reported Outcome Measures; PROMIS, Patient-Reported Outcomes Measurement Information System.

Beck Depression Inventory (BDI)
The 21-item BDI was used to assess depression in the first sample.31 Each item in the scale has four unique response options designed for that item. Depression of pain intensity

The PROMIS average pain intensity over the past 7 days was assessed in the second sample on a 5-point Likert scale where 1= “no pain”, and 5= “very severe pain”, with higher scores indicating more intense pain. The Nepali version of BDI is a reliable measure (Cronbach’s α=0.9030 and ICC =0.8428) and demonstrated excellent internal consistency (Cronbach’s α=0.90) in the current sample.

Patient-Reported Outcomes Measurement Information System (PROMIS) pain intensity
The PROMIS pain intensity scale was assessed in the second sample on a 5-point Likert scale ranging from 1 to 7.31–33 Score 4 represents “no change”, scores >4 indicate improvement, and scores <4 indicate worsening. The PROMIS pain intensity scale is 0.71 in the current second sample.

Global Rating of Change (GRoC)
The Nepali version of GRoC was used to assess global impression of perceived change in their chronic pain-related problems in the second sample on a 7-point Likert scale ranging from 1 to 7.31–33 Score 4 represents “no change”, scores >4 indicate improvement, and scores <4 indicate worsening. The GRoC was used to categorize the sample to “improved” and “stable” groups, considering 1-point change as significant improvement.32,33 We classified participants with a GRoC score of 4 as stable and participants with GRoC scores of 5–7 as improved. Using this classification, participants in the stable group were used for computing the test–retest stability, SEM, MDC, and limits of agreement statistics.34

Other measures
Demographics information collected included age, sex, religion, ethnicity, education, and occupation. Additionally, total duration of chronic pain was recorded in months and pain site was assessed using a pain diagram.

Statistical analyses
Factor analyses
We performed an exploratory factor analysis (EFA) in the first sample (hereafter referred to as the EFA sample) using maximum likelihood for factor extraction. Factor rotation was performed using oblique rotation (Direct Oblimin; delta =0), allowing the factors to correlate with each other. Next, we performed a series of confirmatory factor analyses (CFAs) in the second sample (hereafter referred to as the CFA sample) to compare the results of the EFA from the first sample with 1) a one-factor solution including all the 13 items as a single factor, 2) the two-factor solution obtained from the largest English-speaking sample published,11 and 3) the original three-factor solution found by Sullivan et al’ using AMOS for Statistical Software for Social Sciences (SPSS) 24. Model fit was evaluated using the chi-square goodness-of-fit index, the ratio of chi-square value to degree of freedom, the root mean square error of approximation (RMSEA), comparative fit index (CFI), and parsimony goodness-of-fit index (PGFI). Better fit is indicated by 1) chi-square value and the ratio of chi-square to degree of freedom values closer to zero, 2) lower values of RMSEA, 3) larger values of CFI, and 4) larger value of PGFI, which indicates more parsimonious fit.35

Reliability
To evaluate reliability, we assessed internal consistency, test–retest stability, SEM, MDC, and created Bland–Altman Plots. We computed internal consistencies of the PCS-NP scales for both samples, using Cronbach’s α. We considered values of Cronbach’s α<0.70 as inadequate, values between 0.70 and 0.79 as adequate, values between 0.80 and 0.89 as good, and values ≥0.90 as excellent.24 Two-week test–retest stability was evaluated using the ICC in the stable group (GRoC =4). We considered the values of ICC between 0.40 and 0.59 as fair, values between 0.60 and 0.74 as good, and values ≥0.75 as excellent.24 To further evaluate reliability, we also computed the SEM, which is an indication of how repeated administration of a measure tends to be distributed around the
“true” score. SEM is important to consider because two measurements obtained at different times will not yield exactly the same score, either because of the variation in the participants being assessed or because of the variation in the measurement process, or both. Thus, it is possible that a measure may have large test–retest stability indicating excellent reliability but may also have a large measurement error. Therefore, the SEM compliments test–retest stability and is a recommended parameter to assess and report in clinimetric studies. Larger scores indicate large variability, and smaller scores indicate minimal variability. We calculated SEM using the following formula: SEM = SD change × √(1 - ICC), where SD change is the standard deviation for the mean change of PCS-NP scores. Using the SEM value, we further calculated MDC for 95% confidence interval (MDC95%) of the PCS-NP using the following formula: MDC95% = 1.96 × SD change × SEM. A Bland–Altman plot was created to complement the measurement error and indicates the levels of agreement between the baseline and follow-up assessments. The plot was drawn with change of PCS-NP scores from baseline to follow-up (Y-axis) versus mean score of PCS-NP between baseline and follow-up assessments (X-axis). Limits of agreement with 95% confidence interval (LOA 95%) were computed by using the following formula: mean difference ± 1.96 × SD change.

Validity
We evaluated the concurrent validity of the PCS-NP scales by computing Pearson’s correlation coefficients between the PCS-NP scale scores and the scores from Nepali versions of BDI, BAI administered to participants in the EFA sample, and PROMIS pain intensity short-form 3a measure administered to CFA sample. We considered correlation coefficients (r) <0.30 as weak, coefficients between 0.30 and 0.49 as moderate, and coefficients ≥0.50 as strong. We excluded the participants for all analysis if there were any missing values on the PCS items.

Results
The PCS-NP succeeded in retaining the semantic, idiomatic, experiential, and conceptual equivalence. The reporting of the clinimetric properties is guided by COSMIN recommendations. Scores of PCS-NP were normally distributed in both the EFA and CFA samples. A total of 0.7% (n=1) and 2.6% (n=7) in EFA and CFA samples, respectively, scored 0/52 total score, and 0.7% (n=1) and 1.1% (n=2) in EFA and CFA samples, respectively, scored 52/52 total score.

Demographic characteristics
The majority of study participants were women (65% and 73% in the EFA and CFA samples, respectively) and were Hindu in religion (92% and 85%). Almost half of the participants in both the samples had pain in multiple body parts (42% and 47%), followed by pain in the low back and pelvis (22% and 18%) and knee(s) (21% and 19%), respectively. All the participants in the EFA sample were individuals with chronic musculoskeletal pain; however, 4% of the CFA sample had chronic headache. Descriptive information for the two samples is presented in Table 2.

Of the 144 participants in the EFA sample and 275 participants in the CFA sample who were administered PCS-NP at the baseline assessment, there were missing items in the data completed by one and three participants, respectively.
in the two samples. We excluded the participants with missing items in all analyses from EFA and CFA samples; thus, there were N=143 participants in EFA sample and N=272 participants in the CFA sample for all analyses.

Factor analyses

Results of EFA in the EFA sample (N=143) using maximum likelihood with Oblimin rotation for factor extraction and the Kaiser criterion (ie, eigenvalues >1.0) to determine the number of factors indicated a two-factor solution. As can be seen in the pattern matrix from the EFA analysis (Table 3), Factor 1 comprised items 2, 3, 4, 5, 9, 12, and 13 and Factor 2 comprised items 1, 6, 7, 8, 10, and 11. Given the items that were loaded on each factor, we labeled Factor 1 “helplessness” and Factor 2 “pain focus”.

These two factors explained 50% of variance in the items (Figure 2). The two factors, however, were very strongly associated with one another (r=0.91). The two-factor solution found in the first sample demonstrated an acceptable fit in the CFA sample (Table 4). In addition, the one-factor, another two-factor extracted by Chibnall and Tait,11 and the three-factor extracted by Sullivan et al1 also had adequate fit, comparable to the two-factor solution extracted from the EFA sample. Attempts to improve each of these models further by allowing up to two pairs of error terms to correlate resulted in only modest improvements in the fit of each model (Table 4).

Reliability

Results of the analyses evaluating the reliability of PCS-NP total score and two subscale scores (computed using the results of the factor analyses as a guide) are presented in Table 5. The PCS-NP demonstrated good-to-excellent internal consistency (range 0.83–0.93) in the two samples. Two-week test–retest reliability on the stable sample was excellent (95% CI: 0.85, 0.93). SEM and MDC for the total score of PCS-NP were 2.52 and 6.98, respectively. LOA95% ranged from −15.17 to 16.02 for the total PCS-NP scores. The LOA95% of the subscales is presented in Table 5. Bland–Altman plot, which shows the limits of agreement for the total PCS items and the two factors, separately, is presented as Figure 3.

Validity

Both the total score and subscale scores were positively, moderately, and significantly associated with the criterion variables assessing depression, anxiety, and pain intensity. Concurrent validity of PCS-NP total scores and subscales is presented in Table 6.

Table 3 Pattern matrix after maximum likelihood factor extraction and Oblimin rotation

<table>
<thead>
<tr>
<th>Item number</th>
<th>Original scale</th>
<th>Brief description of item</th>
<th>Factors 1</th>
<th>Factors 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 H</td>
<td>Never get any better</td>
<td>0.78 0.06</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 H</td>
<td>Can’t stand it</td>
<td>0.73 0.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12 H</td>
<td>Nothing I can do</td>
<td>0.73 0.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 H</td>
<td>It’s awful</td>
<td>0.72 0.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 H</td>
<td>Can’t go on</td>
<td>0.62 0.13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9 R</td>
<td>Can’t keep out of mind</td>
<td>0.56 0.24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13 M</td>
<td>Something serious may happen</td>
<td>0.53 0.16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11 R</td>
<td>Want the pain to stop</td>
<td>−0.10 −0.85</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 R</td>
<td>Want the pain to go away</td>
<td>0.03 −0.73</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 R</td>
<td>Thinking how much it hurts</td>
<td>0.28 −0.55</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 H</td>
<td>Worrying whether pain will end</td>
<td>0.37 −0.39</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 M</td>
<td>Thinking of painful experiences</td>
<td>0.25 −0.37</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 M</td>
<td>Afraid if pain may get worse</td>
<td>0.28 −0.34</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Bold items were included in the factors 1 and 2 on the values of the factor loadings.

Abbreviations: H, helplessness; M, magnification; R, rumination.

Figure 2 Path diagram after CFAs, two-factor model (upper), and one-factor model (lower) with adjustment for covariance of error terms in CFA sample (N=272).

Abbreviations: CFAs, confirmatory factor analyses; PCS, Pain Catastrophizing Scale.
Clinimetric properties of Nepali version of PCS

Discussion

The PCS-NP was successfully translated and adapted to Nepali and validated in Nepalese with chronic pain. The findings support its reliability and validity. Specifically, we extracted two-factor solution (Factor 1 = helplessness and Factor 2 = pain focus) from the EFA in the EFA sample and confirmed this solution by CFA in an independent sample. The reliability of the PCS-NP total score and two subscale scores was supported by excellent internal consistencies and test–retest stability over a 2-week period. Concurrent validity was supported via moderate-to-strong associations with measures assessing depression, anxiety, and pain intensity, as hypothesized.

Factor analysis

The factor analysis of the PCS-NP items yielded two factors in our sample. Although three factors are more common,2,14–16,18,20,22,43 two factors sometimes emerge11,12,19,44 in different samples of individuals with chronic pain. The differences in the factor structure of these items across different samples may be related to cultural differences in how catastrophizing thoughts relate to each other (and possibly impact function) in different countries. In our sample, helplessness was largely retained as a unique factor, but distinct “rumination” and “magnification” factors did not emerge as they had in the original English version of the PCS45 or many other samples.2,14–16,18,20,22,43 Instead,

### Table 4 Results of the confirmatory factor analyses for the PCS-NP in the CFA sample (N=272)

<table>
<thead>
<tr>
<th>Model and modifications</th>
<th>$\chi^2$ (df)</th>
<th>$\chi^2$/df</th>
<th>RMSEA</th>
<th>CFI</th>
<th>PGFI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. One-factor model</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. No modifications</td>
<td>229 (65)</td>
<td>3.53</td>
<td>0.097</td>
<td>0.913</td>
<td>0.628</td>
</tr>
<tr>
<td>b. With modification</td>
<td>190 (63)</td>
<td>3.01</td>
<td>0.086</td>
<td>0.933</td>
<td>0.622</td>
</tr>
<tr>
<td>2. Two-factor modela</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. No modifications</td>
<td>190 (64)</td>
<td>2.97</td>
<td>0.085</td>
<td>0.933</td>
<td>0.632</td>
</tr>
<tr>
<td>b. With modifications</td>
<td>166 (62)</td>
<td>2.67</td>
<td>0.079</td>
<td>0.945</td>
<td>0.620</td>
</tr>
<tr>
<td>3. Two-factor modelb</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. No modifications</td>
<td>207 (64)</td>
<td>3.23</td>
<td>0.091</td>
<td>0.924</td>
<td>0.626</td>
</tr>
<tr>
<td>b. With modifications</td>
<td>171 (62)</td>
<td>2.76</td>
<td>0.081</td>
<td>0.942</td>
<td>0.617</td>
</tr>
<tr>
<td>4. Three-factor modelc</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. No modifications</td>
<td>200 (62)</td>
<td>3.23</td>
<td>0.091</td>
<td>0.927</td>
<td>0.608</td>
</tr>
<tr>
<td>b. With modifications</td>
<td>180 (60)</td>
<td>3.00</td>
<td>0.086</td>
<td>0.936</td>
<td>0.596</td>
</tr>
</tbody>
</table>

Notes: Two-factor model based on results from the EFA sample (N=143). Two-factor model based on Chibnall and Tait.11 Three-factor model based on Sullivan et al.1

Abbreviations: CFA, confirmatory factor analysis; CFI, Comparative Fit Index; df, degree of freedom; PCS-NP, Nepali version of Pain Catastrophizing Scale; PGFI, parsimony goodness-of-fit index; RMSEA, root mean square error of approximation.

### Table 5 Reliability of the PCS-NP

<table>
<thead>
<tr>
<th>Sample</th>
<th>N</th>
<th>IC</th>
<th>Test–retest (ICC)</th>
<th>SEM</th>
<th>MDC_{95%}</th>
<th>LOA_{95%}</th>
</tr>
</thead>
<tbody>
<tr>
<td>EFA sample</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCS-NP total</td>
<td>143</td>
<td>0.91</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCS-NP Factor 1</td>
<td></td>
<td>0.87</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCS-NP Factor 2</td>
<td></td>
<td>0.83</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CFA sample</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCS-NP total</td>
<td>272</td>
<td>0.93</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCS-NP Factor 1</td>
<td></td>
<td>0.89</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCS-NP Factor 2</td>
<td></td>
<td>0.85</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CFA-stable group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCS-NP total</td>
<td>122</td>
<td>0.90</td>
<td>0.90 (0.85, 0.93)</td>
<td>2.52</td>
<td>6.98</td>
<td>16.02, –15.17</td>
</tr>
<tr>
<td>PCS-NP Factor 1</td>
<td></td>
<td>0.88</td>
<td>0.88 (0.83, 0.92)</td>
<td>1.67</td>
<td>4.62</td>
<td>9.29, –9.69</td>
</tr>
<tr>
<td>PCS-NP Factor 2</td>
<td></td>
<td>0.89</td>
<td>0.89 (0.84, 0.92)</td>
<td>1.33</td>
<td>3.67</td>
<td>8.47, –7.22</td>
</tr>
</tbody>
</table>

Abbreviations: CFA, confirmatory factor analysis; EFA, exploratory factor analysis; IC, internal consistency (Cronbach’s $\alpha$); ICC, intraclass correlation coefficient (with 95% CI) for 2-week test–retest stability; LOA$_{95%}$, limits of agreement with 95% confidence interval; MDC$_{95%}$, minimal detectable change for 95% confidence interval; PCS-NP, Nepali version of Pain Catastrophizing Scale; SEM, standard error of measurement.
Figure 3 Bland–Altman plots for total PCS (A), PCS Factor 1 (B), and PCS Factor 2 (C).

Note: Y-axis is the change in PCS-NP scores between baseline and follow-up measurements and X-axis is the mean of the PCS-NP scores at the baseline and final measurements. The red line is the mean change of score (\( \bar{d} \)); and green lines are \( \bar{d} \pm Z \times SD_{\text{change}} \) (where \( Z = 1.96 \) for 95% confidence interval).

Abbreviations: GRoC, Global Rating of Change; PCS, Pain Catastrophizing Scale; PCS-NP, Nepali version of Pain Catastrophizing Scale.
the second factor appeared to reflect thoughts related to a focus on pain. At the same time, the two factors that emerged in this study were strongly associated with each other, suggesting that they may both be assessing the same overarching construct. This conclusion is also supported by the similar pattern of associations between the two PCS-NP subscale scores and the criterion variables. Many other studies have also shown strong between-factor associations in both two-factor and three-factor models. As of note, even though there is a large degree of overlap in how the PCS items load onto different factors across studies, rarely if ever do the exact same PCS items load on the same two- or three-factor models in any pair of studies. As a group, these findings indicate that the PCS total score (but not necessarily the subdomain scores) may be most appropriate for cross-cultural research.

Reliability
The internal consistencies of the PCS-NP total score in this study (a range =0.90–0.93) lie within the ranges previously reported. They are higher than the original English and the Catalan versions (a=0.87 in both) and Hindi version (a=0.76), similar to those found in samples who speak other languages such as Malay, Korean, Chinese, Italian, and German (a range =0.90–0.93) and less than those found in a sample of patients from South Africa (a=0.97–0.98). Similarly, we found that the 2-week stability of the PCS-NP was excellent in our test–retest sample, consistent with our a priori hypothesis. Two weeks is considered optimal for the assessment of reproducibility of a measure, because it is long enough to limit potential recall bias and also short enough to limit real change. Additionally, it is important to compute test–retest stability in the sample reporting that did they not perceive a change in their pain problem. Thus, in the current study, we assessed test–retest stability in those participants who reported no change in their pain problem using the GReC as an external anchor. Numerous previous studies either assessed test–retest stability in short periods (range 1–7 days), which may have been resulted in findings that have recall bias, or assessed using longer epochs (3–4 weeks). To our knowledge, only two studies assessed the test–retest reliability of the PCS using a 2-week time period, the results of which are comparable to our study. Additionally, we reported test–retest stability in a stable sample who reported no change in their pain-related problems. Most of the studies evaluating the psychometric properties of the PCS have not computed the test–retest reliability in a stable sample. This is especially true for those studies reporting longer follow-up for retest assessment. The participants in these studies were more likely to have changed with respect to their pain problem (and catastrophizing scores). Thus, the stability of the PCS may be underestimated in these studies. Support for this possibility is found in one study, which computed test–retest stabilities in both a subsample of stable patients with low back pain and the total sample and found better temporal stability in the stable group (ICC =0.92, n=34) versus the total group (ICC =0.85, n=60).

Measurement error
The measurement error parameters we found met our a priori hypothesis for SEM, but we found lower MDC values than previously reported. Both the SEM and MDC are dependent on the test–retest coefficient. As previously indicated, the stability assessed is influenced by the sample chosen (ie, stable sample versus total sample). Not many studies have reported measurement error parameters, which are now recommended by the COSMIN checklist for clinimetric studies. The value of SEM we found for the total score (2.52) is larger than the Hindi version (SEM =1.60) and smaller than other versions (SEM =3.30–4.60). The SEM is influenced by the SD of change of PCS score (SD_{change}) and the ICC (refer to the formula we used to calculate SEM in the “Methods” section). The smaller SEM values in our study are likely related to our
use of the stable group to compute test–retest stability and measurement error, as opposed to other studies that used total group to compute test–retest stability and SEM. Thus, we obtained smaller SD\textsubscript{change} and larger ICC values, giving us more precise value of SEM. The authors of the Hindi version study included the SD of the baseline measurement instead of the SD\textsubscript{change} which may have accounted for their very low SEM.\textsuperscript{22}

The MDC value of 6.98/52 observed in our sample provides a cut-point for determining the reliability of changes in the PCS score. The MDC computed from our sample is smaller than those observed in other samples (range 8.83–13).\textsuperscript{14,17} This is due to the smaller variability in the catastrophizing scores (as represented by the standard deviation) and larger test–retest reliability coefficient (ICC) in our sample, relative to those in previous studies.\textsuperscript{14,17} The MDC statistics is useful as a guide to indicate whether or not a change in PCS score is greater than measurement error; that is, if the change represents true change and not just random error. For the PCS-NP, a change of 7/52 can be viewed as representing true change in the total PCS score beyond measurement error.

The results of LOA\textsubscript{95%} and Bland–Altman plots\textsuperscript{39} (Figure 3) provide further information regarding measurement error.\textsuperscript{38,40} The Bland–Altman plots show agreement between test and retest scores for every study participant, that is, how far each participant deviates from the mean change score between the baseline and final measurements. In the current study, LOA\textsubscript{95%} indicates that for 95% of participants, a measurement at follow-up assessment would be between 15.17 points less and 16.02 points greater (out of a total of 52 points) than the baseline measurement.

**Validity**

The validity of PCS-NP is supported by 1) its comprehensibility and negligible missing items, 2) structural validity interpreted by factor analyses, 3) cross-cultural validity by adherence to high standards of translation guidelines,\textsuperscript{46} and 4) concurrent validity by testing a priori hypotheses. We adapted and adhered to translation and cross-cultural adaptation guidelines by Beaton,\textsuperscript{46} which assured high standards in translation methodology. Importantly, we confirmed the concurrent validity of the PCS-NP based on our a priori hypotheses that PCS-NP would correlate moderately to strongly and significantly with the criterion variables. Association of PCS-NP with measure of depression in the current sample (r=0.56) is within the range previously reported (r=0.40–0.61).\textsuperscript{2,14–17,20} It is larger than those reported by the Chinese (r=0.40),\textsuperscript{2} similar to Korean (r=0.53),\textsuperscript{14} but smaller than Catalan (r=0.61, BDI).\textsuperscript{16} As the association with depression, PCS-NP in the current sample showed moderate negative associations with the measures of anxiety and lies in the range that is previously reported (r=0.53–0.82).\textsuperscript{2,14,15,17,20} Similarly, association of PCS-NP with measure of pain intensity is moderate, which also lies within the range previously reported (r=0.25–0.45).\textsuperscript{2,14–17,20–22} As the a priori hypotheses regarding expected correlations of PCS-NP with criterion variables were met in the current study, concurrent validity of the PCS-NP is supported.

**Strengths and limitations**

One strength of the study is that it was conducted and reported using the COSMIN checklist as a guide.\textsuperscript{37,40} As per the COSMIN recommendations, we used an adequate sample size (>100) and we confirmed participants’ stability during the interim period by the use of GRoC. However, the study also has limitations, which should be considered when interpreting the results. One limitation is the use of cross-sectional design. Given that the correlations of the PCS-NP with the criterion variables presented are based on cross-sectional data, we are not able to draw any causal conclusions regarding the impact of catastrophizing on pain intensity, depression, or anxiety, or vice versa. Instead, the findings support the possibility that catastrophizing may play as large a role in patient functioning in Nepal as found in Western countries. This indicates that research examining the causal relationships in this population is warranted. Furthermore, the study design was not a clinical trial, in which catastrophizing was systematically manipulated. This would have allowed us to evaluate the responsiveness of the PCS-NP to treatment and link changes in the PCS-NP scores with treatment outcomes. Future research to evaluate the responsiveness of the PCS-NP to treatment, as well as to better understand the meaning of change in the PCS-NP scores (eg, the change that represents a minimum important change) would be useful. Finally, we did not assess divergent (discriminant) validity of the PCS-NP, which is another important measure of construct validity. Research to evaluate this validity domain for the PCS-NP would be useful.

**Summary and conclusion**

Here, we developed a comprehensible, culturally appropriate, valid, and reliable PCS-NP for the use in Nepalese with chronic musculoskeletal pain or headache. The availability of this measure will facilitate more cross-cultural comparative study of the role that catastrophizing plays in patient with chronic pain.
For the clinical use of the PCS-NP, clinicians should consider PCS change score ≥7 as a true change beyond the measurement error. Future research should evaluate the discriminant validity of as well as the minimum important changes in the PCS-NP that is meaningful to patients, in order to better understand the validity and clinical significance of the PCS-NP.

Acknowledgments
The authors of the study would like to acknowledge: 1) the International Association for the Study of Pain (IASP) for providing a developing country collaborative research grant to SS and MPJ, which supported this study; 2) the developer of the PCS, Professor Michael Sullivan, for granting permission to translate the PCS to Nepali, reviewing the translation history, and approving the final version; 3) all the translators for volunteering to translate the scale; 4) the physiotherapy students (Anupa Pathak, Dipendra Thakur, Jyoti Jha, Manisha Lakhey, Riju Maharjan, Shree Krishna Shilpakar, and Soniya Maharjan) at Kathmandu University School of Medical Sciences who assisted with the data collection; and 5) all the study participants who volunteered to participate in the study and investing their time for the completion of questionnaires.

This project was supported in part by “Developing Countries Collaborative Research Grant” from the International Association for the Study of Pain awarded to MPJ and SS.

Author contributions
SS contributed to the conception, study design, data collection, data analysis, drafting of the first manuscript, and revision and approval of the final version. PT contributed to the conception, study design, and final approval of the manuscript. JHA contributed to the review of manuscript drafts and final approval of the manuscript. MPJ contributed to the conception, study design, and revisions and final approval of the manuscript. All authors contributed toward data analysis, drafting and critically revising the paper and agree to be accountable for all aspects of the work.

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

References