Cross-cultural adaptation, reliability, and validity of the German version of the Pain Catastrophizing Scale

Meyer, K; Sprott, H; Mannion, A F
Abstract

OBJECTIVE: In patients with chronic pain, catastrophizing is a significant determinant of self-rated pain intensity and disability. The Pain Catastrophizing Scale (PCS) was developed to assist with both treatment planning and outcome assessment; to date, no German version has been validated.

METHODS: A cross-cultural adaptation of the PCS into German was carried out, strictly according to recommended methods. A questionnaire booklet containing the PCS, visual analogue scales (numeric rating scale) for pain intensity and general health, the ZUNG self-rating depression scale, the Modified Somatic Perception Questionnaire (MSPQ), the Fear Avoidance Beliefs Questionnaire (FABQ), and the Roland-Morris (RM) disability questionnaire was completed by 111 patients with chronic low back pain (mean age, 49 years), 100 of which also completed it again 7 days later. RESULTS: Cronbach's alpha (internal reliability) for the three subsections of the PCS—helplessness, magnification, rumination—and for the whole questionnaire (PCSwhole) were .89, .67, .88, and .92, respectively. The intraclass correlation coefficients of agreement for the reproducibility were .81, .67, .78, and .80, respectively. The PCSwhole scores correlated with the other scores as follows: pain intensity r=.26, general health r=.29, ZUNG r=.52, MSPQ r=.53, FABQactivity r=.51, FABQwork r=.61 and RM r=.57. Factor analysis revealed three factors, with an almost identical factor structure to that reported in previous studies. CONCLUSION: The psychometric properties of our German version of the PCS were comparable to those reported in previous studies for the original English version. It represents a valuable tool in the assessment of German-speaking chronic low back pain patients.
Cross-cultural adaptation, reliability and validity of the German version of the Pain Catastrophizing Scale

Running Head: Pain Catastrophizing Scale: German Validation

1) Dept Rheumatology and Institute of Physical Medicine, Gloriatr. 25, University Hospital, 8091 Zürich, Switzerland

Dr Anne Frances Mannion (Senior)², Dr Haiko Sprott¹, Kathrin Meyer MPH (First author)¹

2) Spine Center, Lenghalde 2, Schulthess Klinik, 8008 Zürich, Switzerland

Correspondence:
Ms. Kathrin Meyer, MPH
U GEH 471
Dept Rheumatology and Institute of Physical Medicine
Gloriast. 25
8091 Zürich
Switzerland
Tel: +41 44 255 36 17
Fax: +41 44 255 43 88
Katharina.Meyer@usz.ch

Key words: Catastrophizing, chronic disease, low back pain, psychometrics, reproducibility of results, self assessment
**Abstract**

**Objective** In patients with chronic pain, catastrophizing is a significant determinant of self-rated pain intensity and disability. The Pain Catastrophizing Scale (PCS) was developed to assist with both treatment planning and outcome assessment; to date, no German version has been validated. **Methods** A cross-cultural adaptation of the PCS into German was carried out, strictly according to recommended methods. A questionnaire booklet containing the PCS, visual analogue scales (NRS) for pain intensity and general health, the ZUNG self-rating depression scale, the Modified Somatic Perception Questionnaire (MSPQ), the Fear Avoidance Beliefs Questionnaire (FABQ), and the Roland Morris (RM) disability questionnaire was completed by 111 patients with chronic low back pain (cLBP) (mean age 49 years), 100 of which also completed it again 7 days later. **Results** The Cronbach’s alpha (internal reliability) for the 3 sub-sections of the PCS — “helplessness”, “magnification”, “rumination” — and for the whole questionnaire (PCS\textsubscript{whole}) were 0.89, 0.67, 0.88 and 0.92 respectively. The intraclass correlation coefficients of agreement for the reproducibility were 0.81, 0.67, 0.78 and 0.80, respectively. The PCS\textsubscript{whole} scores correlated with the other scores as follows: Pain intensity \( r = 0.26 \), general health \( r = -0.29 \), ZUNG \( r = 0.52 \), MSPQ \( r = 0.53 \), FABQ\textsubscript{activity} \( r = 0.51 \), FABQ\textsubscript{work} \( r = 0.61 \) and RM \( r = 0.57 \). Factor analysis revealed 3 factors, with an almost identical factor structure to that reported in previous studies. **Conclusion** The psychometric properties of our German version of the PCS were comparable to those reported in previous studies for the original English version. It represents a valuable tool in the assessment of German-speaking cLBP patients.
Introduction

Chronic non-specific musculoskeletal pain is a burden for patients, and is associated with high socio-economic costs (1-3). The underlying construct of chronic pain is complex, and biopsychosocial factors influence both its development and its maintenance: psychological components, such as unhelpful pain cognitions (4), depression and fearful or catastrophizing thoughts, can influence perceived pain, quality of life (5), physical performance (4, 6), and subjective disability (6). Catastrophizing is defined as a maladaptive response to pain and is characterized by an experience of heightened pain intensity and difficulty in disengaging from pain (6); it is an important predictor of pain severity, and of how people cope with pain (7-9), and appears to predict future disability better than do other variables (6). Some studies have indicated that pain catastrophizing predicts depression or even mediates the reduction in depression, the perception of pain and the behaviour in response to cognitive-behavioral or graded-exercise therapy (10-12). Diminishing catastrophizing thoughts can positively influence coping with pain, and behavioral and cognitive traits (13-15). In psychological research, it has been shown that pain catastrophizing behavior can influence those involved with the catastrophizer, leading to over-cautious treatment decisions (16, 17).

For all these reasons, diminishing catastrophizing thoughts should constitute an important ingredient of therapy for chronic low back pain (LBP) (11). The Pain Catastrophizing Scale (PCS) was developed in the English language by Sullivan et al (18) to screen patients with catastrophizing thoughts, and to improve treatment planning, implementation and outcome assessment. The English version of the PCS has been well-investigated and its psychometric properties are good (18-21). A
systematic search of the literature revealed that, to date, no validated German version exists.

The aim of this study was to cross-culturally adapt the English version of the PCS into German and to evaluate its psychometric properties (internal consistency, construct validity, factor structure, reproducibility) in a large group of patients with LBP.

**Methods**

**The Pain Catastrophizing Scale (PCS)**

The PCS is a self-administered questionnaire that consists of 13 items to assess the extent of the patient’s catastrophizing thoughts and behaviors. It comprises three subscales: *helplessness, magnification and rumination*. The questionnaire is completed in relation to the patient’s thoughts and feelings when they are in pain. It contains questions concerning, for example, the degree to which the person worries all the time about whether the pain will end, thinks how awful and overwhelming it is, feels afraid that the pain will get worse, continually thinks how badly they want the pain to stop, etc. Each item is scored on a 5-point scale, with higher values representing greater catastrophizing (18). The scores for the subscales are given by the sum of the corresponding items, and the total score is computed by summation of all items. The PCS score ranges from 0 to 52 points.

**Translation and cross-cultural adaptation**

**Translation**

The translation into German and cross-cultural adaptation of the original English version of the PCS into German was carried out in accordance with previously published guidelines (22, 23). Three native German speakers (T-1, T-2, T-3) carried
out independent translations of the PCS from English to German. T-1 was a psychologist, T-2 was a professional translator and T-3 was a linguist. The forward translations were compared with one another and with the original English version. After discussing any discrepancies, the three versions were synthesized to form one common German version.

**Back-translation**

Two native English speakers with German as their second-language carried out a back-translation of the German version into English. Both back-translators were considered bilingual, according to the definition of Deyo, et al. (24). Neither of the back-translators were familiar with the subject matter of the questionnaire. A third bilingual person highlighted any conceptual errors or gross inconsistencies in the content of the translated versions, in preparation for the expert committee meeting.

**Expert Committee**

An expert committee was formed consisting of all of the translators and back-translators, 1 methodologist, 1 clinical research scientist and the originator of the English version of the PCS (M.J. Sullivan). The task of this expert committee was to assure semantic and idiomatic equivalence and experiential and conceptual equivalence (i.e. to address any peculiarities specific to the cultures examined) between the German and English versions of the questionnaire. A “pre-final” version of the German PCs was produced by the expert committee.

**Test of the pre-final version**

A group of 15 patients with LBP were given the pre-final version of the PCS questionnaire to complete. They were briefly interviewed in order to check what they thought was meant by each question and the chosen response. They were also asked for their general comments on the questionnaire. All the findings were
evaluated by the work-group (to assess face validity), after which two questions were slightly modified to achieve the final German version of the PCS.

Methods to assess the psychometric properties of the German version of the PCS

Questionnaire battery
The patients were asked to complete a questionnaire booklet, which contained the German version of the PCS and a series of other questionnaires or questions intended to assess the PCS’s convergent validity (a subcategory of construct validity). From the literature, interrelationships were expected between pain-related catastrophizing and various other variables. For example, previous studies have found low to moderate positive correlations (ranging from 0.26 to 0.64) between catastrophizing and depression (18), fear of activity (25), disability (26) and poor coping style (5). Further, catastrophizers were shown to have 3 to 5 times greater emotional distress and higher pain intensity than non-catastrophizers (P<0.01) (18).

To cover these constructs we chose the following questionnaires for inclusion in the questionnaire booklet: the ZUNG self-rating depression scale (ZUNG), a screening instrument to assess depression (27, 28); the Modified Somatic Perception Questionnaire (MSPQ) (27, 29, 30), to assess heightened somatic awareness or anxiety; the Fear Avoidance Beliefs Questionnaire (FABQ), to measure fear avoidance beliefs in relation to work and to physical activity (31, 32) and the Roland and Morris Questionnaire (RM) (33-35) to measure disability in everyday activities due to LBP. The latter, 24-item questionnaire (35) was cross-culturally adapted from the original version (36) with a slight modification, to include yes/no categories for each item (rather than just “tick if applicable” in the original English version). The questionnaire booklet also contained 0-10 graphic numeric rating scales (NRS) for
the intensity of back and leg pain and a 5-point ordinal scale for general disability (“none” to “very severe”).

Patients

132 patients were recruited from the Department of Rheumatology of the University Hospital. Inclusion criteria included LBP for > 3 months, with the diagnosis of either non-specific LBP or specific LBP in the ICD10 categories M 40-M 80, and ability to read and understand German. The only exclusion criterion was serious immediate life-threatening diseases. The patients gave their signed informed consent to participate and the study was approved by the local Ethics Committee (Application-No. 588 of the Cantonal Ethic Committee).

Statistical analysis

Floor and ceiling effects were determined in two ways: firstly, in the traditional manner of calculating the number of individuals obtaining, respectively, the lowest (0) or highest (52) possible PCS scores, where a limit of 15% of patients should not be exceeded (37, 38); and secondly, by computing the proportion of individuals obtaining a score within the limits of the minimum detectable change (95% confidence interval) at the two ends of the scale.

Concurrent validity, a component of construct validity, indicates the extent to which the instrument’s scores relate to those of other instruments in the manner expected. The authors hypothesized that the PCS would measure aspects of the patient’s health/complaints that were different from but related to those measured by the other questionnaires (see earlier), which should result in moderate positive correlation coefficients, not exceeding 0.7. The Spearman Rank correlation coefficient was used
to compare the relationships between the PCS and the FABQ activity, FABQ work, RM, MSPQ and ZUNG questionnaires.

*Internal consistency was assessed with factor analysis.* Factor analysis uncovers the latent structure (dimensions) of the items within the instrument. In the present study, two methods were used: 1) a factor analysis with factor extraction by oblique rotation, as performed by Sullivan et al (18) in developing the original PCS; and 2) a principal components analysis with components extraction by varimax rotation. Principal components analysis decomposes the original data into a set of linear variates (39), whereas factor analysis derives a mathematical model from which factors are estimated. Principal components analysis is concerned only with establishing which linear components exist within the data and how a particular variable might contribute to each component. In contrast, factor analysis only estimates the underlying factors (40). The varimax rotation assumes that the extracted factors are orthogonal, that is, not correlated, whereas the oblique rotation allows for such correlation.

Internal consistency was also assessed with Cronbach’s alpha, using the data from the baseline questionnaires. Cronbach’s alpha indicates the strength of the relationship between all the items within the test instrument and indicates whether the items are sufficiently interrelated to justify their combination in an assessment-instrument (41). The PCS has 3 subscales, and the Cronbach’s alpha was reported for each separately; however, for the purposes of comparison with the original English version (18), the internal consistency was also reported for the whole scale, despite the fact it is not theoretically correct to do so, since Cronbach’s alpha indicates the correlation among items that measure one single construct and the PCS is a multidimensional scale.
Reproducibility indicates the extent to which the same results are obtained on repeated administrations of a given instrument when no change is expected. The general health, back pain, and disability questions were used to confirm whether the health status of the study population was the same at baseline and at retest, and this was assessed by means of paired sample t-tests. The intraclass correlation coefficient of agreement (ICC), and the standard error of measurement (SEM) for the repeated trials were computed to examine the reproducibility of measurements (42, 43). To compute the ICC, the single measurement approach was used. SEM was calculated as follows: \[ SEM = s \sqrt{(1-r)} \], where \( s \) is the standard deviation of the PCS and \( r \) is the ICC-value of absolute agreement in the test for reproducibility (44). The SEM was used to indicate the “minimum detectable change” (MDC95%) for the PCS i.e. the degree of change required in an individual’s score, in order to establish it (with a given level of confidence) as being a “real change”, over and above measurement error (45). At the 95% confidence level, this is defined as \( 1.96 \times \sqrt{2} \times SEM \), which is equivalent to \( \pm 2.77 \times SEM \). Statistical significance was accepted at the \( p<0.05 \) level. No adjustments were made for multiple testing, as previously recommended (46).

Results

Cross cultural adaptation of the PCS

The final version of the German PCS is shown in Appendix 1. A few noteworthy difficulties arose during the development of the German version of the PCS, as follows.

The introduction section and response categories of the English PCS: The introduction and the response categories are a mixture of frequency ("all the time") and intensity/amount ("to a slight degree") options. This was discussed with the
originator of the PCS (MJ Sullivan) and the following compromise was found: the translation of the introduction phrase focused on whether “experiencing the listed thoughts and feelings applied to the patient”, with the possible answer options ranging from “doesn’t apply at all” through to “applies always”.

Some expressions (in item 4, 5, 11, 12 and 13) were considered a little vague/ambiguous or too complicated in their structure, and the most appropriate translation into German was hence a matter of finding consensus within the expert group. A more appropriate translation was found by further communication with the translators, back-translators and originator of the PCS (MJ Sullivan).

**Pre-test of the final version**

The general comments of the 15 patients who pre-tested the questionnaire indicated that the wording was easy to understand and the layout was good. No ambiguities prevented the answering of the questions. Some patients with a very low level of catastrophizing mentioned that some phrases of the PCS were a little “over-the-top”, but, interestingly, the catastrophizers did not consider these terms as inappropriate. This indirectly substantiated the validity of the questionnaire.

**Study sample**

Of the 132 patients eligible, 111 patients signed the informed consent letter and returned the baseline questionnaire booklet (84.1% return rate). They had a mean (SD) age of 49 (16) years. 75 (68%) were female and 36 (32%) male. The analyses of internal consistency, factor analysis, and concurrent validity were carried out on these 111 data sets. Of the 111 participants, 6 failed to return the second questionnaire booklet and 5 had a retest interval of more than 21 days. Hence, 100 data sets were available for the reproducibility analysis (90.1%). The mean interval between the two completions of the questionnaire booklet was 7.0 (SD 6.2) days.
The most common diagnosis in the study population was non-specific chronic LBP (n= 73), followed by herniated disc (n= 21) and spinal stenosis (n=5). Generally, the patients showed moderate back pain complaints (5.2, on a 0 to 10 NRS) and moderate disability (10.8, on the 0 to 24 RM) at baseline. PCS scores were low to moderate, with mean scores ranging from 29.8% (magnification subscale) to 39.4% (rumination subscale) of the highest possible score (Table 1).

**Missing data, normality of score distribution at baseline**

The following “missing data” rules (derived from standard principles in questionnaire scoring (Achim Elfering, personal communication)) were applied to the scoring of the PCS: 1 missing value was allowed for the helplessness-subscale, and no missing values were allowed for either the magnification or the rumination subscales. For scoring the total score, 2 missing values were allowed. For the individual items, there were between 0 to 2 missing values, and for the whole scale 0.9 % data were missing. As long as the missing rules were not contravened, the scores for the whole scale or the sub-scales were extrapolated from the mean value of the remaining responses. Scoring the subscales was possible after imputation in 99% cases for the “helplessness” sub-scale, 98% for “magnification”, 96% for “rumination" and 99% for the total scale. Using the traditional approach, moderate floor effects but minimal ceiling effects were found: the lowest possible scores were found in 13.2% of the cases for the “helplessness” sub-scale, 11.2% for “magnification”, 10.4% for "rumination" and 12.3% for the total scale. Highest possible scores were only found for the "rumination" subscale, at a prevalence of 1.9%. However, using the perhaps more relevant approach of examining the proportion of patients with scores lying within the range of the MDC95% at the two ends of the scale, there were 33.7% floor effects for the total scale, 48.1% for “helplessness”, 52.9% for “magnification” and
31.7% for “rumination”; the corresponding ceiling effects were 3.8% for the total scale, 5.8% for “helplessness”, 2.9% for “magnification” and 8.7% for “rumination”.

**Construct validity: Relationship between PCS and other parameters at baseline**

Overall, low to moderate correlations were found between the PCS whole scale scores and the scores of the other scales (Table 2), with positive correlation coefficients ranging from 0.31 (for the correlation with pain) to 0.61 (with FABQ-work), and with a negative correlation (r=-0.38) being recorded with general health. The PCS-helplessness scale correlated better with the scores of the other scales than did the magnification and rumination subscales (Table 2).

To assess whether these correlations were influenced by other factors, several subgroup analyses were carried out: these revealed that the correlation coefficients between the PCS and the other scales were not dependent on gender, the severity of LBP, general health, or the level of fear-avoidance beliefs (i.e., the correlation coefficients were similar for men and women, high pain intensity and low pain intensity groups, etc.; results not shown).

**Exploratory Factor Analysis**

The principal component analysis revealed a 3-factor structure similar to that found by the originator of the PCS (18): Helplessness = items 1 to 5; Rumination = items 8 to 12; Magnification = items 6, 7, 13. Unlike Sullivan (18), item 12 scored higher on the rumination factor than the helplessness factor (see Table 3). The second model, created using factor analysis with oblique rotation, suggested a 2-factor structure similar to that reported by Osman (20).
Confirmatory factor analysis, as described by van Damme (21) was used to evaluate model validity for 4 models: the 3-factor-model reported by Sullivan (18); the 2-factor model, reported by Osman (20); a single factor model and the 3-factor model revealed by the current principal components analysis. To determine which of the models provided the best fit to the data, several goodness-of-fit measures were computed: 1) Bollen's incremental fit index (IFI) (47), in which IFI-values close to 1 indicate a very good fit; 2) RMSEA, the root mean square error (48, 49), a fit-measure based on the population error of approximation. A RMSEA-value of 0.0 indicates a perfect fit and values around 0.8 or less indicate a reasonable error of approximation (48); 3) Bentler’s comparative fit index (CFI) (50), an incremental fit index (51). This represents the proportionate improvement in model fit by comparing the default model with an independence model (a null model in which all the observed variables are uncorrelated). CFI-values close to 1 indicate a very good fit. The 3-factor-model, revealed by the current principal components analysis, elicited the best goodness of fit-measures as judged by the IFI, RMSEA and CFI values (Table 4), closely followed by the 3-main axis factoring model reported by Sullivan (18) and Osman (19).

Internal consistency of the PCS at baseline

Good internal consistency was found, with Cronbach’s alpha values of 0.89 (helplessness), 0.67 (magnification), 0.88 (rumination) and 0.92 (total scale). Reliability testing further revealed good factor stability, with only a negligible gain in Cronbach’s alpha values by omitting items 7 (0.926 instead 0.923) and 12 (0.905 instead 0.890).
Reproducibility of PCS

Reproducibility analysis was conducted on the data from the 100 patients with a return interval for the second questionnaire booklet of 21 days or less. General health, LBP and disability did not differ significantly between the two assessments timepoints. The mean difference between repeated measures for the PCS and its subscales, and ICCs, SEMs, and the MDCs, are shown in Table 5. Acceptable to good reproducibility was found and SEM-values ranged from 1.5 to 4.6 (Table 5).

Reproducibility of FABQ, ZUNG, MSPQ, RM

Good to excellent retest reliability was found for the other scales included in the questionnaire booklet (see Table 6). Several subgroup analyses were conducted to investigate whether there was any influence of factors such as short versus long retest interval, gender, high versus low fear-avoidance groups, on the SEM values derived. No such subgroup differences were found (all p>0.05; results not shown). The results of all the cross-sectional analyses (e.g., the correlation coefficients between the PCS-scores and the scores on the other scales, and the results of the factor analysis), showed good robustness over time: there were no notable differences between the results derived using the retest data as opposed to the baseline data.

Discussion

Translation and cross-cultural adaptation of the PCS

The aim of the present study was to cross-culturally adapt the PCS, for use with German-speaking patients, and to examine the psychometric properties of the German version produced. Overall, the German version of the PCS showed good
psychometric properties. In the following, the translation process and the results concerning validity and reliability will be discussed.

**Adaptation of a German version of the PCS**

The process of translating and back-translating the English PCS was carried out strictly in accordance with established guidelines (23). The only real problems arose in the translation of the response categories for the scale. We chose a more consistent wording in relation to how much each statement (i.e., having particular thoughts or feelings) applied to the patient. This might need to be borne in mind if the questionnaire is translated into other languages, and we recommend that a similar approach is followed as for the German version.

The study was conducted with patients living in the German speaking part of Switzerland. There are very few grammatical or semantic differences in the use of the written German language amongst the German-speaking countries. Thus, we believe that the current version can most likely be used without difficulty in other German-speaking European countries.

**Construct validity of the PCS**

The German PCS showed good construct validity. Convergent validity was examined by investigating the strength of the relationship between the PCS scores and the scores for other pain-related constructs such as fear-avoidance, disability, depression, and somatization. Similar to the findings of previous studies (5, 18, 25, 26) correlation coefficients for these relationships ranged from 0.31-0.61. This represents moderate agreement, which confirms that the PCS assesses a different construct, but one that is related to the above-mentioned constructs, and it can thus be considered suitable as part of the multidimensional battery of assessments in
chronic LBP patients. Nonetheless, the fact that some correlations approached or exceeded 0.5 suggests there may be some redundancy amongst the measures (52). In future studies, the uniqueness of the construct “pain catastrophising” should be investigated using multiple regression models.

Factor analysis as a further tool of construct validity

The principle component factor analysis of the German PCS largely replicated the results of former studies and showed that the 3 factor-solution was reliable in terms of its construct validity. Principle components analysis with varimax rotation revealed almost the same factor structure as that proposed by Sullivan (18). That a 3-factor solution was preferable to either a 1-factor or 2-factor model confirmed the findings of Osman and Van Damme (19-21). Although we found a slightly better fit for the factor structure with item 12 in the subscale rumination, the improvement was very small when comparing the fit indices for the two different models using a confirmatory factor analysis. The incremental fit index, for example changed from 0.942 for our model with item 12 in the rumination scale to 0.936 with the item 12 in the helplessness scale (as in the original model). This small increment would not warrant recommendation of a German version with a different factor structure to that of the English version (i.e. with item 12 in another subscale), especially in view of the consequences regarding comparability with existing and future studies in English and German speaking countries. The items contained in each of the three subscales should therefore remain as proposed by Sullivan (18), with items 1 to 5 and 12 in the subscale helplessness, items 6,7, and 13 in magnification and items 8 to11 in rumination.

The scores for the PCS are given by the sum of specific items for the subscales, or by the sum of all items for the total score. We consider the latter to be somewhat
problematic, because the PCS actually comprises 3 individual sub-domains. Hence, in future studies, it is recommended that the scores for the subscales and the total score always be reported separately.

Internal consistency of the PCS

The internal consistency of the German PCS was examined using Cronbach’s alpha, an item correlation test that reflects the homogeneity of all the items. The Cronbach’s alphas for the sum-scales and the total scale (0.67 to 0.92) were slightly higher than those reported in the original study of Sullivan (between 0.60 and 0.87) (18), and comparable to or a little lower than those previously reported for samples of outpatients and students (19). The similar values in different samples (19) is further endorsement of the good internal consistency of the German version of the questionnaire. All studies showed the lowest values of Cronbach’s alpha for the subscale magnification. This is most likely due to the fewer items in this sub-scale, since Cronbach’s alpha typically increases with increasing number of items in the scale (53); nonetheless, it might also question whether the subscale magnification can be used reliably as an independent instrument. Cronbach’s alphas greater than 0.8 are generally recommended for psychometric scales (44) although for individual patient assessments in the clinical situation, an alpha coefficient of at least 0.9 is recommended (54). Thus, from this perspective, the German PCS for the subscales helplessness and rumination is suitable not only for group analyses but also for the interpretation of individual scores. As mentioned in the methods, it should be noted that the determination of a single alpha coefficient for the 13-item scale as a whole is not theoretically correct, because by definition Cronbach’s alpha indicates the correlation among items that measure one single construct and the PCS is a scale with three dimensions. However, we present it here
for better comparability with the original study (18), where the Cronbach’s alpha for the total scale was also given.

Reproducibility of the PCS

The German version of the PCS showed reasonably good ICCs, ranging from 0.67 to 0.81. ICCs greater than 0.7 are generally considered acceptable, greater than 0.8, good, and over 0.9, excellent (55), although it is also acknowledged that the ICC is highly dependent on the between-subject variance in the group of subjects under investigation (56). Our sample did not show any significant differences from test to retest in general health, disability or back pain, justifying the application of a reproducibility analysis. The ICCs reported in the present study for the whole scale German PCS were slightly higher than those reported for the original English version of the PCS (ICC= 0.75) (18), although in the latter study the longer time interval of 6 weeks between test and retest may have increased the variability and decreased the ICC. The SEM is another expression of the error associated with repeated measurements. To the authors’ knowledge, this measure has not been investigated in previous studies. In the present study, the SEM was 4.6 points for the PCS total. At the 95% confidence level, this would yield a minimal detectable change of 12.8 points. It remains to be seen whether clinical treatments directed at reducing catastrophizing thoughts could effect changes over and above this measurement error. This requires further investigation, since it has implications for the sensitivity (responsiveness) of the instrument.

Limitations of this study

Concerns about floor effects
Using the traditional approach, we detected a floor effect of 12% and no ceiling effect for the PCS-total-score. However, with the more sensible scale width approach, using the minimal detectable change at the two ends of the scale, we detected a noticeable floor effect of 33.7% for the total scale and even greater floor effects for the subscales. This may indicate a scaling problem of the PCS and it suggests that the PCS would not be able to identify/detect changes in catastrophizing thoughts in pain patients with very low initial PCS-scores. Floor effects are highly dependent on the patient group under investigation. Our sample of hospital outpatients had, on average, only moderate pain and disability, and they were not attending any specialized cognitive-behavioural treatment program. For patients with more serious complaints and a longer, more chronic illness history (for example those needing treatment in an intensive pain coping program) a lower proportion of floor effects might be expected. We recommend that further research should focus on more severely disabled German-speaking patients, to evaluate whether there are fewer floor effects in such patient groups or whether the PCS does indeed have a scaling problem. Despite the floor effect, an important aim of the PCS still appears to be feasible, namely the detection of patients with pronounced catastrophizing thoughts and the ability to reliably re-assess them after treatment.

**Conclusion**

The psychometric properties of our German version of the PCS were comparable to those reported by Osman et al and Van Damme et al (19-21), and exceeded those of the original English version (18). The PCS showed good internal consistency, and the 3-factor structure, reported in previous studies, could be replicated. It also showed acceptable to good reproducibility, with a minimal detectable change score of approximately 13 points. Tests of concurrent validity showed that it represents a
different construct compared with existing LBP-related questionnaires, but has the desired overlap. The PCS represents a valuable tool for use in scientific studies and in the clinical setting in patients with chronic LBP in German-speaking countries. Future studies should investigate whether the PCS is sufficiently sensitive to detect a change in catastrophizing thoughts over time, after specific treatment modalities.

**Acknowledgments**

This project was funded by the National Research Programme NRP53 "Musculoskeletal Health - Chronic Pain" of the Swiss National Science Foundation (Project 405340-104787/2). We thank Gabi Umbricht, Astrid Junge, Marco Saltzmann, Sue Huber and Vivien Robinson for their exact and unremitting work in performing the forward and back translations, and Daniel Uebelhart for providing the infrastructure within our Research Unit to carry out this study. Further thanks go to Sven Hoffman and Alex Tschopp, for their statistical advice, Susi Spillmann and Nicole Wassmer for recruiting patients, and Leanne Pobjoy for her help in preparing the referencing for the manuscript. We would also like to thank Mick Sullivan for his helpful advice and guidance during the cross-cultural adaptation procedure.
References


52. Terwee CB, Van der Slikke RM, Van Lummel RC, Benink RJ, Meijers WG, De Vet HC. Self-reported physical functioning was more influenced by pain than performance-based physical functioning in knee-osteoarthritis patients. Journal of Clinical Epidemiology 2006;59:724-31.


Table 1: Questionnaire scores at baseline

<table>
<thead>
<tr>
<th></th>
<th>PCS-help</th>
<th>PCS-magn</th>
<th>PCS-rum</th>
<th>PCS-total</th>
<th>FABQ-act</th>
<th>FABQ-work</th>
<th>ZUNG</th>
<th>MSPQ</th>
<th>RM</th>
<th>LBP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>7.6</td>
<td>3.6</td>
<td>6.3</td>
<td>17.6</td>
<td>11.6</td>
<td>18.4</td>
<td>19.1</td>
<td>10.1</td>
<td>10.8</td>
<td>5.2</td>
</tr>
<tr>
<td>SD</td>
<td>5.4</td>
<td>2.5</td>
<td>4.0</td>
<td>10.5</td>
<td>6.7</td>
<td>11.8</td>
<td>10.6</td>
<td>5.9</td>
<td>5.9</td>
<td>2.5</td>
</tr>
<tr>
<td>max</td>
<td>24</td>
<td>12</td>
<td>16</td>
<td>52</td>
<td>24</td>
<td>42</td>
<td>60</td>
<td>39</td>
<td>24</td>
<td>10</td>
</tr>
<tr>
<td>in %</td>
<td>31.7</td>
<td>29.8%</td>
<td>39.4%</td>
<td>33.8%</td>
<td>48.3%</td>
<td>43.8%</td>
<td>31.8%</td>
<td>25.9%</td>
<td>45.0%</td>
<td>52.0%</td>
</tr>
</tbody>
</table>

PCS-help = PCS helplessness subscale; PCS-magn = PCS magnification subscale; PCS-rum = PCS rumination subscale, PCS-total = PCS total score; FABQ-act = FABQ activity subscale; FABQ-work = FABQ work subscale; ZUNG = ZUNG score; MSPQ = MSPQ-score; RM = RM score; LBP = current back pain; SD = standard deviation; max = highest possible score; in % = mean value in proportion to the highest possible score.
Table 2: Concurrent validity

<table>
<thead>
<tr>
<th></th>
<th>Pchelp</th>
<th>Pcs-magn</th>
<th>Pcs-rum</th>
<th>Pcs-tot</th>
<th>Facact</th>
<th>Facwor</th>
<th>Zung</th>
<th>Mspq</th>
<th>Rm</th>
<th>Lbpav</th>
<th>Lpav</th>
<th>Hilbp</th>
<th>Lolbp</th>
<th>Nowlbp</th>
<th>Genheal</th>
<th>Disabil</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pchelp</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pcs-magn</td>
<td>0.54</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pcs-rum</td>
<td>0.75</td>
<td>0.62</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pcs-tot</td>
<td>0.93</td>
<td>0.73</td>
<td>0.91</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Facact</td>
<td>0.50</td>
<td>0.34</td>
<td>0.50</td>
<td>0.52</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Facwor</td>
<td>0.62</td>
<td>0.42</td>
<td>0.50</td>
<td>0.61</td>
<td>0.55</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zung</td>
<td>0.58</td>
<td>0.23</td>
<td>0.42</td>
<td>0.52</td>
<td>0.35</td>
<td>0.49</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mspq</td>
<td>0.53</td>
<td>0.33</td>
<td>0.47</td>
<td>0.53</td>
<td>0.32</td>
<td>0.49</td>
<td>0.64</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rm</td>
<td>0.62</td>
<td>0.31</td>
<td>0.48</td>
<td>0.57</td>
<td>0.54</td>
<td>0.69</td>
<td>0.55</td>
<td>0.50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lbpav</td>
<td>0.52</td>
<td>0.22</td>
<td>0.30</td>
<td>0.44</td>
<td>0.30</td>
<td>0.53</td>
<td>0.52</td>
<td>0.56</td>
<td>0.57</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lpav</td>
<td>0.37</td>
<td>0.18</td>
<td>0.26</td>
<td>0.35</td>
<td>0.26</td>
<td>0.41</td>
<td>0.34</td>
<td>0.47</td>
<td>0.56</td>
<td>0.54</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hilbp</td>
<td>0.37</td>
<td>0.21</td>
<td>0.28</td>
<td>0.35</td>
<td>0.23</td>
<td>0.46</td>
<td>0.45</td>
<td>0.54</td>
<td>0.51</td>
<td>0.75</td>
<td>0.58</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lolbp</td>
<td>0.40</td>
<td>0.13</td>
<td>0.18</td>
<td>0.31</td>
<td>0.25</td>
<td>0.47</td>
<td>0.35</td>
<td>0.38</td>
<td>0.56</td>
<td>0.75</td>
<td>0.60</td>
<td>0.54</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nowlbp</td>
<td>0.47</td>
<td>0.15</td>
<td>0.28</td>
<td>0.39</td>
<td>0.33</td>
<td>0.52</td>
<td>0.54</td>
<td>0.53</td>
<td>0.59</td>
<td>0.78</td>
<td>0.54</td>
<td>0.68</td>
<td>0.70</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Genheal</td>
<td>-0.41</td>
<td>-0.26</td>
<td>-0.32</td>
<td>-0.38</td>
<td>-0.29</td>
<td>-0.36</td>
<td>-0.50</td>
<td>-0.43</td>
<td>-0.48</td>
<td>-0.31</td>
<td>-0.26</td>
<td>-0.33</td>
<td>-0.23</td>
<td>-0.42</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disabil</td>
<td>0.50</td>
<td>0.25</td>
<td>0.46</td>
<td>0.48</td>
<td>0.42</td>
<td>0.54</td>
<td>0.51</td>
<td>0.54</td>
<td>0.69</td>
<td>0.72</td>
<td>0.60</td>
<td>0.69</td>
<td>0.57</td>
<td>0.69</td>
<td>0.89</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Bivariate Spearman-rank correlation coefficients

Pcs-help= PCS helplessness subscale; Pcs-magn= PCS magnification subscale;
Pcs-rum= PCS rumination subscale; Pcs-total= PCS total score; Fabact= FABQ activity subscale; Fabwor = FABQ work subscale; Zung= ZUNG score; Mspq= MSPQ-score; Rm= RM score; Lbpav=average low back pain last 7 days, Lpav= average leg pain, Hilbp= highest back and leg pain, Lolbp= lowest back and leg pain, Nowlbp=actual back and leg pain, Genheal= general health, Disabil= disability because of pain
Table 3: PCS-Factor structure by principal components analysis

<table>
<thead>
<tr>
<th>Component</th>
<th>4</th>
<th>2</th>
<th>5</th>
<th>3</th>
<th>1</th>
<th>12</th>
<th>10</th>
<th>9</th>
<th>11</th>
<th>8</th>
<th>6</th>
<th>13</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>.286</td>
<td>.269</td>
<td>.327</td>
<td>.238</td>
<td>.186</td>
<td>.745</td>
<td>.676</td>
<td>.648</td>
<td>.633</td>
<td>.587</td>
<td>.265</td>
<td>.401</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>.141</td>
<td>.112</td>
<td>.144</td>
<td>.297</td>
<td>.502</td>
<td>.382</td>
<td>.436</td>
<td>.369</td>
<td>.312</td>
<td>.817</td>
<td>.662</td>
<td>.589</td>
<td></td>
</tr>
</tbody>
</table>

Extraction Method: Principal components analysis; varimax rotation with Kaiser normalization. Rotation converged in 7 iterations; values below 0.1 are suppressed.

The model explained 69.6% of the variance
<table>
<thead>
<tr>
<th>Model</th>
<th>Description</th>
<th>IFI</th>
<th>RMSEA</th>
<th>CFI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>One factor (13 items)</td>
<td>0.832</td>
<td>0.140</td>
<td>0.827</td>
</tr>
<tr>
<td>Model 2</td>
<td>Two oblique factors (7+6 items)</td>
<td>0.855</td>
<td>0.134</td>
<td>0.850</td>
</tr>
<tr>
<td>Model 3</td>
<td>Three oblique factors (6+3+4 items)</td>
<td>0.936</td>
<td>0.090</td>
<td>0.933</td>
</tr>
<tr>
<td>Model 4</td>
<td>Three varimax components (5+3+5 items)</td>
<td>0.942</td>
<td>0.086</td>
<td>0.940</td>
</tr>
</tbody>
</table>

Model 2 = factor structure suggested by Osman [20]; model 3 = model suggested by Sullivan [18]; model 4 = model revealed by the current study. IFI = incremental fit index, RMSEA = Root mean square error of approximation, CFI = comparative fit index.
### Table 5: Reproducibility of the PCS

<table>
<thead>
<tr>
<th></th>
<th>T2-t1</th>
<th>ICC</th>
<th>SEM</th>
<th>MDC</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCS-helplessness</td>
<td>-0.8 (24)</td>
<td>0.81</td>
<td>2.33</td>
<td>6.46</td>
</tr>
<tr>
<td>PCS-magnification</td>
<td>-0.2 (12)</td>
<td>0.67</td>
<td>1.45</td>
<td>4.02</td>
</tr>
<tr>
<td>PCS-rumination</td>
<td>-0.6 (16)</td>
<td>0.78</td>
<td>1.82</td>
<td>5.06</td>
</tr>
<tr>
<td>PCS-total scale</td>
<td>-1.5 (52)</td>
<td>0.80</td>
<td>4.61</td>
<td>12.79</td>
</tr>
</tbody>
</table>

T2-t1 = mean values at t1 subtracted from t2, (in parentheses) = highest possible score for the given attribute; ICC = Intraclass correlation coefficient of agreement (t1*t2); SEM = Standard error of measurement; MDC = minimal detectable change.
<table>
<thead>
<tr>
<th></th>
<th>t2-t1</th>
<th>ICC</th>
<th>SEM</th>
<th>MDC</th>
</tr>
</thead>
<tbody>
<tr>
<td>FABQ-Activity</td>
<td>0.0 (24)</td>
<td>0.82</td>
<td>2.84</td>
<td>7.88</td>
</tr>
<tr>
<td>FABQ-Work</td>
<td>-1.0 (42)</td>
<td>0.88</td>
<td>4.14</td>
<td>11.48</td>
</tr>
<tr>
<td>Zung</td>
<td>-1.0 (69)</td>
<td>0.90</td>
<td>3.42</td>
<td>9.48</td>
</tr>
<tr>
<td>MSPQ</td>
<td>-0.3 (39)</td>
<td>0.87</td>
<td>2.12</td>
<td>5.87</td>
</tr>
<tr>
<td>RM</td>
<td>0.2 (24)</td>
<td>0.90</td>
<td>1.91</td>
<td>5.29</td>
</tr>
</tbody>
</table>

t2-t1= mean values at t1 subtracted from t2, (in parentheses) = highest possible score for the given attribute; ICC= Intraclass correlation coefficient of agreement (t1*t2); SEM= Standard error of measurement; MDC= minimal detectable change.
Appendix: German version of the Pain Catastrophizing Scale

Hier finden Sie verschiedene Fragen vor. Bitte lesen Sie jeweils die Einleitung und füllen Sie alle nachfolgenden Fragen aus.


Die folgenden dreizehn Sätze beschreiben verschiedene Gedanken und Gefühle, die bei Schmerzen auftreten können. Bitte markieren Sie auf der folgenden Skala, wie stark diese Gedanken und Gefühle auf Sie zutreffen, wenn Sie Schmerzen haben.

Wenn ich Schmerzen habe, beschäftigen mich folgende Gedanken...

<table>
<thead>
<tr>
<th></th>
<th>trifft überhaupt nicht zu</th>
<th>trifft eher nicht zu</th>
<th>Teils-teils</th>
<th>trifft eher zu</th>
<th>trifft immer zu</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Ich mache mir ständig Sorgen, ob die Schmerzen wohl jemals wieder aufhören werden.</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>2. Ich denke, ich kann nicht mehr.</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>3. Der Zustand ist schrecklich und ich denke, dass es nie mehr besser wird.</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>
4. Der Zustand ist furchtbar und droht mich zu überwältigen.  

5. Ich habe das Gefühl, ich halte es nicht mehr aus.  

6. Ich bekomme Angst, dass die Schmerzen noch stärker werden.  


8. Ich wünsche mir verzweifelt, dass die Schmerzen weggehen.  

9. Ich kann nicht aufhören, an die Schmerzen zu denken.  

10. Ich denke ständig daran, wie sehr es schmerzt.  


12. Es gibt nichts, was ich tun kann, um die Schmerzen zu lindern.  

13. Ich mache mir Sorgen, dass die Schmerzen auf etwas Schlimmes hindeuten.