Scandinavian Journal of Pain xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Scandinavian Journal of Pain

journal homepage: www.ScandinavianJournalPain.com

Observational study

Item response theory analysis of the Pain Self-Efficacy Questionnaire

Daniel S.J. Costa[a,b,*], Ali Asghari[a,c], Michael K. Nicholas[a]

[a] Pain Management Research Institute, University of Sydney at Royal North Shore Hospital, Sydney, Australia
[b] Sydney Medical School, University of Sydney, Australia
[c] School of Psychology, University of Shahed, Tehran, Iran

HIGHLIGHTS

- The Pain Self-Efficacy Questionnaire (PSEQ) has strong psychometric properties.
- Most PSEQ items perform well in item response theory analysis.
- Item 7 (coping without medication) performed poorly but has clinical utility.
- The PSEQ is a useful tool for assessing self-efficacy in people with pain.

ARTICLE INFO

Article history:
Received 16 June 2016
Received in revised form 1 August 2016
Accepted 2 August 2016
Available online xxx

Keywords:
Pain self-efficacy
Item response theory
Pain Self-Efficacy Questionnaire

ABSTRACT

Background and aims: The Pain Self-Efficacy Questionnaire (PSEQ) is a 10-item instrument designed to assess the extent to which a person in pain believes s/he is able to accomplish various activities despite their pain. There is strong evidence for the validity and reliability of both the full-length PSEQ and a 2-item version. The purpose of this study is to further examine the properties of the PSEQ using an item response theory (IRT) approach.
Methods: We used the two-parameter graded response model to examine the category probability curves, and location and discrimination parameters of the 10 PSEQ items. In item response theory, responses to a set of items are assumed to be probabilistically determined by a latent (unobserved) variable. In the graded-response model specifically, item response threshold (the value of the latent variable for which adjacent response categories are equally likely) and discrimination parameters are estimated for each item. Participants were 1511 mixed, chronic pain patients attending for initial assessment at a tertiary pain management centre.
Results: All items except item 7 ('I can cope with my pain without medication') performed well in IRT analysis, and the category probability curves suggested that participants used the 7-point response scale consistently. Items 6 ('I can still do many of the things I enjoy doing, such as hobbies or leisure activity, despite pain'), 8 ('I can still accomplish most of my goals in life, despite the pain') and 9 ('I can live a normal lifestyle, despite the pain') captured higher levels of the latent variable with greater precision.
Conclusions: The results from this IRT analysis add to the body of evidence based on classical test theory illustrating the strong psychometric properties of the PSEQ. Despite the relatively poor performance of Item 7, its clinical utility warrants its retention in the questionnaire.
Implications: The strong psychometric properties of the PSEQ support its use as an effective tool for assessing self-efficacy in people with pain.

1. Introduction

The Pain Self-Efficacy Questionnaire [1,2] is a 10-item instrument designed to assess the extent to which a person in pain believes s/he is able to accomplish various activities despite their pain. There is evidence for the reliability and validity of the original English-language version, including its ability to predict disability associated with pain, attrition and functional changes after pain management programmes [1] and several translated versions have been used [3–7].

Further evidence for the validity of the PSEQ and its short form may be gleaned from an item response theory (IRT) approach (see [8,9]), in which responses to a set of items are assumed to be probabilistically determined by a latent (unobserved) variable. Whereas classical methods like factor analysis and internal

* Corresponding author at: Pain Management Research Institute, Royal North Shore Hospital, Reserve Rd, St Leonards, NSW 2065, Australia.
E-mail address: daniel.costa@sydney.edu.au (D.S.J. Costa).

http://dx.doi.org/10.1016/j.sjpain.2016.08.001
1877-8860/

consistency focus on inter-item associations, and therefore are most informative at the questionnaire or sub-scale level, IRT is more informative about properties of individual items. In the case of the PSEQ, an IRT model posits that as the level of self-efficacy (operationalised as some aggregate of all of the items) increases, the response to each item increases along its seven-point response scale. Di Pietro, Catley, McAuley, Parkitny, Maher, Costa, Macedo, Williams and Moseley [10] conducted Rasch analysis, a method mathematically equivalent to a one-parameter IRT model, on the PSEQ using 600 individuals with low back pain. They found that the PSEQ performed adequately against several criteria, including unidimensionality, internal consistency and absence of item bias. A two-parameter model, on the other hand, allows examination of the differences in ability to discriminate between high- and low-scoring individuals.

A further question IRT can address is whether respondents use the PSEQ's seven-point response scale as intended. Specifically, consistent use of the response scale would be indicated by the observation that as level of self-efficacy increases, the response to each item changes from a lower response category to a higher one, i.e., those very low on self-efficacy (as an aggregate of all items) should be most likely to choose response option 0 ("Not at all confident"), those very high on self-efficacy should be most likely to choose 6 ("Completely confident"), and each of the other five response options should be the most likely for some value of self-efficacy. This relation may not hold if, for example, respondents have difficulty discriminating between consecutive response options, which may occur when there are many such options. Thus, IRT can provide evidence regarding whether respondents use the response categories in the desired manner, and therefore provide a recommendation about whether fewer categories should be used. Di Pietro et al. [10], using Rasch analysis, found evidence for appropriate category ordering in the PSEQ, but presented average curves rather than curves for individual items, finding a disordered item threshold involving the second response category. Whereas the Rasch model assumes that the extent to which items can discriminate between respondents high and low on the level of the latent variable is constant, there are less constrained IRT models that allow discriminative ability to vary, so it would be instructive to examine whether such a model supports the use of seven response categories.

Nicholas, McGuire and Asghari [11] proposed a short form for the PSEQ, consisting of two items: 'I can do some form of work, despite the pain ("work" includes housework and paid and unpaid work)' and 'I can live a normal lifestyle, despite the pain'. These items were selected for the short form on the basis of item-total correlations, item-disability correlations (using the modified Roland Morris Disability Questionnaire [12]), responsiveness, contribution to total score variance, and construct validity. IRT can provide additional information regarding location and discrimination parameters that can further inform short form development.

The overarching purpose of the present study is re-analyse Nicholas et al.'s [11] data using a two-parameter IRT model, which supplements validation research already conducted. This study extends the work of Di Pietro et al. [10] by examining the threshold and discrimination parameters for each PSEQ item. We also addressed Di Pietro et al.'s [10] call to examine the properties of the PSEQ for patients in tertiary care and with other pain sites, not just back pain.

2. Materials and methods

2.1. Instrument

The PSEQ has 10 items designed to assess the strength and generality of a patient's beliefs about his/her ability to accomplish various activities despite pain. Participants rate each item on a 0 ("Not at all confident") to 6 ("Completely confident") scale. Item scores are summed to provide a score with a possible range of 0–60, where higher scores indicate stronger self-efficacy. Previous psychometric analyses have provided evidence for a single factor with high internal consistency [1].

2.2. Participants

The participants were 1511 patients attending for initial assessment at a tertiary pain management centre in Sydney, Australia. Table 1 presents the demographic and clinical characteristics of the sample. This analysis was conducted on the same data set reported by Nicholas et al. [11].

2.3. Statistical methods

Item response theory (IRT) describes the relation between an unobserved (latent) variable (in this case, pain self-efficacy) and responses to items designed to assess that variable. Specifically, the probability of an individual's response to an item is determined by their value on the latent variable and properties of the item.

For the present analysis we used Samejima's two-parameter graded response model [13]. In this model, the observed responses to polytomous items (i.e., items with more than two response options) are assumed to be a logistic function of the latent variable; the probability of responding with a higher response option increases as the level of the latent variable increases. Two item parameters were estimated for each item in this model: (a) difficulty, or location along the continuum of values of the latent variable; and (b) discrimination, or ability to differentiate between those scoring high and low on the latent variable. (Note that the Rasch model, as used by Di Pietro et al. [10], allows estimation of location, but holds discrimination constant across items.)

IRT analysis was conducted using the grm() function of the ltm() package [14] in R. The analysis allows examination of the location and discrimination parameters of each item and provides category probability curves to determine whether any items exhibited problems with the ordering of item response category thresholds (i.e., the value of the latent variable for which adjacent response categories are equally likely). Problems with threshold ordering suggest that respondents are not using the response scale in the manner expected. The discrimination parameter describes the ability of each item to discriminate between individuals scoring low and high on the latent variable. Information represents the precision of each item across the range of the latent variable. Illustrated graphically, items with higher precision have taller curves, indicating that information is captured with precision at a particular location along the latent variable scale. The fit of the two-parameter model was assessed by comparing this model to the one-parameter model (where the discrimination parameter is held constant between items) using the likelihood ratio test, where a p value of less than .01 was taken to indicate significantly better fit of the two-parameter model.

3. Results

The mean total PSEQ score was 25.83 (SD = 13.96). The fit of the two-parameter model was significantly better than the fit of a one-parameter model ($p < .01$), indicating that the discrimination parameters differed between items (i.e., model fit was relatively poor when they were constrained to be equal). The threshold and discrimination parameters for the 10 items are shown in Table 2, as well their means and standard deviations. Item 9 had the highest discrimination parameter, followed by Items 8 and 6. Item 7 had the lowest discrimination parameter, which most likely arises because

Please cite this article in press as: Costa DSJ, et al. Item response theory analysis of the Pain Self-Efficacy Questionnaire. Scand J Pain (2016), http://dx.doi.org/10.1016/j.sjpain.2016.08.001