



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 28 August 2016

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.

© 2016 Scandinavian Association for the Study of Pain. Published by Elsevier B.V. All rights reserved.

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).

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

Table 1
Demographic data (N = 1511).

		Mean (SD)	Frequency
Age (years)		48.9 (16.1)	
Pain duration (months)		80.7 (119.3)	
Sex	Female		860 (57)
Marital status	Married/de facto		901 (63)
	Separated/divorced/widowed		275 (18)
	Single		264 (19)
Education	University		323 (24)
	Technical education		233 (17)
	Year 12		122 (9)
	Less than Year 12		696 (51)
Work status	Full time/part time/voluntary		454 (33)
	Home duties		129 (9)
	Unemployed because of pain		448 (32)
	Other		375 (27)
Pain sites	Head/face/shoulder		228 (15)
	Lower back or lower limbs		132 (9)
	Lower back and lower limbs		304 (20)
	2 or more sites		825 (55)
How pain began	Work-related injuries		506 (36)
	Accident at home		56 (4)
	Car accident		170 (12)
	After surgery or illness		217 (15)
	Pain just began		294 (21)
	Other reasons		160 (11)

Table 2
Means, standard deviations, and location (threshold) and discrimination parameters for each of the 10 PSEQ items.

Item		Mean	SD	Thresholds						Discrimination
				1	2	3	4	5	6	
1	I can enjoy things, despite the pain	3.19	1.68	−2.01	−1.18	−0.52	0.26	0.95	1.71	2.10
2	I can do most of the household chores (e.g., tidying-up, washing dishes, etc.), despite the pain	3.17	1.86	−1.84	−0.98	−0.44	0.17	0.77	1.62	1.85
3	I can socialise with my friends or family members as often as I used to do, despite the pain	2.85	1.84	−1.69	−0.69	−0.11	0.48	0.97	1.67	2.17
4	I can cope with my pain in most situations	3.03	1.68	−1.83	−0.99	−0.37	0.34	0.99	1.88	2.31
5	I can do some form of work, despite the pain. ("work" includes housework, paid and unpaid work)	3.19	1.90	−1.68	−0.93	−0.40	0.19	0.70	1.39	2.16
6	I can still do many of the things I enjoy doing, such as hobbies or leisure activity, despite pain	2.36	1.81	−1.11	−0.39	0.20	0.74	1.25	1.94	2.61
7	I can cope with my pain without medication	1.57	1.86	−0.26	0.46	0.95	1.67	2.39	3.43	1.05
8	I can still accomplish most of my goals in life, despite the pain	2.19	1.83	−0.88	−0.21	0.26	0.77	1.25	1.96	3.17
9	I can live a normal lifestyle, despite the pain	2.08	1.81	−0.78	−0.17	0.33	0.82	1.35	1.91	3.41
10	I can gradually become more active, despite the pain	2.24	1.78	−1.09	−0.40	0.20	0.89	1.46	2.15	2.11
	Total PSEQ score	25.83	13.96							

of the low rate of endorsement for this item (45% of respondents answered "Not at all confident").

The category probability curves for Items 1 and 7 are shown in Fig. 1. All items except item 7 exhibited item response thresholds in the expected order (the curves for Item 1 are included as a representative example of the other nine items). The item information functions (Fig. 2) show that items 6, 8, 9 and 10 were most precise for slightly higher levels of the latent variables than were items 1–5, as indicated by the location of the peaks of the curves in the horizontal direction. This aligns with the observation from Table 2 that the location parameters for Items 1–5 tended to be lower than those for Items 6–10, i.e., Items 1–5 tend to be more precise for lower levels of self-efficacy, and Items 6–10 more precise for higher levels. Item 7 had low precision for all values of the latent variable. The test information function, which is the sum of the item information functions, is shown in Fig. 3.

4. Discussion

The results of the IRT analysis conducted here confirm the strong psychometric properties of the PSEQ, adding to the body of evidence supporting its validity and reliability [1,3,10]. The category

probability curves provided empirical support for the use of the seven-point response scale, with the exception of Item 7, whose poor performance is consistent with all previous studies examining the validation of the PSEQ [1,10,11]. The large proportion of low responses to this item results in relatively weak relations with the other items. Nonetheless, the item captures clinically important information [1,10], which must be considered alongside statistical considerations in determining its utility. Specifically, the observed low mean item score indicates a large number of respondents are not confident they can cope with their pain without medication. This is important information for clinicians attempting to help their patients to manage their pain without relying on opioids, for example [15]. It is also an important metric for evaluating the outcomes of interventions aimed at enhancing pain self-management. For these reasons we believe the item should not be removed.

It is noteworthy that the mean PSEQ score in this sample (25.83) was much lower than those reported in Di Pietro et al. [10] (range 39.8–44.6). The participants in this study were recruited from a tertiary pain management centre, and their pain was not limited to back pain. Thus, the present results illustrate the psychometric properties of the PSEQ for a sample that is both more impaired (with

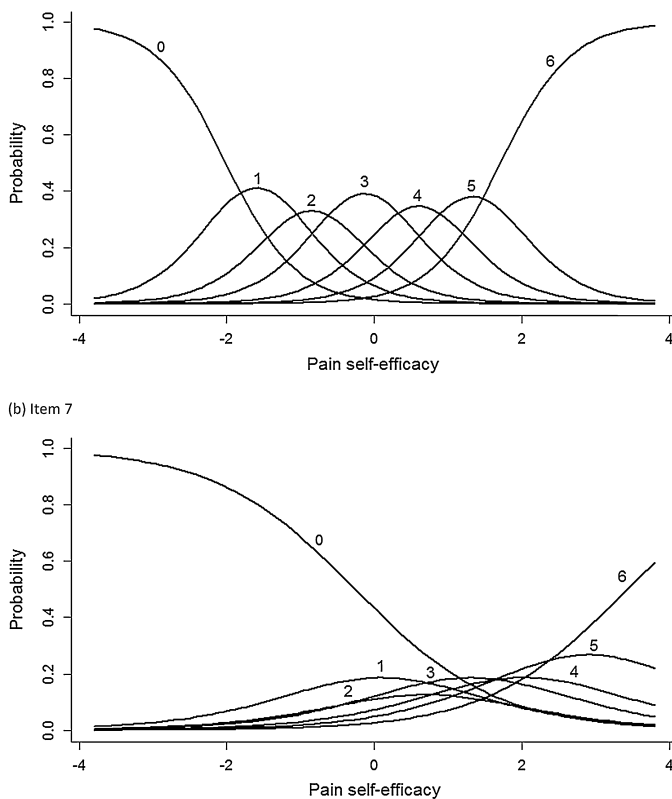


Fig. 1. Category probability curves for Items 1 and 7 of the PSEQ. The numbers on each graph represent the response options (0 = “Not at all confident”, 6 = “Completely confident”). The horizontal axis represents the value of the pain self-efficacy latent variable, and the vertical axis represents the probability of selecting each of the seven response options. (a) Item 1 (appropriate threshold ordering, representative of all items except Item 7). (b) Item 7

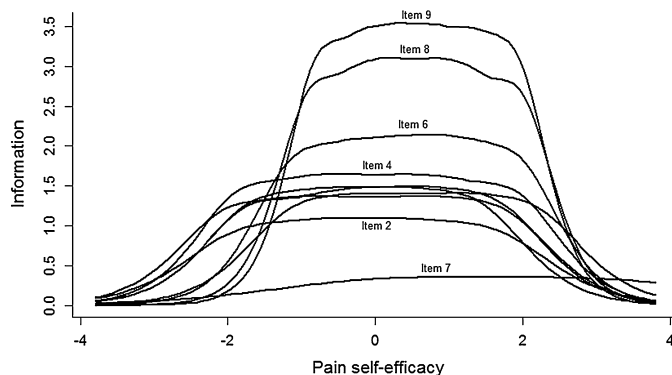


Fig. 2. Item information functions for each PSEQ item, which represents the precision of each item across the range of the latent variable. The horizontal axis represents the value of the pain self-efficacy latent variable. For clarity of presentation, only items with the highest (items 4, 6, 8 and 9) and lowest (items 2 and 7) information are labelled.

respect to self-efficacy) and more heterogeneous than in Di Pietro et al. [10], highlighting the wide applicability of the PSEQ.

With respect to the two items included in Nicholas et al.'s [11] short form, Item 9 was the best discriminating item and had greatest precision for higher values of the latent trait. Item 5 was not among the best-discriminating items, but it exhibited no problems. Notably, the precision of Item 9 was greatest for higher levels of self-efficacy, whereas the precision of Item 5 was greatest for lower levels; thus, these two items between capture a range of levels of self-efficacy. Also worth noting is that the four best discriminating items were those used in a 4-item short form proposed recently

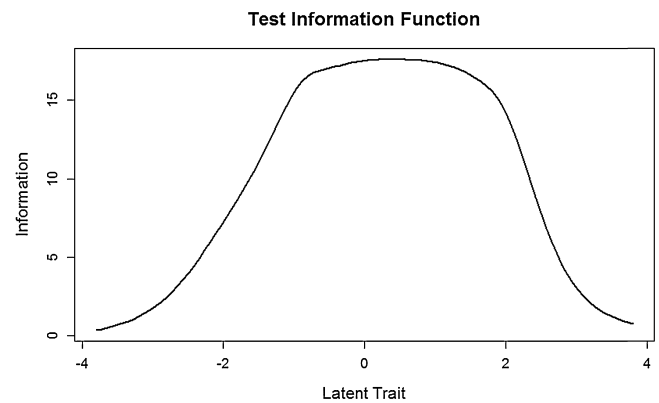


Fig. 3. Test information function, which represents the sum of the item information functions. The horizontal axis represents the value of the pain self-efficacy latent variable.

[16]. In combination with the analyses reported by Nicholas et al. [11], which complement the current analysis, we suggest that no changes are required to the 2 item PSEQ short form.

Item response theory has been applied in a wide variety of research areas. It provides information about the measurement properties of an instrument that complements that of more traditional methods, e.g., factor analysis, internal consistency, and is particularly useful in developing short-form instruments. Thus, its application in pain research has great potential to assist in reducing the burden on patients, particularly when patients are administered a battery of questionnaires.

A limitation of this study is that participants were drawn from a single pain clinic. Nonetheless, the sample was large and heterogeneous with respect to origin and site of pain, as well as various demographic characteristics, thus supporting the generalisability of the results. In summary, this study adds to the growing body of evidence supporting the psychometric properties of the PSEQ as a tool for assessing self-efficacy in people with pain.

Ethical issues

The data used in this study were derived from previous studies at the Pain Management Research Institute, all of which had approval from the Northern Sydney Area Human Ethics Committee. All participants in this study provided written informed consent to use their de-identified data in this research.

Conflict of interest

The authors have no conflicts of interest to declare.

References

- [1] Nicholas MK. The pain self-efficacy questionnaire: taking pain into account. *Eur J Pain* 2007;11:153–63.
- [2] Nicholas MK. Self-efficacy and chronic pain. St Andrews, Scotland: Annual conference of the British Psychological Society; 1989.
- [3] van der Maas LCC, de Vet HCW, Köke A, Bosscher RJ, Peters ML. Psychometric properties of the Pain Self-Efficacy Questionnaire (PSEQ): validation, prediction, and discrimination quality of the Dutch version. *Eur J Psychol Assess* 2012;28:68–75.
- [4] Lim HS, Chen PP, Wong TCM, Wong E, Chan ISF, Chu J. Validation of the Chinese version of Pain Self-Efficacy Questionnaire. *Anesth Analg* 2007;104:918–23.
- [5] Mangels M, Schwarz S, Sohr G, Holme M, Rief W. Der Fragebogen zur Erfassung der schmerzspezifischen Selbst-wirksamkeit (FESS): Eine Adaptation des Pain Self-Efficacy Questionnaire für den deutschen Sprachraum. *Diagnostica* 2009;55:84–93.
- [6] Sardá Júnior JJ, Nicholas MK, Pimenta CAM, Asghari A. Biopsychosocial predictors of pain, incapacity and depression in Brazilian chronic pain patients. *Revista Dor* 2012;13:111–8.

- [7] Asghari A, Nicholas MK. An investigation of pain self-efficacy beliefs in Iranian chronic pain patients: a preliminary validation of a translated English-language scale. *Pain Med* 2009;10:619–32.
- [8] Baker F. The basics of item response theory. University of Maryland: ERIC Clearinghouse on Assessment and Evaluation; 2001.
- [9] Embretson SE, Reise SP. Item response theory for psychologists. Mahwah: Erlbaum; 2000.
- [10] Di Pietro F, Catley MJ, McAuley JH, Parkitny L, Maher CG, Costa LCM, Macedo LG, Williams CM, Moseley GL. Rasch analysis supports the use of the Pain Self-Efficacy Questionnaire. *Phys Ther* 2014;94:91–100.
- [11] Nicholas MK, McGuire BE, Asghari A. A 2-item short form of the Pain Self-Efficacy Questionnaire: development and psychometric evaluation of PSEQ-2. *J Pain* 2015;16:153–63.
- [12] Roland MO, Morris RW. A study of the natural history of back pain. Part 1: development of a reliable and sensitive measure of disability in low back pain. *Spine* 1983;8:141–4.
- [13] Samejima F. Estimation of latent ability using a response pattern of graded scores. *Psychom Monogr Suppl* 1969:34.
- [14] Rizopoulos D. ltm: an R package for latent variable modeling and item response theory analyses. *J Stat Softw* 2006;17:1–25.
- [15] Ballantyne JC, Sullivan MD. Intensity of chronic pain – the wrong metric? *N Engl J Med* 2015;373:2098–9.
- [16] McWilliams LA, Kowal J, Wilson KG. Development and evaluation of short forms of the Pain Self Efficacy Questionnaire. *Eur J Pain* 2016;5(19):1342–9.