Development and Psychometric Properties of a Joint Protection Self-Efficacy Scale *Running headline:* ***Developing a Joint Protection Self-Efficacy Scale***

Niedermann K[1,2,6], Forster A[3], Ciurea A[1], Hammond A[4], Uebelhart D[1], de Bie R[5,6]

[1]Department of Rheumatology and Institute of Physical Medicine, University Hospital Zurich, Switzerland

[2]Department of Health, University of Applied Sciences, Winterthur, Switzerland

[3]Rehabilitation Clinic St. Katharinental, Diessenhofen, Switzerland

[4]Centre for Rehabilitation & Human Performance Research, University of Salford, Greater Manchester, UK

[5]Department of Epidemiology, Faculty of Health Sciences, Maastricht University, NL

[6]CAPHRI Research Institute, Maastricht University, NL

Corresponding author

Karin Niedermann, MPH

University Hospital Zurich

Department of Rheumatology and Institute of Physical Medicine

Gloriastrasse 25

CH-8091 Zurich, Switzerland

Phone: +41 44 255 21 49

Fax: +41 44 255 43 88

E-mail: karin.niedermann@usz.ch

Abstract

Introduction: Self-efficacy is one of the most powerful determinants of behaviour change. To increase effectiveness of joint protection (JP) education, it may be important to address perceptions of JP self-efficacy directly. The aim of this study was to develop a scale to measure JP self-efficacy (JP-SES) in people with RA

Methods: Instrument development included item generation, construct validity and reliability testing. Rasch analysis was applied to determine construct validity and the revised JP-SES was tested again to confirm validity and establish test-retest reliability and internal consistency.

Results: 46 items were generated by literature review, occupational therapists and people with RA. After semi-structured interviews and field-testing with RA participants, a 26-item questionnaire draft was constructed and tested. Rasch analysis to determine construct validity reduced the JP-SES to 13 items with good overall fit values. Rasch analysis of confirmatory validity resulted in a final 10-item version of the JP-SES. Test-retest results supported the validity of the scale, with high internal consistency (α=0.92) and good test-retest reliability (r_s= 0.79; p<0.001).

Conclusions: The JP-SES is a valid and reliable scale to assess perceived ability of people with RA to apply JP methods. The JP-SES could help stimulate the use of efficacy-enhancing methods in JP education.

Key words: occupational therapy, rehabilitation, validation, joint protection, self-efficacy, Rasch analysis, rheumatoid arthritis,

Introduction

Self-efficacy is one of the most powerful determinants of behaviour. According to Bandura (1) the confidence of a person to successfully execute a future specific behaviour or task, i.e. (self)-efficacy expectation, and the person's belief that the desired behaviour has a positive effect, i.e. outcome expectation, determine the initiation of a process to perform a behaviour, continue with it and persevere when difficulties arise. Self-efficacy refers to perceived ability in specific domains of activities, i.e. it is a specific state, not a general trait, although a variety and range of positive mastery experiences may lead to a general sense of self-efficacy. For example, a person with rheumatoid arthritis (RA) could potentially have high self-efficacy to follow a drug prescription correctly, but low self-efficacy for using joint protection (JP) methods correctly, i.e. in accordance with JP principles. Self-efficacy is a *belief* that one is able to perform a specific behaviour or task, rather than indicating that one actually does perform it. However the perceived ability to perform a given behaviour is strongly related to one's actual performance of that behaviour (1, 2).

In people with RA, higher self-efficacy has been shown to be associated with better ability to cope with their disease, i.e. manage pain (3) and daily living with RA (4), as well as with better current (5) and future (2 and 5 year) health status (6, 7). Applying self-efficacy approaches within education programmes increased use of exercise, relaxation, cognitive symptom and fatigue management and joint protection (4, 8-10) Self-efficacy is thus a strong mediator of behaviour change.

Hands are the commonest joints affected and hand involvement is one of the major problems from the RA patients' perspective (11). In consequence, occupational therapists (OTs) providing joint protection (JP) education to people with RA mainly focus on the hand and wrist joints (12). JP education aims to reduce pain and maintain functional ability. It includes applying ergonomics, altering working methods, balancing activity and rest and using assistive devices.

In line with progress in drug therapy for RA, JP education has evolved from increasing knowledge about how to preserve joint structures and joint function to a self-management approach to cope with pain, reduce functional limitations and thus improve daily task and role performance. Studies consistently demonstrate that JP improves function and pain in the short- and long-term, given that psycho-educational methods are applied (13-15). However this approach is not commonly adopted. In Switzerland JP education typically consists of oral and written information about RA and JP principles (16); demonstrations and supervised practice of hand JP methods, mostly in a kitchen activity; and demonstration of appropriate assistive devices. This is provided in a one-to-one setting for between 1-4 hours over 1-9 sessions, dependent on the therapist's evaluation of patient need and staffing levels.

In a psycho-educational JP approach, practice under supervision and at home, goal setting and feedback, observing and exchange with peers in groups, discussions and verbal persuasion are important strategies. These learning and practice situations to acquire JP behaviours are also the four sources for acquiring self-efficacy: 1) direct personal mastery experience (skills performance with increasing difficulty and complexity); 2) vicarious experience (role modelling), 3) verbal persuasion (counselling, suggestions, reinforcement and 4) emotional arousal (re-interpretation of physiological signals). Direct experience is far more effective than indirect experience (1). However, self-efficacy should also be addressed directly, by evaluating and supporting the patients' belief in their ability to acquire, perform and adhere to JP behaviour.

The Arthritis Self-Efficacy Scale (ASES) (2) is the most recognized instrument for assessing self-efficacy in people with RA. The ASES pain and other symptoms subscales have also been used in JP studies (13, 15). As self-efficacy is domain-specific, using a JP self-efficacy instrument may be more specific than the ASES in this context. Domain-specific self-efficacy scales, such as those for back pain (17) and physical activity (18), can contribute to determine which situations are important for the perceived ability to perform the target behaviour and for

understanding how to tailor interventions to meet individual needs. A JP specific self-efficacy scale could offer similar benefits. Consequently, the aim of this study was to develop and evaluate the psychometric properties of a JP self-efficacy scale (JP-SES) for use in research and clinical practice.

Materials and Methods

Design

The development of the JP-SES consisted of six steps (outlined in Table 1).

Participants

Patients: Different groups of people with RA attending the outpatient facility of a hospital's rheumatology department participated in the scale development. The Local Research Ethics Committee approved the study protocol and patients provided informed consent prior to participation.

Occupational therapists (OTs): Therapists experienced in JP education, working in four different rheumatology outpatient facilities, participated in the item generation process.

Steps of scale development

Step 1: Item generation

The first set of items was composed by a group of five OTs, led by the principal investigator (KN). Items were selected and adapted through consensus from the contents of self-efficacy scales for back pain prevention behaviours (17) and physical activity (18), as items included in these scales are potentially also relevant to JP behaviour and thus JP self-efficacy. Items from these scales that were not adaptable for hand JP specifically were therefore omitted, primarily to ensure that the JP-SES content would be applicable to the majority of people with RA. Other items were generated from the OT's clinical experience. Semi-structured interviews were

conducted with a convenience sample of 10 participants with RA who had previously attended "typical" individual JP education (see Introduction).

Firstly, they were asked to judge the relevance of each item (yes/no), and then to consider if, for these tasks or situations, JP would be difficult. Secondly, patients were asked to suggest additional items.

Step 2: Construction of a draft questionnaire

Version 1 of the JP-SES draft was developed containing all items from the list judged relevant by at least half of the RA participants, as well as the newly suggested items. A short explanation of self-efficacy and JP methods was included to enable participants' understanding of the underlying concepts of JP. The root statement "I *am confident that I can care for my joints when I am..." (17)* was combined with each of the selected items. These items were conceptually allocated to five areas: 'physical activities' (11 items), 'affect' (8 items), 'time constraints' (3 items), 'social activities' (7 items) and 'pain' (2 items). A 5-point response scale (0=not at all confident; 4=very confident), was applied as in the back pain prevention self-efficacy scale (17).

Step 3: Preliminary testing

The principal investigator asked participants with RA, currently attending JP education, to complete the version 1 of the JP-SES draft. They were then each interviewed regarding adequacy and comprehensiveness of the selected items (cognitive debriefing).

Step 4: Construct validity study

Version 2 of the JP-SES draft was distributed by treating rheumatologists to a consecutive sample of German-speaking RA patients. A sample size of 150 (n range 108-243) and of 100 (n range, 64-144) was identified as appropriate based on providing a 99% and 95% respectively confidence of item calibration of +/-0.5 logits (19).

Most or all of these participants would have received "typical" individual JP education as referral to OT was part of usual care at the participating hospital. They were asked to return the

completed questionnaire within a week. Non-responders received a reminder after approximately two weeks. Additionally, disease activity (using the DAS28) (20), pain level (on a 0-10 visual analogue scale) and functional disability (Health Assessment Questionnaire, HAQ) (21) were assessed.

Steps 5 and 6: Confirmatory construct validity study and test-retest reliability

Version 3 of the JP-SES draft was sent to German-speaking people with RA taking part in a clinical quality management project at the participating hospital. Again most or all would have received "typical" JP education. Disease-specific data, pain and HAQ disability were also assessed. The patient information letter was non-specific about completing the questionnaire a second time (i.e. to test stability of questions) so as to reduce the likelihood of patients memorizing or recording answers. All responding patients were re-sent the questionnaire for reliability testing approximately three weeks later.

Statistics

The psychometric properties of the JP-SES drafts were tested with Rasch analysis (22-24), in order to examine how well observed data fit the Rasch model's expectations. For this, various error estimates and fit statistics are calculated and response category ordering and differential item functioning (DIF) (i.e. do items work the same way irrespective of group, e.g. gender) are examined. Additionally, if data fit the model, linear transformation of ordinal raw data is obtained, allowing for valid parametric statistics.

For construct validation (step 4) and confirmatory analysis (step 5), data from the Version 2 of the JP-SES draft were fitted to the Rasch Partial Credit Model ($\ln(p_{nij} / 1\text{-}p_{nij-1})=\theta_n-\delta_{ij}$). Overall summary fit statistics were calculated to test the model fit of the JP-SES data: item-person interaction statistics (transformed to standardized normal distribution with expected overall item and person fit means of 0 and standard deviations of 1; and item-trait interaction chi square

statistic, which, if significant ($p<0.05$), would indicate violation of the rule of invariance across trait. In addition, individual item fit was calculated, with significance level of 0.05. Before removing misfitting items, response category ordering was checked. If disordered, collapsing of adjacent categories may improve overall fit to the model.

DIF assessed bias for gender, younger/older age, short/long disease duration, low/high pain level and low/high disease activity. Bias may manifest as uniform DIF (consistent ability difference between groups) and/or non-uniform DIF (inconsistent ability difference, referred to as class intervals, between groups). Items displaying multiple uniform and non-uniform DIF at the 0.05 significance level (i.e. <0.05) were deleted.

On the final solution with all fitting items free of DIF, unidimensionality was formally tested using Principal Component Analysis PCA by examining principal component loadings of residuals (25, 26). Positively loading items were equated with negatively loading items and a series of independent t-tests was then performed to compare person location estimates by using differing item subsets. Less than 5% of these tests should be significant to confirm unidimensionality. Internal consistency (reliability) was assessed by the person separation index (PSI), which is equivalent to Cronbach's alpha. A value of at least 0.70 would be adequate for demonstrating internal consistency at group level, or 0.85 at individual level (27).

The Spearman rank correlation coefficient was calculated for test-retest correlation (using person locations) and Pearson correlation coefficients to explore associations between JP-self efficacy and disease status. Rasch analysis was performed using RUMM2020 software package (RUMM Laboratory, Dungraig, Western Australia). Statistical testing and reliability calculations were performed using SPSS software, version 12.0 (SPSS, Chicago, IL).

Results

Step 1: Item generation

A set of 46 items was identified: 33 from the two self-efficacy questionnaires (17, 18), nine by therapists and four by people with RA. Interviews were conducted with eight female and two male participants RA (mean age 60.40 years (SD 12.79), with mean disease duration of 17.40 years (SD 9.97). Interviews lasted between 15 to 30 minutes.

Step 2: Construction of a draft questionnaire

The version 1 of the JP-SES draft contained 31 items, after removing 12 items judged as not relevant by the majority of RA participants during the interviews and three items due to duplication.

Step 3: Preliminary testing

Cognitive debriefing of the version 1 of the JP-SES draft was performed with nine other RA patients. It took 5-10 minutes for them to complete the 31-item JP-SES draft. Five items, all from the area 'social activities' were perceived as either too general or redundant and removed: '*at a party', 'on vacation', 'travelling", 'having guests', 'eating out'*. No new items were suggested. After this process two response categories of *'not possible because of RA'* and *'not applicable for me'* were added to reduce missing data. This version 2 of the JP-SES draft, consisting of 26 items, was used in the next step.

Step 4: Construct validity

A total of 114 people with RA agreed to participate and 101 (89%) questionnaires were returned. Participants' disease-specific data in both validation studies (steps 4 and 5) are presented in Table 2. The additional response category 'not applicable' was reported by 30% or more of respondents for several items: '*not supported by colleagues'* (59.4%; n=60); '*not supported by employer'* (53.5%; n=54); '*gardening'* (44.6%; n=45); '*assistive device gets dirty whilst using'* (30.7%; n=31); and '*driving'* (29.7%, n=30). These were therefore removed.

Rasch analysis was performed on the remaining 21 items to identify misfit to model expectations and DIF. The response options of several items were found to be disordered. Two

adjacent categories were collapsed and then eight items removed due to misfit (‘*in pain’, ‘not in pain’, ‘lifting heavy objects’, ‘carrying heavy objects’, ‘eating’, ‘cleaning’, ‘stressed’, and ‘feeling tense’)* in a stepwise procedure. This resulted in a 13-item JP-SES (draft version 3) with four response categories (0=not at all confident to 3=very confident).

Step 5: Confirmatory construct validity

The version 3 of the JP-SES draft was sent to 175 German-speaking RA patients of whom 126 (72%) responded (see Table 2). Rasch analysis was performed on complete data from 116 participants with RA (87 women and 29 men).

Rasch modeling identified two further items with misfit (‘*feeling depressed’ and ‘using an assistive device’*) and one item with DIF for gender and pain (‘*feeling anxious’*). After removing them, the analysis resulted in the final 10-item JP-SES with good overall fit (Table 3): item fit mean of 0.0 (SD 0.70), person fit mean of -0.10 (SD 2.33) and total item chi square of 24.32 (df = 20; p=0.23). Six of the final items were identical with ones derived from the two self-efficacy scales for exercise (4 items) and back pain prevention (2 items).

Individual item fit statistics ranged from p values of 0.92 to 0.07, i.e. items did not significantly differ from the model. Individual item difficulty level ranged from -0.80 (lowest requirement of self-efficacy) to 0.99 (highest requirement of self-efficacy). The person-item threshold distribution (Figure 1) shows that the items’ difficulty targeted the persons’ ability well. For some people with very high (or very low) self-perceived ability, there were no items in the scale to capture this.

Internal consistency: Person separation index (PSI), (which corresponds to Cronbach’s alpha, i.e. reliability) was high (0.95), allowing use of the scale at the individual level. This step confirmed construct validity of the final version of the JP-SES (Appendix 1). Principal Component Analysis PCA of the residuals identified two subsets of items consisting of the highest positive and negative loading items. The positive subset (PC loadings >0.3) comprised

four items related to 'affect' (*in a bad mood, not feel like, no time, busy*) and the negative subset (PC loadings <-0.3) comprised five items related to physical activities (*cook, shop, write, lift light weight, carry light weight*). The equating/independent t-tests of the person estimates for the positive loading and negative loading subsets showed 4.08% of the tests were significant, which supported the assumption of unidimensionality of JP self-efficacy across the areas of 'affect' and 'physical activities"

Correlations: Pearson's correlations of the total JP-SES score (from the confirmatory study) with disease related factors were low although significant: 0.19 (p=0.03) for pain; 0.22 (p=0.02) for DAS28; and 0.27 (p=0.008) for the HAQ disability index.

Step 6: Test-retest – reliability

The test re-test sample consisted of all responding participants in step 5 (n=126). Of these, 110 (79 women: 31 men; 87%) returned the second questionnaire with complete data. Rasch summary fit statistics were: item fit mean 0.0 (SD 0.61), person fit mean 0.18 (SD 2.24) and total item chi-square 21.56 (df =20; p=0.37). Individual item fit statistics ranged from p values of 0.92 to 0.08. Individual item difficulty level ranged from -0.80 to 0.99 (lowest to highest requirement of self-efficacy) (Table 3). PSI of the re-test scale data was also 0.95, once again high enough for use in individuals and groups.

Mean total JP-SES scores of the confirmatory construct validity study and retest study were 14.94 (SD 8.08) (n=126) and 16.22 (SD 8.52) respectively (n=110). Cronbach's alpha for internal consistency of the items in both of the two samples was α = 0.92. The test-retest reliability Spearman rank correlation was 0.78 (p<0.001).

Discussion

The JP-SES was developed to assess how people with RA rate their perceived ability to perform hand JP methods, as these are the commonest methods used by people with RA and

taught by OTs. Overall, it demonstrated good validity and reliability and can be used for research purposes to evaluate effectiveness of JP education or changes in JP-SE over time.

Item reduction of the JP-SES items initially generated was necessary to determine the final set valid for assessing JP self-efficacy. Items removed were all within the areas 'physical activities' and "affect'. Items related to 'affect' generally appeared more difficult, almost acting as 'barriers' and thus requiring more self-efficacy than activity items (Table 3). Some difficult 'physical activity' items were deleted due to misfit or serious DIF (e.g. *'lifting heavy objects', 'carrying heavy objects', 'cleaning')* and items representing (negative) affect in a more general way (e.g. *'stressed'*, *'feeling tense'*). These situations may practically make performance of JP methods more difficult, rather than affect perceived ability of performance per se. All items, within the areas 'time constraints' and 'social activities', as well as the two pain items, showed misfit and DIF and were therefore discarded. Although pain has been found to strongly determine the performance of JP methods (28, 29), being '*in pain* or *'not in pain"* was not related to perceived ability to apply JP methods. Six of the ten final items were identical with the self-efficacy scales for exercise and back pain prevention. Although SE is domain-specific, it may be covered by identical or similar items across related domains.

The set of the final, unidimensional items, i.e. fitting the Rasch model, allowed for an interval scale transformation of the (ordinal) raw scores. This is especially important in an evaluative instrument where correct change scores have to be calculated (30). Test-retest reliability was good. For estimating response stability over time, stability of the disease condition has to be assumed. RA is a disease with unpredictable daily changes, although the disease variables pain, DAS28 and HAQ were similar at both time points. Also correlations between the disease variables with the psychological construct of JP self-efficacy were significant but low, indicating these have less effect on self-perceived ability of JP performance.

Rasch analysis, considered state-of-the-art in the development of new scales, was applied ensuring unidimensionality of the JP-SES as a fundamental requirement for construct validity (20). A potential limitation of the study thus was its relatively small sample size. We originally planned to obtain a sample size of 100 -150 as this would provide between 95-99% confidence of item calibration and we achieved this. For tests that are not well targeted, larger sample sizes are required for adequate item location precision (19). The well-targeted JP-SES reduced its sample sized requirements, which conversely, could have challenged the validity of the JP-SES. No floor or ceiling effects were present (31). Whilst further development work on the scale may improve the JP-SES' coverage by adding very easy and very difficult items, it should be considered that items that everybody achieves (i.e. very easy ones) or items that everyone misses (i.e. too difficult ones) may not add much information to a scale. We used a 5-point response scale initially. Many self-efficacy scales use 0-10 numeric rating scales (2, 18, 32) and some use Likert-type scales of five response options (17, 33). Generally, the extension of response categories to improve gain of information (sensitivity) and reliability is advocated (34). However, there is conflict between improving precision and reliability by raising the number of response categories and people's limited channel capacity, i.e. their (in)ability to discriminate between narrow scales (35, 36). Reducing the number of response categories can, in some circumstances, result in instruments with lower cognitive burden without loss of sensitivity or reliability (37, 38). Rasch analysis addresses response category ordering and participants' inability to discriminate between response categories may result in disordered response options that are breaking the rule of scale additivity. This tension - between precision and ability to discriminate in the JP-SES scales - was demonstrated in our analysis, where several items had disordered categories. The 5-response categories in these items were not distinguishable for all patients and therefore provided no additional useful information. Interestingly, reducing to a 4-

point response scale was the best solution and provided a better item fit than performing the same Rasch modelling procedure with a trichotomous or dichotomous scale.

The focus of the study lied on the distinct steps of a development process and use of Rasch analysis, to determine validity. A limitation of the study is that no other psychological constructs were used in this validation process. Future research to support the construct validity of the JP-SES should investigate associations with other psychological constructs or self-efficacy subscales, e.g. the Rheumatoid Arthritis Self-Efficacy scale (39).

The JP-SES was primarily developed and validated for use in research in group studies. If the aim is change scores are calculated parametric statistics are used, the (ordinal) raw scores have to be transformed into interval scores. Theoretically, scales with raw scores developed using Rasch analysis can be used clinically on an individual level if a score-to-measure conversion table is developed, translating ordinal (raw) scores to interval data, which would give a person location, i.e. information about the patient's ability (in this context in terms of JP self-efficacy). A valid conversion table must be produced with a large patient sample (40) and this would only then be applicable to people of similar characteristics, e.g. diagnosis and age range. Such conversion tables may be of limited use for individual clinical assessment, as well as making scoring more complex. However, because the person separation index reliability was high, the JP-SES can be used on an individual level and raw scores used to create scale totals.

Clinical use of the JP-SES in individuals may be more appropriate for screening patients' tasks and perceptions of JP use in the assessment prior to JP education. To improve patients' self-efficacy, it may also be important to identify outcome expectations (i.e. beliefs in JP effectiveness), which may also be preconditions for improving perceived ability to cope (2). The JP-SES could help stimulate consciously applying efficacy enhancing methods (e.g. individual goal setting, specific instructions and practice, contracting, modeling, reinterpretation of physiological signals) in JP education (4).

Self-efficacy is not an easy concept to understand for patients. Some participants would have preferred the statement *'I am confident to care for my joints when I am ...'* to be changed into *'I care for my joints when I am …'*. However, this would be much more than a semantic change. The difference between the statement about performing JP during an activity and the 'hypothetical' statement of perceived ability to perform JP in certain situations should be emphasized in instructions.

The modern concept of JP education as a self-management approach requires psycho-educational teaching methods. Such methods, among them self-efficacy enhancing strategies should be more widely adopted. In this context, a more appropriate term, such as 'ergonomic education' would better reflect the modern JP concept.

The value of the JP-SES is to provide information about the patient's perceived ability to perform JP and to help plan treatment, e.g. by goal setting and conscious use of self-efficacy enhancing methods. Further research is needed to determine the scale responsiveness and its usefulness as an evaluative instrument, as well as to understand how JP self-efficacy and JP performance are related.

Acknowledgements

See acknowledgments page

Competing interests

None

Funding

None

References

1. Bandura A. Self-efficacy: towards a unifying theory of behaviour change. Psychol Rev 1977;84:191-215.
2. Lorig K, Chastain RL, Ung E, Shoor S, Holman HR. Development and evaluation of a scale to measure perceived self-efficacy in people with arthritis. Arthritis Rheum 1989;32(1):37-44.
3. Keefe FJ, Lefebvre JC, Kerns RD, Rosenberg R, Beaupre P, Prochaska J, et al. Understanding the adoption of arthritis self-management: stages of change profiles among arthritis patients. Pain 2000;87(3):303-13.
4. Taal E, Rasker J, Wiegman O. Patient education and self-management in the rheumatic diseases: a self-efficacy approach. Arthritis Care Res 1996;9:229-38.
5. Taal E, Rasker JJ, Seydel ER, Wiegman O. Health status, adherence with health recommendations, self-efficacy and social support in patients with rheumatoid arthritis. Patient Educ Couns 1993;20(2-3):63-76.
6. Brekke M, Hjortdahl P, Kvien TK. Self-efficacy and health status in rheumatoid arthritis: a two-year longitudinal observational study. Rheumatology (Oxford) 2001;40(4):387-92.
7. Brekke M, Hjortdahl P, Kvien TK. Changes in self-efficacy and health status over 5 years: a longitudinal observational study of 306 patients with rheumatoid arthritis. Arthritis Rheum 2003;49(3):342-8.
8. Barlow JH, Turner AP, Wright CC. Long-term outcomes of an arthritis self-management programme. Br J Rheumatol 1998;37(12):1315-9.
9. Hammond A, Bryan J, Hardy A. Effects of a modular behavioural arthritis education programme: a pragmatic parallel-group randomized controlled trial. Rheumatology (Oxford) 2008;47(11):1712-8.
10. Lorig KR, Sobel DS, Ritter PL, Laurent D, Hobbs M. Effect of a self-management program on patients with chronic disease. Eff Clin Pract 2001;4(6):256-62.
11. Heiberg T, Finset A, Uhlig T, Kvien TK. Seven year changes in health status and priorities for improvement of health in patients with rheumatoid arthritis. Ann Rheum Dis 2005;64(2):191-5.
12. Hammond A, Lincoln N. Development of the Joint Protection Behavior Assessment. Arthritis Care Res 1999;12:200-7.
13. Hammond A, Lincoln N, Sutcliff L. A crossover trial evaluating an educational-behavioural joint protection programme for people with rheumatoid arthritis. Patient Educ Couns 1999;40:1044-51.
14. Hammond A, Freeman K. The long-term outcomes from a randomized controlled trial of an educational-behavioural joint protection programme for people with rheumatoid arthritis. Clin Rehabil 2004;18:520-8.
15. Hammond A, Freeman K. One-year outcomes of a randomized controlled trial of an educational-behavioural joint protection programme for people with rheumatoid arthritis. Rheumatology (Oxdford) 2001;40:1044-51.
16. Cordery J, Rocchi M. Joint Protection and Fatigue Management. In: Rheumatologic Rehabilitation: Assessment and Management. Bethesda: The American Occupational Therapy Association, Inc; 1998. p. 279-322.
17. Basler H, Herda C, Keller S, Ridder K. Selbstwirksamkeitserwartung, Entscheidungsbalance und rückenbewusstes Verhalten zur Prävention von Rückenschmerz. Z Med Psychol 1998:149-57.
18. Marcus BH, Selby VC, Niaura RS, Rossi JS. Self-efficacy and the stages of exercise behavior change. Res Q Exerc Sport 1992;63(1):60-6.
19. Linacre J. Sample size and item calibration stability. Rasch Mes Trans 1994;7:328.

20. Prevoo M, van 't Hof M, Kuper H, van Leeuwen M, van de Putte L, van Riel P. Modified disease activity scores that include twenty-eight-joint counts. Development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. Arthritis Rheum 1995;38:44-48.
21. Brühlmann P, Stucki G, Michel B. Evaluation of a German version of the physical dimension of the Health Assessment Questionnaire in patients with rheumatoid arthritis. J Rheumatol 1994;21:1245-9.
22. Andrich D. Rasch Models for Measurement. Newbury Park, London, New Delhi: Sage Publications, Inc; 1988.
23. Rasch G. Probabilistic Models for Some Intelligence and Attainment Tests. Chicago: University of Chicago Press; 1960 (reprint 1980).
24. Tennant A, Conaghan PG. The Rasch measurement model in rheumatology: What is it and why use it? When should it be applied, and what should one look for in a Rasch paper? Arthritis Rheum 2007;57(8):1358-62.
25. Tennant A, Pallant J. Unidimensionality matters! (a tale of two Smiths?). Rasch meas trans 2006;20:1048-51.
26. Smith E. Detecting and evaluating the impact of multidimensionality using item fit statistics and principal components analysis of residuals. J Appl Meas 2002;3:205-231.
27. Streiner D, Norman G. Health Measurement and Scales. Oxford: Oxford University Press; 1995.
28. Hammond A. Joint protection behavior in patients with rheumatoid arthritis following an education program: a pilot study. Arthritis Care Res 1994;7:5-9.
29. Niedermann K, Forster A, Hammond A, Uebelhart D, de Bie R. Development and validation of a German version of the joint protection behavior assessment in patients with rheumatoid arthritis. Arthritis Rheum 2007;57(2):249-55.
30. Bezruczko N. Raw score nonlinearity obscures growth. Rasch Measurement Transactions 2004;18(2):973-974.
31. McHorney CA, Tarlov AR. Individual-patient monitoring in clinical practice: are available health status surveys adequate? Qual Life Res 1995;4(4):293-307.
32. Mueller A, Hartmann M, Mueller K, Eich W. Validation of the arthritis self-efficacy short-form scale in German fibromyalgia patients. Europ J Pain 2002;7(2):163-71.
33. Tonge B, King N, Klimkeit E, Melvin G, Heyne D, Gordon M. The Self-Efficacy Questionnaire for Depression in Adolescents (SEQ-DA). Development and psychometric evaluation. Eur Child Adolesc Psychiatry 2005;14(7):357-63.
34. Sallis J, Pinski R, Patterson T, Nader P. The development of self-efficacy scales for health-related diet and exercise behaviors. Health Edu Res 1988;3:283-92.
35. Miller GA. The magical number seven plus or minus two: some limits on our capacity for processing information. Psychol Rev 1956;63(2):81-97.
36. Molenaar N, editor. Response-effects of 'formal' characteristics of questions. New York: Academic 1982.
37. Hawthorne G, Mouthaan J, Forbes D, Novaco RW. Response categories and anger measurement: do fewer categories result in poorer measurement?: development of the DAR5. Soc Psychiatry Psychiatr Epidemiol 2006;41(2):164-72.
38. Katz JN, Larson MG, Phillips CB, Fossel AH, Liang MH. Comparative measurement sensitivity of short and longer health status instruments. Med Care 1992;30(10):917-25.
39. Hewlett S, Cockshott Z, Kirwan J, Barrett J, Stamp J, Haslock I. Development and validation of a self-efficacy scale for use in British patients with rheumatoid arthritis (RASE). Rheumatology (Oxford) 2001;40(11):1221-30.
40. Siegert RJ, Tennant A, Turner-Stokes L. Rasch analysis of the Beck Depression Inventory-II in a neurological rehabilitation sample. Disabil Rehabil 2009:1-10.

Appendix

German Version of the JP-SES and *Corresponding English Translation*

Fragebogen zur Zuversicht für Gelenkschutz

Joint Protection Self-Efficacy Scale

Dieser Fragebogen erfasst Ihre Zuversicht, d.h. Ihr Vertrauen in Ihre Fähigkeit, bei verschiedenen Tätigkeiten und in verschiedenen Situationen auf Ihre Hand- und Fingergelenke zu achten und Gelenkschutz-Methoden anzuwenden. Gelenkschutz heisst:

- die Hand- und Fingergelenke in möglichst natürlichen und gelenkschonenden Stellungen belasten
- kraftraubendes Heben und Tragen vermeiden
- bei Tätigkeiten möglichst beide Hände zu gebrauchen
- für bestimmte Tätigkeiten ein Hilfsmittel benützen

Bitte kreuzen Sie für jede Situation das entsprechende Feld an, auch wenn etwas nicht vorkommt.

This questionnaire assesses your confidence in your ability to care for your wrist and finger joints in different activities and situations using joint protection methods. Joint protection means

- *using your wrist and finger joints in their most stable natural and joint-friendly position*
- *avoiding lifting and carrying weights*
- *using both hands for tasks whenever possible*
- *using an assistive device for certain tasks*

Please tick the appropriate answer or tick if not applicable.

Ich bin zuversichtlich, auf meine Hand- und Fingergelenke zu achten, (auch dann) wenn ich … ***I am confident that I can care for my joints, (even) when I …***	Sehr überzeugt *Very confident*	Eher überzeugt *Quite confident*	Eher nicht überzeugt *A little confident*	Gar nicht überzeugt *Not at all confident*	Kommt nicht vor *Not applicable*
einen leichten Gegenstand hebe *lift a light object*					
schlecht gelaunt bin *am in a bad mood*					
koche / *cook*					
beobachtet werde *am being watched*					
einkaufe / *shop*					
keine Zeit habe *have no time*					
schreibe / *write*					
Nicht danach zumute sein *do not feel like it*					
einen leichten Gegenstand trage *carry a light object*					
sehr beschäftigt bin *am very busy*					

The authors give permission for the Joint Protection Self Efficacy Scale to be reproduced for clinical and research use, provided that acknowledgement is made both to the authors and to the ***name of publishing journal.***

Table 1: Overview of steps in the development of the Joint Protection Self-Efficacy Scale

Step (n = number of people with RA included)	Procedures
1. Item generation	1. 33 items derived from related self-efficacy scales 2. 9 additional items generated by occupational therapists 3. 4 additional items suggested by patients ➔ 46 items
2. Construction of draft questionnaire	1. 12 items removed (judged *not relevant* by patients) 2. 3 duplicate items removed ➔ Draft version 1: 31 items
3. Preliminary testing / cognitive debriefing (n=9)	1. Minor changes / refinements in item formulation 2. 5 items removed (perceived too general or redundant) 3. Adding 2 additional response options ➔ Draft version 2: 26 items, 5 response categories (0-4)
4. Construct validity study (n=101)	1. 5 items removed (not applicable for > 30% of patients) 2. Application of Rasch analysis for construct validity 3. Collapsing two response categories as disordered 4. 8 items removed due to misfit ➔ Draft version 3: 13 items, 4 response categories (0-3)
5. Confirmatory construct validity study (n=125)	1. Rasch Analysis 2. 3 items removed: 2 due to misfit; 1 due to DIF* ➔ Final version: 10 items, 4 response categories (0-3)
6. Test-retest reliability study (n=110)	1. Rasch analysis calculating summary fit statistics 2. Calculating internal consistency and test-retest correlation

* DIF= Differential Item Functioning

Table 2: Demographic and clinical characteristics of study participants

	Construct Validity Study Step 4 (n = 101)	Confirmatory Study Step 5 (n = 126)
Female (%)	78 (78%)	93 (74%)
Age (years)	58.6 (13.7)	59.0 (12.9)
Disease duration (years)	12.8 (8.3)	13.0 (9.3)
Rheumatoid nodules (%)	25 (25%)	28 (22%)
Rheumatoid factor (%)	85 (85%)	97 (77%)
Anti-CCP antibodies, where available (%)	NA	43/56 (77%)
ANA, where available (%)	48/84 (40%)	57/109 (52%)
Erosions (%)	73 (73%)	96 (76%)
DMARDs (no. of patients and %)	80 (79%)	122 (97%)
Steroids (no. of patients)	36 (36%)	48 (38%)
NSAIDs (no. of patients)	19 (19%)	28 (22%)
DAS28, mean (SD)	3.3 (1.5)	3.2 (1.2)
Hand pain (VAS 0-10)	3.1 (2.6)	2.9 (2.4)
HAQ score (0-3)	1.0 (0.6)	0.9 (0.7)

Values are the mean (SD), unless stated otherwise; NA = not available.

ANA = anti-nuclear antibodies, Anti-CPP = anti-cyclic citrullinated peptide, DMARDs = disease-modifying anti-rheumatic drugs; NSAIDs = non-steroidal anti-inflammatory drugs; DAS28 = Disease Activity Score in 28 joints; ESR = erythrocyte sedimentation rate; HAQ = Health Assessment Questionnaire

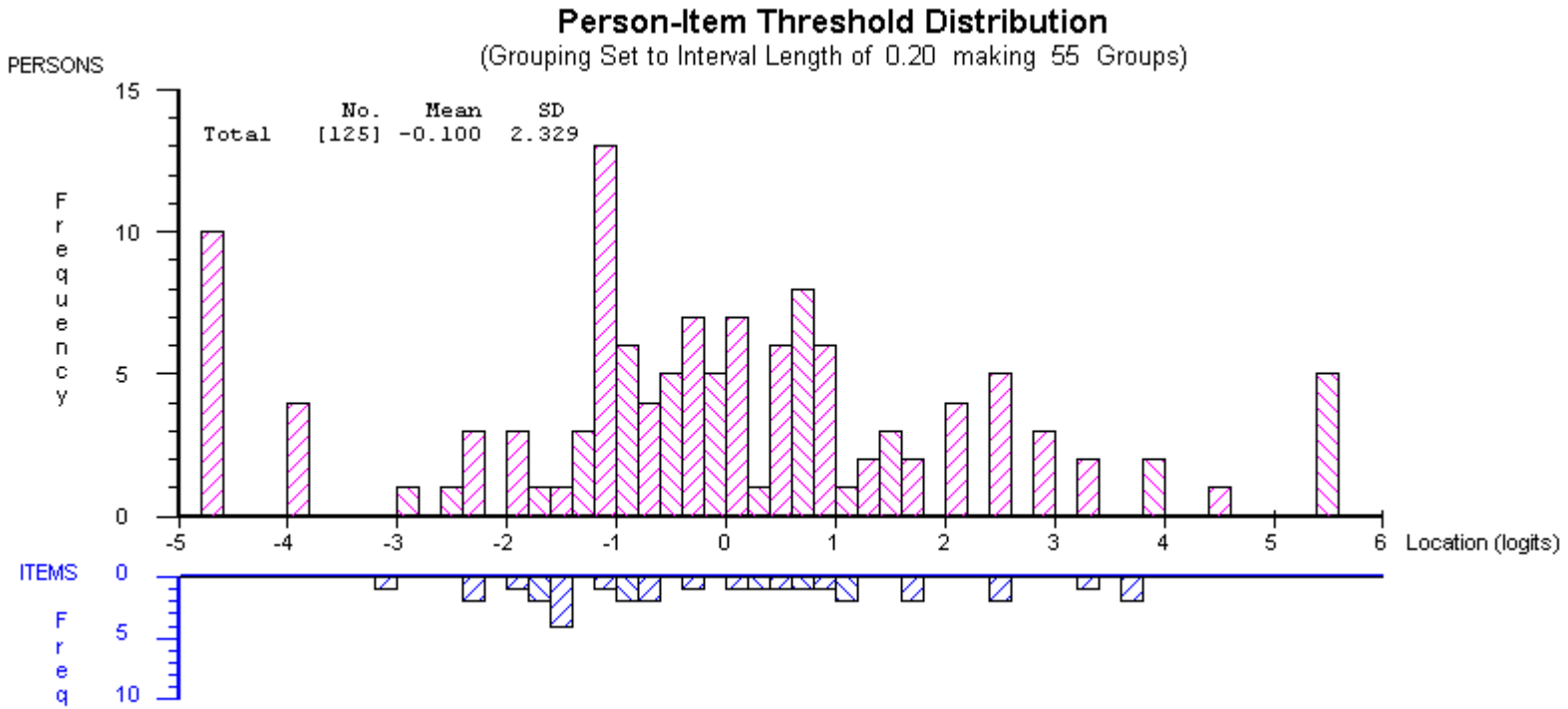
Table 3: Fit of the final 10 items of the Joint Protection –Self Efficacy Scale (JP-SES) (in descending order of p-values)

I *am confident that I can care for my joints when I …*	Mean Location = item difficulty	Fit Residual	X^2 value	P value
Items representing 'physical activities'				
lift a light object	-0.27	-0.24	0.55	0.76
carry a light object	-0.49	-0.43	0.96	0.62
cook *	-0.80	1.12	2.69	0.26
shop	-0.71	3.18	4.63	0.10
write	-0.76	0.33	5.05	0.08
Items representing 'affect'				
am in a bad mood §	0.99	0.95	0.16	0.92
am being watched	0.26	-1.38	1.73	0.42
do not feel like it	0.75	-1.03	2.43	0.30
am very busy	0.18	-0.26	2.64	0.27
have no time	0.83	-1.46	3.49	0.17

§ Most difficult item = requiring most self-efficacy, * easiest item = requiring least self-efficacy (mean location)

Figure 1: Person / Item-Threshold Targeting Graph of the JP-SES (persons n=125; items n=10)

Locations of persons (= person abilities) and of each item threshold on the interval scale, representing the measure of JP self-efficacy.



Easiest item threshold is from ‘not at all confident’ to ‘a little confident’ for the item ‘writing’ with mean logit of –2.75 (on the left). Most difficult item thresholds are from ‘quite confident’ to ‘very confident’ for the tasks ‘lacking time’ and ‘in a bad mood’ with mean logits of 3.47 and 3.49 respectively (on the right).