Development and Validation of a German version of the Joint Protection Behavior Assessment in people with rheumatoid arthritis

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

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

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

[3] School of Health Professions, University of Brighton, Eastbourne, UK

[4] Department of Epidemiology and Caphri Research Institute, Faculty of Health
Sciences, Maastricht University, the Netherlands

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

Objective: Joint protection (JP) is an important part of the treatment concept for patients with rheumatoid arthritis (RA). The Joint Protection Behavior Assessment short form (JPBA-S) assesses the use of hand JP methods by RA patients while preparing a hot drink. The aim of this study was to develop a German version of the JPBA-S (D-JPBA-S) and to test its validity and reliability.

Methods: A manual was developed through consensus with 8 occupational therapists (OT) experts as the reference for assessing the patients' JP behavior. Twenty-four RA patients and 10 healthy individuals were videotaped while performing ten tasks reflecting the activity 'preparing instant coffee'. Recordings were repeated after 3 months for test-retest analysis. One rater assessed all available patient recordings (n=23, recorded twice) for test-retest reliability. Ten randomly selected patients and all healthy individuals' video recordings were independently assessed by 6 OTs for inter-rater reliability explicitly asked to follow the manual. Rasch analysis was performed to test construct validity and transform ordinal raw data to interval data for reliability calculations.

Results: Nine of the ten tasks fit the Rasch model. The D-JPBA-S, consisting of nine valid tasks, had an intra-class correlation coefficient (ICC) of 0.77 for inter-rater reliability and 0.71 for test-retest reliability.

Conclusions: The D-JPBA-S provides a valid and reliable instrument for assessing JP behavior of RA patients for use in German language speaking countries.

Introduction

People with rheumatoid arthritis (RA) experience physical impairment and functional limitations, even though impressive advances in drug treatment have been achieved (1). A multidisciplinary approach in the management of RA is important, with physiotherapy and occupational therapy aiming at maintaining or improving independence and quality of life (2,3).

Hand involvement during the course of the disease occurs inevitably. Within five years of onset, finger and wrist joints are affected (4) and destruction of the dominant hand is more frequently observed (5). Joint protection (JP) is therefore an important intervention. Principles of JP have been developed based on anatomical and biomechanical research to guide occupational therapists (OTs) in their work with RA patients (6), e.g. altering working methods (use of proximal joints, dynamic activities), energy conservation (balance between activity and rest) and using assistive devices.

JP has beneficial short-term effects on pain and function in patients with established RA and moderate functional problems (7,8). Using assistive devices reduces pain during task performance (7) and altering working methods reduces difficulties in activities of daily living (ADL) (8). If JP is taught using behavioral education methods, it can also have long-term impact on reducing pain and maintaining function for those with less than 5 years disease duration (13-15). JP is also taught as a preventive intervention to patients with recent disease onset. However, its effectiveness at this early stage has not yet been convincingly demonstrated (9).

Discrepancy has been found between self-reported and observed JP use (8). An assessment instrument to systematically and objectively evaluate effectiveness of JP interventions is called for. Hammond et al. developed and evaluated the Joint Protection Behavior Assessment (JPBA) (10) assessing JP use while preparing instant coffee and a snack meal, since most JP methods taught clinically focus on protecting hand and wrist

joints during kitchen activities. In the JPBA five JP principles (6) are assessed while performing finger-wrist activities: 1) Reducing effort, by using labor saving gadgets, assistive devices, avoiding lifting, and good workplace organization, 2) Distributing load over several joints, 3) Using joints in stable positions, 4) Use of stronger, larger (proximal) joints and 5) Avoiding positions of deformity.

The original JPBA consists of 20 tasks integrating these JP principles. Several aspects of validity and reliability of the JPBA have been extensively examined (10,11) and the JPBA has been used in clinical studies (12-16). A short version (the JPBA-S), consisting of 10 tasks for the activity 'preparing instant coffee' has been tested as reliable compared to the full-length JPBA (Spearman correlation of 0.94) (11). The cultural adaptation and validation of a German version was therefore based on the short form. 'Preparing instant coffee' is also a common activity in Switzerland and requires little time, which minimizes the risk of fatiguing people with RA.

The aims of this study were to cross-culturally adapt the original JPBA-S to a German (Deutsch) version (D-JPBA-S) and to develop an assessment manual and assess the psychometric properties in a German speaking Swiss RA population. Before assessing inter-rater, intra-rater and test-retest reliability, special attention was paid to the construct validity of the D-JPBA-S.

Participants and Methods

Participants

Eight Occupational Therapists (OTs) from different hospital rheumatology departments in Zurich, Switzerland, experienced in treating people with RA and teaching JP, were invited to participate in the development of the German manual.

Twenty-four patients were consecutively recruited from the outpatient facility of the Department of Rheumatology, University Hospital Zurich, between June and July 2004, to video record their JP performance. They all fulfilled the 1987 ACR classification criteria for

RA (17), were on stable DMARD treatment, including anti-TNF treatment, steroids and non steroidal anti-inflammatory drugs (NSAIDs) for at least 4 weeks and had mild to moderate disease activity (DAS28 < 5.1). All participants had received at least one JP instruction session since onset of the disease. People with severe RA and functional limitations preventing JP behavior or independent task performance were excluded. All but one person (due to exacerbation of co-morbidities) participated in the test-retest recordings after 2-3 months (n=23). All patients had stable disease activity during this period. Self-perceived disease activity, measured with RADAI (Rheumatoid Arthritis Disease Activity Index), and erythrocyte sedimentation rate (ESR) remained unchanged (Table 1). However, in two patients the DMARD dosage was slightly increased. Ten non-health professional employees of the University Hospital, without health problems, age and gender matched to the RA group, participated as controls (Table 1). The Local Research Ethics Committee approved the study protocol and all individuals provided informed consent prior to participation.

Six therapists (4 OTs and 2 physiotherapists (PTs)), recruited from different rheumatology departments in Zurich, assessed the video recordings.

Manual Development

Face validity: The tasks contained in the UK JPBA-S were checked for cultural applicability, as there might be differences in equipment used in Switzerland. Hence, 'putting in an electric plug' was removed. This is not a physically difficult task for Swiss people with RA due to different plug design. It was replaced by 'opening a milk pack' – a convenient alternative applying several JP principles and frequently used in JP education. Filling, carrying and pouring kettle tasks were replaced with holding, carrying and pouring a pan (optionally an electric kettle) as saucepans are more commonly used in Switzerland to boil water. We anticipated both men and women performed all 10 tasks as routine daily activities. To determine face validity, JP literature and the UK JPBA manual were reviewed

to identify which JP principles were being applied during task performance. Several principles can be applicable to each task dependent on method of performance. There are some differences to the UK JPBA-S as some tasks are performed differently in Switzerland (Table 2).

Content validity: All methods of performing the 10 tasks were described in behavioral codes. These either described normal hand use (i.e. as performed by healthy people) or hand use consistent with joint protective adaptation in people with RA. We translated all codes of the selected tasks from the UK JPBA-S and added new codes found in German leaflets or books about JP, reported by experienced rheumatology OTs and identified in the video recordings of individuals with and without RA. In total 91 behavioral descriptions were generated for the 10 tasks - between 6 and 11 for each. A draft manual containing these (illustrated with photographs to ensure understanding) was developed and sent to eight OT experts. They were asked individually to score each code as correct, partially correct or incorrect JP behavior for people having mild to moderate RA with wrist and hand involvement, but without severe finger, hand, elbow or shoulder deformities, as these can lead to difficulty performing common JP methods and require more idiosyncratic solutions. Final scores allocated to the behavioral codes were based on the preliminary decision that consensus about being a 'correct' or an 'incorrect' method by at least 6 of the 8 expert OTs was necessary. Descriptions that not achieved this level of consensus were scored as 'partially correct'. The manual was used as reference for assessing the video recordings.

Video recordings and additional measures

Video recordings were completed in kitchen facilities of the University Hospital Zurich. Participants were asked to use the same styles of faucet, container for boiling water, milk pack and assistive devices they normally used, to ensure the assessment situation was as similar as possible to their home. All utensils were weighty enough to offer sufficient resistance to require a JP response from participants with RA. Participants were asked to

make two cups of coffee in the same sequence and manner they would normally do at home. They were kept unaware about the true purpose of the video recordings in order to reduce socially desirable responses. They were informed that the video camera would only focus on their hands and not their faces to preserve anonymity. Light conversation continued during the video recording to distract participants from consciously paying attention to their hand movements. The assessment was repeated after 3 months. Video recordings were transferred to Pinnacle Instant CD/DVD 8.0 software and edited on compact discs for assessments.

The following parameters were measured for people with RA:

Physical functional ability: using the Hochberg functional classes (18) and the Health Assessment Questionnaire (HAQ), a disease specific self-administered 20-item questionnaire (19).

Self-perceived disease activity and typical RA symptoms such as pain and morning stiffness: using the Rheumatoid Arthritis Disease Activity Index (RADAI), a self-administered 5-item questionnaire (20).

General health status: using a 10cm visual analogue scale (VAS) with the endpoints of 'bad' and 'excellent'.

Impairment of the dominant hand: Range of motion (ROM) for active wrist joint motion was measured using a goniometer. Finger and wrist joint deformity were assessed using the Joint Alignment and Motion Scale (JAM) (21). Grip strength was measured using a Jamar hand dynamometer (22).

Disease activity: using the Disease Activity Score (DAS28), calculated from the results of a 28 tender joint count, a 28 swollen joint count and erythrocyte sedimentation rate (23).

Assessment procedures

Discriminant validity

Assessments were performed with RA patients and healthy people to determine if their behavior differed regarding JP.

Cross-sectional validity

JP behavior within the RA patient group was correlated with functional impairment (assessed with HAQ) and hand pain (assessed with RADAI pain items).

Reliability assessments

Four OTs and two PTs independently assessed JP performances of ten randomly selected RA patients and all ten healthy subjects (inter-rater reliability). Two random duplicate video recordings of two patients A and B, were included to determine intra-rater reliability. Raters were blinded to the presence of duplicates. These duplicates were reassessed four weeks later by all raters, thus simulating the clinical situation of OTs reassessing their patients. One of the six raters assessed the video recordings of all 23 patients at both time points. The raters were asked to strictly follow the manual to minimize observer drift whilst assessing.

Statistics

Rasch analysis

The Rasch model reverses the traditional view of data-model relationship, i.e. data must fit the model, meaning that the observed frequencies should not differ too much from expected values (24,25). Rasch model theory states that response probabilities change as a function of participant ability and item difficulty (expressed as 'logits'), i.e. the probability of a person with a logit score of 1.0 is 50% to pass an item with a difficulty of 1.0 logits, but less (or more) to pass one with a difficulty of > 1 logit (or < 1 logit respectively). Rasch models provide various error estimates and fit statistics, especially for testing unidimensionality (i.e. if indeed a single dominant trait is being measured) and scale additivity (i.e. the probability that difficult items are only passed by high-scoring participants whereas less able participants only pass easier items).

This particularly allows gathering of further evidence of the construct validity of a measure. Each item and person are calibrated to provide an difficulty estimate and an ability estimate respectively of the location on an abstract linear continuum from 'less' to 'more', thus providing an equal interval scale representing the variable, in this case JP behavior.

Statistical analysis

The Rasch Partial Credit Model was applied (26), as the steps (thresholds) between the adjacent scores (incorrect / partially correct / correct = 0/1/2) might be different across tasks. The raters were also accounted for as a 'person factor' to control for bias. Complete data for all 120 ratings were available, 90 ratings without extreme scores (zero points) (27) were analyzed for construct validity.

Individual item fit to the model was examined with alpha at 5%. To reach overall probability in the 10-item-D-JPBA-S testing, Bonferroni's correction was used throughout and thus the significance values were set at 0.005 (28).

All reliability tests were performed for the D-JPBA-S (O) (i.e. using ordinal raw scores for all 10 tasks) and the D-JPBA-S (R) (i.e. using linear data of all tasks fitting to the Rasch model). Intra-class correlation coefficients (ICCs) were calculated using two-way random effect models and consistency definition for all reliability measures. The ICC provides information on the ability of 2 or more observers to differentiate between subjects. For inter-rater reliability (model 2,6) we expected an ICC of around 0.80. For intra-rater reliability (model 2,1) we expected an ICC of around 0.80. To evaluate real changes in clinical practice and research, a test-retest change determined by a specific measurement must be at least the Smallest Detectable Difference SDD (=$1.96\sqrt{2}SEM^2$, where SEM (Standard Error of Measurement) is $SD\sqrt{(1-r)}$ and r the reliability coefficient) (29). Pearson's correlation coefficients were calculated to measure associations between the D-JPBA-S (R) data and disease specific data; Mann Whitney U test was used to test differences between healthy individuals and people with RA. Rasch analysis was

performed using the Rasch Unidimensional Measurement Model RUMM2020 software package. All ICC calculations and statistical testing were performed using the SPSS 12.0 release.

Results

Content validity of the D-JPBA-S

Agreement on the scores between six or more of the eight OT experts was achieved for 53 of the 91 behavioral descriptions (58%), scoring 22 descriptions as 'correct' and 31 as 'incorrect.' There was insufficient agreement on 38 descriptions, therefore scored as 'partially correct'.

Construct validity of the D-JPBA-S using Rasch analysis

Examining fit of the 10-tasks-D-JPBA-S data to the Rasch model revealed that task 1 (turn on tap), task 2 (hold pan) and task 4 (carry pan) were significant at $p < 0.005$ (Chi Sq probabilities, all after Bonferroni's correction), i.e. the observed values of these three tasks were significantly different from the expected values and thus did not fit the model (Table 3). Additionally, the thresholds for task 2 (hold pan), task 4 (carry pan) and task 8 (open milk pack) were disordered, i.e. their scoring categories were not progressing in a logical order. It can be expected that as a person's ability increases, it will be more likely for him/her to obtain a higher score, however, in case of disordered thresholds, the items are not working in this way. Subsequently, scoring categories 1 and 2 for the three disordered tasks were collapsed, resulting in dichotomous data of 0 (for 'incorrect' and 'partially correct') and 1 ('correct') for tasks 2, 4 and 8. After rescoring, task 8 (open milk pack) still did not fit the model at the 0.5% significance level. No uniform differential item functioning (DIF) was found, meaning that no task was biased by raters, gender or age. Therefore task 8 was removed, resulting in model fitting of all remaining items, i.e. a valid assessment was obtained (Table 3). Test-of-fit statistics shows a item mean location (i.e.

difficulty) of 0 (SD 2.2) and a person mean location (i.e. ability) of -3.4 (SD 1.3), implying participants' ability was too low in relation to items' difficulty (Table 4). The formal test of invariance (item-trait interaction) revealed a total-item chi square of 46.4 (p=0.001), indicating significant deviation between the observed data and what was expected from the model at group level. Reliability Indices were 0.79 (Person-Separation-Index, indicative of the power of the D-JPBA-S to discriminate amongst the respondents) and 0.77 (Cronbach Alpha).

Person logits of the D-JPBA-S (R) were transformed into an arbitrarily chosen 0 – 18 interval scale for further calculations.

Reliability

The demographic characteristics of the healthy participants and those with RA were comparable (Table 1). Mean values of the six raters' scorings were between 3.5 and 5.4 on the D-JPBA-S (R) scale and differences were not significant (Kruskall Wallis H; p=0.50).

Inter-rater reliability

Overall inter-rater reliability for the D-JPBA-S (O) was 0.79 (CI 0.74 – 0.85), ranging between 0.84 (CI 0.76 - 0.91) and 0.70 (CI 0.54 - 0.82) for each pair of raters. Reliability values slightly decreased when calculated for the D-JPBA-S (R), being 0.77 (0.70 - 0.83) across all raters and ranging between 0.84 (0.75 - 0.90) and 0.65 (0.47-0.80) for each pair of raters.

Intra-rater reliability

Intra-rater reliability of each rater was generally higher in the assessments of time point 1 than of time point 2, which was true for both patients A and B. Intra-rater agreements' range for patient A were 80 – 100% (mean 95%,SD 8.4) at time point 1 and 50-100% (mean 75%, SD 20.7) at time point 2; for patient B values were 100% at time point 1 and 70-100% (mean 90%, SD 11.7) at time point 2. Results were the same for D-JPBA-S (O)

and D-JPBA-S (R). As raters scored this sample within a very restricted range, resulting in low variability, ICC calculations were not applicable.

Test-retest reliability

Patients repeated the kitchen activity after approximately 11 (SD 2.5) weeks. ICC was 0.65 (CI 0.27 - 0.87) for the D-JPBA-S (O) and 0.71 (CI 0.31-0.88) for the D-JPBA-S (R). On the 18-point linear scale, the median D-JPBA-S (R) score on test 1 was 7.7 points (IQR 3.4-9.3) and 5.8 (IQR 3.35-7.7) on test 2. Score changes over the two tests were between 0 and 12.2, mean score change was 1.1 point (SD 3.7). SDD was 5.5 points on the linear scale.

Disease related factors and JP behavior

Calculations in this section were performed with D-JPBA-S (R) linear data. The D-JPBA-S (R) scores of the RA participants were negatively correlated with grip strength (r=−0.63; p<.001). Correlations with all other disease related factors were significantly positive. No correlation was found with range of motion at the wrist joint (Table 1).

Discriminant validity

The D-JPBA-S (R) discriminated significantly (Mann Whitney U; Z=-8.215; p<.0001) between healthy people and those with RA regarding JP behavior. The median score was 0 points (IQR 0–2.9) for healthy people and 6.5 (IQR 5.0-9.0) for people with RA (Table 1).

Cross-sectional validity

Pearson's correlation coefficient of (D-JPBA-S (R) measured) JP behavior with functional impairments in RA patients was r=0.42 (p<.0001) and correlation with hand pain was r=0.57 (p<.0001) (Table 1).

Discussion

The final version of the D-JPBA-S, as obtained with Rasch modeling, i.e. consisting of nine valid tasks, is suitable for measuring JP behavior.

Traditionally, analysis of outcome data focused on summing and dividing raw scores that are ordinal, however calculations with such data may not be justified. Performing Rasch analysis is far more than a conceptual issue and its results had practical implications for the construct of the D-JBPA-S. Task 2 (hold pan) and task 4 (carry pan) were difficult to assess whether the assisting hand held the pan's weight (scored as correct) or was only supporting (partially correct). Therefore, raters may have randomly assigned scores, not perceiving a substantial difference. Collapsing incorrect and partially correct scores in this case is advantageous without losing information. Task 8 (open milk pack) was not an appropriate item, as it did not discriminate between JP performance of healthy people and RA patients. Both groups had trouble opening it, irrespective of health status or awareness of caring about joints.

Reliability calculations were based on linear scores of the D-JPBA-S (R) as well as on summed raw scores of the D-JPBA-S (O). Reliability for the D-JPBA-S (R) was slightly lower as one item was deleted. Since the D-JPBA-S will be used as an evaluative assessment, it is of no use to collapse all ordered polytomous into dichotomous scales to raise reliability, as this would diminish precision.

More important is the accurate measurement of change between two time points by transforming raw scores to linear data as raw score changes might be misinterpreted. Every linear difference (test 2 – test 1) corresponds to a range of raw score differences, which differ depending on test 1 initial status (30). Test-retest reliability integrates variability within the patients' group and within the rater, i.e. a change of the rater's assessment might reasonably be due to different JP performances of some patients.

The period of time between test and retest may appear long. However, we anticipated that noticeable changes in habitual JP behavior could happen due to unpredictable daily pain changes. This was also identified in an earlier study (8) and confirmed in our video recordings. Different JP performances due to large pain changes in (few) individuals explain our SDD of 5.5 points (about 30 % of the total range), even though most patients were in a stable condition and the overall correlation between pain and JP behavior was moderate. Although the usually low initial scores promote large improvements, it might be difficult to detect true differences in individual patients, when disease dependant changes interfere with real changes.

The discrepancy between items' difficulty and persons' ability also illustrates that individuals without RA have no reason to perform JP and people with RA perform less JP than might be expected, as they do not recall JP instructions. Participants stated that the effective drug treatment had lowered their perceived need and their motivation to apply JP during daily activities. However, items' difficulty levels are very different and there is a large gap within the scale. Further development of the scale should take this into account, e.g. by weighting the items, to improve its appropriateness (31).

The manual describes the application of the D-JPBA-S and all possible JP methods, illustrated with pictures. This is essential to ensure reliable assessments. The decision about how much agreement was required to assign the scores correct – partially correct - incorrect for the manual was arbitrary. Requiring a higher level of agreement for 'correct'/'incorrect' scores means more 'partially correct' scores. People with RA often show partially correct behavior, having been told about JP behavior without fully understanding the principles behind it, or having developed their own idiosyncratic methods, and thus the potential for improvement is quite substantial. For example, a common easy-to-learn principle is to work bilaterally.

Intra-rater agreement between time points 1 and 2 was almost 100%. Raters may have

recognized the individuals and, being convinced of their first scorings, persisted in these. Intra-rater agreement after 4 weeks was considerably lower for some raters, suggesting they performed assessments in a more unbiased and critical fashion. This second value might therefore be more accurate and nearer to the reality of clinical practice in which assessments are repeated some weeks later in the course of OT intervention.

Our patients may be considered representative for the RA population on relevant characteristics such as gender, age, and disease severity and there is evidence that measurement constructs are stable across samples from a common population regardless of sample size (32). This validation provided the prerequisites for using the D-JPBA-S in research. Further analysis (e.g. using generalizability theory) is needed to allow estimations of change on an individual patient level.

References

1.   Scott DL. Prognosis and Clinical Course. In: St. Clair EW, Pisetsky DS, Haynes BF, editors. Rheumatoid Arthritis. 1st edition. Philadelphia: Lippincott Williams & Wilkins; 2004. p. 26-41.

2.   Stenström CH. Physical Therapy. In: St. Clair EW, Pisetsky DS, Haynes BF, editors. Rheumatoid Arthritis. 1st edition. Philadelphia: Lippincott Williams & Wilkins; 2004. p. 425-30.

3.   Backman CL, Fairleigh A, Kuchta G. Occupational Therapy. In: St. Clair EW, Pisetsky DS, Haynes BF, editors. Rheumatoid Arthritis. 1st edition. Philadelphia: Lippincott Williams & Wilkins; 2004. p. 431-9.

4.   Eberhardt KB, Fex E. Functional impairment and disability in early rheumatoid arthritis - development over 5 years. J Rheumatol 1995;22:1037-42.

5.   Mody GM, Meyers OL, Reinach SG. Handedness and deformities, radiographic changes, and function of the hand in rheumatoid arthritis. Ann Rheum Dis 1989;48:104-7.

6.   Cordery J, Rocchi M. Joint Protection and Fatigue Management. In: Melvin J, Jensen G, editors. Rheumatologic Rehabilitation: Assessment and Management. Bethesda: The American Occupational Therapy Association, Inc; 1998. p. 279-322.

7.   Nordenskiold U, Grimby G, Dahlin-Ivanoff S. Questionnaire to evaluate the effects of assistive devices and altered working methods in women with rheumatoid arthritis. Clin Rheumatol 1998;17:6-16.

8.   Hammond A. Joint protection behavior in patients with rheumatoid arthritis following an education program: a pilot study. Arthritis Care Res 1994;7:5-9.

9.   Freeman K, Hammond A, Lincoln NB. Use of cognitive-behavioral arthritis education programmes in newly diagnosed rheumatoid arthritis. Clin Rehabil 2002; 16:828-836.

10. Hammond A, Lincoln N. Development of the Joint Protection Behavior Assessment. Arthritis Care Res 1999;12:200-7.

11. Klompenhouwer PJ, Lysack C, Dijkers M, Hammond A. The Joint Protection Behavior Assessment: a reliability study. Am J Occup Ther 2000;54:516-24.

12. Hammond A, Lincoln N. The effect of a joint protection education programme for people with rheumatoid arthritis. Clin Rehabil 1999;13:392-400.

13. Hammond A, Lincoln N, Sutcliffe L. A crossover trial evaluating an educational-behavioural joint protection programme for people with rheumatoid arthritis. Patient Educ Couns 1999;37:19-32.

14. 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 (Oxford) 2001;40:1044-51.

15. 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.

16. Hammond A, Jeffreson P, Jones N, Gallagher J, Jones T. Clinical applicability of an educational-behavioral joint protection programme for people with rheumatoid arthritis. Br J Occup Ther 2002;65:405-12.

17. Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 1988;31:315–24.

18. Hochberg MC, Chang RW, Dwosh I, Lindsey S, Pincus T, Wolfe F. The American College of Rheumatology 1991 revised criteria for the classification of global functional status in rheumatoid arthritis. Arthritis Rheum 1992;35:498-502.

19. Brühlmann P, Stucki G, Michel BA. 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.

20. Fransen J, Langenegger T, Michel BA, Stucki G. Feasibility and validity of the RADAI, a self-administered Rheumatoid Arthritis Disease Activity Index. J Rheumatol 2000;39:321-27.

21. Spiegel TM, Spiegel JS, Paulus HE. The joint alignment and motion scale: a simple measure of joint deformity in patients with rheumatoid arthritis. J Rheumatol 1987;14:887-92.

22. Jones E, Hanly JG, Mooney R, Rand LL, Spurway PM, Eastwood BJ. Strength and function in the normal and rheumatoid hand. J Rheumatol 1991;18:1313-8.

23. Prevoo ML, van 't Hof MA, Kuper HH, van Leeuwen MA, van de Putte LB, van Riel PL. 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;38:44-8.

24. Rasch G. Probabilistic Models for some intelligence and attainment tests. Expanded ed. Chicago: University of Chicago Press; 1960/1980.

25. Cialdella P, Figon G, Haugh MC, Boissel JP. Prescription intentions in relation to therapeutic information: a study of 117 French general practitioners. Soc Sci Med 1991;33:1263-74.

26. Tesio L. Measuring behaviours and perceptions: Rasch analysis as a tool for rehabilitation research. J Rehabil Med 2003;35:105-15.

27. Betemps EJ, Smith R, Baker DG, Rounds-Kugler BA. Measurement precision of the clinician administered PTSD scale (CAPS): a RASCH model analysis. J Applied Meas 2003;4:59-69.

28. Gross Portney L, Watkins MP. In: Foundations of Clinical research. Applications to Practice. Norwalk, Connecticut: Appleton&Lange. 1993. p. 406-7

29. Gyatt G, Walter S, Norman G. Measuring Change over time: assessing the usefulness of evaluative instruments. Chronic Dis. 1087; 40:171-8

30. Bezruczko N. Raw score nonlinearity obscures growth. Rasch Measurement Transactions 2004;18:973-74.

31. Tennant A, Hillman M, Fear J, Pickering A, Chamberlain MA Are we making the most of the Stanford Health Assessment Questionnaire?. Br J Rheumatol. 35(6):574-8, 1996

32. Fisher WP, Jr. Physical disability construct convergence across instruments: towards a universal metric. J Outcome Meas 1997;1;87-113.

Table 1: Demographic and clinical characteristics of study participants

	Healthy controls (n=10)	RA patients (n = 23)		Correlation with D-JPBA-S (R) at baseline
		Baseline	3 months	
Gender, women / men (no.)	7 / 3	18 / 5		
Age, years, median (IQR)	57 (47-63)	63 (47-70)		
Disease duration median (IQR)	NA	11 (7-18)		
Functional class (Hochberg)	NA	2 (1-4)	NM	0.44**
DMARDs (no. of patients)	NA	22	22	
Steroids (no. of patients)	NA	11	11	
NSAIDs (no. of patients)	NA	10	10	
DAS28, mean (SD)	NA	3.2 (1.5)	NM	0.45**
ESR, mean (SD)	NA	14 (11.8)	17 (16.5) [◐]	0.42*
RADAI, mean (SD)	NA	2.6 (1.6)	2.6 (1.7) [◐]	0.43**
Hand pain (RADAI), mean (SD)	NA	1 (0.95)	0.8 (0.97) [◐]	0.57**
HAQ score, mean (SD)	NA	1.3 (0.6)	1.4 (0.6) [◐]	0.42**
General health (HAQ), median (IQR)	NA	7 (5-8)	6 (5-9) [◐]	-0.55**
JAM (§) median (IQR)	NA	2 (1-3)	1 (1.3) [◐]	0.48**
Grip strength (§) median (IQR)	NA	16 (6-24.5)	13 (7-26) [◐]	-0.63**
ROM (§) wrist flexion, mean (SD)	NA	48.2 (26.1)	57.6 (23.1) [◐]	0.03
ROM (§) wrist extension, mean (SD)	NA	36.3 (17.7)	43.9 (19.9) [◐]	-0.05

[◐] no significant change between baseline and 3 months   **p< 0.001   *p<0.05

(§) of dominant hand; SD = standard deviation; IQR = interquartile range

NA = not applicable; NM = not measured

DMARDs = disease-modifying antirheumatic drugs; NSAIDs = nonsteroidal anti-inflammatory drugs; DAS28 = Disease Activity Score in 28 joints; ESR = erythrocyte sedimentation rate; HAQ = Health Assessment Questionnaire; RADAI = RA Disease Activity Index; JAM = Joint Alignment and Motion scale; ROM = range of motion

Table 2: Face and Content Validity of the D-JPBA-S

D-JPBA-S tasks	Joint protection principles – Face Validity				
	1	2	3	4	5
Turn on water tap	√	√	√	√	√
Hold pan	√	√			√
Turn off water tap	√	√	√	√	√
Carry full pan	√	√			√
Open coffee jar	√	√		√	√
Close coffee jar	√	√		√	√
Pour hot water into cups	√	√	√		√
Open milk pack *	√				√
Hold milk pack to pour milk	√	√		√	√
Carry full cup(s)		√	√		√

Joint protection principles:

1) Reducing effort, through the use of aids, assistive devices and avoiding lifting, as well as good organization of workplace

2) Distributing load over several joints

3) Using joints in stable positions

4) Use of strongest, largest (proximal) joints

5) Avoiding positions of deformity

* omitted after Rasch Analysis
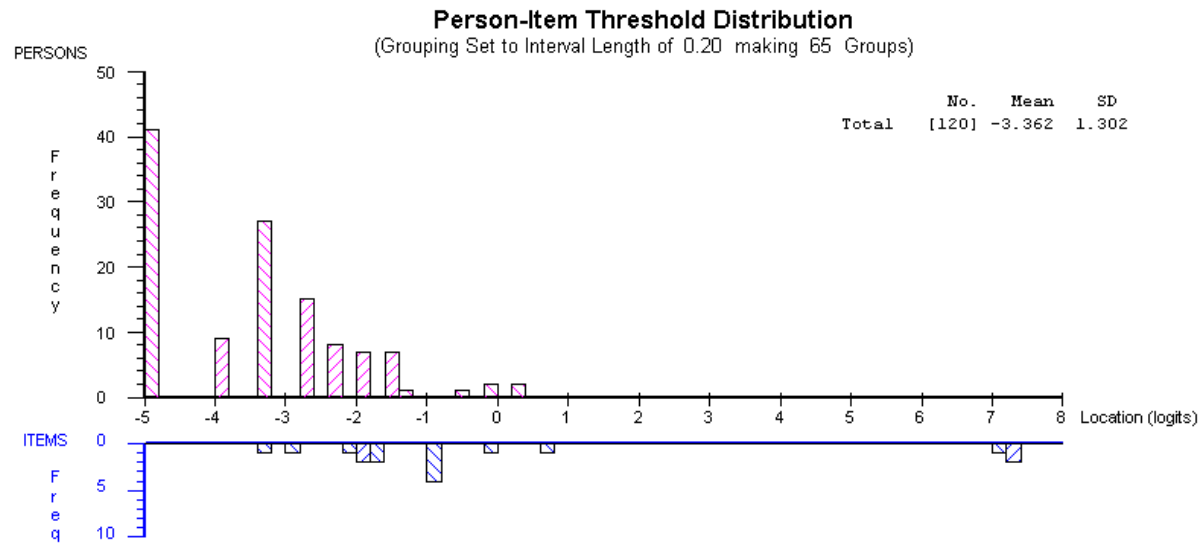
Table 3: Individual item fit of the D-JPBA-S (R)

Initial values are presented in serial order; values after re-scoring and after removing non-fitting task are presented in Chi Square Probability order

	Initial values unchanged			**After rescoring tasks 2,4,8**			**After removing task 8**		
	Item	ChiSq values	ChiSq prob.	Item	ChiSq values	ChiSq prob.	Item	ChiSq values	ChiSq prob.
Task 1	Turn on tap	12.34	0.002+	Turn on tap	1.04	0.569	Turn off tap	0.373	0.830
Task 2	Hold pan	20.18	0.000+	Carry pan	1.59	0.452	carry pan	0.418	0.812
Task 3	Turn off tap	1.11	0.574	Hold pan	1.60	0.450	Carry cups	3.500	0.174
Task 4	Carry pan	11.51	0.003+	Turn off tap	2.00	0.369	Pour water	5.322	0.070
Task 5	Open jar	6.45	0.040	Pour water	3.21	0.201	Turn on tap	5.598	0.060
Task 6	Close jar	8.50	0.014	Carry cups	3.78	0.151	Hold pan	6.707	0.035
Task 7	Pour water	7.00	0.030	Open jar	9.52	0.009	Pour milk	6.799	0.033
Task 8	Open milk	10.73	0.005	Pour milk	9.93	0.007	Close jar	6.995	0.030
Task 9	Pour milk	8.53	0.014	Close jar	10.05	0.007	Open jar	10.67	0.005
Task 10	Carry cups	4.11	0.128	Open milk	27.00	0.000+			

[+] significant at p < 0.005 (after Bonferroni's correction)

Table 4: Person / Item-Threshold Targeting Graph of the D-JPBA-S (R) ((ratings of) persons n=120; items n=9)

Locations of persons (= person abilities) and of each item threshold (one threshold for the dichotomous tasks 2 and 4, two thresholds for all other polytomous items) on the interval scale, representing the measure of JP behaviour.

Person-Item Threshold Distribution
(Grouping Set to Interval Length of 0.20 making 65 Groups)

	No.	Mean	SD
Total	[120]	-3.362	1.302

Easiest item thresholds are from 'incorrect' to 'partially correct' for the tasks 'pour milk' and 'turn on tap' with mean logits of –3.3 and 2.8 respectively (on the left). Most difficult item thresholds are from 'partially correct' to 'correct' for the tasks 'turn on tap', 'open jar' and 'pour water' with mean logits of 7.2 and 7.4 respectively (on the far right).