The Consultation and Relational Empathy measure: A modern psychometric evaluation using Rasch analysis

Paula Kersten, PhD

Person Centred research Centre, School of Rehabilitation and Occupation Studies

Auckland University of Technology, Private Bag 92006, Auckland 1142, New Zealand

Peter J White, PhD

Faculty of Health Sciences

University of Southampton, Highfield Campus, Southampton SO17 1BJ, United Kingdom

Alan Tennant, PhD

Department of Rehabilitation Medicine

Faculty of Medicine and Health, University of Leeds, D Floor, Martin Wing, Leeds General Infirmary, Leeds LS1 3EX, United Kingdom

Corresponding author: Paula Kersten

Keywords: patient-reported outcome measures, empathy, validity, Rasch analysis

The Consultation and Relational Empathy measure: an investigation of its scaling structure

Abstract

Purpose

The Consultation and Relational Empathy (CARE) measure is recommended to evaluate the quality of care. However, there is no evidence it is valid in rehabilitation. Aims 1) to examine the internal construct (factorial) validity of the CARE in the assessment of the patient-therapist relationship.

Method

CARE data were part of an experimental study of acupuncture and different currently used acupuncture placebo controls, including 213 patients (age 66.8, SD 8.3, 58% female) with chronic stable hip or knee pain of mechanical origin, waiting for a joint replacement. CARE was completed two weeks into the study and on completion two weeks later. Data analysis: Cronbach alpha, factor analysis and Rasch analysis.

Results

Internal construct validity was supported (82% of variance explained by the first factor; fit to the Rasch model χ^2=18.2, P=0.57). CARE was unidimensional, had local independence of items, good item fit, absence of Differential Item Functioning and invariance over time. Three percent of people did not complete items 9&10.

Conclusions

CARE satisfied strict criteria for internal construct validity. An interval scale transformation is available that can be used in clinical practice and research. Further work is required to investigate item non-response and how this may be dealt with in clinical settings.

Introduction

Following recent reviews of health care in the United Kingdom (UK) a shift has occurred from the emphasis upon increasing the quantity of healthcare, to improving the quality of healthcare [1]. Quality was defined as clinically effective, personal and safe. Consequently, organisations are placing increasing emphasis on person-centred health care and UK government directives enforce the measurement of 'patient-reported outcome measures' in certain areas of clinical practice such as hip and knee replacements [2]. For these procedures the Oxford Hip Score, Oxford Knee Score and the Euroqol (EQ-5D) are collected [2]. However, these measures do not seek patients' perceptions of the quality of healthcare in terms of the processes and interactions with clinicians. This is a missed opportunity, since there is now some evidence from empirical work that the relationship between clinicians and patients impacts not only on the satisfaction with healthcare but also on the outcomes of healthcare [3-5]. Whilst much of this research has been conducted to explore the relationship between doctors and patients, more recent studies are beginning to focus on the relationship between rehabilitation therapists and their patients [6-9]. A recent systematic review concluded that there is an association between therapeutic alliance and treatment outcomes in physical rehabilitation [7]. Therapeutic alliance is reportedly a multi-dimensional concept, including concepts such as rapport, trust, communication, empathy, mutual understanding, compassion and respect [4, 5, 7, 8, 10, 11], causing difficulty in determining what aspect is responsible for the desired enhanced outcome, or what to measure. We would argue that the concept of 'empathy' is a distinct construct, which is important in the therapeutic alliance [12, 13]. Empathy in the clinical context involves an ability to (i) understand the patient's situation, perspective and feelings (and their attached meanings); (ii) communicate that understanding and check its accuracy; and (iii) act on that understanding with the patient in a helpful (therapeutic) way [14, 15]. Interventions have been developed to improve empathy in

clinical encounters [16, 17]. Nevertheless, recent research has shown that physicians do not always demonstrate empathy to patients [18] and it is important therefore to continue improving and measuring empathy in clinical encounters. Several measures have been developed to measure empathy, for example, the Reynolds Empathy Scale [19], developed for use in nursing care, and the Jefferson Scale of Physician Empathy [20], a scale that was clinically led in its development.

A more recent patient derived measure, developed for use in primary care initially, is the Consultation and Relational Empathy measure (CARE). The CARE is a short patient-assessed measure that has been shown to provide doctors with direct feedback of their strengths and weaknesses in terms of relational empathy, as perceived by their patients [15, 21-25]. It is recommended for use in physiotherapy by the UK Chartered Society of Physiotherapy [12]. The CARE measure was developed following a review of existing measures [15], qualitative research and pilot studies [14] in primary care. It has been shown to be valid and reliable in primary and secondary care [22-25] and more recently also as part of a composite measure, the Consultation Quality Index or CQI of holistic interpersonal care in primary care consultations [4, 26]. To date validation studies have employed procedures, such as factor analysis, which is a commonly used approach to examine the construct validity of health status measures [27]. However, there are known limitations of using factor analysis on ordinal scales, including its parametric basis and the emergence of 'difficulty factors', which may spuriously indicate multidimensionality [28]. By contrast, Rasch analysis makes no assumptions that data are of interval nature and has more stringent criteria for testing unidimensionality. In addition, there is no evidence that the empathy measure is valid in assessing a patient-therapist relationship. The aim of this paper is therefore to examine the

internal construct (factorial) validity of the CARE in the assessment of the patient-therapist relationship, using factor analysis and Rasch analysis.

Methods

The CARE data were part of an experimental study, which aimed to investigate the relative effects of acupuncture and different currently used acupuncture placebo controls. The study took place in a primary care setting in Southern England. Patients were recruited via the Orthopaedic Consultant's operation waiting list and were consented on their first visit. Patients were eligible if they were waiting for a hip or knee joint replacement, had chronic stable pain predominantly from a single joint (hip or knee) of mechanical origin and were not on active treatment (apart from their normal analgesia). Those with serious co-morbidity, pregnant, prolonged or current steroid use, or waiting for a joint revision were excluded.

The patients' views on the therapist's empathy were measured with the CARE, which consists of 10 questions with five response options each (ranging from one to five, table 1). The scale ranges from 10-50, with higher scores reflecting more empathy [29]. Each question also has a 'does not apply' response option and when this is ticked the data for this question is entered as 'missing'. The scale's unidimensionality has been supported using factor analysis [24-26]. Since patients would not be able to comment upon the therapists' empathy before the start of the study these data were collected two weeks into the study (after which patients would have had four sessions with the therapists) and at the end (after a total of eight sessions). For the purpose of this paper we will refer to these two sets of data as 'mid point data' and 'end point data'. Information on gender, age, practitioner, treatment group, consultation type (empathic or not empathic), previous experience of acupuncture or joint affected was also collected.

Data analysis

Score distribution was explored using frequency tables. Item homogeneity was examined with a Cronbach alpha. There is debate in the literature what Cronbach alpha values are acceptable [30]. In the main it is accepted that values >0.80 are acceptable for group comparisons and values >0.90 for individual use [31]. A Factor analysis (Principal Component Analysis), conducted in SPSS15 [32], was used to confirm the unidimensional structure of the scale, with parallel analysis to determine how many factors were significant [33]. Data were then fitted to the Rasch measurement model (partial credit) [34, 35], using RUMM2020 software [36]. Detailed description of the Rasch model are given in key texts [37, 38].

In brief, Rasch analysis tests whether a set of items from a questionnaire conform to the Rasch measurement model: Each item within a scale has its own level of difficulty on the trait (item parameter) and a scale should consist of items that are easier and items that are harder to endorse. Similarly, respondents can experience different levels of empathy (person parameter). During the Rasch analysis item parameters are estimated independently from the person parameters, and once estimated can be placed along the same interval scaled ruler. The Rasch model is a mathematical algorithm [35]. It specifies that the probability of a correct response or endorsement of an item is a logistic function of the difference between the person and item parameter. Thus, converting the ordinal data to interval data results in a logarithmic interval scale. Therefore, the probability that a person will endorse an item is related to his or her level of experienced empathy and the item's level of difficulty. In other words, the probability of a positive response is a logistic function of the difference on the interval scaled ruler between the person and item parameter (figure 1).

Insert figure 1 about here

Fit to the Rasch model is considered acceptable when the observed data fit the predetermined Rasch model: Each item should have non-significant chi-square fit statistics and the item-trait interaction chi-square should be non-significant. Further requirements for fit to the model include non- substantial deviation of individual items and respondents from the Rasch model (individual item and person residuals should be within the range of +/- 2.5, average fit residual statistics should be close to a mean of zero and standard deviation of one).

Log-transformed item scores generated from the response choices should reflect the increasing or decreasing latent trait to be measured (threshold ordering). Thresholds are the points where the probabilities of a response of either 1 or 2, and 2 or 3 (and so forth) are equally likely. If the item responses options indeed reflect increasing amount of experienced empathy, then thresholds defining the categories should be ordered along the trait of empathy likewise. When a given level of empathy is not confirmed by the expected response option to an item, disordered thresholds will be observed.

The person separation index (PSI) is an indicator of the number of statistically different strata (groups) that the test can identify in the sample [39]. The PSI ranges from 0 to 1 [40]. Values ≥0.70 allow for group comparisons but for individual clinical use values should be ≥0.85.

The items should be unbiased (invariance) across key groups such as gender or age [41, 42]. Observed variance is termed Differential Item Functioning (DIF). We used analysis of variance of the residuals with the key group as the main factor to examine for DIF by age, gender, practitioner, treatment group, consultation type, previous experience of acupuncture,

or joint affected. Stability of the scale should show invariance across time points when fitting both the mid and end point data to the Rasch model.

An assumption of the Rasch model is that the answers to one item should not be dependent on the responses to another item, conditional upon the trait being measured. This conditional independence is examined by exploring the correlations between items' residuals, which should be smaller than 0.30 (data is said to be locally independent) [43].

Another key assumption of the Rasch model is that the scale is unidimensional. This is examined by creating two subsets of items. These are identified by a principal component analysis of the item residuals, with those loading negatively forming one set and those positively loading the second set [44]. Strict unidimensionality is then examined using an independent t-test on the two estimates derived from the subtests for each respondent. If the 95% confidence interval of t-tests includes 5%, unidimensionality is supported [44, 45].

Sample size

It is recommended that a 10:1 ratio of subjects to items is adequate for factor analysis [31]. For Rasch analyses, a range of sample sizes has been recommended depending on how well the items and persons are targeted. To have 99% confidence that the estimated item difficulty is within +/- ½ logit of its stable value the minimum sample size range is 108 to 243 (best to poor targeting) [46, 47]. Therefore, our sample size of 213 was deemed adequate for the purposes of the proposed analyses. Where data fit the Rasch model the observed raw total score can be transformed into interval scale measurement [40].

The Southampton & South West Hampshire and the Salisbury & South Wiltshire Research ethics Committees granted ethical approval (number 170/03/t).

Results

Two hundred and thirteen patients were entered into the study (mean age 66.8 SD 8.3, 58% female). Forty one percent were waiting for a hip replacement and 59% for a knee replacement. They experienced significant pain as measured on a visual analogue scale over a one-week period at the beginning of the study (median 59.3mm, interquartile range 48.0 to 68.3).

Median CARE scores were 44.0 (IQR 37 to 50). Frequency counts of responses to individual items indicate high scores on all items (table 1). In addition, 27% of participants attained the highest score of 50 on the measure, suggesting a ceiling effect. There was some missing data, especially for items 9 and 10.

Insert table 1 about here

Factor analysis showed a strong unidimensional set of items (82% of variance attributable to the first factor and one significant eigenvalue). When the data were fitted to the Rasch model the correlation matrix of the residuals gave support to the assumption of local independence of items and good summary fit to the model with strict unidimensionality (table 2, analysis 1).

Insert table 2 about here

There were no disordered thresholds in any of the items and all items were found to fit the model with item fit residuals between -1.4 and 1.9, and no significant chi-squares (table 3). There was no DIF for any of the items for gender, age category, practitioner, treatment group, consultation type, previous experience of acupuncture or joint affected. The PSI and Cronbach Alpha values both were 0.97. This indicates the scale is able to identify eight statistically different strata (groups) in the sample [48].

Insert table 3 about here

The Person-Item threshold distribution map shows that the item thresholds cover a wide range of the empathy construct and that many patients scored at the high end of the scale, that is, they perceived the consultation to be empathic (figure 2).

Insert figure 2 about here

To examine whether the CARE measure was invariant over time points, data from the mid point were put together with the data at the end point. Combined mid and end data fitted the Rasch model (table 2, analysis 2). An additional person factor of time ('mid' versus 'end') was created to enable the testing of invariance using DIF analysis, which was satisfactory (i.e. no DIF by time).

In the absence of DIF over time and resolution of non-unidimensionality we were able to produce a conversion table (table 4) that can be used to convert the raw ordinal scores to interval equivalent scores. These can be used when data are complete, and can be used in parametric data analyses given appropriate distributions.

Insert table 4 about here

Using the conversion table the mid point and end point data total scores were transformed to interval scaling. Mean (SD) mid point data was 42.2 (6.8) and mean end point data was 42.0 (7.2), which showed no difference in the level of perceived empathy across time (t=0.48, P=0.63).

Discussion

This analysis suggests that the CARE measure satisfies strict standards for internal construct validity, demonstrated by a fit to the Rasch model, so allowing for an interval scale transformation when required. It appears to be free of bias, invariant across time, and has a level of reliability consistent with individual use. The high Cronbach alpha and findings from the factor analysis are in line with results from previous studies [14, 24, 25]. Therefore, this measure appears a good candidate for the evaluation of the empathic nature of the consultation process between therapists and patients. This is important as it has been suggested that this process may be an important mediator for health outcomes [21, 49, 50] and one that is poorly addressed in practice [9, 51]. Using routine CARE data, collected in practice, can assist therapists in evaluating their practice, in communicating with their patients and improving the quality of care provided.

The general consistency between factor analysis and Rasch analysis is encouraging as is the consistency with pervious work on this measure [24, 25]. The factor analysis simply determines if a set of items form a unidimensional construct, which also is an assumption of the Rasch Model. We included this for direct comparison with previous work. However,

factor analysis gives no guarantee that such a set of items can be added together to enable transformation to interval scaling, which is the province of the Rasch model [52]. The latter makes more demands upon data than factor analysis, requiring a parametric form of probabilistic Guttman scaling in addition to the unidimensionality assumption.

A simple transformation of the manifest ordinal raw score into a latent interval scale metric is provided, and this can be used for calculating change scores in routine clinical practice, as well as research settings [53]. One limitation of this transformation is that it is valid only in the presence of complete data. Yet missing data were found on some items (especially in items 9 and 10), although this was lower to that found in other studies [22, 24]. Patients have the option to tick 'does not apply' to each CARE question. This may be one reason why people did not commit to providing a response to some questions. This response option may need to be reconsidered as missing values are a threat to the validity of the scale, as well as making it difficult for everyday use. Some solution to these problems should be found. For example, the current 'does not apply' option may need to be reconsidered.

Study limitations

This study included patients with osteoarthritis who were waiting for a joint replacement and further research needs to explore the scale in other populations. Our results build on previous reports on the scale's validity, though those were mainly by the originators and using traditional psychometric approaches [14, 22-24]. The main limitation of the scale appears to be, at least in the sample reported, a ceiling effect. This may have occurred as a result of our recruitment of subjects who were also taking part in a study of complementary medicine; it is possible that consent was given by those who view empathy more favourably than would

otherwise be the case, resulting in more prominent results. There is a need for replication in different samples.

The current sample size proved adequate for the study, although the higher levels of empathy meant that the sample was off-target, which requires a larger sample size to achieve the same degree of precision as a well targeted sample. As the extent of targeting is not known prior to the analysis, this has implications for determining the sample size in advance, and suggests that studies ought to consider sample size in the context of off-target samples (as we did), otherwise the degree of precision for item and person estimates may be much lower than required.

Conclusions

The 10 item patient-report scale, the Consultation and Relational Empathy measure, has been found to satisfy the Rasch model, a strict model for testing internal construct validity, and be free of Differential Item Functioning. It is a unidimensional scale, and an interval scale transformation is available. It is suitable to evaluate the empathic nature of the consultation process between therapists and patients. Some further work is required to investigate item non-response and how this may be dealt with in routine clinical settings.

Acknowledgements

The authors thank the participants in this study without whom this study would not have been possible.

Declaration of interest

The authors report no declarations of interest

This project was funded by a Department of Health Post Doctoral Fellowship Award (NCC RCD - CAMs 03/12). The funders played no role in the design, data collection and analysis or interpretation of the data. Similarly the funders were not involved in the production of this manuscript.

Authors' contribution

PW conceived of the study, was responsible for its design and coordination and helped to draft the manuscript. PK performed the analyses and drafted the manuscript. AT supported the analysis and manuscript draft. All authors read and approved the final manuscript.

Figure 1 Item Characteristic curve for item 10 of the CARE measure, demonstrating the probability that a person will endorse an item is related to his or her level of experienced empathy and the item's level of difficulty.

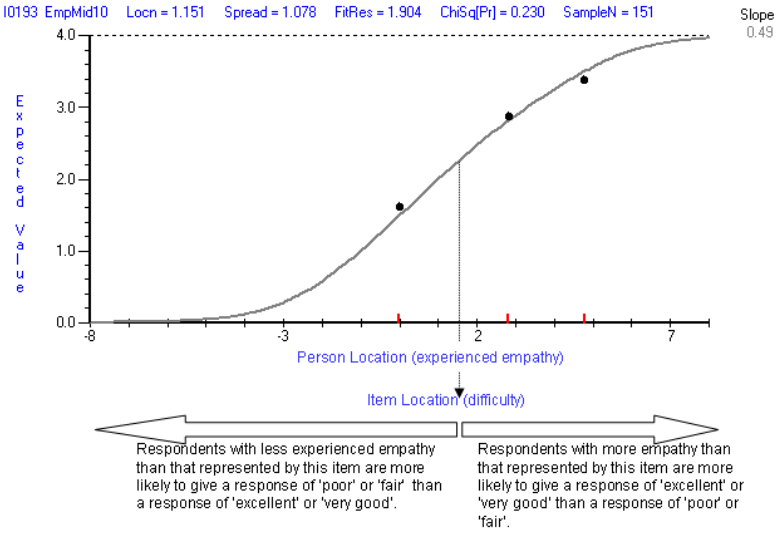
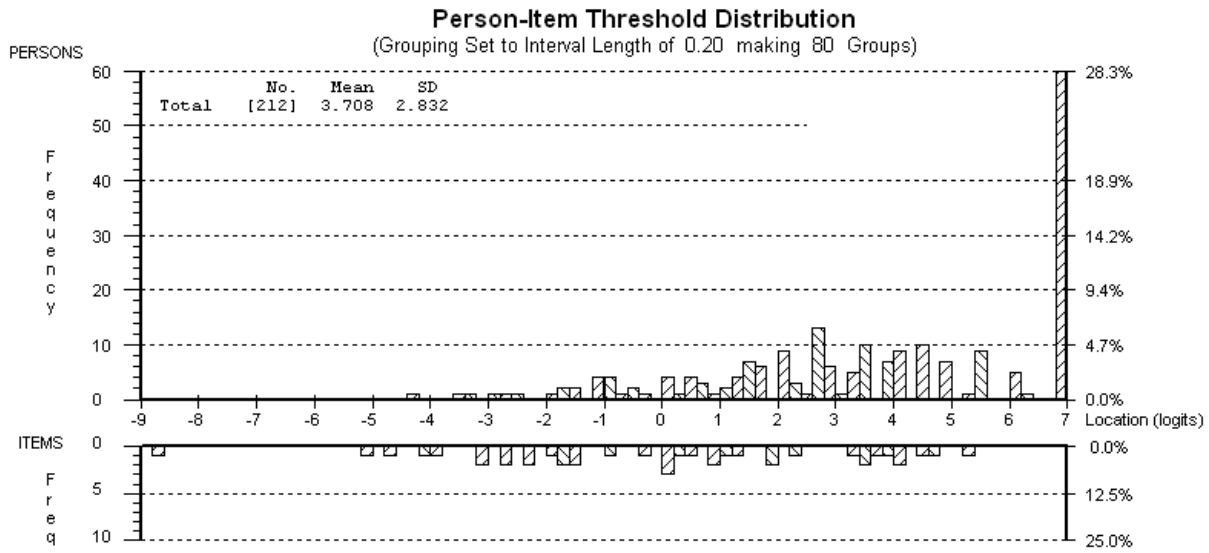


Figure 2 Person-Item Threshold Distribution Map CARE measure (mid point data)



Legend to Figure 2: The graph displays the person-item threshold distribution map with the x-axes displaying item threshold locations or difficulty (lower half) and level of empathy (person location) expressed by participants (upper half). The y-axes display the frequencies of item thresholds (lower half) and participants (upper half).

References

[1] Lord Darzi of Denham. High quality care for all: NHS Next Stage Review final report. Department of Health; 2008.

[2] Department of Health. Guidance on the routine collection of Patient Reported Outcome Measures (PROMs). For the NHS in England 2009/10; 2008.

[3] Dibbelt S, Schaidhammer M, Fleischer C, Greitemann B. Patient-doctor interaction in rehabilitation: The relationship between perceived interaction quality and long-term treatment results. Patient Educ Couns 2009;76:328-35.

[4] Howie JGR, Heaney DJ, Maxwell M, Walker JJ, Freeman GK, Rai H. Quality at general practice consultations: Cross sectional survey. BMJ 1999;319:738-43.

[5] Mercer SW, Watt GCM. Empathy is important for enablement. BMJ 2001;322:865.

[6] Greenfield B, Keough E, Linn S, Little D, Portela C. The meaning of caring from the perspectives of patients undergoing physical therapy: A pilot study. J Allied Health 2010;39:e43-e47.

[7] Hall AM, Ferreira PH, Maher CG, Latimer J, Ferreira ML. The influence of the therapist-patient relationship on treatment outcome in physical rehabilitation: a systematic review. Phys Ther 2010;90:1099-110.

[8] MacLeod R, McPherson KM. Care and compassion: part of person-centred rehabilitation, inappropriate response or a forgotten art? Disabil Rehabil 2007;29:1589-95.

[9] Besley JC, Kayes NM, McPherson KM. Assessing Therapeutic Relationships in Physiotherapy: Literature Review. New Zeal J Physiotherapy, 2011;39:81-91.

[10] Beattie P, Turner C, Dowda M, Michener L, Nelson R. The MedRisk instrument for measuring patient satisfaction with physical therapy care: A psychometric analysis. J Orthop Sports Phys Ther 2005;35:24-32.

[11] Cole MB, McLean V. Therapeutic relationships re-defined. Occup Ther Ment Health 2003;19:33-56.

[12] Chartered Society of Physiotherapy. Measuring for quality improvement in physiotherapy. [cited 28 Jan 2011]. Available from: http://www.csp.org.uk/topics/measuring-quality-improvement-physiotherapy.

[13] French S, Sim J. Physiotherapy. A psychosocial approach. 3 ed. Butterworth-Heinemann; 2004.

[14] Mercer SW, Maxwell M, Heaney D, Watt GC. The consultation and relational empathy (CARE) measure: development and preliminary validation and reliability of an empathy-based consultation process measure. Fam Pract 2004;21:699-705.

[15] Mercer SW, Reynolds WJ. Empathy and quality of care. Br J Gen Pract 2002;52(Suppl.):S9-S12.

[16] Leiblum SR, Schnall E, Seehuus M, DeMaria A. To BATHE or not to BATHE: Patient satisfaction with visits to their family physician. Fam Med 2008;40:407-11.

[17] David L, Freeman G. Improving consultation skills using cognitive-behavioural therapy: A new 'cognitive-behavioural model' for general practice. Educ Prim Care 2006;17:443-52.

[18] Morse DS, Edwardsen EA, Gordon HS. Missed opportunities for interval empathy in lung cancer communication. Arch Int Med 2008;168:1853-8.

[19] Reynolds W. The measurement and development of empathy in nursing. Aldershot: Ashgate Publishing Limited; 2000.

[20] Hojat M, Gonnella JS, Nasca TJ, Mangione S, Veloksi JJ, Magee M. The Jefferson Scale of Physician Empathy: Further psychometric data and differences by gender and specialty at item level. Acad Med 2002;77(10 Suppl.):S58-S60.

[21] MacPherson H, Mercer SW, Scullion T, Thomas KJ. Empathy, enablement, and outcome: an exploratory study on acupuncture patients' perceptions. J Altern Complement Med 2003;9:869-76.

[22] Mercer SW, Hatch DJ, Murray A, Murphy DJ, Eva KW. Capturing patients' views on communication with anaesthetists: The CARE Measure. Clin Governance 2008;13:128-37.

[23] Mercer SW, McConnachie A, Maxwell M, Heaney D, Watt GC. Relevance and practical use of the Consultation and Relational Empathy (CARE) Measure in general practice. Fam Pract 2005;22:328-34.

[24] Mercer SW, Murphy DJ. Validity and reliability of the CARE Measure in secondary care. Clin Governance 2008;13:269-83.

[25] Mercer SW, Neumann M, Wirtz M, Fitzpatrick B, Vojt G. General practitioner empathy, patient enablement, and patient-reported outcomes in primary care in an area of high socio-economic deprivation in Scotland-A pilot prospective study using structural equation modeling. Pat Educ Couns 2008;73:240-5.

[26] Mercer SW, Howie JG. CQI-2--a new measure of holistic interpersonal care in primary care consultations. Br J Gen Pract 2006;56:262-8.

[27 Matsudaira T, Igarashi H, Kikuchi H, Kano R, Mitoma H, Ohuchi K, et al. Factor structure of the Hospital Anxiety and Depression Scale in Japanese psychiatric outpatient and student populations. Health Qual Life Outcomes 2009;7:Article Number 42.

[28] Wright BD. Comparing Rasch measurement and factor analysis. Structural Equation Modeling 1996;3:3-24.

[29] Mercer SW. The CARE measure. [cited 28 Jan 2011]. Available from http://www.gla.ac.uk/departments/generalpracticeprimarycare/research/caremeasure/downloads/.

[30] Streiner DL, Norman GR. Health Measurement Scales: a practical guide to their development and use. 4 ed. Oxford: Oxford University Press; 2008:87-89.

[31] Nunnally JC, Bernstein IH. Psychometric theory. 3rd ed. New York: McGraw-Hill; 1994.

[32] Spss Inc, SPSS 15.0 for Windows. Release 15.0.1. 2007.

[33] Horn JL. A rationale and test for the number of factors in factor analysis. Psychometrica 1965;30:179-85.

[34] Masters G. A Rasch model for partial credit scoring. Psychometrica 1982;47:149-74.

[35] Rasch G. Probabilistic models for some intelligence and attainment tests (revised and expanded ed.). Chicago: The University of Chicago Press; 1960/1980.

[36 Andrich D, lyne A, Sheridan B, Luo G. RUMM 2020. Perth: RUMM Laboratory; 2003.

[37] Andrich D. Rasch models for measurement series: quantitative applications in the social sciences no. 68. London: Sage Publications; 1988.

[38] Bond TG, Fox CM. Applying the Rasch model. Fundamental measurement in the human sciences. London: Lawrence Erlbaum Associates; 2001.

[39] Wright BD. Reliability and separation. Rasch Measurement Transactions 1996;9:472.

[40] Wright BD, Stone MH. Best test design. Chicago: Mesa press; 1979.

[41] Grimby G. Useful reporting of DIF. Rasch Measurement Transactions 1998;12:651.

[42] Holland PW, Wainer H. Differential Item Functioning.NJ: Hillsdale. Lawrence Erlbaum; 1993.

[43] Sherridan B, Andrich D. Extending the RUM2030 analysis. Perth: RUMM Laboratory; 2009.

[44] Smith EV. Detecting and evaluation the impact of multidimensionality using item fit statistics and principal component analysis of residuals. J Appl Meast 2002;3:205-31.

[45] Tennant A, Pallant JF. Unidimensionality matters! (a tale of two Smiths?). Rasch Measurement Transactions 2006;20:1048-51.

[46] Linacre JM. Sample size and item calibration [or Person Measure] stability. Rasch Measurement Transactions 1994;7:328.

[47] Wright BD, Tennant A. Sample size again. Rasch Measurement Transactions 1996;9:468.

[48] Fisher WJr. Reliability Statistics. Rasch Measurement Transactions 1992;6:238.

[49] Mercer SW. Practitioner empathy, patient enablement and health outcomes of patients attending the Glasgow Homoeopathic Hospital: a retrospective and prospective comparison. Wiener Medizinische Wochenschrift 2005;155:498-501.

[50] Renzi C, Abeni D, Picardi A, Agostini E, Melchi CF, Pasquini P, et al. Factors associated with patient satisfaction with care among dermatological outpatients. Br J Dermatol 2001;145:617-23.

[51] Cooper K, Smith BH, Hancock E. Patient-centredness in physiotherapy from the perspective of the chronic low back pain patient. Physiotherapy 2008;94:244-52.

[52] Christensen KB, Kreiner S. Monte Carlo tests of the Rasch model based on scalability coefficients. Br J Math Stat Psychol 2010;63:101-11.

[53] Tesio L, Simone A, Bernardinello M. Rehabilitation and outcome measurement: where is Rasch analysis-going? Europa Medicophysica 2007 Sep;43:417-26.

Implications for rehabilitation

- The Consultation and Relational Empathy (CARE) measure CARE satisfies strict criteria for internal construct validity
- A transformation table is available, which can be used to convert the raw ordinal data into interval data
- The CARE is suitable to evaluate the empathic nature of the consultation process between therapists and patients

Table 1 Distribution of response frequencies of the Consultation and Relational Empathy measure (mid point data)

	Response					
	Poor (1)	Fair (2)	Good (3)	Very good (4)	Excellent (5)	Missing
Question	Freq (%)	Freq (%)	Freq (%)	Freq (%)	Freq (%)	Freq (%)
1. Making you feel at ease	-	6 (2.8)	23 (10.8)	69 (32.5)	114 (53.8)	-
2. Letting you tell your "story"	2 (0.9)	10 (4.7)	28 (13.2)	67 (31.6)	105 (49.5)	-
3. Really listening	3 (1.4)	9 (4.2)	21 (9.9)	65 (30.7)	113 (53.3)	1 (0.5)
4. Being interested in you as a whole person	7 (3.3)	17 (8.0)	34 (16.0)	56 (26.4)	98 (46.2)	-
5. Fully understanding your concerns	2 (0.9)	16 (7.5)	26 (12.3)	71 (33.5)	97 (45.8)	-
6. Being compassionate	4 (1.9)	10 (4.7)	32 (15.1)	58 (27.4)	108 (50.9)	-
7. Being positive	1 (0.5)	9 (4.2)	34 (16.0)	66 (31.1)	102 (48.1)	-
8. Explaining things clearly	1 (0.5)	10 (4.7)	20 (9.4)	61 (28.8)	119 (56.1)	1 (0.5)
9. Helping you to help yourself	5 (2.4)	24 (11.3)	36 (17.0)	62 (29.2)	79 (37.3)	6 (2.8)
10. Deciding on a treatment plan with you	7 (3.3)	17 (8.0)	34 (16.0)	60 (28.3)	87 (41.0)	7 (3.3)

Table 2 Fit of the Consultation and Relational Empathy measure data to the Rasch Model

Analysis number	Item fit residual		Person fit residual		χ^2 interaction		PSI	Unidimensionality Independent t-test (95% CI)
	Mean	SD	Mean	SD	Value (df)	P		
Mid Point*								
1	-0.011	1.241	-0.455	1.461	18.23 (20)	0.572	0.97	7.3% (3.8 to 10.8)
Combined mid and end point data**								
3	-0.203	0.152	-0.595	0.822	19.17 (12)	0.085	0.95	3.3% (0.7 to 5.8)

* Analysis 1: Fit to the Rasch model of data at the mid point.

** Analysis 2: Fit to the Rasch model of mid point and end point data combined.

Table 3 Item fit statistics (mid point data)

Item number	Item Location	Standard Error	Item Fit Residual	χ^2	Probability
1	-1.986	0.161	0.484	0.208	0.901
2	-0.252	0.148	-1.355	1.240	0.538
3	-0.265	0.149	-1.162	0.575	0.750
4	0.901	0.132	0.982	1.012	0.603
5	0.176	0.146	-1.392	4.909	0.086
6	0.088	0.141	-0.543	3.266	0.195
7	-0.381	0.15	-0.896	1.142	0.565
8	-0.774	0.152	1.669	1.701	0.427
9	1.342	0.138	0.195	1.241	0.538
10	1.151	0.133	1.904	2.938	0.230

Table 4 Conversion table for the total CARE scores

Raw total score (range 10 to 50)	Interval equivalent total score (range 10 to 50)
10	10.00
11	16.85
12	20.11
13	21.64
14	22.80
15	23.78
16	24.63
17	25.40
18	26.12
19	26.80
20	27.45
21	28.05
22	28.63
23	29.21
24	29.77
25	30.33
26	30.88
27	31.42
28	31.98
29	32.53
30	33.07
31	33.62
32	34.21
33	34.79
34	35.37
35	35.99
36	36.67
37	37.34
38	38.06
39	38.81
40	39.57
41	40.34
42	41.10
43	41.87
44	42.64
45	43.45
46	44.29
47	45.26
48	46.40
49	47.93
50	50.00