Available online at www.sciencedirect.com

ScienceDirect

journal homepage: http://ees.elsevier.com/hsag/default.asp

The BETA® nursing measure: Calibrating construct validity with Rasch analyses☆

Hendrik J. Loubser [a,b,*], *Daleen Casteleijn* [c,1], *Judith C. Bruce* [d,2]

[a] South African Database for Functional Medicine, Johannesburg, South Africa
[b] Department of Nursing Education, University of the Witwatersrand, South Africa
[c] Department of Occupational Therapy, University of the Witwatersrand, Johannesburg, South Africa
[d] School of Therapeutic Sciences, University of the Witwatersrand, Johannesburg, South Africa

A R T I C L E I N F O

Article history:
Received 21 February 2015
Accepted 26 February 2015
Available online 19 August 2015

Keywords:
Beta
Nursing
Routine
Measure
Validity
Rasch
Activities of daily living
Restorative nursing
Rehabilitation

A B S T R A C T

Background: The BETA nursing measure has been introduced as a tool to routinely measure and monitor the outcomes of patients' activities of daily living in a restorative nursing care context.

Objectives: To investigate the BETA's construct validity using the Rasch model with specific reference to the BETA's potential to be used as an interval scale providing metric or interval data.

Method: A quantitative analytical design was followed using Rasch analyses whereby BETA raw data was collected from patients ($n = 4235$) receiving nursing care in 28 South African sub-acute and non-acute nursing facilities. The data was prepared for Rasch analyses and imported into WINSTEP® Software version 3.70.1.1 (2010). Final results were shown by means of figures and graphs.

Results: A successful outcome was achieved by dividing the BETA into four subscales. In this process one of the original BETA items was omitted and seven other items required collapsing of their categories before the four subscales achieved a satisfactory fit to the Rasch model.

Conclusion: The four BETA subscales achieved “very well” to “excellent” levels of fit to the Rasch model. This finding thus creates an opportunity to convert the BETA's Likert qualities into an interval measure to calculate change in patients' activities of daily living metrically as a direct result of effective restorative nursing.

☆ Research significance: The BETA nursing scale was developed to provide ordinal scores on patients' activities of daily living in a restorative nursing context. In this study the BETA's construct validity is tested to confirm the extent to which the BETA can function as a standardised measure. If successful, routine BETA measurements can be used in valuable empirical calculations such as change in activities of daily living during nursing care, nursing performance and efficiencies of service delivery.

* Corresponding author. Box 2356, Houghton, 2041, Johannesburg, South Africa. Tel.: +27 82 574 5129 (mobile).
E-mail addresses: hennie@sadfm.co.za (H.J. Loubser), Daleen.casteleijn@wits.ac.za (D. Casteleijn), Judith.bruce@wits.ac.za (J.C. Bruce).
[1] 7 York Road, Parktown, 2193, Johannesburg, South Africa. Tel.: +27 21 717 3701.
[2] 7 York Road, Parktown, 2193, Johannesburg, South Africa. Tel.: +27 21 717 2063.

Peer review under responsibility of Johannesburg University.
http://dx.doi.org/10.1016/j.hsag.2015.02.001

1. Introduction

1.1. Background

Nurses in the specialised fields of rehabilitation, gerontology and long-term care share the same objective - they have to maximise the functional ability of people living with temporary or permanent disability, commonly known as restorative nursing. There is a lack of objectively validated nursing tools to routinely and empirically establish a patient's functional ability (Loubser, 2012). This means that neither the caregivers nor the nursing staff working routinely with these patients have any method for accurately measuring, communicating, monitoring or calculating their patients' restorative nursing care plans. To bridge this incongruity, a routine nursing scale to measure patients' activities of daily living, named the BETA, has been developed. The development was done with South African nurses and its utility in nursing was tested (Loubser, Bruce, & Casteleijn, 2013). Before implementing the BETA as a routine nursing measure, its construct validity, including its ability to be successfully converted from an ordinal scale with basic Likert scoring qualities into an interval measure with metric measurement qualities had to be tested. Once validated and successfully converted, the BETA measurements can be used in metric analyses to calculate changes in patients' activities of daily living as a result of effective restorative nursing. This article reports on the BETA's construct validity properties using the Rasch measurement model.

1.2. The Rasch measurement model (RMM)

The concept of internal construct validity refers to how well a scale correlates with the construct that it purports to measure in order to be successfully operationalised (Linacre, 2010). With this in mind, there is a strong tendency to move from qualitatively-ordered scales, e.g. those producing Likert-related ordinal scores not useful in inferential statistics, towards quantitatively-ordered interval measures that can be applied in metric outcomes analyses (Bond & Fox, 2007). The RMM is suited to perform this transformation. Although the RMM has been widely used in the education sciences over the last 40 years, this method of validating scales only became popular in the health sciences in the last decade with the reporting of a variety of health care measures being validated by the RMM (Tennant & Conaghan, 2007).

In the 1960s, George Rasch, a Danish mathematician, tried to find a solution for a particular problem the Danish Department of Defence experienced with educational tests. He discovered the relationships between human ability versus item difficulty and concluded a logic that became popular. Rasch detected an underlying probability principle in a data matrix of a well-constructed dichotomous test: "*a person having a greater ability than another person should have the greater probability of solving any item of the type in question, and similarly, one (test) item being more difficult than the other means that for any person the probability of solving the second (test) item is the greater one*" (Rasch, 1960, p. 117). This principle led him to devise a mathematical model to develop rules for a hypothetically perfect fundamental measure for social scientists, today known as the Rasch Measurement Model (RMM).

The original RMM was invented for dichotomous (yes/no response options) measures; and the Rasch relationship equation of the simple dichotomous formula is as follows:

$$B_n - D_i = \log(P_{ni}/(1 - P_{ni}))$$

where

B_n = Ability measure of person n
D_i = Difficulty calibration measure of item i
P_{ni} = Probability of a correct response from person n on item i
$1\text{-}P_{ni}$ = Probability of an incorrect response from person n on item i

In non-mathematical terms the logarithm of the odds ratio between the probability of passing an item and the probability of failing an item equals the difference between the ability of the person and the difficulty of the item. More explicitly, the Rasch analysis enables the calibration of item difficulty (e.g. where *Di* is placed on the straight line) and person ability (e.g. where B_n is placed on the same straight line). As both these calibrations are expressed in logits (log-odds probability units), they are additive in nature (Kottorp, 2003). The perfection of the RMM lies in its simplicity which also renders it applicable to all human sciences and is "*currently the closest generally assessable approximation of fundamental measurement principles in the human sciences*" (Bond & Fox, 2007, p. 14).

As the BETA has a polyotomous design (three or more response options) and characteristics, the Rasch-Masters Partial Credit Model was used in this study (PCM). Masters (1982) devised this Rasch derivative in an attempt to give partial credit for achieving a partially correct score when the "partial-correctness" structure differs from item to item in the same scale. His solution was that the Partial Credit Model recognises a partial-credit ratings scale as being specific to each item (Linacre, 2010).

$$\log_e\left(P_{nij}/P_{ni(j-1)}\right) = B_n - D_{ij}$$

The Partial Credit Model specifies the probability, Pnij, that person *n* of ability Bn is observed in category j of a rating scale specific to item i of difficulty Di as opposed to the probability Pni (j-l) of being observed in category (j-l) of a rating scale with categories j = 0. The rating scale structure (Fij) is now specific to item i. This means that partial credit items with the same number of categories and the same raw marginal scores, taken by the same people, can have different difficulties if the pattern of category usage differs between the items (Masters, 1982).

Rasch analyses provide a formal procedure to test scales against a mathematical formula for its construct validity. The results of the series of analyses guide the researcher in refining the scale to perfection.

This process of refining scale structure is referred to as scale calibration (Bond & Fox, 2007). If poor fit is achieved, poor measurement qualities are reported. However, the RMM will guide the analyst along a diagnostic pathway to identify under- and over-fitting characteristics in the scale and, if