Te Waka Kuaka, Rasch analysis of a cultural assessment tool in traumatic brain injury in Māori

Elder, H., Czuba, K., Kersten, P., Caracuel, A. and McPherson, K.

Elder H, Czuba K, Kersten P et al. Te Waka Kuaka, Rasch analysis of a cultural assessment tool in traumatic brain injury in Māori [version 1; peer review: 1 approved with reservations, 1 not approved]. *F1000Research* 2017, 6:1034 (https://doi.org/10.12688/f1000research.11500.1)

RESEARCH ARTICLE

Te Waka Kuaka, Rasch analysis of a cultural assessment tool in traumatic brain injury in Māori [version 1; peer review: 1 approved with reservations, 1 not approved]

Hinemoa Elder[1], Karol Czuba[2], Paula Kersten[3], Alfonso Caracuel[4], Kathryn McPherson[5]

[1]Te Whare Wānanga o Awanuiārangi, Whakatane, New Zealand
[2]Auckland University of Technology, Auckland, New Zealand
[3]University of Brighton, Brighton, UK
[4]University of Granada, Granada, Spain
[5]Health Research Council of New Zealand, Auckland, New Zealand

v1 **First published:** 30 Jun 2017, **6**:1034
https://doi.org/10.12688/f1000research.11500.1
Latest published: 30 Jun 2017, **6**:1034
https://doi.org/10.12688/f1000research.11500.1

Abstract

Background: The aim was to examine the validity of a new measure, Te Waka Kuaka, in assessing the cultural needs of Māori with traumatic brain injury (TBI).
Methods: Māori from around Aotearoa, New Zealand were recruited. 319 people with a history of TBI, their whānau (extended family members), friends, work associates, and interested community members participated. All completed the 46-item measure. Rasch analysis of the data was undertaken.
Results: All four subscales; Wā (time), Wāhi (place), Tangata (people) and Wairua practices (activities that strengthen spiritual connection) were unidimensional. Ten items were deleted because they did not fit the model, due to statistically significant disordered thresholds, non-uniform differential item functioning (DIF) and local dependence. Five items were re-scored in the fourth subscale resulting in ordered thresholds.
Conclusions: Rasch analysis facilitated a robust validation process of Te Waka Kuaka.

Keywords

traumatic brain injury, Māori, Rasch analysis, measurement

Open Peer Review

Approval Status ✗ ?

	1	2
version 1 30 Jun 2017	✗ view	? view

1. **Stephen McKenna**, Manchester University, Manchester, UK

2. **Sara Simblett**, King's College London, London, UK

Any reports and responses or comments on the article can be found at the end of the article.

Corresponding author: Hinemoa Elder (hinemoa@xtra.co.nz)

Competing interests: No competing interests were disclosed.

Grant information: This research was funded by the Health Research Council of NZ [14-060].
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

How to cite this article: Elder H, Czuba K, Kersten P *et al.* **Te Waka Kuaka, Rasch analysis of a cultural assessment tool in traumatic brain injury in Māori [version 1; peer review: 1 approved with reservations, 1 not approved]** F1000Research 2017, **6**:1034 https://doi.org/10.12688/f1000research.11500.1

First published: 30 Jun 2017, **6**:1034 https://doi.org/10.12688/f1000research.11500.1

Introduction

Traumatic brain injury (TBI) in Māori is a significant health problem. Recent population data shows that Māori youth are three times more likely to sustain clinically significant TBI compared to non-Māori (Feigin *et al*., 2013). A complicating factor in responding to Māori with TBI has been the lack of understanding of the cultural importance of injury to the brain and head to Māori, given the primacy placed on the head in Māori culture. For instance, 'he tapu te upoko' is a well-known saying from Te Ao Māori (the Māori world) which means, the head is sacred, clearly indicating the important 'place' of brain injury from a cultural perspective (28th Māori Battalion, 2011). Recent work has explored these concepts and developed a Māori theory and praxis of TBI (Elder, 2013a; Elder, 2013b). What that research found was that the concepts of wā (time), wahi (place), tangata (people) and wairua practices (activities that strengthen the unique connection between Māori people and the universe) were central to Māori in navigating recovery. Indeed, how much time is taken, where assessment and treatment takes place, who is present at assessments, and what culturally salient activities are embedded in these assessments and treatment are well understood by practitioners as being critical to the engagement of Māori whānau, although formal research in these areas has not been conducted. Practice-based evidence also shows that without these factors being implemented, Māori whānau disengage from services and therefore miss out on rehabilitation interventions, leading to compromised outcomes. Allocating enough time when working with Māori has recently been identified as vital to ensuring cultural practices are undertaken and therefore more accurate assessment and recommendations are provided (Elder *et al*., 2016). These aspects of comprehensive assessment of Māori may be in tension with clinical imperatives that emphasize efficiencies of time and prioritize brevity of assessment and treatment.

While some needs of patients and relatives after a TBI are held trans-culturally, others depend on the specific social and cultural context in which people live. As tools for the assessment of these needs are influenced by culture, measures adapted from other cultures have shown substantial differences between countries, even if they share historical roots and language (Norup *et al*., 2015).

Despite some, albeit variable, awareness by health practitioners and researchers of these cultural issues (Harwood, 2010; Harwood *et al*., 2012), no measures have been developed that might help conceptualize the magnitude and nature of the cultural needs associated with Māori TBI. Such measures should enable tailored responses to these needs and thereby improve recovery outcomes and communication between whānau and clinicians, and therefore improve the quality of assessments. The lack of such measures means that Māori cultural needs in the context of TBI lack recognition and attention, or if there is some awareness of these on the part of clinicians the approach is not systematically provided or monitored (Elder, 2012).

Measures used to monitor recovery and needs post-TBI, such as neuropsychological tests, have been developed elsewhere; and Māori cultural norms and validation in the Māori community have not been carried out, although such work is now underway in the context of the ageing brain (Dudley, 2016). This issue is well recognized as contributing to difficulties in interpreting scores for Māori (Ogden & McFarlane-Nathan, 1997). Experts in cross-cultural neuropsychology warn that adaptations of tools across cultures has serious drawbacks that affect all stages of the assessment: review of records; interviews; neuropsychological testing; and interpretation of results (Puente *et al*., 2013). Having measures developed by Māori for Māori is therefore a critical issue in ensuring cultural validity. Indeed, there continues to be some debate about what can be measured and how this could occur in a culturally authentic way, given the experience of historical measures being used as a means of cultural marginalization of Māori (Durie, 2004). Developing such measures aligns with the literature on Patient Reported Outcome Measures (PROM) that recognizes these measures as a central component to improving multiple facets of care and support, raising the quality of outcomes from illness and injury including in TBI (Friedly *et al*., 2014; Reeve *et al*., 2013). The need for a dual purpose tool which serves to assess both cultural needs and also measure outcomes with cultural salience for Māori is apparent from clinical experience, and is frequently requested by Māori whānau seeking tools they feel reflect their realities. The lack of such measures in the literature indicates this is a significant gap that needs to be addressed.

This study aimed to examine the internal construct validity of a post-TBI assessment of Māori cultural needs and outcome measures by Rasch analysis.

Methods

Study procedures and data collection

A 46 item draft scale was developed from verbatim quotes taken from transcripts of an earlier phase of the study (see Supplementary File 1), and refined using a culturally responsive method (Elder & Kersten, 2015). Rangahau Kaupapa Maori (Māori research approaches determined and conducted by Māori, with the goal of supporting Māori health advancement) were utilized.

The statements used in the first iteration of the tool came from Māori participants in marae wānanga (traditional learning fora). The items were then refined via four focus groups, with the final group of participants having experienced TBI. This was to ensure the items were acceptable to those with direct experience, and that the items had face validity in addressing the sub-scale areas and were easily understood. The measure was then completed by 319 participants from a range of settings in the North Island of Aotearoa, New Zealand, between June and November 2015. They included attendees at Kura Reo; a week-long total immersion Te Reo Māori wānanga (Māori language learning environment). The attendees had a range of proficiencies in speaking Te Reo Māori, from beginner to expert level.

People were invited to participate in two ways. First, via Māori health service providers, appointments were set up with the first author. Second, wānanga groups were offered participation and the first author provided a presentation about the project, answered questions and provided oversight of completion of the tool. Inclusion criteria were Māori with TBI, or non-Māori who were part of Māori whānau (extended families), for example by marriage,

whānau members, friends of Māori with TBI, those with work connections with Māori with TBI and Māori community members concerned about TBI. TBI was defined by self-reporting, as either confirmed, possible or unknown. Information was collected about TBI severity and placed into mild, moderate, severe and unknown categories, however given the questionable accuracy of self-reporting, this data was not included in our analysis. The emphasis here was on offering participation to whānau as well as to individuals affected by TBI. This reflects the centrality of whānau as a health and wellbeing construct, which is well recognised in Māori scholarship (Durie, 2001) and tikanga (cultural lore) (Moko-Mead, 2003). Indeed, the theoretical basis of this tool proposes that TBI affects the whole whānau and that the whole whānau needs to considered as "the patient" (Elder, 2013a). All 319 participants provided written informed consent. The research was approved by the Health and Disability Ethics Committee of NZ (14/CEN/17) and by Te Whare Wānanga o Awanuiārangi, the first author's institution (EC14 034HE). Participants were supervised by the first author or a research assistant, when completing the draft 46-item outcome measure. These data were then entered into the Rasch analysis software programme, RUMM2030 (Andrich *et al*., 2010).

The instrument resulting from the earlier research (Elder, 2013a) contained four subscales and 46 items. The four subscales were labeled Wā (time), Wāhi (place), Tangata (people) and Wairua practices (Wairua is defined here as an aspect of health and well-being characterized as a unique connection between Māori people and all aspects of the universe). The participants were invited to score each of the items as strongly agree, agree, disagree or strongly disagree. While debate continues around whether or not to include a neutral response option in surveys or assessment tools, the rationale used here aligns with others who have shown absence of a neutral option encourages mental effort to engage with the item and negates the effect of social desirability bias (Krosnick *et al*., 2002). Other demographic information was collected about each participant as presented in Table 1.

Table 1. Demographic characteristics of study participants.

	Category	Frequency	Percent of total
Age	11–25	81	25.4
	26–35	78	24.5
	36–50	86	27
	51–76	74	23.2
Gender	Male	118	37
	Female	200	62.7
	Trans	1	0.3
Relationship	Whānau	176	55.2
	Friend	48	15
	Job related	32	10
	Community member	63	19.7
Main Iwi of origin by Maori Electorate	Tāmaki Makaurau/ Te Tai Tokerau	183	57.4
	Hauraki Waikato	47	14.7
	Ikaroa-Rāwhiti	18	5.6
	Te Hauāuru	6	1.9
	Waiariki	56	17.6
	Te Tai Tonga	1	0.3
	Other	8	2.5
TBI type	Confirmed	183	57.4
	Possible	87	27.3
	Unknown	49	15.4

Data analysis

All analyses of each of the subscales were carried out using RUMM2030 (Andrich *et al*., 2010) in order to determine the fit of the data to the Rasch model. Rasch analysis is a probabilistic mathematical model that draws on item response theory with the advantage of estimating the item difficulty and the person ability separately, which is not possible using measures based on classical test theory (Hays *et al*., 2000). The 1-parameter logistic function enables item difficulty to vary but assumes all items discriminate equally. Before Rasch analysis is used to transform ordinal observation data into linear measures, the Rasch fit statistics are examined to enable assessment of any threats to linear measurement (Haigh *et al*., 2001; Whiteneck *et al*., 2011).

Rasch analysis is used to assess the measurement properties of existing measures and to guide the development of new ones (Czuba *et al*., 2016). The Rasch model states that the outcome of an encounter between a person and an item is governed by the product of the construct of interest of the person, together with the easiness of the item (Bond & Fox, 2001). The person's estimate of cultural needs is derived by dividing the percentage of items that scored highly, by the percentage of items scored in the low range, and then by taking the natural log.

Scalable items are important because they capture difficulty, and make the measurement useful in a practical sense. For instance, an item with high difficulty means it more urgently needs to be acted upon, and fluctuations can be monitored. Likewise, items which capture low, and intermediate levels of need are important in a measure, so that both lower and intermediate levels of need can be identified, and changes over time can be monitored and responded to. In the Rasch model, the item difficulty is estimated by calculating the odds of success in identifying those who scored highly and those who scored in the low range.

Each item within the scale has its own level of difficulty on the trait (item parameter), and every person has his or her own level of "ability/trait". Item parameters are estimated independently from the person parameters and once they are identified they can be placed along the same interval scaled ruler. The item and person performance probabilities determine the interval sizes on the "ruler" of the measure.

A number of tests were performed to assess the fit of the subscales to the Rasch model. Fit to the assumptions of the model can have a number of contributing factors which are explained in detail elsewhere (Andrich, 1988; Bond & Fox, 2001; Kersten & Kayes, 2011; Tennant & Conaghan, 2007). It is important to note that 'misfit' should not be taken to mean that the item has no merit or is of no interest, but rather that it does not fit the unidimensional structure of a measure (or in this case domain). If this is the case, collapsing scores or moving an item to a different domain is considered, for items that do not fit but add discriminatory information. Table 2 presents a brief overview of the central Rasch analytical concepts and the actions that can be taken in the case of conditions not being met for the transfer from ordinal to linear scores.

For Rasch analyses, reasonably well-targeted samples of 150 are reported to have 99% confidence that the estimated item difficulty is within ± ½ logit, and n=243 for poorly targeted samples (Linacre, 1994). Our sample of 319 was therefore optimal for the purpose of this analysis.

Table 2. Brief overview of Rasch analysis concepts (adapted from Czuba *et al.*, 2016).

Concept	Test used	Expected results[24–26,39–41]	Strategies to deal with misfit
Item threshold ordering[A]	Examination of the threshold location and their 95% confidence intervals to determine significance of disordering if observed visually.	Logical progression across the trait being measured	Disordered category responses might have to be collapsed into one
Person fit	Mean fit residuals (SD); range	Mean close to zero and SD close to 1; range -2.5 to 2.5 χ^2 non-significant with a Bonferroni correction	Person(s) might have to be deleted from the dataset[B]
Item fit	Mean fit residuals (SD); range	Mean close to zero and SD close to 1; range -2.5 to 2.5 χ^2 non-significant with a Bonferroni correction	Item might have to be deleted from the subscale
Local dependency[C]	Residual item correlation matrix between all items	Correlations between the residuals >0.20 above the average residual correlation	Locally dependent items to be combined into testlets
Unidimensionality	Principal component analysis (PCA) of the residuals[D]	The 95% CI of the proportion of significant tests should include 5%	
Reliability index	Person Separation Index	Values of ≥0.70 good for group comparisons (e.g. in research trials); ≥0.85 for individual clinical use.	Not applicable
Overall fit to the Rasch model	Item-trait interaction χ^2	Non-significant with a Bonferroni correction	Not applicable
Targeting of the scale[E]	Logit value; visual inspection of person-item distribution map	Logit value above that of the highest item on the subscale	Not applicable
Differential item functioning (DIF) by person factor (e.g. gender)[F]	ANOVA	Non-significant with a Bonferroni correction	If DIF is uniform, items to be combined into testlets[G] or split by person factor. If DIF is non-uniform items to be deleted.

Key:

[A]: Thresholds represent points where the probability of scoring either of the two adjacent categories is 50%. If it is not the case, one would observe disordered thresholds where the individual score cannot be reliably interpreted.

[B]: Extreme scores (much lower than -2.5, or much higher than 2.5) indicate issues with response pattern which may include: responding according to a socially desired norm, carelessness with responding or low motivation in responding. As such data would not add any meaningful information to the calibration process, it has been suggested to consider excluding extreme persons from the sample (Bond & Fox, 2001; Tennant & Conaghan, 2007).

[C]: Local dependency occurs when a person's response to one item is reflected in their response to another item

[D]: Two subsets of items are identified by PCA: one with positively loading items and one with negatively loading items. Two estimates derived from these subtests are then tested by using an independent t-test. If the result is insignificant at p≤0.05, the unidimensionality is supported.

[E]: Targeting of the scale to the latent trait allows identification of floor and ceiling effects.

[F]: DIF occurs when people from different groups (for example, males and females) with equal amounts of the underlying trait do not respond to items in a similar manner.

[G] : A testlet is a bundle of items that share a common stimulus.

Results

This section reports the analysis of results for each Te Waka Kuaka subscale separately. There were no missing data in the dataset. Please see Supplementary File 2 for the complete final version of Te Waka Kuaka, and Supplementary File 1 for the draft version, from which items were deleted.

Wā (time)

The proposed subscale had 9 initial items all concerned with the broad concept of time. These items were not specifically linked to issues such as time to access treatment or time since injury. Rather, time in this subscale is concerned with what needs to happen first in time, the role of time in facilitating healing, taking time for a range of purposes and flexibility of time schedules.

The initial analysis of the Wā subscale showed that there were no items with statistically significant disordered thresholds, and the scale was unidimensional. However, the scale did not fit the Rasch model because of a significant (p=0.0005) item-trait interaction chi-square, and a particularly high mean persons location (2.8; SD=1.5). 23% (n=60) of the sample had extreme scores, and so were deleted from the analysis, the remainder of n=259 provided a robust sample to analyse. Deletion of the subgroup improved the mean persons location (2.3; SD=1.2), but did not result in an improvement in item-trait interaction.

Further examination of the items revealed three (items 3, 5 and 9) that were misfitting the model. Item 3, "whakawhanaungatanga (the process of making connections with others) at the beginning sets the scene for the journey" functioned differently according to iwi (tribe), with the "other" group being an outlier. Also, the item did not fit the Rasch model, with item fit residual of -2.825, and chi-square probability of 0.006. Importantly, the item seemed to identify issues already captured by items 1 (Starting the process of wairua healing is the first thing that needs to happen for our whānau), 2 (The journey of wairua healing is enhanced with time), and 8 (whakawhanaungatanga time builds, to keep hope and dreams alive). Hence, it was deleted from this subscale.

Item 5, "It is important that kaimahi (health workers) are flexible in their schedules of work", had a high fit residual (2.722; p=0.0006) indicating the item does not fit the scale. It was also deleted from the subscale.

Lastly, Item 9, "Whānau unity and strength builds healing" showed local dependency problems with item 3 "whakawhanaungatanga at the beginning sets the scene for healing". It also displayed non-uniform differential item functioning (DIF) for relationship (see Table 1). A number of possible solutions described in Table 3 were tested, however, only deletion of the item led to solving the local dependence with item 3.

These modifications improved the fit of the Wā subscale and provided the final solution (see Table 3). The resulting 6-item scale was unidimensional and the item-trait interaction was non-significant (p= 0.1237). The reliability of the subscale is relatively low (PSI=0.56). The targeting of the subscale Wā was skewed, suggesting people on average scored towards the upper end of the scale (Figure 1)

Wāhi (place)

The proposed Wāhi subscale included 10 items concerned with aspects to do with places, such as those of cultural significance as well as clinics and hospitals.

Upon initial examination of the Wāhi subscale, it was found that the item-trait interaction chi-square was highly significant (p<0.00001) and the scale was not unidimensional. None of the items showed disordered response category thresholds. Further analysis of DIF and fit statistics revealed four items that required specific attention: items 10, 11, 16 and 17.

Table 3. Summary statistics of Rasch analysis. PSI: Person Separation Index; Alpha: Cronbach's alpha. 'First' refers to the analysis of results of the raw ordinal data; 'Final' refers to the analysis of results of the Rasch-transformed data.

Analysis*		Item Fit residual[1]	Person Fit residual[1]		Chi Square interaction			PSI			Tests of Unidimensionality 95% CI [%]		
		Mean	SD	Mean	SD	Value	DF	p	With extremes	No extremes	alpha	Lower bound	Higher bound
Wā	First	-0.703	1.755	-0.676	1.836	44.6	18	0.0005	0.689	0.681	0.858	2.0	7.3
	Final	-0.335	1.042	-0.645	1.746	17.7	12	0.1237	0.560	0.530	0.721	-0.3	5.0
Wahi	First	-0.217	2.953	-0.352	1.154	173.0	40	0	0.772	0.743	0.853	1.4	6.2
	Final	-0.216	1.2326	-0.421	1.000	27.1	24	0.2976	0.777	0.739	0.851	3.3	8.0
Tangata	First	-0.076	1.556	-0.587	1.711	107.0	60	0.0002	0.790	0.759	0.841	3.3	8.0
	Final	-0.062	1.289	-0.480	1.361	52.1	27	0.0026	0.733	0.718	0.862	-0.5	4.3
Wairua	First	-0.284	1.598	-0.565	1.560	102.5	48	0.00001	0.784	0.787	0.898	4.8	9.6
	Final	-0.159	1.249	-0.482	1.371	54.4	27	0.0013	0.733	0.718	0.862	-0.2	4.8

Key:

[1]: Ideally, mean fit residual statistics should be close to a mean of zero with a standard deviation of one.

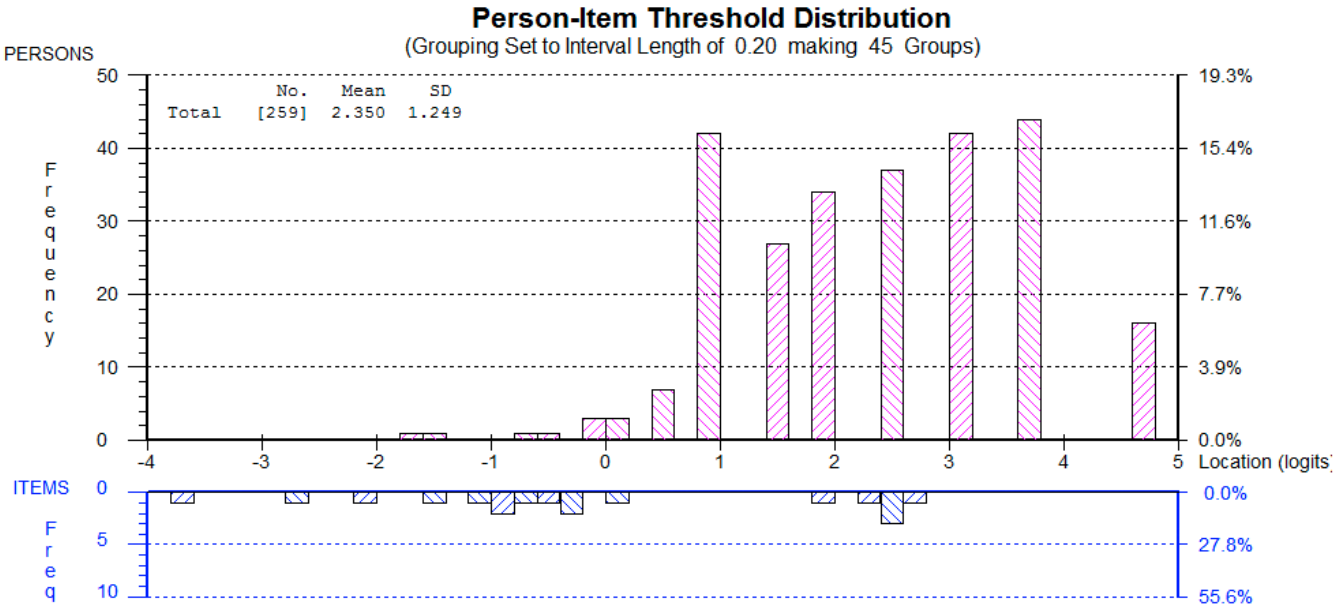

Figure 1. Person-Item Threshold Distribution for the Wā (time) subscale.

Items 10, "The use of pepeha within treatment would support the healing", and item 17, "Whānau from home are an essential link with home", had uniform DIF by TBI severity. These items were combined into a testlet with item 13, "Whakaairo (carvings) teach important lessons that help with healing", which had showed non-significant DIF in an opposite direction. This resulted in these opposing directional DIF cancelling each other out.

Item 11, "Being inside buildings like hospitals does not help me", had a very high fit residual of 7.785 ($p<0.00001$), demonstrating it did not fit the subscale. This item was therefore removed.

Item 16, "Gathering, preparing and eating food from home is an important part of healing", showed uniform DIF by location and TBI. This item was combined into a testlet with item 19 as this item visually showed to have DIF in the opposite direction (non-significant), "being on the marae is a good place to start to feel strong again", and therefore these DIF in opposite directions cancelled each other out.

These modifications improved the fit of the subscales to the Rasch model and provided the final solution (Table 3). The final subscale had 9 items and was unidimensional, the item-trait interaction was non-significant, and no DIF was observed. The reliability of subscale Wāhi was good (PSI=0.78) and the targeting acceptable (Figure 2).

Tangata (people)

This subscale is concerned with people involved with the person with TBI and their whānau and had a total of 15 statements.

The initial analysis of the Tangata subscale showed that the scale was unidimensional and none of the items had statistically significant disordering of response categories thresholds. However, the scale did not fit the Rasch model with statistically significant item-trait interaction chi-square ($p=0.0002$).

Further examination revealed three pairs of items with high residual correlations. Item 22, "Within whānau there are a lot of resources", was locally dependent (residual correlation = 0.25) on item 23, "within the whānau is the rongoā" (rongoā is the Māori word for medicine). From a theoretical point of view, these two items consider two very similar concepts. However, item 23 is focused more specifically on the healing process, whereas item 22 (Within whānau there are a lot of resources) is much less specific as to what sort of resources might be available, when and for what purpose. Furthermore, item 23 showed a better spread on the latent trait of interest (3.8 versus 2.8 logits). Therefore, item 22 was deleted from this subscale.

Item 26, "Māori have a different point of view from Pākehā (non-Māori of European ancestry)", was locally dependent upon item 27 (0.405) "Māori cultural needs are different from Pākehā". Theoretically, cultural needs secondary to the culturally determined injury to wairua are critical to the functioning of this tool, in order to best understand how whānau conceptualise these needs. While asking about similar issues, item 27 more specifically asks about cultural needs, whereas item 26 was more vague, referring only to a different point of view. Hence, the decision was made to delete item 26.

Item 28, "When health workers relate to the culture of the whānau outcomes are improved", was locally dependent (residual correlation = 0.444) on item 29, "When health workers support whānau to address wairua outcomes are improved". Item 29, was deemed to be theoretically more important, because it more directly

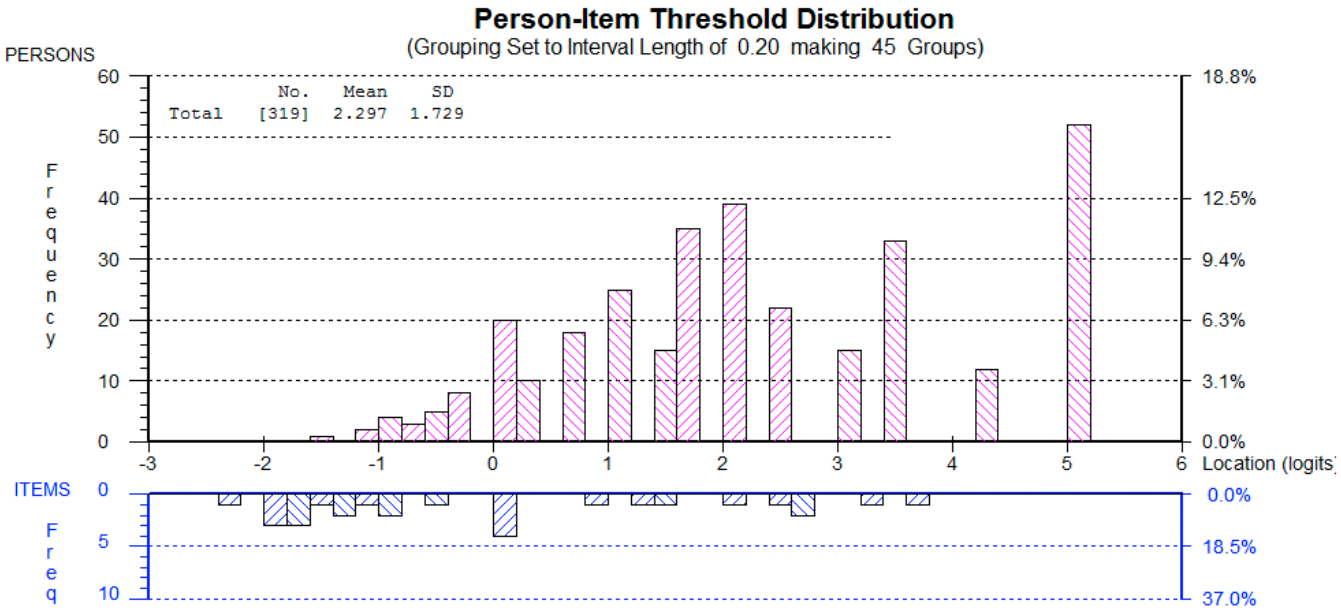

Figure 2. Person-Item Threshold Distribution for the Wāhi (place) subscale.

measures the issue of wairua, which is central to the theory of the cultural aspect of injury. Therefore, item 28 was deleted from the subscale.

Deletion of these three items improved fit of data to the model and provided the final solution (Table 3). The item-trait interaction chi-square was non-significant, the scale was unidimensional and no DIF was observed. The reliability of the subscale Tangata was good (PSI=0.740) and the targeting was acceptable (Figure 3).

Wairua practices

Wairua practices is a phrase used to describe activities that strengthen wairua. Wairua is an area of hauora (health and wellbeing) that conveys the unique connection between Māori and all aspects of the universe. While wairua is mentioned in other subscales, wairua is the primary focus of this subscale. This subscale consisted of 12 items.

The initial analysis of the Wairua subscale found that the scale was unidimensional, but it did not fit the Rasch model ($p<0.0001$). Moreover, there was one misfitting item, one item showed non-uniform DIF, two items were locally dependent and a number of items had statistically significantly disordered response category thresholds.

Item 35, "Practices that strengthen wairua are as important as clinical interventions", was found to be misfitting with chi-square p=0.00014. The item was deleted and fit to the model improved.

Examination of item 46, "Use of Te Reo Māori means wairua is being strengthened", identified non-uniform DIF by location and statistically significant disordering of response category thresholds. The decision was made to delete this item and this improved fit to the model.

Items 43, "Romiromi (type of massage) can be a powerful healing tool", and 42, "Mirimiri (type of massage) can be a powerful healing tool", were found to be locally dependent (residual correlation = 0.638). Because these types of massage are very similar and mirimiri (massage) is more commonly known, item 42 was retained and item 43 was deleted.

Five items, 36, 38, 39, 44 and 45 showed statistically significant disordered thresholds. The lower two response categories ("strongly disagree" and "disagree") of these items were collapsed into one category. This modification further improved fit of data to the model and provided the final solution for the Wairua subscale. The scale fit the model with non-significant item-trait interaction and was unidimensional. The reliability of the scale was good (PSI=0.733) and the targeting was acceptable (Figure 4). Scoring was modified accordingly.

Item difficulty

Table 4 presents the relative difficulty of each item of the Te Waka Kuaka subscales. The easier the item, the higher the expected scores are for people with high levels of investigated construct.

Dataset 1. Data file containing responses from all participants that completed Te Waka Kuaka

http://dx.doi.org/10.5256/f1000research.11500.d166175

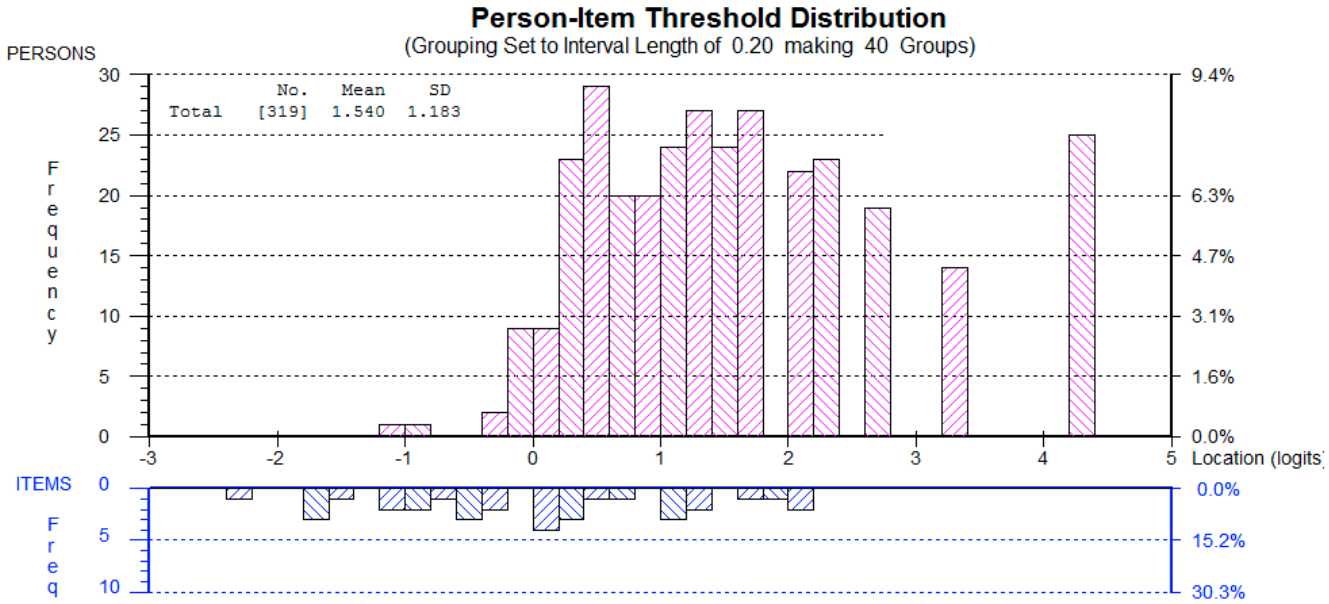

Figure 3. **Person-Item Threshold Distribution for the Tangata (people) subscale.**

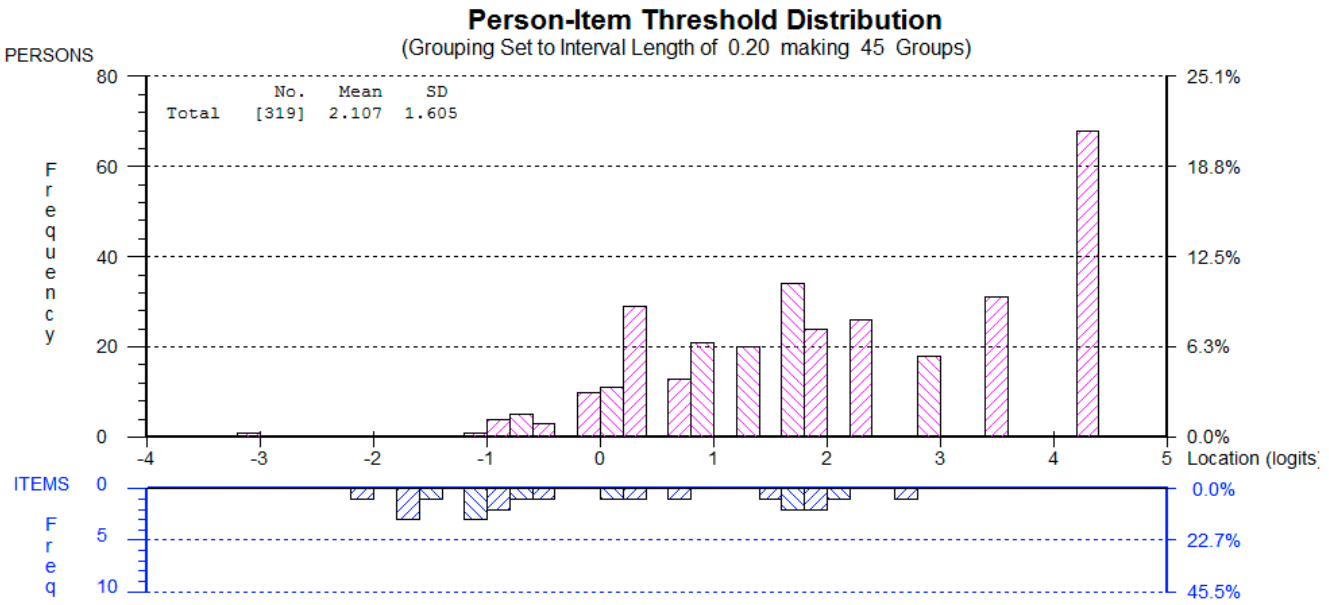

Figure 4. **Person-Item Threshold Distribution for the Wairua practices (activities that strengthen spiritual connection) subscale.**

Table 4. Rasch estimates of true item difficulty.

Item Difficulty	Wā		Wāhi		Tangata		Wairua	
Less	Item	Location (logits)	Item	Location (logits)	Item	Location (logits)	Item	Location (logits)
↕	Q8	-0.753	Q18	-0.583	Q34	-0.56	Q38	-0.882
	Q2	-0.165	Q15	-0.394	Q27	-0.28	Q36	-0.468
	Q7	0.03	*Q16&19	-0.312	Q30	-0.271	Q40	-0.292
	Q6	0.137	*Q10&13&17	0.265	Q31	-0.15	Q42	-0.246
	Q4	0.277	Q12	0.4	Q33	-0.005	Q39	0.469
	Q1	0.474	Q14	0.623	Q29	0.015	Q44	0.549
					Q23	0.06	Q41	0.559
					Q21	0.067	Q45	0.581
					Q32	0.233		
					Q24	0.35		
More					Q25	0.541		

Key:

*: Indicates a testlet made of two or more original NFI items. Testlet is scored by summing up the scores from included items.

Discussion

This study presents the Rasch analysis of a new measure, Te Waka Kuaka, for use in assessment of Māori cultural needs following traumatic brain injury. Given the over representation of Māori with TBI (Feigin *et al.*, 2012) alongside Māori beliefs about the sacred quality of the head, 'he tapu te upoko', (Moko-Mead, 2003) this scale is much needed. This investigation was done to examine the validity of Te Waka Kuaka. Our analysis identified ten items that did not fit the Rasch model and were deleted. The resulting four subscales fit the Rasch model and were unidimensional.

Very few measures developed to assess Māori specific aspects of health exist. One that has been used in the area of mental health and addictions is called "Hua Oranga" (Durie & Kingi, 1997). The Hua Oranga operates a well-known framework called "Te Whare Tapa Whā" (the four walled house). This framework does not have an underpinning theory. It presents four constructs, whānau (extended family), wairua (spirituality), hinengaro (mind) and tinana (body). While some analyses of the psychometric properties of this measure have been made, we are not aware of any previous measure being developed using Rasch analysis (Harwood *et al.*, 2012; McClintock *et al.*, 2011). Overall, the Hua Oranga measure was developed in a different manner to Te Waka Kuaka and measures a construct of hauora, without theoretical basis, rather than four subscales based on a theory of brain injury.

From a clinical perspective, responses to a number of the items were interesting. Item 3 highlighted that there were a range of groups for whom the item functioned differently, by iwi (tribal group) and with the "other" group being an outlier. One interpretation of this is that the small non-Māori "other" group had a different understanding of whakawhaunaungatanga. This is not unexpected, given that the concept is Māori-specific. Also, it is possible that differing iwi (tribal) groups conceptualise this activity in different ways. This finding added to a richer understanding of whanaungatanga itself. Item 11, "being inside buildings like hospitals does not help me", was a statement that came from the preliminary research. While this statement may have assisted in considerations about the location of rehabilitation processes, the item did not have explicit theoretical salience regarding the wairua aspects of the injury. These were considered better assessed by item 19 "being on marae is a good place to start to feel strong again". The negative frame of the statement ("does not help me") was thought to contribute to a different perception of the item by participants, compared to the positively framed items.

The lower PSI (0.56) of the Wā subscale indicates that most of the participants scored those items highly. The heterogeneity of the participants resulted from the wide range of iwi (tribal) affiliations represented, arising from different parts of Aoeatoa, New Zealand. This also meant a range of competencies in Te Reo Māori were represented. Despite this heterogeneity, the importance of the concept of time was evident from the likelihood that these items would be highly endorsed. Recognition of responses to items in this subscale, especially those most strongly endorsed, is of clinical importance and can directly inform priorities in subsequent management strategies.

The spread of difficulty of Te Waka Kuaka items was relatively narrow: between -1 and 1 (see Table 4). Including deleted items did not affect the spread. Similarly, it is possible that because the method of deriving the items was culturally conservative, that is,

developed on marae (traditional meeting houses), albeit urban, rural and remote, the items do not address Māori cultural needs that are either very easy and or very difficult to endorse. Given the positive skew in this sample, further testing could be undertaken with people who are less in touch with their Māori cultural identity. We hypothesise that the sample would score more towards the lower end of the Te Waka Kuaka subscales.

One of the limitations of the study was that the wider sample of possible participants is unknown, so no response rate can be calculated. However, given the large sample size the analysis itself remains robust.

Dissemination of the findings of the analysis to research partners, namely health and education providers in the Māori community, has led to widespread requests for use of Te Waka Kuaka in settings outside of TBI rehabilitation. This is an unexpected development. One approach being considered is to develop a further study protocol to collect this data. Analysis would then enable better understanding of the scope of the tool's application.

Clinical implications of the use of the tool are significant. By being able to clearly and quickly identify the immediate needs of the whānau means that the whānau themselves and the health workers can focus on addressing those needs without delay. How these needs change can be easily reviewed and this can in turn guide further tailoring of supports. Given the theoretical importance of addressing the cultural aspect of TBI, namely the injury to wairua, it is vital to ensure these cultural needs are thoroughly monitored and responded to. In this way, healing the cultural injury is likely to improve the recovery process, as well as outcomes for the whānau.

Conclusions

Te Waka Kuaka is a new measure that has been in development to assess the cultural needs of Māori with TBI. This paper reports the Rasch analysis phase. Our findings show that the revised subscales are unidimensional and fit the Rasch model. Te Waka Kuaka can now enable valid and accurate measurement of Māori cultural needs following TBI. Future research examining the responsiveness of Te Waka Kuaka would be a useful addition to better understanding the applicability of this measure.

Data availability

Dataset 1. Data file containing responses from all participants that completed Te Waka Kuaka. DOI, 10.5256/f1000research.11500.d166175 **(Elder *et al.*, 2017).**

Author contributions

HE devised and carried out research, analysis and writing of paper. KC assisted with analysis and writing up of paper. PK assisted in design, analysis support and writing of paper. AC assisted in analysis, support and writing up of paper. KM assisted in writing of paper.

Competing interests

No competing interests were disclosed.

Grant information

This research was funded by the Health Research Council of NZ [14-060].

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgments

The authors thank Te Hiku Hauora, Whakawhiti Ora Pai, Te Kura Reo ki Ōtaki, Te Kura Reo ki Kahungunu, Te Pīnakitangi ki te Reo Kairangi, Kearoa Mokaraka, Raumāhoe Ani Morris, Whaea Moe Milne, Health Research Council of NZ.

Supplementary material

Supplementary File 1: Te Waka Kuaka: original 46-item draft measure.

Click here to access the data.

Supplementary File 2: Te Waka Kuaka: final 36-item bilingual version.

Click here to access the data.

References

28th Māori Battalion: **C Company haka.** http://www.28maoribattalion.org.nz/. 2011.
Reference Source

Andrich D: **Rasch Models for Measurement Series: Quantitative Applications in the Social Sciences No. 68**. London, England: Sage Publications. 1988.
Reference Source

Andrich D, Sheridan B, Luo G: **RUMM2030 (Comupter Software and Manual).** Perth, Australia: RUMM Laboratory 2030. 2010.
Reference Source

Bond TG, Fox CM: **Applying the Rasch Model. Fundamental Measurement in the Human Sciences.** London, England: Lawrence Erlbaum Associates. 2001.
Reference Source

Czuba KJ, Kersten P, Kayes NM, *et al.*: **Measuring Neurobehavioral Functioning in People With Traumatic Brain Injury: Rasch Analysis of Neurobehavioral Functioning Inventory.** *J Head Trauma Rehabil.* 2016; **31**(4): E59–68.
PubMed Abstract | Publisher Full Text | Free Full Text

Dudley M: **[Developing a Māori theory of dementia as the basis of culturally robust neuropsychological testing for ageing Māori].** 2016.

Durie M: **Mauri ora, dynamics of Māori health.** Auckland: Oxford University Press. 2001.
Reference Source

Durie M: **Understanding health and illness: research at the interface between science and indigenous knowledge.** *Int J Epidemiol.* 2004; **33**: 1138–1143.
PubMed Abstract | Publisher Full Text

Durie M, Kingi TK: **A framework for measuring Māori mental health outcomes.** Retrieved from Wellington; 1997.
Reference Source

Elder H: **An examination of Maori tamariki (child) and taiohi (adolescent) traumatic brain injury within a global cultural context.** *Australas Psychiatry.* 2012; **20**(1): 20–23.
PubMed Abstract | Publisher Full Text

Elder H: **Indigenous theory building for Māori children and adolescents with traumatic brain injury and their extended family.** *Brain Impairment.* 2013a; **14**(3): 406–414.
Publisher Full Text

Elder H: **Te Waka Oranga. An indigenous intervention for working with Māori children and adolescents with traumatic brain injury.** *Brain Impairment.* 2013b; **14**(3): 415–424.
Publisher Full Text

Elder H, Kersten P: **Whakawhiti Kōrero, a Method for the Development of a Cultural Assessment Tool, Te Waka Kuaka, in Māori Traumatic Brain Injury.** *Behav Neurol.* 2015; **2015**: 8, 137402.
PubMed Abstract | Publisher Full Text | Free Full Text

Elder H, Czuba K, Kersten P, *et al.*: **Dataset 1 in: Te Waka Kuaka, Rasch analysis of a cultural assessment tool in traumatic brain injury in Māori.** *F1000Research.* 2017.
Data Source

Elder H, Kersten P, McPherson K, *et al.*: **Making time: deeper connection, fuller stories, best practice.** *Annals of Psychiatry and Mental Health.* 2016; **4**(6): 1079–1082.
Reference Source

Feigin VL, Theadom A, Barker-Collo SL, *et al.*: **Incidence of traumatic brain injury in New Zealand: a population-based study.** *Lancet Neurol.* 2013; **12**(1): 53–64.
PubMed Abstract | Publisher Full Text

Friedly J, Akuthota V, Amtmann D, *et al.*: **Why disability and rehabilitation specialists should lead the way in patient-reported outcomes.** *Arch Phys Med Rehabil.* 2014; **95**(8): 1419–1422.
PubMed Abstract | Publisher Full Text

Haigh R, Tennant A, Biering-Sorensen F, *et al.*: **The use of outcome measures in physical medicine and rehabilitation within Europe.** *J Rehabil Med.* 2001; **33**(6): 273–278.
PubMed Abstract | Publisher Full Text

Harwood M: **Rehabilitation and indigenous peoples: the Māori experience.** *Disabil Rehabil.* 2010; **32**(12): 972–977.
PubMed Abstract | Publisher Full Text

Harwood M, Weatherall M, Talemaitoga A, *et al.*: **An assessment of the Hua Oranga outcome instrument and comparison to other outcome measures in an intervention study with Maori and Pacific people following stroke.** *N Z Med J.* 2012; **125**(1364): 57–67.
PubMed Abstract

Hays RD, Morales LS, Reise SP: **Item response theory and health outcomes measurement in the 21st century.** *Med Care.* 2000; **38**(9 Suppl): II28–42.
PubMed Abstract | Publisher Full Text | Free Full Text

Kersten P, Kayes NM: **Outcome measurement and the use of Rasch analysis, a statistics-free introduction.** *New Zealand Journal of Physiotherapy.* 2011; **39**(3): 92–99.
Reference Source

Krosnick JA, Holbrook AL, Berent MK, *et al.*: **The impact of "no opinion" response options on data quality: Non-attitude reduction or an invitation to satisfice?** *Public Opin Quart.* 2002; **66**(3): 371–403.
Publisher Full Text

Linacre JM: **Sample size and item calibration [or Person Measure] stability.** *Rasch Measurement Transactions.* 1994; **7**(4): 328.
Reference Source

McClintock K, Mellsop G, Kingi T: **Development of a culturally attuned psychiatric outcome measure for an indigenous population.** *International journal of Culture and Mental Health.* 2011; **4**(2): 128–143.
Publisher Full Text

Moko-Mead H: **Tikanga Māori, living by Māori values.** Wellington: Huia. 2003.
Reference Source

Norup A, Perrin PB, Cuberos-Urbano G, *et al.*: **Family needs after brain injury: A cross cultural study.** *NeuroRehabilitation.* 2015; **36**(2): 203–214.
PubMed Abstract | Publisher Full Text

Ogden JA, McFarlane-Nathan G: **Cultural bias in the neuropsychological assessment of young Māori men.** *New Zealand Journal of Psychology.* 1997; **26**: 2–12.
Reference Source

Puente AE, Perez-Garcia M, Vilar-Lopez R, *et al.*: **Neuropsychological Assessment of Culturally and Educationally Dissimilar Individuals.** In F. En Paniagua & AM. Yamada (Eds.), *Handbook of Multicultural Mental Health: Assessment and Treatment of Diverse Populations.* (Second ed). San Diego CA, USA Academic Press, 2013; 225–241.
Publisher Full Text

Reeve BB, Wyrwich KW, Wu AW, *et al.*: **ISOQOL recommends minimum standards for patient-reported outcome measures used in patient-centered outcomes and comparative effectiveness research.** *Qual Life Res.* 2013; **22**(8): 1889–1905.
PubMed Abstract | Publisher Full Text

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): 1358–1362.
PubMed Abstract | Publisher Full Text

Whiteneck GG, Dijkers MP, Heinemann AW, *et al.*: **Development of the participation assessment with recombined tools-objective for use after traumatic brain injury.** *Arch Phys Med Rehabil.* 2011; **92**(4): 542–551.
PubMed Abstract | Publisher Full Text

Open Peer Review

Current Peer Review Status: ✗ ?

Version 1

Reviewer Report 06 October 2017

https://doi.org/10.5256/f1000research.12420.r24646

?

Sara Simblett
King's College London, London, UK

This paper reports a psychometric validation of a new measure to assess cultural needs of Māori with traumatic brain injury. It is an important piece of research that has embedded engagement with local Māori communities and their specific social and cultural context at the heart of designing the new measure. It draws on qualitative methods to generate items that are then rigorously tested using Rach analysis techniques, presenting a very thoughtful and through piece of research, which has clear practical application. This type of translational research could make a significant contribution to closing gaps in service provision and shaping new developments for clinical services that meet the needs of their service users including minority groups. There are a number of other specific strengths to this work including a very clear introduction of the Rasch analysis techniques that are described well for an audience less familiar with these approaches; and the inclusion of the raw data file as well as copies of the questionnaire items in an open access format for readers to view. Additionally, the comprehensive dataset without any missing data, which must have taken a considerable effort to collect.

Below are a few aspects that could be addressed as means of improving the paper further:

1. The characteristics of the sample could be expanded upon. More specifically, it is stated that participants self-reported whether or not they had experienced a traumatic brain injury (TBI). It is not clear about the severity of these injuries and, for some, the TBI was only 'possible'. I think more clarity on how this was assessed would be useful and I wonder if 'head injury' might be more accurate than 'TBI', especially if there is no medical evidence to support a head injury leading to a TBI in all cases. If this further data could be gathered and TBI diagnoses established this would considerably enhance the validity of the results reported in this paper in relation to TBI. Variables such as time post injury may moderate responses to the measure and so would also be useful contextual information.

2. The sample size would allow for a secondary analysis on only participants with a confirmed diagnosis of TBI. Alternatively, a clinical sample of people using TBI health services could be recruited to test the reliability of the results.

3. Although the supplementary file provides some further details about the qualitative component of the initial development of the items, some more information would be useful within the paper itself. For example, the content of the initial interviews and the characteristics of the people including in the focus groups.

4. Similarly, further details about the methods of recruiting participants to answer the questionnaire would help the reader to understand the context relating to the sample selected to validate the measure.

5. Details about the four subscales are described in the results section but I would have liked to have read this information earlier in the methods section, when the questionnaire is introduced.
6. I wonder if the reader may benefit from a glossary so that they can more easily understand words that they are unfamiliar with. For example, in Table 1.

7. I could not find reference to a Rasch analysis of the questionnaire as a whole. It might be worth adding that the psychometric properties of the subscales were explored individual because responses to the questionnaire as a whole did not meet the requirements of the Rasch model.

8. In the analysis of the 'time' subscale a relatively large subgroup were removed from the analysis. I would be interested in understanding more about the characteristics of this subgroup.

9. The low internal consistency (PSI) for the 'time' subscale may indicate that it is measuring a number of underlying constructs. Did you perform a Rasch factor analysis to explore whether it is best divided into two dimensions? In figure 1 there seems to be two clusters of items. Only one of which is well targeted to the sample.

10. As a general point, the questionnaire as it stands is very long and the exploration of psychometric properties of the items individually provides an opportunity to select the 'best' items to contribute to the measure of several underlying constructs. From looking at Figures 1 and 3, for example, there seems to be a number of items where they is a consensus of endorsement – are these items best removed because they do not contribute to separating out any of the abilities of the persons in the sample?

11. In the discussion the clinical implications are rightly addressed. This measure could be very useful for establishing people's cultural preferences and values. However, I question whether there is likely to be change on the constructs measured as values are more likely to be disposition and not so strongly related to state, unless moderated by factors such as outlook on life more generally. I believe this is, however, scope for future research.

Thank you for the opportunity to review this paper, which describes a mixed methods approach that have enormous potential for creating valid, reliably and appropriate measure for using in rehabilitation after head injury/TBI.

Is the work clearly and accurately presented and does it cite the current literature?

Yes

Is the study design appropriate and is the work technically sound?
Partly

Are sufficient details of methods and analysis provided to allow replication by others?
Partly

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Partly

Competing Interests: No competing interests were disclosed.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Reviewer Report 05 July 2017

https://doi.org/10.5256/f1000research.12420.r23969

Stephen McKenna
Manchester University, Manchester, UK

I like the approach of generating statements from local people, rather than adapting measures developed in other cultures and assuming they are meaningful to the indigenous population. Also, the authors are open about the weaknesses in the scales identified by the Rasch analyses.

However, the article is very unclear about the nature of the measure they are testing or what its purpose is. Reading the literature about the instrument development the same lack of clarity is found. Is it intended for children and adolescents with TBI or for all people with TBI? The items represent health beliefs so this is not an outcomes measure. If it is measuring health beliefs do these differ dependent on the condition? As the items were not generated from people who had experienced TBI, why should the measure be relevant to a TBI population? This lack of relevance appears to explain why the four scales are poorly targeted to the study population. Clinicians and family members are notoriously poorly informed on the impact of disease on others.

As some of the study population did have a TBI it would be important to test for DIF by the

presence of the condition.

To produce an instrument specific to TBI, fundamental work would need to be done with a representative sample of people who had experienced TBI.

As well as being poorly targeted, the reliability of the scales is average to poor. The poor targeting suggests that the instrument would not be responsive to changes in health beliefs, if this was an intention of the instrument.

The article covers Rasch analysis of the data. While fit to the Rasch model is fundamental to the validity of an instrument, additional analyses are required to show that the measure works as expected. For example, data should be presented showing that scores are related to perceived severity of TBI.

Much of the information in the Methods section covers work previously conducted, rather than that conducted in the present study. This should be included in the Introduction if it is deemed relevant.

Is the work clearly and accurately presented and does it cite the current literature?
No

Is the study design appropriate and is the work technically sound?
Partly

Are sufficient details of methods and analysis provided to allow replication by others?
No

If applicable, is the statistical analysis and its interpretation appropriate?
Partly

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
No

Competing Interests: No competing interests were disclosed.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to state that I do not consider it to be of an acceptable scientific standard, for reasons outlined above.