Journal of Clinical Epidemiology 151 (2022) 18−28

ORIGINAL ARTICLE

# Rasch Measurement Theory's contribution to the psychometric properties of a co-created measure of health and wellness for Indigenous children and youth

Skye P. Barbic[a,b,c,*], Nancy L. Young[d], Koyo Usuba[d], Elizabeth Stankiewicz[d]

[a]*Department of Occupational Science and Occupational Therapy, Faculty of Medicine, University of British Columbia (UBC), Vancouver, British Columbia, Canada V6T 2A1*

[b]*Centre for Health Evaluation and Outcome Sciences, Vancouver, British Columbia, Canada V6Z 1Y6*

[c]*Foundry, Vancouver, British Columbia, Canada V6Z 2A9*

[d]*Evaluating Children's Health Outcomes (ECHO) Research Centre, Laurentian Univeristy, Sudbury, ON P3E 2C6, USA*

Accepted 18 July 2022; Published online 2 August 2022

## Abstract

**Objectives:** To determine how Rasch Measurement Theory (RMT) methods can be used to assess the psychometric properties of the Aaniish Naa Gegii: the Children's Health and Wellbeing Measure (ACHWM) and Qanuippit.

**Study Design and Setting:** Indigenous children aged 8−18 years, from five communities, completed the 62-item ACHWM. We applied RMT methods to ACHWM data from 401 children (mean age 13.4 ± 3.4 years; 51% male) from across Ontario to examine how well the items captured the full range (±3 logit) of the concept of interest in each domain, targeted the needs of Indigenous children, and met the criteria for unidimensional and invariant measurement.

**Results:** RMT results indicated moderate-fit overall fit (raw $\chi^2 = 809$, $P < 0.001$). This model was further improved by aggregating the five response categories into three categories. All four domains showed excellent overall fit to the Rasch model ($P > 0.05$), with items covering between −4.51 and 6.02 logit, with no gaps along the theoretical continua.

**Conclusion:** This study provides evidence that a set of conceptually derived items was able to produce a measure that fits the Rasch model. These results aid our understanding of wellness by establishing the clinical meaning of the scale scores. 

*Keywords:* Indigenous; Children; Youth; Rasch; Measurement; Validity

## 1. Introduction

Patient-reported outcomes (PROs) are important for decision-making in health services [1,2]. PRO measures (PROMs) have been traditionally used in health research and clinical trial designs [3,4]. PROMs play important roles in guiding healthcare delivery and clinical practice decisions [5]. The process of PROM development and evaluation has changed significantly in the last decade, to ensure PROMs reflect patients' perspectives, are clinically useful, inform evidence-based practice, and empower patients in the evaluation process [3,6]. Patient engagement at all stages of the research is a vital part of current national research initiatives such as Strategy for Patient Oriented Research [7] and Patient- Centered Outcome Research Institute [8].

The health field increasingly has examples of conceptually driven patient-centered measures in areas such as rheumatology [9,10], cardiology [11] oncology [12], and plastic surgery [13−15]. Patient-engagement guidelines for PROM

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. However, the original datasets were gathered with the support of several grants.

Declaration of interests: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this article.

* Corresponding author. Faculty of Medicine, Department of Occupational Science and Occupational Therapy, The University of British Columbia | St. Paul's Hospital, 2255 Wesbrook Mall, Vancouver, British Columbia, Canada V6T 2A1. Tel.: +1 604 822 0770; fax: +1 604 822 7624.

*E-mail address:* Skye.barbic@ubc.ca (S.P. Barbic).

https://doi.org/10.1016/j.jclinepi.2022.07.010

**What is new?**

**Key findings**

- A self-reported health assessment approach that was conceptually derived with and for Indigenous children and youth fits the Rasch model.

**What this adds to what was known?**

- The overall and domain ordering of items is now available to inform discussions of health services for children along the journey from illness to wellness.

**What is the implication and what should change now?**

- This new information will support community and culturally based services to meet children where they are at (rather than limiting services to respond only to illness) and enhance our strength-based approaches to support health equity.

development now exist [16,17] and are widely used across different populations. One population which has been quietly absent from the evolution of evidence-based practice and health measurement is Indigenous people in Canada. In an age of truth and reconciliation (www.trc.ca), it is critical that we address the health inequities experienced by this population.

Indigenous people in Canada include the First Nations, Inuit, and Métis. The Indigenous population in Canada was 1,673,785 as per the national census (2016), accounting for 4.9% of the total population. Indigenous children are the fastest-growing segment of the paediatric population in Canada [18] and experience serious health inequities [18,19]. For example, the rate of youth suicide is five to six times the national average [18,20] and suicide/self-inflicted injury is the leading cause of death among Indigenous youth [21]. These statistics reflect the impact of historical trauma and underscore the need for new research to generate evidence and guide programs to address these inequities. A continual focus on crisis intervention without time for research and/or reflection has permitted health challenges to continue. Indigenous health leaders are eager to improve children's health. However, a mismatch exists between the needs of youth and the scope and capacity of local services [22]. One key factor identified by Indigenous health leaders was the lack of tools to give Indigenous children and youth a voice to inform their needs and reflect cultural-specific ways of knowing. The focus of this article is on a metric that is culturally relevant and solution-focused to support the reclamation of both culture and health within Indigenous communities.

A community-led initiative began in 2009 to develop an approach to measuring the needs of this population. At the onset, the team recognized that it was crucially important to look for opportunities to blend Indigenous ways of knowing with Western science. This has been referred to as "two-eyed seeing" [23]. The team also acknowledged the potential to apply several available approaches to health measurement psychometric evaluation. Increasingly, methods guided by Rasch Measurement Theory (RMT) [24–27] are supporting the development of person-centered and psychometrically sound PROMs that are designed to inform individual-level screening and/or evaluation. Specifically, the principles of the Rasch model are powerful for evaluating the way in which items capture the full range of a construct being measured (e.g., exploring for indicators of construct under or over representation and understanding which items cover which part of the construct being measured). This approach can support an understanding of whether existing PROMs require changes or modifications to fit to the Rasch model (e.g., response option, scoring, deleting items, etc.). For the purposes of this program of research, this information is important to guide refinements of the Children's Health and Well-being Measure (ACHWM) and help initiate conversations with Indigenous children, youth, service providers, and elders about where measurement strengths and areas for improvement exist.

These methods are yet to be applied to Canadian Indigenous health assessment tools.

As such, the purpose of this study was to measure the extent to which RMT methods can be used to examine the psychometric properties of the Aaniish Naa Gegii: the Children's Health and Wellbeing Measure (ACHWM)[1]. The ACHWM is developmentally and culturally relevant and grounded in the Medicine Wheel, providing an overall score and scores in four domains mapped back to the Medicine Wheel: physical, emotional, mental, and physical. Our specific question is "How well do the items within the ACHWM work collectively and within each domain to estimate the health and wellness of Indigenous children and youth?" This information is critical to create a robust understanding of the instrument to inform clinical practice, policy, and future measurement refinement that is fit for purpose in Indigenous communities in Canada.

## 2. Methods

### *2.1. Aaniish Naa Gegii: the Children's Health and Wellbeing Measure*

The ACHWM is a measure of health and wellbeing from the perspectives of Indigenous children and youth. It was

[1] In the original research the A in ACHWM represented Aboriginal; however, under the direction of the children this was changed to Aanissh Naa Gegii, meaning how are you in Anishnabemowin.

developed through collaboration between First Nation health leaders, children and youth, and academic researchers. Its initial purpose was to enable health leaders to assess the health of children in their community, evaluate local program delivery, and generate evidence to guide local service planning. It was developed from the perspectives of children living in the Wiikwemkoong Unceded Territory using photo voice techniques [28] and is an exemplar of youth and community engagement, then adapted and validated for Inuit, Métis, and children in care [29]. The ACHWM contains 62 items that are aggregated into an overall summary score and four domain scores, based on the underlying Medicine Wheel conceptual model [30]: spiritual, emotional, physical, and mental (or intellectual) health. There are two different 5-point categorical response sets (frequency and importance). The intended age range is 8 to 18 years. More information on the instrument and copies of previous publications [28,29,31−36] are available at www.ACHWM.ca.

### 2.2. Design and context

This study was a secondary analysis of existing data. These data were gathered from a general community sample between August 20, 2013 and August 8, 2015. These data are controlled by communities [37,38] and were included in this analysis with their permission. These communities included three rural First Nations and two Indigenous agencies. Participants were recruited within communities by local community members, primarily through the local health centres and recreation facilities.

### 2.3. Dataset

The dataset included the 62 ACHWM items, a global rating of self-reported health with five response options (ranging from poor to excellent) and age and gender data for each respondent ($n = 402$). All data were collected with Research Ethics Board approval and the consent of participants and their parents when required.

### 2.4. Analysis procedure

Before this study began, the instrument development team was asked to review the items and hypothesize the item structure of the overall measure and of each domain (e.g., which items belonged in each domain and which items were easiest and hardest). Next, the lead author (S.B.) reviewed the dataset for obvious inconsistencies and missing data on a case-wise basis and excluded cases when data quality was poor (Section 3.1). We described the sample in terms of age, gender, and health and well-being scores and calculated an overall summary score and domain scores and examined floor and ceiling effects based on domain scores. We identified a priori that some items would have ceiling and floor effects based on the level of difficulty of the item (i.e., its location within the construct). We examined the distribution of the ACHWM summary score and domain scores. We also assessed the relationship between each item, the ACHWM summary score, and each of the domain scores, using Spearman's correlations. Finally, we assessed the internal consistency of all 62 items as a single construct and the items within each of the four domains using Cronbach's alpha [39].

### 2.5. Rasch measurement approach

We used methods guided by RMT to elucidate the item structure within the overall measure (all 62 items) [40]. This approach seeks to identify the item hierarchy within the construct of health and wellness, although simultaneously maximizing the ability to discriminate between individuals based on person logit. Item hierarchy arranges items as per their ability to detect very healthy (referred to as easy items) to very unhealthy (hard items), although person discrimination tests whether people are reliably responding to these items as expected for the amount of health they have. We used the Masters' partial credit Rasch polytomous model [41], an appropriate mathematical derivation of the Rasch model suitable for investigating items with ordinal response options. All analyses were performed using RUMM 2030 [42].

Once the main analysis was complete, we repeated these methods for the items within each of the domain scores. The domain scores were conceptually derived based on the cultural understandings of health grounded in the Anishinaabek Medicine Wheel teachings. The domain scores are relevant at the individual level to inform a strength-based conversation to promote wellness. Thus, it was critical to optimize item behaviour while retaining the ability to generate domain scores.

#### 2.5.1. Thresholds for item response options

For each item and response scale, we looked at the thresholds between the response options to determine if the successive raw response options were ordered as intended, meaning that a "1" on a 5-point scale was lower in the continuum than a "2" and so on [43]. We recoded the scores if disordered thresholds were detected.

#### 2.5.2. Item fit statistics

We tested the extent to which the participants' responses to each item met the expectations of the Rasch model. Misfit of an item implies that the item is not working as intended and warrants further investigation. We applied three indicators of fit: (a) log residuals (item-person interaction) (b) Chi-square values (item-trait interaction), and (c) item characteristic curves. Using graphical and statistical methods, we specifically looked for fit residuals between −2.5 and +2.5 and Chi-square values that are nonsignificant after Bonferroni adjustment. Rather than using absolute criteria for interpreting fit, these

indicators of fit were interpreted separately to understand the contribution of each item to capturing the targeted concept of interest within each domain (spiritual, emotional, physical, and mental).

#### 2.5.3. Structural validity

We also examined the extent to which ACHWM items in each domain covered the latent spectra of each dimension of health in an order expected by the expert team. Specifically, we statistically and graphically examined how items in the summary and domain scores were distributed along the proposed respective latent continuum as experienced by the sample (from low to high). We flagged items in similar locations as potentially redundant, perhaps out of expected order, and also we identified where gaps existed throughout the estimated targeted range (−3 to +3 logit).

#### 2.5.4. Dependency

We tested to see whether the responses to any of the items in the scales directly influenced the response to other items by examining whether item residual correlations between pairs of items were more than 0.25 [44].

#### 2.5.5. Stability/differential item functioning

Using analysis of variance, we determined whether each item in the scales performed the same across subgroups (gender, age, and community) within the sample ($P < 0.05$). We also used item characteristic curves and Chi-square values significant after Bonferroni adjustment to indicate items with potential differential item functioning [45].

#### 2.5.6. Person separation index

We computed the person separation index for each scale as an internal consistency statistic [46] that is analogous to Cronbach's alpha [39] to measure errors associated with measurement of participants in the sample and test the extent to which scale scores in the sample can be separated. High values are reflective of greater reliability. This was an important test to the community to ensure that there was a potential to identify subgroups (e.g., low, medium, high) to support future achievement of norms for communities using the ACHWM.

## 3. Results

### 3.1. Sample details

Data were available from 402 children across five sites. After a detailed review of response patterns, we identified one case with multiple irregularities (e.g., suggesting a pattern of response that they were not paying attention) and this case was excluded. Thus, we analyzed a pooled sample of 401 respondents. In this sample, 50.6% were male. The mean age of the respondents was 13.4 years (standard deviation [SD] = 3.4) with a range of 6.9 to 21.8 years. Self-reported global health ratings were available for 397 and the distribution of responses was 19% excellent, 25% very good, 41% good, 12% fair, and 3% poor health. The data were gathered at schools (60%), a youth center (19%), community events (12%), and local health centers (9%).

### 3.2. ACHWM score distributions

In this sample, the mean of the ACHWM summary score was 73.7 (SD = 12.9; range: 39.0 to 98.8). Note that these scores were scaled to a 0–100 range where 100 indicates optimal wellness. The means for each of the scaled domain scores were as follows: Spiritual 80.0 (SD = 14.4; range: 31.7 to 100), Emotional 72.9 (SD = 15.6; range: 24.4 to 100), Physical 76.1 (SD = 14.34; range: 34.4 to 100), and Mental 61.9 (SD = 14.0; range: 22.2 to 100). Distributions for the raw scores and the traditional psychometric scale analysis findings are reported in Table 1.

### 3.3. Overall Rasch results

The initial analysis of the 62-item set with five response categories revealed moderate-fit overall fit [47] (raw $\chi^2 = 809$, $P < 0.001$, degrees of freedom = 310) as seen in Table 2 (column 2 under Initial Model), with six items that showed both graphical and statistical evidence of misfit: 4, 28, 40, 53, 61, and 62. These were worthy of further consideration and were carefully monitored throughout the remaining sections of the analysis.

#### 3.3.1. Thresholds

We observed that 32 of 62 items showed evidence of disordered thresholds. This was corrected by collapsing the five response categories into three as follows: responses 1 and 2 (on the low or poor health end of the response scale) were recoded as 0, responses 3 and 4 were recoded as 1, and the highest (best) response (5) was recoded as 2. For consistency purposes, we recoded all 62 items in this way. All subsequent analyses used the collapsed response categories (0, 1, and 2). This recoding resulted in improved overall fit to the Rasch model as seen in Table 2 (column 3) with excellent-fit overall fit [47] (raw $\chi^2 = 351$, $P < 0.054$, degrees of freedom = 310), with four items showing both graphical and statistical evidence of misfit: 4, 53, 61, and 62. The model was repeated after removing these two items and the resulting fit was mildly improved as shown in Table 2 (column 4).

#### 3.3.2. Structural validity

As shown in Table 3, the items in each domain were clinically cohesive and covered the latent spectra of each dimension.

**Table 1.** Traditional psychometric based on raw scores of the original 62 items ($n$ = 401 youth)

| Property | Physical (13 items) | Emotional (24 items) | Mental (9 items) | Spiritual (16 items) |
|---|---|---|---|---|
| Data quality | | | | |
| Missing data (%) | <1% | <1% | <1% | <1% |
| Computable scale scores | 400 | 399 | 400 | 399 |
| Scale assumptions | | | | |
| Average item mean scores (range) (based on 0 to 4 scaling where 4 is best) | 3.04 (2.76–3.60) | 2.13 (0.46–3.56) | 2.50 (1.76–3.50) | 3.07 (2.32–3.54) |
| Item SD: range | 0.75–1.25 | 0.81–1.21 | 0.86–1.17 | 0.72–1.23 |
| Targeting (match between items and respondents) | | | | |
| Raw Mean score (SD) | 47.81 (7.19) | 68.67 (12.45) | 22.25 (4.58) | 47.70 |
| Possible score range[a] | 0–52 | 0–96 | 0–36 | 0–64 |
| Observed score range | 15–52 | 25–96 | 8–36 | 18–64 |
| Floor/ceiling effect (%)[b] | 0/2 | <1/<1 | <1/<1 | 0/2 |
| Internal Consistency | | | | |
| Cronbach's alpha | 0.82 | 0.87 | 0.60 | 0.87 |
| Interitem correlation: range | 0.01–0.25 | 0.04–0.44 | 0.01–0.33 | 0.04–0.37 |
| Readability statistics (FK) | 3.2 | 2.7 | 1.7 | 3.2 |

*Abbreviations*: FK, Flesch Kincaid.
[a] Higher scores indicate higher on each domain.
[b] Floor effect = % scoring lowest score; ceiling effect = % scoring maximum possible score.

### 3.3.3. Targeting

The overall summary score and domain scores all had excellent targeting with the exception of one gap at the low end of the mental domain scale between −1.4 and −2.4. The items collectively covered a range from −4.51 to 6.02 logit, with no gaps along the theoretical continua. These results are presented in Table 3.

### 3.3.4. Item fit statistics

Details of the item-fit statistics for each domain are presented in Table 4 and Figure 1.

*3.3.4.1. Spiritual domain*. The Spiritual scale was based on 16 items. These had a good overall distribution and none were identified as misfitting. However, small gaps in the distribution were apparent among the items at the upper (higher functioning) end of the scale.

*3.3.4.2. Emotional domain*. The Emotional scale was based on 24 items. Four had misfit suggesting further review. Items 4 (bullying), 28 (stay home from school), and 40 (ending one's life) had modest misfit but were retained because of their conceptual and clinical relevance. Item 62 (importance of long-time friendships) was identified as misfitting and was removed after confirmation with the ACHWM team.

*3.3.4.3. Physical domain*. The Physical scale was based on 13 items. Two items 53 (enough to eat) and 61 (importance of drinking water) were misfitting. The first was retained due to its unique location on the continuum and item 61 was removed due to overlap. Item 30 (sports) was more

**Table 2.** Summary of Rasch model fit statistics by phase of analysis

| Measure of fit | Initial model (five response categories) | Overall model adjusted for thresholds (three response categories) | Revised overall model[a] with two misfitting items removed |
|---|---|---|---|
| Number of items | 62 | 62 | 60 |
| Item fit residual (SD) | 0.00 (0.63) | 0.01 (0.63) | 0.01 (0.64) |
| Person fit residual (SD) | 1.07 (0.69) | 1.01 (0.61) | 1.03 (0.64) |
| Total item $\chi^2$ | 809.16 | 351.11 | 340.68 |
| $\chi^2$ $P$ value[b] | < 0.001 | 0.054 | 0.022 |
| Person Separation Index | 0.93 | 0.93 | 0.93 |

[a] Subtest analysis: subtest items excluding items 4, 53, 61, 62 for reasons of graphical and statistical misfit and using the three response categories.
[b] Bonferroni adjusted Chi-square 0.005.

**Table 3.** Summary of Rasch model fit statistics for the refined ACHWM domain scores

| Measure of fit | Physical ($n$ = 12) | Mental ($n$ = 9) | Emotional ($n$ = 22) | Spiritual ($n$ = 16) |
|---|---|---|---|---|
| Item fit residual (SD) | 0.55 (1.55) | −0.02 (0.81) | −0.12 (1.64) | 0.01 (1.29) |
| Person fit residual (SD) | −0.28 (1.25) | 0.11 (0.98) | 0.94 (1.07) | 1.39 (1.25) |
| Total item $\chi^2$ | 111.67 | 40.11 | 127.84 | 72.13 |
| $\chi^2$ $P$ value[a] | < 0.001[b] | 0.68 | 0.12 | 0.72 |
| Person Separation Index | 0.78 | 0.59[c] | 0.85 | 0.82 |

[a] Bonferroni adjusted Chi-square 0.005.
[b] Item fit improves with the removal of item #53 for a $\chi^2$ = 56.13 and $P$ = 0.43.
[c] Moderate PSI was related to a gap in the conceptual range between −1.4 and −2.4 logit.

often endorsed positively by boys than girls, but this was not unexpected and will be monitored. It is the only item with slightly differential functioning related to gender.

*3.3.4.4. Mental domain (cognition).* The Mental scale is based on nine items related to thinking and problem solving (mental processing). These items generally had a good distribution and good fit to the model with one gap identified at the lower end of the continuum (Fig. 1D).

### *3.3.5. Local dependency*

Local dependence was observed between items 33 (self-harm) and 40 (ending one's life) which had a Spearman's correlation of 0.65 and a local item dependence of 0.37 after accounting for the latent factor. Our a priori standard set the acceptability at < 0.25 [25]; however, both were retained because of their importance to the community.

### *3.3.6. Stability/differential item functioning*

Differential item functioning for gender was apparent on one item (30, Sports) and was not unexpected. We will continue to monitor this item.

### *3.3.7. Person separation index*

Scale reliability was high (person separation index = 0.93), indicating the ACHWM items adequately separated the participants along the measurement continuum.

## 4. Discussion

As health systems shift their attention to measuring patient-centered outcomes, systematically capturing the needs and priorities of Indigenous children and youth is critical. The aim of this study was to psychometrically evaluate the ACHWM and assess its ability to capture the wellness of a sample of Canadian Indigenous youth. The main finding of this study is the refinement of the ACHWM to be a 60-item scale with a 3-point response scale. In this format, the ACHWM scores can be interpreted confidently as reliable and valid when scored as indicators of the domains. Specifically, the results lend evidence towards sets of items within the ACHWM that work together both conceptually and psychometrically to cover the full ranges of the domains (spiritual, physical, emotional, and mental). The logit scores allow for meaningful interpretation of the ACHWM scores and can support communities to map services along the spectrum of items capturing the domains being measured. However, the results do suggest the minor improvements may be achieved by the removal of items 61 (importance of drinking water) and 62 (importance of long-time friendships), which would reduce the domains of physical and emotional health, respectively, by one item each. Removal of these items has been approved by community leaders. We have demonstrated that the ACHWM items are reliable, unidimensional, and meet the criteria for objective rigorous measurement as outlined by the Rasch model.

Our results also suggested some "statistically significant" anomalies that warrant further monitoring. For example, local dependence was observed between items 33 (self-harm) and 40 (suicide). Community leaders, children, and youth underscored the importance of both questions in this population. Stakeholders validated the ordering of the items in the hierarchy, and in this case, advocated strongly for the items to be included despite statistical misfit. It was also felt that interventions to support young people who endorse these items would be different, which is also supported by the literature [48]. As such, for conceptual reasons, both of these items have been retained in the final measure.

It is important to put the ACHWM scores into context. The best data available for comparison come from the First Nations Regional Health Survey, which found that 22.8% of First Nations youth (aged 12 to 17 years) in Ontario reported their health as "excellent" [49]. In the pooled sample reported here, 18.6% reported their health as "excellent". Furthermore, among those who were aged 12 to 17 years ($n$ = 188) in this same sample, 15.4% reported their health to be "excellent". This comparison suggests that the provincial results may reflect some sampling bias and reinforces the need for local data to inform health planning.

Communities, Indigenous service agencies, primary care professionals, families, educators, and government may be able to use precise, interpretable measures such as the ACHWM to engage in dialogue with Indigenous youth and

**Table 4.** RMT statistical indicators of fit for each domain score of the ACHWM

| Scale | Item # | Description | Location | SE | FitResid | Chi$^2$ | Prob |
|---|---|---|---|---|---|---|---|
| Physical | | | | | | | |
| | 46 | Access to clean water | −1.122 | 0.107 | 0.957 | 12.782 | 0.026 |
| | 50 | Being active | −0.38 | 0.098 | −2.412 | 13.714 | 0.018 |
| | 8 | Energy level | −0.327 | 0.106 | −1.644 | 7.373 | 0.194 |
| | 6 | Exercise | −0.264 | 0.097 | −1.31 | 4.928 | 0.425 |
| | 29 | Active outdoors | 0.005 | 0.096 | −2.027 | 11.217 | 0.047 |
| | 19 | Staying safe | 0.021 | 0.097 | 1.355 | 6.125 | 0.294 |
| | 12 | Access to adult support | 0.073 | 0.087 | 1.897 | 6.629 | 0.250 |
| | 43 | Eating healthy food | 0.139 | 0.111 | −1.683 | 10.378 | 0.065 |
| | 5 | Healthy choices | 0.188 | 0.12 | −0.61 | 2.812 | 0.729 |
| | 30 | Sports | 0.266 | 0.083 | 2.476 | 10.472 | 0.063 |
| | 2 | Physically fit | 0.419 | 0.1 | −2.003 | 13.193 | 0.022 |
| | 53 | Enough to eat[a] | 0.98 | 0.118 | **5.408** | 75.024 | 0,000 |
| Emotional | | | | | | | |
| Missing 62 | 42 | Happiness | −1.092 | 0.103 | 1.083 | 8.023 | 0.155 |
| | 40 | Ending my life | −0.975 | 0.102 | **−3.047** | 27.462 | < 0.001 |
| | 24 | Feeling safe at home | −0.797 | 0.121 | −0.359 | 7.497 | 0.186 |
| | 1 | Laughter and fun | −0.776 | 0.101 | 0.94 | 9.978 | 0.076 |
| | 49 | Happiness | −0.596 | 0.097 | −0.833 | 6.4 | 0.269 |
| | 33 | Hurt myself | −0.443 | 0.089 | −2.002 | 13.228 | 0.022 |
| | 44 | Supportive family | −0.324 | 0.092 | −0.78 | 2.551 | 0.769 |
| | 45 | Hope | −0.314 | 0.138 | −0.922 | 2.576 | 0.765 |
| | 52 | Feel loved | −0.153 | 0.09 | −0.717 | 5.442 | 0.364 |
| | 23 | Breaking things | −0.113 | 0.085 | −0.036 | 11.336 | 0.045 |
| | 51 | Feeling safe in community | −0.11 | 0.09 | −1.28 | 10.802 | 0.055 |
| | 28 | Skip school | −0.095 | 0.094 | **3.142** | 30.841 | < 0.001 |
| | 48 | Not bored | 0.231 | 0.088 | −0.775 | 7.244 | 0.2031 |
| | 18 | Hurting others | 0.279 | 0.087 | 1.614 | 11.262 | 0.0464 |
| | 4 | Bullied | 0.425 | 0.084 | **3.962** | 27.208 | < 0.001 |
| | 38 | Moody | 0.448 | 0.102 | −0.511 | 6.267 | 0.281 |
| | 31 | Easily upset | 0.476 | 0.082 | 1.258 | 6.36 | 0.272 |
| | 37 | Encouraged | 0.728 | 0.085 | −0.893 | 1.463 | 0.917 |
| | 21 | Lonely | 0.744 | 0.084 | **−2.742** | 19.588 | 0.002 |
| | 26 | Bounce-back quickly | 0.877 | 0.097 | −0.644 | 5.643 | 0.342 |
| | 41 | Worried in my body | 0.93 | 0.088 | 0.721 | 5.1 | 0.404 |
| | 3 | Afraid or scared | 1.114 | 0.102 | −0.173 | 1.164 | 0.948 |
| Mental | | | | | | | |
| | 20 | Proud of my community | −1.67 | 0.10 | −1.23 | 8.36 | 0.137 |
| | 55 | Learning is important | −1.53 | 0.10 | −0.32 | 7.23 | 0.204 |
| | 25 | Have alone time | −1.38 | 0.09 | 1.10 | 10.95 | 0.052 |
| | 39 | I get a good night's sleep | −0.18 | 0.08 | −0.78 | 8.58 | 0.127 |
| | 16 | Succeeding at school | 0.40 | 0.09 | −0.11 | 4.77 | 0.444 |
| | 7 | Focusing during school | 0.72 | 0.09 | −0.20 | 3.26 | 0.660 |
| | 36 | Too much to do | 0.73 | 0.10 | 0.41 | 8.81 | 0.117 |
| | 9 | Forget things | 1.23 | 0.10 | −0.30 | 6.56 | 0.256 |
| | 47 | I miss doing things that used to be fun | 1.67 | 0.10 | 1.23 | 5.49 | 0.359 |

(*Continued*)

**Table 4.** Continued

| Scale | Item # | Description | Location | SE | FitResid | Chi$^2$ | Prob |
|---|---|---|---|---|---|---|---|
| Spiritual | | | | | | | |
| | 22 | Grateful for what I have | −1.194 | 0.112 | −0.917 | 7.385 | 0.194 |
| | 14 | Respecting others | −0.844 | 0.109 | 2.409 | 10.351 | 0.066 |
| | 59 | Importance of traditional medicine | −0.694 | 0.105 | 0.057 | 6.782 | 0.237 |
| | 13 | Seeing the beauty in nature | −0.689 | 0.104 | 1.972 | 10.097 | 0.073 |
| | 57 | Importance of language | −0.666 | 0.104 | −0.179 | 2.483 | 0.779 |
| | 58 | Importance of culture | −0.497 | 0.103 | −0.727 | 7.959 | 0.159 |
| | 54 | Hope for the future | −0.494 | 0.149 | 0.663 | 2.222 | 0.818 |
| | 34 | Good support person | −0.285 | 0.095 | 1.244 | 9.383 | 0.095 |
| | 56 | Time with elders | −0.114 | 0.104 | −1.269 | 7.242 | 0.203 |
| | 11 | Time with family | 0.069 | 0.094 | 0.691 | 0.683 | 0.984 |
| | 60 | Importance of a higher power | 0.433 | 0.088 | −0.024 | 2.049 | 0.842 |
| | 27 | Choosing a good path | 0.490 | 0.102 | 0.488 | 5.190 | 0.393 |
| | 32 | Connection to land | 0.710 | 0.093 | **−2.941** | 14.849 | 0.011 |
| | 35 | Learning native language | 1.018 | 0.089 | −0.389 | 4.803 | 0.440 |
| | 10 | Time with elders | 1.162 | 0.088 | −0.075 | 5.541 | 0.353 |
| | 15 | Connecting with a higher power | 1.594 | 0.088 | −0.974 | 16.751 | 0.005 |

Bold = Fit residuals that exist outside of the −2.5 and +2.5 expected range.

[a] When removing item #53 from the model, fit improves. The item was chosen to stay because of its conceptual importance and clinical relevance.

measure the effectiveness of interventions/programs over time. We anticipate that rigorous and meaningful assessment of the health and wellbeing of this population may help: (1) identify early warning signs of illness, (2) prioritize the need for integrated youth-centered services, (3) create a road map for the types of responsive and sustainable services that are needed to optimize the health and wellness of Indigenous youth, and (4) act as a culturally appropriate means for stakeholders to communicate. As well, given the rigour and interpretability of the findings, policy makers may also use the ACHWM to systematically evaluate youth-oriented Indigenous health services across Canada.

During the development of this article, we identified one publication in which RMT methods were applied a Native American sample to refine a quality of life tool related to vision in the United States [50] and one publication related to numeracy testing of aboriginal and nonaboriginal children in Australia [51]. These studies had sample sizes of 181 and 290 children, respectively, and support the utility of this approach in Indigenous populations. This article extends that work to Indigenous children in Canada with a focus on ‘‘wholistic’’ health and wellness.

### *4.1. Limitations*

The intended context of use for the ACHWM is communities where Indigenous youth reside. For exploratory purposes, we began our conceptual and measurement testing with young adults in one geographical location, Ontario. Our preliminary results show that the ACHWM targets this sample well, but we recognize that this sample may not generalize to other Indigenous young people in Canada. Future testing with other samples is needed to understand the extent to which the ACHWM items are fit for purpose and measure what they purport to measure in diverse contexts.

## 5. Conclusion

It is time to listen to the voices of Indigenous children: a point made in several recent reports including the Feathers of Hope [52]. The perspectives of youth (patients) have been engrained in the ACHWM's development from inception. This study not only provides evidence to further support the ACHWM as a reliable and valid tool to measure health and wellness but also suggests how communities can work together to improve outcomes for Indigenous youth. Opportunity exists to build capacity for stakeholders to use the ACHWM to align services with the needs of Indigenous youth and provide the impetus to fund and support the relevant services where and when youth need it.

## Acknowledgments

We are grateful for the contributions made by many First Nations and Inuit children and youth, whose wisdom led to the creation of the Aaniish Naa Gegii and Qanuippit, and to their peers who contributed to the data reported here. We are grateful to the Indigenous community leaders who played a critical role in gathering the data and are using the results to

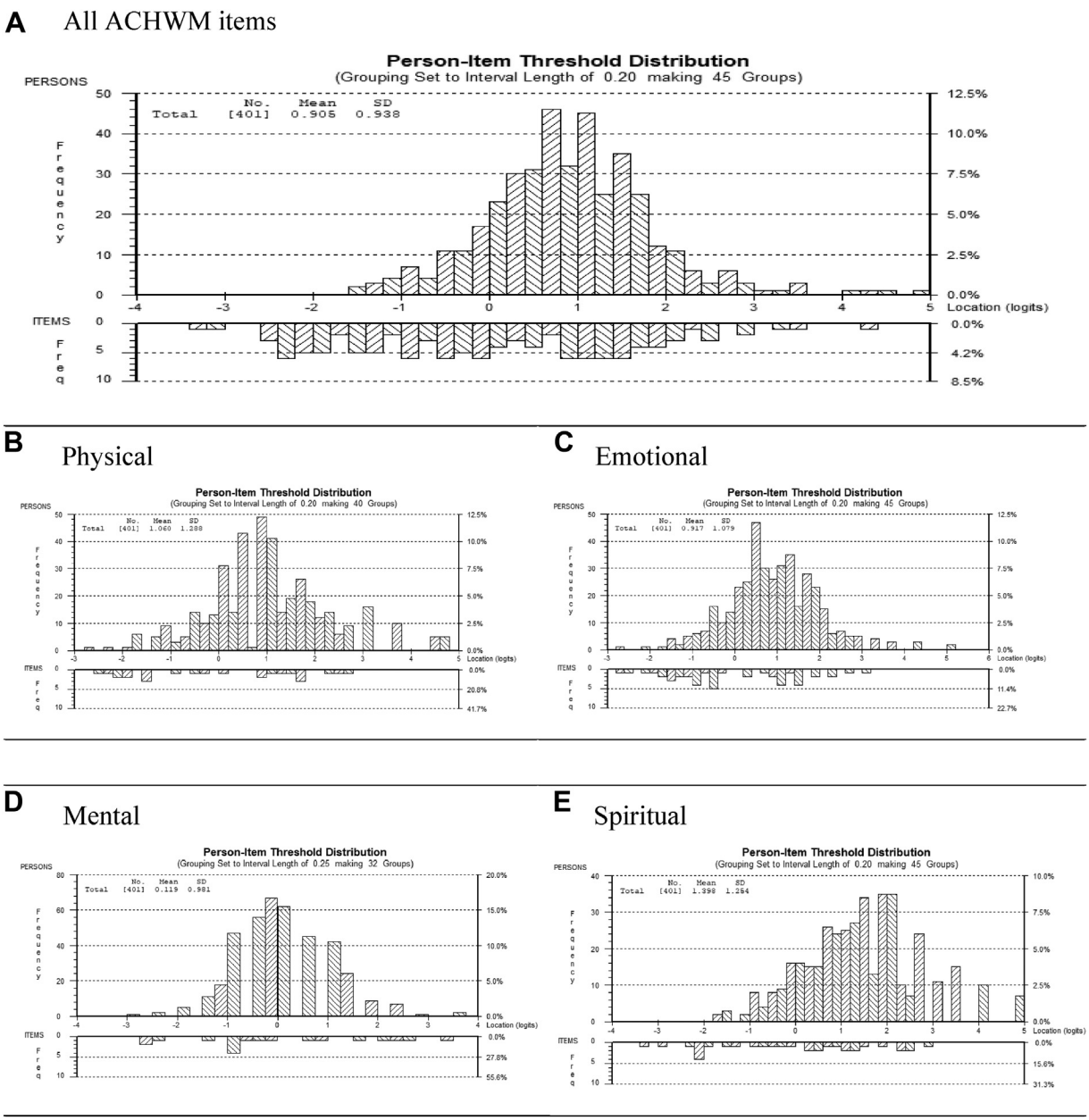


**Fig. 1.** Distribution of 62 items together (A) and by domain (B-physical, C-emotional, D-mental, and E-spiritual), obtained by converting total raw scores into linear measurements across the measurement continuum in the ACHWM. The distribution above the horizontal axis represents the distribution of the sample (children and youth, aged 8–18 years). The distribution below the horizontal axis represents the distribution of the items from low to high.

inform programming within their communities. These leaders have elected to not be named in this article to keep the identities of their communities and their children confidential. In doing so, they have enabled important lessons to be shared to enhance the wellbeing of many more Indigenous children, youth, and communities. Miigwetch/Nakurmiik.

## References

[1] Reeve BB, Wyrwich KW, Wu AW, Velikova G, Terwee CB, Snyder CF, et al. ISOQOL recommends minimum standards for patient-reported outcome measures used in patient-centered outcomes and comparative effectiveness research. Qual Life Res 2013;22:1889–905.

[2] Patel AA. Patient-reported outcome measures: the promise of PROMIS. J Am Acad Orthop Surg 2016;24(11):743.

[3] Cano SJ, Hobart JC. The problem with health measurement. Patient Prefer Adherence 2011;5:279–90.

[4] Patrick D. Reporting of patient-reported outcomes in randomized trials: the CONSORT PRO extension. Value Health 2013;16(4):455–6.

[5] Belvedere SL, de Morton NA. Application of Rasch analysis in health care is increasing and is applied for variable reasons in mobility instruments. J Clin Epidemiol 2010;63:1287–97.

[6] Hobart J, Cano S, Baron R, Thompson A, Schwid S, Zajicek J, et al. Achieving valid patient-reported outcomes measurement: a lesson

from fatigue in multiple sclerosis. Mult Scler 2013;19(13):1773−83.
[7] Canadian Institutes for Health Research. Strategy for Patient-Oriented Research. 2016. Available at http://www.cihr-irsc.gc.ca/e/41204.html. Accessed September 6, 2016.
[8] Frank L, Basch E, Selby JV. The PCORI perspective on patient-centered outcomes research. JAMA 2014;312:1513−4.
[9] Bartlett SJ, Barnes T, McIvor RA. Integrating patients into meaningful real-world research. Ann Am Thorac Soc 2014;11 Suppl 2:S112−7.
[10] Bartlett SJ, Bykerk VP, Cooksey R, Choy EH, Alten R, Christensen R, et al. Feasibility and domain validation of rheumatoid arthritis (RA) flare core domain set: report of the OMERACT 2014 RA flare group plenary. J Rheumatol 2015;42:2185−9.
[11] Anker SD, Agewall S, Borggrefe M, Calvert M, Jaime Caro J, Cowie MR, et al. The importance of patient-reported outcomes: a call for their comprehensive integration in cardiovascular clinical trials. Eur Heart J 2014;35:2001−9.
[12] Basch E, Snyder C, McNiff K, Brown R, Maddux S, Smith ML, et al. Patient-reported outcome performance measures in oncology. J Oncol Pract 2014;10(3):209−11.
[13] Cohen WA, Mundy LR, Ballard TN, Klassen A, Cano SJ, Browne J, et al. The BREAST-Q in surgical research: a review of the literature 2009-2015. J Plast Reconstr Aesthet Surg 2016;69(2):149−62.
[14] Klassen AF, Cano SJ, Alderman A, Soldin M, Thoma A, Robson S, et al. The BODY-Q: a patient-reported outcome instrument for weight loss and body contouring treatments. Plast Reconstr Surg Glob Open 2016;4(4):e679.
[15] Klassen AF, Cano SJ, Grotting JC, Baker SB, Carruthers J, Carruthers A, et al. FACE-Q eye module for measuring patient-reported outcomes following cosmetic eye treatments. JAMA Facial Plast Surg 2017;19(1):7−14.
[16] Kirwan JR, de Wit M, Frank L, Haywood KL, Salek S, Brace-McDonnell S, et al. Emerging guidelines for patient engagement in research. Value Health 2017;20(3):481−6.
[17] Calvert M, Blazeby J, Altman DG, Revicki DA, Moher D, Brundage MD. Reporting of patient-reported outcomes in randomized trials: the CONSORT PRO extension. JAMA 2013;309:814−22.
[18] Canadian UNICEF Committee. Canadian Supplement to the State of the World's Children. Aboriginal Children's Health: Leaving No Child Behind. 2009. Available at www.unicef.ca/sites/default/files/.../Leaving%20no%20child%20behind%2009.pdf. Accessed May 12, 2010.
[19] Health Canada. First Nations and Inuit Health Strategic Plan: A shared path to improved health. Ottawa, Ontario, Canada: Health Canada; 2012.
[20] Advisory Group on Suicide Prevention (Canada). Acting On What We Know: Preventing Youth Suicide in First Nations. 2003. Available at http://www.hc-sc.gc.ca/fniah-spnia/alt_formats/fnihb-dgspni/pdf/pubs/suicide/prev_youth-jeunes-eng.pdf. Accessed November 23, 2015.
[21] Adelson N. The embodiment of inequity: health disparities in aboriginal Canada. Can J Public Health 2005;96:S45−61.
[22] MHASEF Research Team. The Mental Health of Children and Youth in Ontario: A Baseline Scorecard. Toronto, Ontario, Canada: Ontario Institute of Clinical Evaluative Sciences; 2015.
[23] Barlett C, Marshall M, Marshall A. Two-Eyed seeing and other lessons learned within a co-learning journey of bringing together Indigenous and mainstream knowledges and ways of knowing. J Environ Stud Sci 2012;2:331−40.
[24] Andrich D. Rasch Models for Measurement. Beverly Hills, CA: Sage Publications; 1988.
[25] Andrich D. Rating scales and Rasch measurement. Expert Rev Pharmacoecon Outcomes Res 2011;11(5):571−85.
[26] Rasch G. Probabilistic Models for Some Intelligence and Attainment Tests. Copenhagen: Danish Institute for Education Research; 1980.
[27] Rasch G. On General Laws and the Meaning of Measurement in Psychology. Proceedings of the IV Berkeley Symposium on Mathematical Statistics and Probability, 4, 321-333. Publications Ltd; 1961. Available at https://dx.doi.org/10.4135/9781446262481.n15.
[28] Young NL, Wabano MJ, Burke TA, Ritchie SD, Mishibinijima D, Corbiere RG. A process for creating the aboriginal children's health and well-being measure (ACHWM). Can J Public Health 2013;104(2):e136−41.
[29] Young NL, Wabano MJ, Blight S, Baker-Anderson K, Beaudin R, McGregor LF, et al. Relevance of the aboriginal children's health and well-being measure (ACHWM) beyond wikwemikong. Rural and Remote Health 2017;17(2):394−404.
[30] Dumont J. First Nations regional longitudinal health survey (RHS) cultural framework. Ottawa (ON): RHS; 2005.
[31] Young NL, Wabano MJ, Usuba K, Pangowish B, Trottier M, Jacko D, et al. Validity of the aboriginal children's health and well-being measure: aaniish Naa Gegii? Health Qual Life Outcomes 2015;13(1):148−56.
[32] Young NL, Wabano MJ, Ritchie SD, Burke TA, Pangowish B, Corbiere RG. Assessing children's interpretations of the aboriginal children's health and well-being measure (ACHWM). Health Qual Life Outcomes 2015;13(105):1−7.
[33] Young NL, Wabano MJ, Usuba K, Mishibinijima D, Jacko D, Burke TA. Reliability of the aboriginal children's health and well-being measure (ACHWM). Springer Plus 2016;5(1):2082−7.
[34] Young NL, Jacko D, Wabano MJ, Hawthorne L, Seabrook S, Wabanosse S, et al. A screening mechanism to recognize and support aboriginal children at-risk: based on a child-centric survey. Can J Public Health 2016;107(4−5):e399−403.
[35] Young NL, Wabano MJ, The ACHWM Effectiveness Team. Beyond the patient − lessons from community engagement in rural and remote first Nations. Can Med Assoc J 2018;190(S1):S16−8.
[36] Wabano MJ, McGregor LF, Beaudin R, Jacko D, McGregor LE, Kristensen-Didur S, et al. Health profiles of first Nations children living on-reserve in Northern Ontario. CMAJ Open 2019;7(2):E316−22.
[37] Schnarch B. Ownership, control, access and possession (OCAP) or self-determination applied to research. A critical analysis of contemporary first Nations research and some options for first Nations communities. J Aboriginal Health 2004;1(1):80−95.
[38] First Nations Centre. OCAP: Ownership, Control, Access and Possession Sanctioned by the First Nations Information Governance Committee, Assemby of First Nations. Ottawa: National Aboriginal Health Organization; 2007.
[39] Cronbach LJ. Coefficient alpha and the internal structure of tests. Psychometrika 1951;16:297−334.
[40] Cano SJ, Mayhew A, Glanzman AM, Krosschell KJ, Swoboda KJ, Main M, et al. Rasch analysis of clinical outcome measures in spinal muscular atrophy. Muscle Nerve 2014;49(3):422−30.
[41] Masters GN. A Rasch model for partial credit scoring. Psychometrika 1982;47(2):149−74.
[42] RUMM laboratory. RUMM2030. Rasch unidimensional measurement models software. Perth: Western Australia; 2007.
[43] Massof RW. A general theoretical framework for interpreting patient-reported outcomes estimated from ordinally scaled item responses. Stat Methods Med Res 2014;23(5):409−29.
[44] Andrich D, Kreiner S. Quantifying response dependence between two dichotomous items using the Rasch model. Appl Psychol Meas 2010;34:181−92.
[45] Cameron IM, Scott NW, Adler M, Reid IC. A comparison of three methods of assessing differential item functioning (DIF) in the Hospital Anxiety Depression Scale: ordinal logistic regression, Rasch analysis and the Mantel chi-square procedure. Qual Life Res 2014;23:2883−8.
[46] Andrich D. An index of person separation in latent trait theory, the traditional KR20 index, and the Guttman scale response pattern. Educ Psychol Res 1982;9(1):95−104.

[47] Andrich D. Controversy and the Rasch model: a characteristic of incompatible paradigms? Med Care 2004;42:I7−16.
[48] Ougrin D, Tranah T, Stahl D, Moran P, Asarnow JR. Therapeutic interventions for suicide attempts and self-harm in adolescents: systematic review and meta-analysis. J Am Acad Child Adolesc Psychiatry 2015;54(2):97−107.e2.
[49] Chiefs of Ontario. First Nations Regional Health Survey (RHS) Ontario Region Phase 2 (2008/10): People's Report. Toronto, Ontario: RHS; 2014.
[50] Crescioni M, Messer DH, Warholak TL, Miller JM, Twelker JD, Harvey EM. Rasch analysis of the student refractive error and eyeglass questionnaire. Optom Vis Sci 2014;91(6):624−33.
[51] Hippisley J, Douglas G, Houghton S. A cross-cultural comparison of numeracy skills using a written and an interactive arithmetic test. Educ Res 2005;47(2):205−15.
[52] Provincial Advocate for Children & Youth. Together we are ... Feathers of Hope: a first Nations youth action plan. Toronto: Office of the Provincial Advocate for Children and Youth; 2014.