Received: 17 April 2024 | Accepted: 6 June 2024

DOI: 10.1002/hcs2.106

ORIGINAL ARTICLE

HEALTH CARE SCIENCE

Confirmatory analysis of the 13-MD and ranking of its meta-dimensions and items

Moustapha Touré[1,2] | **Thomas G. Poder**[2,3]

[1]École de gestion, Université de Sherbrooke, Sherbrooke, Quebec, Canada

[2]Centre de recherche de l'IUSMM, CIUSSS de l'Est de l'île de Montréal, Montreal, Quebec, Canada

[3]École de santé publique, Université de Montréal, Montreal, Quebec, Canada

Correspondence
Thomas G. Poder, School of Public Health, University of Montreal, 7101 Park Ave, Montreal, QC H3N 1×9, Canada.
Email: thomas.poder@umontreal.ca

Funding information
Unité Soutien SSA du Québec

Abstract

Background: The 13-MD is a new generic instrument developed to measure general health-related quality of life (GHRQoL). This instrument considers all aspects of health (i.e., physical, mental, and social) in a balanced way. A previous study led to minor changes in the original version of the 13-MD. The objective of this study was to confirm the validity of the modified 13-MD.

Methods: Validity was assessed with recent data from the general population of Quebec, Canada. The meta-dimensions and items composing the 13-MD were also subjected to a ranking procedure, which allowed to determine the most important aspects for respondents.

Results: A total of 1337 French-speaking participants were recruited with 1099 completing the 13-MD for validation purposes and 1084 completing the ranking procedure. The 13-MD showed very satisfactory results and confirmed to be a valid instrument. The ranking of the meta-dimensions revealed that "Well-being" received the most points, followed by "Sleep and energy" and "Body functioning."

Conclusion: These results will be very useful in the continuous improvement of the 13-MD, ultimately leading to the valuation stage (i.e., development of a value set).

KEYWORDS
cost-utility analysis, health-related quality of life, generic instrument

Abbreviations: 13-MD, 13 meta-dimensions; 15D, 15 dimensions; AIC, Akaike's information criterion; AQoL-7D, Assessment of Quality-of-Life 7D; AQoL-8D, Assessment of Quality-of-Life 8D; BIC, Bayesian information criterion; CEGEP, Collège d'enseignement général et professionnel; CFI, comparative fit index; CORR, correlation value; CUA, cost-utility analysis; CUMLOR, cumulative log-odds ratio in logits; DIF, differential item functioning; EFA, exploratory factor analysis; EXP, expected value; GHRQoL, general health-related quality of life; HRQoL, health-related quality of life; IRT, item response theory; KMO, Kaiser Meyer Olkin; MNSQ, mean-square; PCA, principal component analysis; PTMEASUR-AL, point-measure correlation; QALY, quality-adjusted life-year; QWB-SA, Quality of Well Being Self-Administered; RMSEA, root mean square error of approximation; SE, standard error; SEM, structural equation model; SRMR, standardized root mean square residual; TLI, Tucker and Lewis Index; WHO, World Health Organisation; ZSTD, standardized as a *z*-score.

1 | INTRODUCTION

As technologies undergo continuous improvements, health systems are faced with a growing demand for health services, despite limited resources [1]. The quality-adjusted life-year (QALY) is a concept that is used in cost-utility analysis (CUA) to help decision-makers achieving efficiency in the use of health resources [1–3]. It combines the quality (morbidity) and the quantity (mortality) of life into a single measure to better compare interventions or programs [4, 5]. Several instruments have been developed to help in cost-utility analysis and specifically in determining the Q in QALYs [6]. In this regard, generic instruments are extremely useful in capturing the respondent's general health-related quality of life (GHRQoL) [6, 7]. By allowing the conversion of patient's health state into utility value, this concept places the patient or health services' user at the heart of health policies, representing a fundamental aspect of the effectiveness of public health policies [8]. However, most generic instruments do not encompass all aspects of health and primarily focus on physical aspects, disregarding social and mental aspects [6]. In fact, the systematic review conducted by Touré et al. [6] showed that all 12 generic instruments listed in this study considered dimensions related to physical health at the expense of dimensions related to mental and social health. However, the World Health Organisation (WHO) defines health as "a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity." Embracing this definition, a new instrument, the 13-MD, was developed to palliate the observed imbalance in the existing generic instruments [9–12]. Through a Delphi procedure, an instrument with 15 meta-dimensions and 36 dimensions or items was originally developed [12]. This version of the instrument was administered via a survey to assess its psychometric validity and determine which meta-dimensions and items should be included in the final version [12]. After conducting various tests, a version with 13 meta-dimensions and 33 items, known as 13-MD was retained due to its strong psychometric properties and better fit with the data collected with the older version of the questionnaire (15-MD). This final version of the 13-MD was derived by removing and reorganizing misfitting items from the older version (the 15-MD) and has since been translated from French to English. Linguistic validation was then conducted in both French and English versions through cognitive debriefing interviews [13]. Although the 13-MD is considered a valid instrument that comprehensively addresses all important aspects of health [12], it required to be validated in its final form with directly collected data using its final version. This study aimed to validate the 13-MD by presenting its final French version to the general population. The specific data collected in this regard allowed for the validation process, during which various tests were conducted to analyze and validate the 13-MD's psychometric properties.

The 13-MD can be considered as a lengthy instrument due to its attempt in considering all important aspects of health. Therefore, a reduced version would be beneficial in certain situations. In this case, it would be helpful to determine which meta-dimensions/items are the most important to the population. Thus, in addition to the confirmatory validation procedure of the latest version of the 13-MD, this study also presents the steps and results for ranking its meta-dimensions and items using data from the general population of Quebec, Canada.

2 | METHODS

2.1 | The 13 meta-dimensions (13-MD) classification system

The 13-MD is a measure of GHRQoL consisting of 33 items that fall into 13 meta-dimensions. A meta-dimension is a set of related components (items) designed to gauge a particular facet of health. The 13-MD measures different aspects of health, including five meta-dimensions for physical health, four meta-dimensions for mental health, three meta-dimensions for social health, and one meta-dimension for sexuality and intimacy. These meta-dimensions include "Body Functioning," "Cognitive abilities, senses, and language," "Sleep and energy," "Self-esteem and self-acceptance," "Physical pain and discomfort," "Mobility and physical disability," "Daily activities and work," "Social and leisure activities," "Social and interpersonal relationships," "Citizenship, inclusion, and autonomy," "Depression, anxiety, and anger," "Well-being," and "Sexuality and intimacy" [12]. For example, the meta-dimension "Body functioning" is composed of items like "Breathing," "Eating," and "Relieving oneself (urine and fecal matter)." This meta-dimension is one of the five meta-dimensions that aim to measure the respondent's physical health. Each item offers response options ranging from 5 to 7 levels, allowing for a vast number of potential health states (i.e., 1.42×10^{26} possible health states). The measure was originally developed in French through a Delphi procedure, with an initial version containing 15 meta-dimensions. This version underwent psychometric validation through a survey targeting the general population of Quebec [12]. The final version with 13 meta-dimensions was established after observing misfitting items. An English version of the 13-MD is also available [13].

2.2 | Study population

The final version of the 13-MD was submitted to a survey and data from the general French-speaking population of Quebec were collected through an online design survey in 2023. Participants were recruited by Dynata Inc., a company specialized in online surveys. To ensure representativity across gender, age, and education, a quota sampling method was used. To be eligible, participants must reside in Quebec (Canada) and be over 18 years old. Participants rated items on a scale from 1 (best level) to 5–7 (worst level) [12] in the final version of the 13-MD [13]. To date, no utility score can be derived from the instrument and a value set is planned to be developed to make it usable in QALY calculation for CUA. Only participants who provided complete responses were included in the analysis, and any incomplete observations were removed. Considering the Quebec population, a minimum of 1067 participants was required to ensure representativity with a confidence level of 95% and a risk of error below 3%.

2.3 | Psychometric tests

To ensure better comparativeness, the same aspects of the previous validation process were considered [12]. The degree to which items are connected and measure the same construct, known as internal reliability or consistency, was evaluated. Cronbach's α was computed to measure internal reliability (or consistency) which informs the extent to which the different items of an instrument measure a same construct or are interconnected [14]. The Cronbach's α is generally between 0 and 1 and a value between 0.70 and 0.95 was targeted (i.e., a higher coefficient means that items are correlated and measure the same concept) [15, 16]. Additionally, items that if eliminated would increase the Cronbach's α score by 2% were identified [15]. We also computed the item-test correlation, which measures the correlation between each item and the overall test correlation. The item-rest correlation was calculated to measure the correlation between one item and the other items. Furthermore, we examined the interitem correlation to assess the level of correlation between the items. We expected the interitem coefficients to range from 0.2 to 0.4 [17], item-test coefficients to be high and approximately constant across items, and item-rest coefficients to be greater than 0.2 [18].

All items were examined for floor and ceiling effects, which indicate the proportion of respondents who always selected the worst and best possible answers, respectively (i.e., the lowest and best levels describing health). This was done to ensure that the instrument's reliability was maintained and that the response levels allowed us to record variations between users [19]. To rule out the absence of ceiling or floor effects, less than 15% of responses should always indicate the best/worst achievable score [19].

A principal component analysis (PCA) was implemented to conduct an exploratory factor analysis (EFA) which helps in identifying the underlining factor structure or model for a set of variables [20, 21]. Using directions called principal components, the PCA helps in minimizing the dimensionality of the data and preserving most of the variation within it [22]. This exploratory analysis was conducted to determine if the same distribution pattern of the different items across the various components would be observed when using the new collected data as in previous analyses [12]. If this pattern holds true, we can then confirm the stability of its structure even when using different data sets.

The Bartlett test of sphericity and the Kaiser Meyer Olkin (KMO) test were also performed to assess the data's suitability for factor analysis. If results were satisfying, the PCA was performed. In our case, the varimax rotation, which is the most popular rotation method was applied [23]. It consists in searching for a rotation of the initial components to maximize the loadings' variance [23]. The PCA helped analyze the loading patterns and confirm items grouped together in the same meta-dimension (i.e., groups consisting of related items that assess a specific health characteristic) [24]. Items that contributed less than 20% to the components were considered problematic.

The EFA conducted through the PCA helps to identify and confirm the factor structure previously obtained including the number of factors and pattern of factors loadings. In addition to this exploratory analysis, a confirmatory factor analysis (CFA) was also performed. Unlike EFA which is more used as a descriptive technique, CFA, conducted through a structural equation model (SEM), allows to specify the structure in advance and assess how well the previously obtained structure replicates the sample covariance matrix [21, 25]. The CFA also provides evidence in supporting the convergence of the 13-MD's structure [25]. Various statistics, including the root mean square error of approximation (RMSEA), standardized root mean square residual (SRMR), comparative fit index (CFI), Tucker and Lewis Index (TLI), Akaike's information criterion (AIC), and Bayesian information criterion (BIC) were compiled and compared to the results of our previous analyses. We expected the RMSEA to be lower than 0.08 (the closer to 0, the better), SRMR to fall between 0.05 and 0.1, CFI and TLI to exceed 0.9, and AIC and BIC to be low [24, 26–29].

The item response theory (IRT) model was used to assess the questionnaire's ability to distinguish between

individuals in different regions on the latent continuum. IRT aims to understand or forecast the relationship between unobservable qualities or attributes (latent traits) and their observable manifestations (results or reactions) [30–32]. The IRT calculates the likelihood that individuals with varying levels of the latent trait will respond to the items in different ways. Among its many benefits, IRT enables the identification of how the level of precision may vary across different items or levels of the instrument, unlike classical test theory, which uses a single estimate to describe instrument's reliability [31]. The discriminant coefficient was of interest as a coefficient of 1.0 indicates good discrimination by the item; individuals with low ability (health) are much less likely to achieve a higher score compared to those with average or high ability. Therefore, a higher discrimination coefficient indicates better ability of the item to distinguish and provide accurate information. It is important to avoid a negative discrimination coefficient as it would imply that the probability of achieving a higher score decreases as the respondent's ability (health) increases [33].

Instrument development frequently employs the Rasch analysis method which is part of the IRT models and is probabilistic. The Rasch procedure is an analytical approach that aids in the development of new instruments and in assessing the quality of existing questionnaires [34]. The Rasch analysis allows us to analyze the fit of the questionnaire (unidimensionality) using various measures such as person/item separation reliability, person/item separation index, standard error of measurement (SEm), infit and outfit statistics, and others [6, 34, 35]. The person/item separation index and person/item separation reliability were expected to be greater than 0.2 and 0.7, respectively [36]. As for the SE, the smaller the value the better it is [36]. Ideally, infit and outfit statistics should fall between 0 and 1.5, indicating that the item functions as intended within this range [37]. Additionally, the point-measure correlation (PTMEASUR-AL) was calculated to determine the correlation between the item and the entire set of items for all observations. Generally, there is a close and positive relationship between the model's predicted expected value (EXP) and correlation value (CORR) [38]. Furthermore, the Rasch analysis helped identify any potential differential item functioning (DIF), which is used to assess the fairness of the items. DIF techniques evaluate whether different respondent subgroups have varying probabilities of responding to items in different ways [19, 39]. The Rasch-Welch, Mantel, and cumulative log-odd in logits (CUMLOR) statistics are commonly used to evaluate DIF. According to the guidelines provided on the Winsteps website, an item with a CUMLOR absolute value of less than 0.43 is considered to have a negligible difference in functioning. A value between 0.43 and 0.64 indicates a slight to moderate difference in functioning. A value above 0.64 indicates the presence of a moderate to large difference in functioning.

2.4 | Meta-dimensions and items' ranking

The meta-dimensions and items of the 13-MD were ranked to determine their relative importance. The same participants who completed the 13-MD also performed the ranking tasks within the same survey. They were given 100 points and asked to distribute them among the 13 meta-dimensions. After that, they were asked to distribute another 100 points among the items of each meta-dimension. There was no intention to compare items not belonging to the same meta-dimension, as ranking 33 items would have been too cognitively demanding. If a respondent used all their points for a meta-dimension or item, leaving some with missing values, those missing values were replaced with zeros, assuming empty scores did not carry any weight for the respondent. Standardized weights were calculated to account for respondents who gave values exceeding the initial 100 points. For the ranking of meta-dimensions, each meta-dimension's score was divided by the sum of all meta-dimensions scores and then multiplied by 100. The same process was applied to the ranking of items, where the items' scores were divided by the total score of the corresponding meta-dimension. This ensured that all weights fell within a range of 0–100.

Participants with incomplete or unusable responses were excluded. Unusable responses referred to respondents who did not seem to understand the task and instead provided comments rather than the required scores.

For all analyses, Winsteps Rasch 5.1.5.2 and Stata BE 17 (Stata Corp) were utilized.

3 | RESULTS

3.1 | Psychometric tests

The confirmatory analysis of the 13-MD was conducted using data collected through an online survey involving the general population of Quebec. 1337 French-speaking participants were invited to participate in this survey. During the screening process, 238 observations were excluded, due to incompleteness. The remaining 1099 observations were included for psychometric tests.

Descriptive statistics collected from the respondents can be found in Table 1.

Cronbach's α was 0.9524, which falls within the desired range. When we eliminated any of the items, none of them resulted in an increase in Cronbach α's value by 2% or more. Additionally, all items in the item-test, item-rest and interitem correlations displayed good and desired coefficients (see Table 2). Approximately 2.91% of participants always selected the best option available, while only 0.09% always chose the worst option for each response choice. Based on these results, we can conclude that the 13-MD does not suffer from floor or ceiling effects, as the percentages are significantly below the previously mentioned cut-off point of 15%.

PCA was conducted to validate the pattern described by the correlation analysis. Before conducting PCA, the KMO test and Barlett test of sphericity were performed. The KMO test indicated a sampling adequacy of 0.95 (i.e., the higher the value, the better it is). The Barlett test revealed a significant χ^2 value of 22257.455 ($p < 0.001$) indicating sufficient intercorrelations in our data to proceed with factor analysis. To determine the components retained in the varimax rotation, an unrotated PCA was carried out. Six components were retained and accounted for 64.27% of the total variance. The contribution of each item to the explained variance of each component was ascertained using PCA with varimax rotation. All items appeared to have coefficients greater than 0.2. It was observed that almost all items belonging to the same meta-dimension were in the same component (see Table 3).

Regarding the CFA, the estimated SEM showed good statistics, very close to those observed in our previous results using a different data set [12]. The statistics obtained in Model 2 (compared to those of our previous analysis named Model 1) are presented in Table 4. In both models, the goodness of fit of the 13-MD is confirmed. All variables in our model also appeared to be statistically significant ($p < 0.05$).

Using the IRT model with graded responses and Rasch analysis, the model fit was tested. When evaluating the discriminant capacity of the items, all discriminant coefficients showed signs that were of high value (>1) and significant. Rasch analysis yielded similar results (see Table 5). The item separation reliability (0.99), item separation index (12.46), person separation reliability (0.87), and person separation index (2.57) were all satisfactory. Additionally, the SE was extremely low at 0.2.

One item, specifically "Anxious or stressed," appeared to be associated with moderate to large DIF. However, its coefficient did not exhibit any significant issues in the end (see Table 6).

TABLE 1 Descriptive statistics of the participants.

Characteristics	Participants $n = 1099$ (%)
Median age (year) [range]	56 [19–94]
BMI (kg/m^2) [range]	25.79 [7.57–94.33]
Gender	
Women	558 (50.77)
Men	540 (49.14)
Intersex/other	1 (0.09)
Status	
Single	317 (28.84)
Married/spouse	635 (57.78)
Separated/divorced	99 (9.01)
Widowed	48 (4.37)
Occupation	
At home	50 (4.55)
Employed or independent worker	534 (48.59)
Retired	431 (39.22)
Sick leave/maternity	21 (1.91)
Student	39 (3.55)
Unemployed	24 (2.18)
School termination	
Pursued study after minimum age (16 years old)	891 (81.07)
Got a diploma/certification	922 (83.89)
Education (highest level)	
Primary school	29 (2.64)
Secondary school (high school)	240 (21.84)
Diploma of professional study	146 (13.28)
College (CEGEP)	311 (28.30)
University certificate	94 (8.55)
Bachelor's degree	187 (17.02)
Master	76 (6.92)
PhD	16 (1.45)
Household income	
<$5000	31 (2.82)
$5000–9999	7 (0.64)
$10,000–14,999	10 (0.91)
$15,000–19,999	23 (2.10)
$20,000–24,999	94 (8.56)
$25,000–34,999	85 (7.74)
$35,000–44,999	96 (8.74)

(Continues)

TABLE 1 (Continued)

Characteristics	Participants $n = 1099$ (%)
$45,000–54,999	96 (8.74)
$55,000–64,999	128 (11.66)
$65,000–74,999	95 (8.65)
$75,000–84,999	85 (7.74)
$85,000–99,999	106 (9.66)
$100,000–119,999	78 (7.10)
$120,000–149,999	96 (8.74)
>$150,000	68 (6.20)
Dwelling	
Urban	789 (71.19)
Rural	310 (28.21)
Homeowner	716 (65.15)
Tenant	383 (34.85)
Country of birth	
Canada	988 (89.90)
Other	111 (10.10)
Health problem	
Have been at least once confronted with a serious illness	280 (74.52)
Self-reported problems affecting HRQoL	333 (30.30)
Self-rated health state	
Excellent	135 (12.28)
Very good	431 (39.22)
Good	402 (36.58)
Fair	114 (10.37)
Bad	17 (1.55)

The excellent results obtained confirmed the goodness of fit of our questionnaire. No item revealed any unwanted score during the validation process. It is thus confirmed that the 13-MD is a valid tool designed to measure what it is intended to.

3.2 | Meta-dimensions and items' ranking

In the psychometric validation, 1099 respondents provided complete responses, with 15 respondents giving unusable responses. Therefore, the ranking procedure was based on the responses of the remaining 1084 participants. The ranking of meta-dimensions and the ranking of items within each meta-dimension are presented in Table 7.

Ranking scores were found to be evenly distributed among meta-dimensions indicating that each meta-dimensions holds significance. However, the meta-dimension "Well-being" received the highest score of 11.45 suggesting it is the most crucial aspect of health for the respondents. This was followed by "Sleep and energy" with a score of 8.86, and "Body functioning" with a score of 8.65. The items "Fulfilled," "Sleep," and "Breathing" were the top-ranked within their respective meta-dimensions. On the other hand, the meta-dimensions "Sexuality and intimacy," "Citizenship, inclusion, and autonomy," and "Depression, anxiety, and anger" were deemed as the least important, scoring at 6.50, 6.03, and 5.07, respectively.

4 | DISCUSSION

Instruments play a crucial role in both clinical practice and research, providing valuable insights into the impact of health conditions and interventions on a person's overall quality of life. Therefore, they should be as accurate as possible in measuring people's HRQoL in a global manner. Inaccurate evaluations of various interventions and health conditions may occur if important domains are missing from the measurement tool used [40]. This is why the 13-MD was developed, considering the imbalance in the structure of existing instruments regarding their composition, to allow for a complete measurement of GHRQoL [6, 12]. As shown by some studies, generic instruments often fail to capture important aspects of health, especially those related to specific diseases [6, 41–43]. The 13-MD offers a broader range of meta-dimensions to achieve more precision and specificity in GHRQoL measurement [12].

This study provided an analysis of the psychometric properties of the 13-MD's latest version using recent data from the general population of Quebec. The 13-MD demonstrated very satisfying results for all criteria considered. Both EFA and CFA confirmed the good structure of the 13-MD. The various psychometric tests performed also proved the 13-MD to measure what it is intended to. As our previous analysis suggested, the 13-MD is confirmed to be a valid instrument [12].

The 13-MD was developed to address the lack of coverage of health aspects in existing instruments [6, 12]. It aims to balance and depict the dimensions of health by addressing important aspects that can affect one's physical, mental, or social well-being. Compared to

TABLE 2 Detailed Cronbach's α results.

Items	Item-test correlation	Item-rest correlation	Interitem correlation	Cronbach's α
Breathing	0.5428	0.5075	0.3811	0.9517
Eating	0.5774	0.5439	0.3797	0.9514
Relieving oneself	0.5055	0.4684	0.3827	0.9520
Cognitive abilities	0.6381	0.6082	0.3771	0.9509
Senses	0.5224	0.4861	0.3820	0.9519
Language	0.5907	0.5580	0.3791	0.9513
Sleep	0.5990	0.5668	0.3788	0.9512
Energy	0.7144	0.6896	0.3739	0.9503
Confidence or self-esteem	0.6910	0.6645	0.3749	0.9505
Accepting myself	0.6998	0.6740	0.3745	0.9504
Pain	0.5744	0.5408	0.3798	0.9514
Discomfort	0.6988	0.6729	0.3746	0.9504
Performing strenuous activities	0.5526	0.5178	0.3807	0.9516
Performing moderate activities	0.6428	0.6132	0.3769	0.9509
Selfcare	0.5685	0.5346	0.3800	0.9515
Daily activities	0.6973	0.6713	0.3746	0.9504
Work- or school-related activities	0.6721	0.6444	0.3757	0.9506
Social	0.7239	0.6998	0.3735	0.9502
Leisure	0.7422	0.7194	0.3728	0.9500
Accepted and listened	0.6996	0.6737	0.3745	0.9504
Affection and support	0.6768	0.6494	0.3755	0.9506
Engaged in my role as a citizen	0.5124	0.4756	0.3824	0.9520
Integrated into society	0.6339	0.6038	0.3773	0.9510
Autonomous	0.6313	0.6010	0.3774	0.9510
Sad or depressed	0.7287	0.7048	0.3733	0.9502
Anxious or stressed	0.6956	0.6695	0.3747	0.9504
Angry or irritated	0.6206	0.5897	0.3778	0.9511
Fulfilled	0.6774	0.6500	0.3755	0.9506
Useful	0.7374	0.7142	0.3730	0.9501
Satisfied with life	0.7564	0.7346	0.3722	0.9499
Sex life	0.4821	0.4439	0.3837	0.9522
Intimacy	0.5248	0.4886	0.3819	0.9519
Sexual identity	0.4445	0.4046	0.3852	0.9525
Test scale			0.3775	0.9524

similar length instruments like the Assessment of Quality-of-Life 7D (AQoL-7D—26 items), the Assessment of Quality-of-Life 8D (AQoL-8D—35 items), the 15D (15 items), or the Quality of Well Being Self-Administered (QWB-SA—77 items), the 13-MD not only allows more possibilities (i.e., 1.42×10^{26} possible health states) but covers dimensions that are not included in those instruments such as affection and

TABLE 3 Items ponderation for PCA with varimax rotation.

Variable	Comp1	Comp2	Comp3	Comp4	Comp5	Comp6	Unexplained
Body functioning							
Breathing	−0.0099	0.1231	**0.3305**	−0.0188	−0.0086	−0.0475	0.5453
Eating	0.0746	−0.0446	**0.3133**	0.1932	−0.0996	−0.0684	0.4085
Relieving oneself	−0.0596	0.0075	**0.4104**	0.0209	0.0102	0.0857	0.5233
Cognitive abilities, senses, and language							
Cognitive abilities	0.0587	0.0301	**0.405**	−0.1133	0.057	0.0012	0.4091
Senses	−0.0632	−0.0061	**0.4869**	−0.0486	0.0703	0.0301	0.4288
Language	0.0386	−0.0737	**0.3607**	0.1644	0.0095	−0.0897	0.3679
Sleep and energy							
Sleep	**0.1978**	0.1841	0.1543	−0.2085	−0.0911	0.0123	0.4768
Energy	0.1782	**0.2258**	0.0767	−0.1004	−0.0439	0.0195	0.3903
Self-esteem and self-acceptance							
Confidence or self-esteem	**0.3694**	−0.0483	−0.0154	0.0124	−0.0433	−0.0449	0.3137
Accepting myself	**0.3504**	−0.055	−0.0292	0.1056	−0.0559	−0.0712	0.3181
Physical pain and discomfort							
Pain	0.0198	**0.4272**	0.0525	−0.1086	−0.1316	0.0837	0.326
Discomfort	0.0724	**0.3757**	0.0334	0.0268	−0.1424	0.023	0.2916
Mobility and physical disability							
Performing strenuous activities	−0.0918	**0.4222**	−0.0169	−0.0649	0.1136	0.0523	0.3302
Performing moderate activities	−0.1008	**0.3799**	−0.0352	0.1693	0.0827	−0.0165	0.2653
Selfcare	−0.0116	0.1164	−0.0319	**0.4841**	−0.082	−0.0835	0.3093
Daily activities and work							
Daily activities	0.0062	0.2779	−0.1017	**0.2785**	0.0547	−0.077	0.2932
Work- or school-related activities	0.0424	0.1584	−0.0303	**0.245**	0.0975	−0.1015	0.4304
Social and leisure activities							
Social	0.0808	0.1821	−0.0599	0.0725	**0.2072**	−0.0387	0.3881
Leisure	0.0536	**0.2182**	−0.0407	0.082	0.1964	−0.0392	0.35
Social and interpersonal relationships							
Accepted and listened	0.1778	−0.0712	0.0095	0.1224	**0.1842**	0.0178	0.4181
Affection and support	0.1012	−0.0934	0.0281	0.1257	**0.2406**	0.1535	0.4179
Citizenship, inclusion, and autonomy							
Engaged in my role as a citizen	−0.0695	0.0224	0.0307	−0.1356	**0.6136**	0.0035	0.3061
Integrated into society	0.03	−0.0816	0.0483	0.0941	**0.4362**	−0.0335	0.3833
Autonomous	0.0581	−0.0577	0.1458	**0.2767**	0.075	−0.0501	0.4436
Depression, anxiety, and anger							
Sad or depressed	**0.3787**	0.0193	−0.0302	−0.0816	−0.051	0.0301	0.2347
Anxious or stressed	**0.3631**	−0.0255	0.0412	−0.0591	−0.0699	−0.0005	0.3099
Angry or irritated	**0.3185**	0.019	−0.0065	0.0479	−0.1622	−0.0075	0.4697

TABLE 3 (Continued)

Variable	Comp1	Comp2	Comp3	Comp4	Comp5	Comp6	Unexplained
Well-being							
Fulfilled	**0.2678**	0.0011	−0.0587	−0.1099	0.1686	0.0847	0.3284
Useful	**0.2288**	−0.027	−0.0416	0.041	0.2166	−0.0189	0.3301
Satisfied with life	**0.2527**	0.0125	−0.0745	0.0023	0.1432	0.0983	0.2848
Sexuality and intimacy							
Sex life	−0.0088	0.0455	0.0113	−0.023	−0.0065	**0.6478**	0.1661
Intimacy	0.0063	−0.0094	−0.0076	0.0673	0.0086	**0.6359**	0.1676
Sexual identity	−0.0214	−0.1209	0.0198	**0.5096**	−0.1194	0.2505	0.3633

Note: Highest coefficients in bold.

TABLE 4 Statistics resulting from the CFA.

	Models	
Statistical test	**Model 1**[a]	**Model 2**[b]
RMSEA	0.059	0.062
SRMR	0.068	0.059
CFI	0.914	0.919
TLI	0.891	0.898
AIC	235072.414	95983.04
BIC	236086.421	96868.42
n	2273	1099

Abbreviations: CFI, comparative fit index; RMSEA, root mean square error of approximation; SRMR, standardized root mean square residual; TLI, Tucker and Lewis Index.

[a]Model 1 is the estimated model from our previous analyses [12].

[b]Model 2 is the actual estimated model using specific data collected when the actual structure of the 13-MD was presented to respondents.

support, self-acceptance and role in society/citizenship [6, 12]. This further demonstrates the 13-MD's potential as a more complete questionnaire.

However, the 13-MD will always be subject to review and improvement throughout its development and use. In fact, its widespread adoption by the research community may uncover limitations and areas for improvement that were not initially recognized during its creation. For instance, the EuroQol Group introduced the EQ-5D-5L, an enhanced version of the EQ-5D-3L, two decades after its inception [44, 45]. The previous version was criticized for its lack of sensitivity and issues with ceiling effects [44, 45]. This illustrates that refining and perfecting an instrument can be a lengthy process. With its 13 meta-dimensions and 33 items, the 13-MD appears to be a lengthy instrument [12]. Therefore, being able to rank its items in order of importance is a great asset for further improvement. In fact, we are considering the development of a shorter version that still maintains the psychometric properties of the full version. In this context, in addition to its psychometric properties assessment, the 13-MD also underwent a ranking procedure to determine which meta-dimensions/items were most prominent to the respondents. The "Well-being" meta-dimension received the highest number of points out of 100. Within this meta-dimension, being "Satisfied with life" and "Fulfilled" were the most important items. This result aligns with the reasoning behind the development of the 13-MD, which emphasizes that social and mental aspects of health are as important as the physical aspect and should be adequately represented in instruments measuring GHRQoL [6, 46, 47].

The ranking exercise conducted in this study will provide some insights for a potentially shortening of the 13-MD, allowing for quick and effective responses from the respondents when necessary. This can be achieved by identifying key meta-dimensions/items, testing them, and determining which items should be included in the reduced instrument [48]. Additionally, it is important to ensure that such a reduced version will be a valid substitute for the full one by testing its discriminative abilities and comparing them to the full instrument [48]. Therefore, the developers must determine if the benefits of reducing respondent burden (i.e., instrument's length) outweigh the potential drawbacks of using a less complete instrument if measurement properties are compromised [48].

This study may have some limitations related to information gathering and the study sample itself. First, the responses of many participants, nearly 18% of the sample, were not considered in this study as they were incomplete or unusable. This reduced the number of observations by 238 but ensured to have the best quality

TABLE 5 Rasch analysis results.

Item	MODEL S.E.	INFIT		OUTFIT		PTMEASUR-AL		Discrimination
		MNSQ	ZSTD	MNSQ	ZSTD	CORR.	EXP.	
Breathing	0.04	1.24	3.61	1.48	5.59	0.41	0.44	1.32
Eating	0.05	1.60	7.60	1.17	1.91	0.39	0.39	1.77
Relieving oneself	0.04	1.64	8.53	1.28	3.28	0.38	0.42	1.27
Cognitive abilities	0.04	1.03	0.49	0.85	−2.10	0.49	0.45	1.76
Senses	0.04	1.25	3.98	1.29	3.80	0.42	0.46	1.19
Language	0.05	1.32	4.26	0.95	−0.53	0.40	0.38	1.78
Sleep	0.03	1.04	0.71	1.15	2.45	0.53	0.55	1.48
Energy	0.03	0.71	−6.69	0.79	−4.14	0.62	0.57	2.00
Confidence or self-esteem	0.03	1.17	3.05	0.92	−1.36	0.53	0.51	2.28
Accepting myself	0.03	1.13	2.24	0.89	−1.75	0.53	0.50	2.37
Pain	0.03	0.90	−2.18	1.11	2.02	0.53	0.57	1.28
Discomfort	0.03	0.72	−6.10	0.77	−4.30	0.59	0.55	1.80
Performing strenuous activities	0.03	1.45	8.45	1.50	8.25	0.51	0.60	1.21
Performing moderate activities	0.03	1.22	3.81	1.07	1.11	0.50	0.50	1.65
Selfcare	0.05	1.73	8.55	1.03	0.39	0.38	0.37	1.79
Daily activities	0.04	1.35	4.97	0.82	−2.44	0.48	0.42	2.51
Work- or school-related activities	0.04	1.54	7.67	1.01	0.12	0.47	0.45	2.33
Social	0.03	0.86	−2.73	0.84	−2.90	0.59	0.53	2.19
Leisure	0.03	0.76	−5.16	0.79	−3.83	0.60	0.54	2.19
Accepted and listened	0.04	0.65	−6.79	0.59	−6.95	0.55	0.47	2.32
Affection and support	0.03	0.59	−8.69	0.63	−6.60	0.57	0.50	2.01
Engaged in my role as a citizen	0.03	1.32	6.04	1.84	9.90	0.47	0.58	1.18
Integrated into society	0.03	1.00	−0.06	1.07	1.15	0.51	0.51	1.75
Autonomous	0.04	1.10	1.64	0.89	−1.53	0.47	0.43	1.89
Sad or depressed	0.03	0.57	−9.90	0.64	−6.91	0.61	0.54	2.33
Anxious or stressed	0.03	0.71	−6.66	0.76	−4.76	0.61	0.57	2.05
Angry or irritated	0.03	0.67	−7.35	0.85	−2.7	0.57	0.55	1.64
Fulfilled	0.03	0.75	−5.72	0.9	−1.85	0.62	0.59	2.20
Useful	0.03	0.69	−7.05	0.72	−5.67	0.63	0.57	2.57
Satisfied with life	0.03	0.56	−9.90	0.58	−9.22	0.66	0.58	2.68
Sex life	0.02	1.70	9.90	1.74	9.90	0.52	0.64	1.12
Intimacy	0.03	1.91	9.90	1.75	9.90	0.52	0.62	1.28
Sexual identity	0.04	2.09	9.90	1.79	7.97	0.33	0.41	1.14
MEAN	0.03	1.12	0.62	1.04	−0.05	–	–	–
P.SD	0.01	0.41	6.46	0.35	5.12	–	–	–

TABLE 6 Differential item functioning (DIF) analysis.

Items	Rasch-Welch		Mantel	Size	
	t	Prob.	χ^2	Prob.	CUMLOR
Breathing	−2.76	0.0059	5.4825	0.0192	−0.35
Eating	−1.31	0.1919	0.0024	0.9606	−0.01
Relieving oneself	−3.04	0.0024	3.2172	0.0729	−0.30
Cognitive abilities	0.93	0.3518	2.2253	0.1358	0.23
Senses	−3.14	0.0017	4.2784	0.0386	−0.30
Language	−1.54	0.1233	0.5103	0.475	−0.13
Sleep	0.97	0.3343	0.9133	0.3392	0.13
Energy	1.67	0.0957	1.9839	0.159	0.18
Confidence or self-esteem	3.80	0.0002	5.6046	0.0179	0.35
Accepting myself	4.56	0	12.5099	0.0004	0.54
Pain	−2.98	0.0029	6.0725	0.0137	−0.30
Discomfort	−1.92	0.0556	4.1613	0.0414	−0.27
Performing strenuous activities	−0.55	0.5853	0.1319	0.7165	−0.05
Performing moderate activities	−0.51	0.6133	0.1277	0.7209	−0.05
Selfcare	−0.92	0.3561	0.7744	0.3788	−0.18
Daily activities	0.75	0.4532	0.0261	0.8717	0.03
Work- or school-related activities	0	1	0.2684	0.6044	0.10
Social	1.59	0.111	1.6814	0.1947	0.18
Leisure	0	1	0.4617	0.4968	0.09
Accepted and listened	1.49	0.1369	1.1683	0.2798	0.16
Affection and support	−0.35	0.7254	0.5429	0.4612	−0.10
Engaged in my role as a citizen	0	1	0.0255	0.873	0.02
Integrated into society	−0.33	0.7428	0.3309	0.5651	−0.08
Autonomous	−1.44	0.149	0.4447	0.5048	−0.1
Sad or depressed	3.98	0.0001	17.6867	0	0.57
Anxious or stressed	4.95	0	33.1443	0	0.75
Angry or irritated	0	1	1.0139	0.314	0.12
Fulfilled	0.85	0.3982	0.3084	0.5786	0.07
Useful	0	1	0.0611	0.8047	−0.03
Satisfied with life	0.93	0.3534	0.4609	0.4972	0.09
Sex life	−4.75	0	15.8245	0.0001	−0.50
Intimacy	0	1	0.2129	0.6445	−0.06
Sexual identity	−3.50	0.0005	3.554	0.0594	−0.36

data to perform analyses and provide reliable results. Second, the 13-MD was developed and formatted in a way that would assist the respondent in filling out the questionnaire effectively. The final version includes 13 tables that reflect the 13 meta-dimensions of the questionnaire. In that way, the official format of the 13-MD would offer more clarity and better comprehension to the respondents while minimising filling errors.

TABLE 7 Items and meta-dimensions of the 13-MD ranking.

Meta-dimension/item	Ranking	Score/100
Body functioning	***3***	***8.65***
Breathing	1	38.30
Eating	2	33.71
Relieving oneself	3	27.99
Cognitive abilities, senses, and language	***5***	***8.36***
Cognitive abilities	1	36.15
Senses	2	32.66
Language	3	31.19
Sleep and energy	***2***	***8.86***
Sleep	1	51.02
Energy	2	48.98
Self-esteem and self-acceptance	***4***	***8.50***
Confidence or self-esteem	1	51.86
Accepting myself	2	48.14
Physical pain and discomfort	***10***	***6.93***
Pain	1	52.90
Discomfort	2	47.10
Mobility and physical disability	***8***	***7.27***
Performing strenuous activities	3	27.01
Performing moderate activities	2	35.00
Selfcare	1	38.00
Daily activities and work	***9***	***7.34***
Daily activities	1	57.87
Work- or school-related activities	2	42.13
Social and leisure activities	***7***	***7.35***
Social	2	48.22
Leisure	1	51.78
Social and interpersonal relationships	***6***	***7.69***
Accepted and listened	2	49.36
Affection and support	1	50.64
Citizenship, inclusion, and autonomy	***12***	***6.03***
Engaged in my role as a citizen	3	27.72
Integrated into society	2	28.23
Autonomous	1	44.05
Depression, anxiety, and anger	***13***	***5.07***
Sad or depressed	2	33.71
Anxious or stressed	1	35.97
Angry or irritated	3	30.33
Well-being	***1***	***11.45***
Fulfilled	2	32.65
Useful	3	31.09
Satisfied with life	1	36.25
Sexuality and intimacy	***11***	***6.50***
Sex life	3	30.89
Intimacy	2	32.13
Sexual identity	1	36.97

Note: Information related to meta-dimensions are in bold and italic. Ranking was first conducted among meta-dimensions and then within each meta-dimension.

However, it was not possible to reproduce this format in the administered online survey due to technical issues. This difference in format between the original version of the 13-MD and the one used in the survey may possibly cause some biases that were not observed in this present study. Third, since the 13-MD is expected to be used in various contexts, it would be beneficial to gather data from a more linguistically and culturally diverse sample to test its cross-cultural validity [49, 50]. However, we specifically targeted French Canadian speaking respondents because the 13-MD was initially developed in this language. Now that it has been translated and validated in Canadian English, we can expand the target population and gather more data to assess the cross-cultural aspect [13]. Fourth, the study design did not allow for the assessment of another aspect of the instrument's reliability, known as results reproducibility. The current version of the 13-MD was only administered in one survey. If we were able to submit it to the same group twice, in different time and under the same conditions (i.e., test–retest), we could further confirm the instrument's reliability [51]. However, the scope of the study and the sample size did not allow the performance of these tests.

With its psychometric properties confirmed, the 13-MD is now ready to undergo its final development phase which is the valuation stage. This stage will allow for the development of an algorithm or value set that can convert any health state into a utility value/score. Once completed, the 13-MD will be among the instruments that can be used for assessing the Q in QALY. The ranking procedure outlined in this study will play a crucial role in this final step. A combination of the items will be arranged based on their relative rank to create choice cards. These cards will be presented to participants to elicit their

preferences using a best–worst scaling exercise [6, 52]. Value sets for both English and French versions of the 13-MD are expected to be created. This final step in the instrument development will make the 13-MD fully operational for use in CUA.

5 | CONCLUSION

The 13-MD is a well-balanced instrument designed to measure the GHRQoL. In this study, we used the latest version of the 13-MD with recent data to confirm the instrument's validity and reliability. Additionally, we conducted a meta-dimensions/items ranking exercise to determine which aspects are the most important for respondents. This valuable information will help in improving the 13-MD and guide further steps, such as developing a shorter version of the instrument and establishing a value set for converting its health states into utility scores.

AUTHOR CONTRIBUTIONS

Moustapha Touré: Data curation (equal); formal analysis (equal); methodology (equal); validation (equal); writing—original draft (equal); writing—review and editing (equal). **Thomas G. Poder**: Conceptualization (equal); formal analysis (equal); funding acquisition (equal); investigation (equal); methodology (equal); project administration (equal); supervision (equal); validation (equal); visualization (equal); writing—original draft (equal); writing—review and editing (equal).

ACKNOWLEDGMENTS

We thank all participants and the Fondation de l'IUSMM. T. G. P., which is member of the FRQS-funded CR-IUSMM. This study was funded by a grant from Unité Soutien SSA du Québec.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

Data are available upon reasonable request as per ethics committee requirement.

ETHICS STATEMENT

This study was approved by the ethics committee of the CIUSSS de l'Est de l'île de Montréal (#2021-2389).

INFORMED CONSENT

All participants provided informed consent to participate in the study. Information was collected anonymously. Participants consented to publication.

ORCID

Moustapha Touré https://orcid.org/0000-0002-7640-0356

Thomas G. Poder http://orcid.org/0000-0001-7017-096X

REFERENCES

1. Le Pen C, Lévy P. L'évaluation Médico-Économique—Concepts et Méthodes. Available from: https://www.ces-asso.org/sites/default/files/levaluation_medico-economique_levy_et_le_pen.pdf (accessed 2 September 2020).
2. Touré M, Poder TG. Évaluation économique en santé mentale: concepts et enjeux 2020. Int J Health Prefer Res. 2020;1:12–24. https://doi.org/10.21965/IJHPR.2020.002
3. Kind P, Lafata JE, Matuszewski K, Raisch D. The use of QALYs in clinical and patient decision-making: issues and prospects. Value Health. 2009;12(Suppl 1):S27–30. https://doi.org/10.1111/j.1524-4733.2009.00519.x
4. Brazier JE, Yang Y, Tsuchiya A, Rowen DL. A review of studies mapping (or cross walking) non-preference based measures of health to generic preference-based measures. Eur J Health Econ. 2010;11(2):215–25. https://doi.org/10.1007/s10198-009-0168-z
5. Olsen JA, Misajon R. A conceptual map of health-related quality of life dimensions: key lessons for a new instrument. Qual Life Res. 2020;29(3):733–43. https://doi.org/10.1007/s11136-019-02341-3
6. Touré M, Kouakou CRC, Poder TG. Dimensions used in instruments for QALY calculation: a systematic review. Int J Environ Res Public Health. 2021;18(9):4428. https://doi.org/10.3390/ijerph18094428
7. Revicki DA, Kleinman L, Cella D. A history of health-related quality of life outcomes in psychiatry. Dialogues Clin Neurosci. 2014;16(2):127–35. https://doi.org/10.31887/DCNS.2014.16.2/drevicki
8. The WHOQOL Group. The World Health Organization quality of life assessment (WHOQOL): position paper from the World Health Organization. Soc Sci Med. 1995;41(10):1403–9. https://doi.org/10.1016/0277-9536(95)00112-k
9. International Health Conference. Constitution of the World Health Organization. Bull World Health Organ 2002. 1946;80(12):983–984.
10. Vernooij-Dassen M, Jeon YH. Social health and dementia: the power of human capabilities. Int Psychogeriatr. 2016;28(5):701–3. https://doi.org/10.1017/s1041610216000260
11. Holt-Lunstad J, Smith TB, Layton JB. Social relationships and mortality risk: a meta-analytic review. PLoS Med. 2010;7(7):e1000316. https://doi.org/10.1371/journal.pmed.1000316
12. Touré M, Lesage A, Poder TG. Development of a balanced instrument to measure global health-related quality of life: the 13-MD. Front Psychiatry. 2022;13:837510. https://doi.org/10.3389/fpsyt.2022.837510
13. Poder TG, Touré M. Canadian English translation and linguistic validation of the 13-MD to measure global health-related quality of life. Expert Rev Pharm Out Res. 2024;24(2):267–72. https://doi.org/10.1080/14737167.2023.2268275
14. Tavakol M, Dennick R. Making sense of Cronbach's alpha. Int J Med Educ. 2011;2:53–5. https://doi.org/10.5116/ijme.4dfb.8dfd

15. Bédard SK, Larivière C, Poder TG. Processus de validation du questionnaire IPC65: un outil de mesure de l'interdisciplinarité en pratique clinique (Validation of the IPC65 questionnaire: A tool for measuring interdisciplinarity in clinical practice). Sante Publique. 2013;25(6):763–73. https://doi.org/10.3917/spub.136.0763
16. Bland JM, Altman DG. Statistics notes: Cronbach's alpha. BMJ. 1997;314(7080):572. https://doi.org/10.1136/bmj.314.7080.572
17. Briggs SR, Cheek JM. The role of factor analysis in the development and evaluation of personality scales. J Pers. 1986;54(1):106–48. https://doi.org/10.1111/j.1467-6494.1986.tb00391.x
18. Zijlmans EAO, Tijmstra J, van der Ark LA, Sijtsma K. Item-score reliability in empirical-data sets and its relationship with other item indices. Educ Psychol Meas. 2018;78(6):998–1020. https://doi.org/10.1177/0013164417728358
19. Terwee CB, Bot SDM, de Boer MR, van der Windt DAWM, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol. 2007;60(1):34–42. https://doi.org/10.1016/j.jclinepi.2006.03.012
20. Izquierdo I, Olea J, Abad FJ. Exploratory factor analysis in validation studies: uses and recommendations. Psicothema. 2014;26(3):395–400. https://doi.org/10.7334/psicothema2013.349
21. Stapleton CD. Basic concepts in exploratory factor analysis (EFA) as a tool to evaluate score validity: a right-brained approach. 1997. Available from: https://www.researchgate.net/publication/234592298_Basic_Concepts_in_Exploratory_Factor_Analysis_EFA_as_a_Tool_To_Evaluate_Score_Validity_A_Right-Brained_Approach. Accessed 3 April 2420.
22. Ringnér M. What is principal component analysis? Nature Biotechnol. 2008;26:303–4. https://doi.org/10.1038/nbt0308-303
23. Abdi H. Factor rotations in factor analyses. Encyclopedia for research methods for the social sciences. In: Lewis-Beck M, Bryman A, Futing T, editors. Thousand Oaks, CA: Sage; 2003. p. 792–5. https://personal.utdallas.edu/~herve/Abdi-rotations-pretty.pdf
24. Brazier JE, Mulhern BJ, Bjorner JB, Gandek B, Rowen D, Alonso J, et al. Developing a new version of the SF-6D health state classification system from the SF-36v2:SF-6Dv2. Med Care. 2020;58(6):557–65. https://doi.org/10.1097/mlr.0000000000001325
25. Brown TA, Moore MT. Confirmatory factor analysis. In: Hoyle RH, editor. Handbook of structural equation modeling. New York, NY: Guilford Publications; 2012. p. 361–79.
26. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Model. 1999;6(1):1–55. https://doi.org/10.1080/10705519909540118
27. Kline RB. Principles and practice of structural equation modeling. 4th ed. New York, NY: Guilford Publications; 2005.
28. Tucker LR, Lewis C. A reliability coefficient for maximum likelihood factor analysis. Psychometrika. 1973;38(1):1–10. https://doi.org/10.1007/BF02291170
29. Fabozzi FJ, Focardi SM, Rachev ST, Arshanapalli BG. Appendix E: model selection criterion: AIC and BIC. in the basics of financial econometrics. Hoboken, New Jersey: John Wiley & Sons, Ltd; 2014. p. 399–403. https://doi.org/10.1002/9781118856406.app5
30. Hambleton RK, Swaminathan H, Rogers HJ. Fundamentals of item response theory; measurement methods for the social sciences series. Vol 2. Newbury Park, Calif: Sage Publications; 1991.
31. Nguyen TH, Han HR, Kim MT, Chan KS. An introduction to item response theory for patient-reported outcome measurement. Patient. 2014;7(1):23–35. https://doi.org/10.1007/s40271-013-0041-0
32. Zanon C, Hutz CS, Yoo H, Hambleton RK. An application of item response theory to psychological test development. Psicologia: Reflexão e Crítica. 2016;29(1):18. https://doi.org/10.1186/s41155-016-0040-x
33. Yang FM, Kao ST. Item response theory for measurement validity. Shanghai Arch Psychiatry. 2014;26(3):171–7. https://doi.org/10.3969/j.issn.1002-0829.2014.03.010
34. Tennant A, McKenna SP, Hagell P. Application of Rasch analysis in the development and application of quality of life instruments. Value Health. 2004;7(Suppl 1):S22–6. https://doi.org/10.1111/j.1524-4733.2004.7s106.x
35. An X, Yung Y. Item response theory: what it is and how you can use the IRT procedure to apply it. SAS Institute Inc. 2014;10(4)364–2014. https://support.sas.com/resources/papers/proceedings14/SAS364-2014.pdf
36. Ahorsu DK, Lin CY, Imani V, Saffari M, Griffiths MD, Pakpour AH. The fear of COVID-19 scale: development and initial validation. Int J Ment Health Addict. 2022;20(3):1537–45. https://doi.org/10.1007/s11469-020-00270-8
37. Che Musa NA, Mahmud Z, Baharun N. Exploring students' perceived and actual ability in solving statistical problems based on Rasch measurement tools. J Phys: Conf Ser. 2017;890:012096. https://doi.org/10.1088/1742-6596/890/1/012096
38. Yusup M. Using Rasch model for the development and validation of energy literacy assessment instrument for prospective physics teachers. J Phys: Conf Ser. 2021;1876(1):012056. https://doi.org/10.1088/1742-6596/1876/1/012056
39. Shanmugam SKS. Determining gender differential item functioning for mathematics in coeducational school culture. Malays J Learn Instr. 2018;15(2):83–109. https://doi.org/10.32890/mjli2018.15.2.4
40. Kuspinar A, Mayo NE. Do generic utility measures capture what is important to the quality of life of people with multiple sclerosis? Health Qual Life Outcomes. 2013;11:71. https://doi.org/10.1186/1477-7525-11-71
41. Kuspinar A, Mate K, Lafontaine AL, Mayo N. Evaluating the content validity of generic preference-based measures for use in Parkinson's disease. Parkinsonism Rel Disord. 2019;62:112–6. https://doi.org/10.1016/j.parkreldis.2019.01.014
42. Mayo NE, Aburub A, Brouillette MJ, Kuspinar A, Moriello C, Rodriguez AM, et al. In support of an individualized approach to assessing quality of life: comparison between Patient Generated Index and standardized measures across four health conditions. Qual Life Res. 2017;26(3):601–9. https://doi.org/10.1007/s11136-016-1480-6
43. Brazier J, Connell J, Papaioannou D, Mukuria C, Mulhern B, Peasgood T, et al. A systematic review, psychometric analysis and qualitative assessment of generic preference-based measures of health in mental health populations and the estimation of mapping functions from widely used specific

measures. Health Technol Assess. 2014;18(34): 1–188. https://doi.org/10.3310/hta18340
44. Thompson AJ, Turner AJ. A comparison of the EQ-5D-3L and EQ-5D-5L. Pharmacoeconomics. 2020;38(6):575–91. https://doi.org/10.1007/s40273-020-00893-8
45. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res. 2011;20(10):1727–36. https://doi.org/10.1007/s11136-011-9903-x
46. Breslow L. A quantitative approach to the World Health Organization definition of health: physical, mental and social well-being. Int J Epidemiol. 1972;1(4):347–55. https://doi.org/10.1093/ije/1.4.347
47. Ware JE. Standards for validating health measures: definition and content. J Chronic Dis. 1987;40(6):473–80. https://doi.org/10.1016/0021-9681(87)90003-8
48. Guyatt GH, Feeny DH, Patrick DL. Measuring health-related quality of life. Ann Intern Med. 1993;118(8):622–9. https://doi.org/10.7326/0003-4819-118-8-199304150-00009
49. Corless IB, Nicholas PK, Nokes KM. Issues in cross-cultural quality-of-life research. J Nurs Scholarsh. 2001;33(1):15–20. https://doi.org/10.1111/j.1547-5069.2001.00015.x
50. Robitail S, Ravens-Sieberer U, Simeoni MC, Rajmil L, Bruil J, Power M, et al. Testing the structural and cross-cultural validity of the KIDSCREEN-27 quality of life questionnaire. Qual Life Res. 2007;16(8):1335–45. https://doi.org/10.1007/s11136-007-9241-1
51. Berchtold A. Test–retest: agreement or reliability? Methodol Innov. 2016;9:2059799116672875. https://doi.org/10.1177/2059799116672875
52. Weernink MGM, Janus SIM, van Til JA, Raisch DW, van Manen JG, IJzerman MJ. A systematic review to identify the use of preference elicitation methods in healthcare decision making. Pharm Med. 2014;28(4):175–85. https://doi.org/10.1007/s40290-014-0059-1

How to cite this article: Touré M, Poder TG. Confirmatory analysis of the 13-MD and ranking of its meta-dimensions and items. Health Care Sci. 2024;3:249–63. https://doi.org/10.1002/hcs2.106