



OPEN ACCESS

EDITED BY

Marcela Agudelo-Botero,
National Autonomous University of Mexico,
Mexico

REVIEWED BY

Soo Jin Kang,
Daegu University, Republic of Korea
Ervin Toci,
University of Medicine, Tirana, Albania

*CORRESPONDENCE

Yanhong Wang
✉ wyhong826@pumc.edu.cn

RECEIVED 19 January 2023

ACCEPTED 10 April 2023

PUBLISHED 17 May 2023

CITATION

Wang Y, Jia Q, Wang H, Zou K, Li L, Yu B,
Wang L and Wang Y (2023) Revised Chinese
resident health literacy scale for the older
adults in China: simplified version and initial
validity testing.

Front. Public Health 11:1147862.

doi: 10.3389/fpubh.2023.1147862

COPYRIGHT

© 2023 Wang, Jia, Wang, Zou, Li, Yu, Wang and
Wang. This is an open-access article distributed
under the terms of the [Creative Commons
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,
distribution or reproduction in other forums is
permitted, provided the original author(s) and
the copyright owner(s) are credited and that
the original publication in this journal is cited,
in accordance with accepted academic
practice. No use, distribution or reproduction is
permitted which does not comply with these
terms.

Revised Chinese resident health literacy scale for the older adults in China: simplified version and initial validity testing

Yilin Wang, Qiaoling Jia, Haiyan Wang, Kaiwen Zou, Lu Li,
Bing Yu, Li Wang and Yanhong Wang*

Department of Epidemiology and Biostatistics, Institute of Basic Medical Sciences, Academy of Medical Sciences and School of Basic Medicine Peking Union Medical College, Beijing, China

Objective: This study aimed to develop a short version of the Chinese Resident Health Literacy Scale focused on older adults in China, and further assess the reliability and validity of this short version.

Methods: The data was from a cross-sectional community-based older adults health survey conducted in 2020. The total of 5,829 older adults were randomly divided into two parts using for the simplification and assessment of the scale, respectively. Item Response Theory (IRT) and Differential Item Functioning (DIF) were used for item analysis and scale simplification. Cronbach's alpha and McDonald's omega were used to assess the reliability and three factors Confirmatory Factor Analysis (CFA) was used to assess the validity, which were compared to the original version. Moreover, Multi-group Confirmatory Factor Analysis (MCFA) was used to test the model invariance of the short version across groups of gender, age groups, level of education, and cognitive status.

Results: The simplified version consisted of 27 items taken from 50 original items, of them 11 items from the dimension of knowledge and attitudes, 9 items from the dimension of behavior and lifestyle, and 7 items from the dimension of health-related skills. The overall Cronbach's alpha and McDonald's omega were both 0.87 (95%CI: 0.86–0.88). The goodness-of-fits of CFA in simplified version were still acceptable in CFI, TLI, GFI, and RMSEA, even improved in CFI, TLI, and GFI compared to those of original version. Also, the model was stable and invariant in MCFA across gender, cognitive status, and educational level groups.

Conclusion: In this study, we formed a simplified instrument for measuring health literacy focused on older adults in China. This short version might be more suitable for the priority recommendation in extended tracking of the dynamic changes on the levels of health literacy in the whole life cycle in public health settings. Further research might be to identify the cut-off values to distinguish the older adults with different levels of health literacy.

KEYWORDS

health literacy, item response theory, differential item functioning, confirmatory factor analysis, older adults

Introduction

Among community-dwelling older adults, inadequate health literacy was independently linked to poorer physical and mental health (1). The previous studies found that lower health literacy was related to less understanding health information, a lack of basic knowledge of diseases, and/or poorer medication adherence (2), which could further increase the hospitalization and mortality rates (3, 4). Especially for those suffering from chronic diseases, over 40% of them took a grave risk to misunderstand, forget, or ignore healthcare advice (5). Therefore, promoting health literacy might be the most effective and affordable strategy for dealing with Non-communicable Chronic Disease (NCD) challenges (6, 7). In order to develop interventions to improve health literacy, it needed to give priority to measure the level of health literacy, especially among older adults, which was fundamental for the evaluation and surveillance.

Until now, over 150 instruments of health literacy have been developed, ranging from traditional tools focused on individual skill and health education to updated instruments from multidimensional perspectives (8–11). In China, health literacy also attracted more and more attention and some instruments were developed and applied in clinical or public health settings (12–17). Of these instruments, the Chinese Resident Health Literacy Scale (CRHLS) was developed on the manual of “Basic Knowledge and Skills of People’s Health Literacy” published by the National Health and Family Planning Commission and firstly released in 2008 (16). It was widely used in the National Health Literacy Surveillance among Chinese residents aged 15–69 years old (17, 18). According to data from National Health Literacy Surveillance, the health literacy levels (CRHLS scores of 80% or above) increased to 23.15% in 2020 and social development index, age, and education level were highly associated with health literacy (19). However, rare data of the health literacy for the older adults were released and in previous published study only 5.31–7.74% older adults aged 60–69 were found to reach this level (20).

CRHLS might be unsuitable for older adults in China. Firstly, the time consuming might be a huge challenge for older adults. Most of the young and middle-aged adults might spend about 30 min to complete measurement (16), while older adults might need to spend more time to do it, even longer for those with limited literacy (21). Secondly, some of items in CRHLS might be floor or ceiling effects for older adults. For example, the items of “reading and understanding OTC drug facts label” or “description of the liver function” might be beyond the scope of reading and understanding ability especially for those with lower level of education, or some items (e.g., item of “national unified toll-free hotline number”) deviated from the areas of concern for some older adults, which might led to the lower awareness rates. Lastly, CRHLS score 80% or more as the cutoff value with health literacy might be too strict for older adults and lack of evidence (22). Developing a brief version adapted from the original CRHLS for older adults might have several advantages, including brevity, availability of normative data, ease of administration, and lower cost. It would be more suitable for tracking the dynamic changes in the whole life cycle and evaluating the relationship between health literacy and health outcomes or health behaviors in the Chinese society context. Therefore, it was necessary to investigate how to simplify original scale reasonably and maintain the original reliability and validity as much as possible.

Item response theory (IRT) was a complex approach that attempted to explain the relationship between latent traits (unobservable characteristic or attribute) and their manifestations (i.e., observed outcomes, responses or performance) (23–31). Recently, more and more researchers used IRT models to analyze the latent properties of the scales and furtherly revise or simplify the scales (32). The advantage of using IRT was that each item in the scale was paid attention during the simplification of scale. It was assumed that the latent construct (e.g., knowledge, attitudes, *et al*) and items measured were organized in an unobservable continuum (24), which established a link between the properties of items on an instrument, individuals responding to these items and the underlying trait being measured (23).

Because that the level of education and cognitive status were the important factors related to the health literacy in previous studies (33, 34), and the health literacy level of older adults associated positively with the education level or cognitive function. The equivalence in the measurement of health literacy should be considered across groups of cognitive ability or education. The differential item functioning (DIF) to assess the equivalence of items in the scales was more and more applied in some studies of educational or health assessment (35). The measurement invariance (MI) was defined as statistical property of a measurement that the same underlying construct was measured across groups or across time (36). MI should be considered in the simplification of scale and lack of MI (meaning DIF) indicated that the populations with the same latent ability do not have equal probability of getting an item correct, regardless of group membership (37). DIF also examined the relationship between the item response and another group variable, after controlling for the underlying construct (38). When groups had different probabilities of response to a given item in DIF analysis, that indicated DIF occurred and this item was labeled as DIF-item (37).

Therefore, the purpose of this study was to simplified CRHLS based on IRT and DIF methods to form a more suitable instrument of health literacy for the Chinese older adults, and further assess the reliability and validity of simplified version.

Methods

Participants

The data was from a cross-sectional health survey carried out in Beijing among the community-based older adults in 2020, in which 5,829 participants were used for analysis in this study (39).

The simple random sampling method was used to select the 50% of total older adults as the sample for simplification (marked as Sample A) and the other 50% as the sample for validation (marked as Sample B). There were no differences in demographic and original version health literacy scores found between the two samples.

Measures

The survey collected information on sociodemographic characteristics (including age, gender, ethnicity, education level, marital status, and so on), health literacy, and cognition impairment.

Health literacy was measured by the CRHLS for the 2018 edition, which consisted of 50 items to assesses health knowledge, attitudes, behaviors and skills that are necessary to address real-world health problems and three dimensions: (1) knowledge and attitudes: basic knowledge and concepts related to health (22 items); (2) behavior and lifestyle: health-related behaviors and lifestyle (16 items); and (3) health-related skills: basic health-related skills (12 items). There are four types of questions in the scale: true-or-false; single-answer (only one correct answer in multiple-choice questions); multiple-answer (more than one correct answer in multiple-choice questions); and situational questions. With multiple-answer questions, a correct response had to contain all the correct answers and no wrong ones. Situational questions were given following a paragraph of instruction or medical information. Of these items, according to the scoring rules of this scale 2018 edition, total 50 items were included in the scores of health literacy with correct response allocated 1 point (except for the multiple-answer items in which correct response was allocated 2 points). The total score ranged from 0 to 66 points. Cronbach's α coefficient and the McDonald's omega of the Chinese Resident Health Literacy Scale were both 0.90 in our study.

Older adults might experience subtle cognitive changes associated with aging. The Ascertain Dementia 8 (AD8) questionnaire was used for screening cognitive impairment in this study (40). AD8 was a brief informant-based measure with only eight questions and performed well in distinguishing cognitive impairment from normal cognition (41). The person with an AD8 score ≥ 2 was suspected of having a cognitive impairment and needed further testing to be diagnosed. Cronbach's α coefficient of AD8 was 0.87 in our study.

Data analysis

The first phase of simplification in sample A

There were two stages included in this study. The first phase was used for simplification of the CRHLS in Sample A. The two-parameter logistic model (2PLM), one of the IRT models, was used to predict the probability of a successful answer for each item in each dimension (40). Difficulty and discrimination were two important parameters in 2PLM. The formula was as follows (42):

$$P(x_j = 1 | \theta) = \frac{1}{1 + \exp[-a_j(\theta - b_j)]}$$

Where x_j was the response to item j in the scale; a_j was the discrimination parameter; b_j was the difficulty parameter, which indicated the point where an individual would have a 50% chance of endorsing item j ; and θ is the ability value being measured. The basic assumptions of IRT included unidimensionality, local independence, and monotonicity (25–29), which were assessed, respectively, by the eigenvalue examination in factor analysis (30), Yen's Q3 statistic (29), splines modeling via flexible IRT models and Pearson χ^2 statistics (31) in this study.

The information about individuals was based on their responses to items and the properties of all the items (43). The maximum likelihood method was used for parameter estimations. According to the parameters of discrimination and difficulty, the information function for each item was calculated (44). Item information functions for the 2PLM would have their maximum value at the value of the

threshold (difficulty) (16). And slopes (discrimination) control how peaked the item information function was, and the higher the slope value was, the more information that item provided, the smaller the measurement error (45). In our study, the item response probability was chosen as 2/3 to calculate θ value in order to estimate the maximum amount of information for each item labeled as $I(\theta)$, which was one of the most popular standard setting methods of response probability in the Bookmark method (46–48).

Furthermore, in order to present how each dimension of scale was functioning as a whole, we also combine the item-level information functions to create a test information function, shown as test information curves (TIC). The standard errors were the square root of the inverse of information at a given level of ability. The item, with a discrimination parameter of 0.5 to 2.0 and a difficulty parameter of -3.0 to 3.0 , was considered to provide the most information (16), which was also used in this study. Furthermore, $I(\theta)$ was also the criteria of item selection in this study, that was: removed items ($I(\theta) \leq 0.10$), modified items ($0.10 < I(\theta) \leq 0.20$), and kept items ($I(\theta) > 0.20$) (49).

Moreover, Differential Item Functioning (DIF) was used to test the equivalence of items. We applied logistic regression for identification of DIF using the ability (θ) derived from IRT analysis as the matching variable. This method has been used in previous published studies (50–52). The group variables were cognitive status (Normal or Impairment) or the levels of education (Illiterate or Primary school, Junior high school, and Senior high school or above). If the determination coefficient (*value of p*) of DIF for a given item was significant (item with DIF), the item was labeled as DIF-item, which was considered to be removed from the scales (53).

In brief, if any of the following criteria was met in this phase, the item would be removed from original CRHLS: (a) discrimination (α) < 0.5 or > 2.0 (16); or (b) difficulty (b) < -3.0 or > 3.0 (16); or (c) $I(\theta) \leq 0.20$ (49); or (d) items exhibiting DIF on cognitive status or education level (54).

The second phase of validation in sample B

The second phase was validation of the simplified version as above in Sample B. Compared to the original version, the reliability, the construct validity, as well as Goodness-of-fit was estimated in the simplified version scale. Both Cronbach's alpha and McDonald's omega, as well as their 95%CI, were used to assess the reliability of the simplified version scale. The three factors Confirmatory Factor Analysis (CFA) was used to assess the convergent validity and discriminant validity. The goodness-of-fit was examined via structural equation modeling using the maximum likelihood method (55). The indicators of comparative fit index (CFI), Tucker-Lewis index (TLI), goodness-of-fit index (GFI), and root of the mean square residual (RMSEA) were calculated and compared between models. It was considered to be acceptable for the goodness-of-fit when CFI > 0.90 , TLI > 0.90 , GFI > 0.90 , and RMSEA < 0.08 (56).

In addition, Multi-group CFA (MCFA) was used to test the invariance of the simplified version across groups of gender, age, level of education, and cognitive status. The four models were used in MCFA analysis, that was: (1) configural invariance: the basic factor structure was the same across groups; (2) metric invariance: constraining factor loadings to be equal between groups based on the basic factor structure; (3) strong invariance: constraining the intercepts of the items to be the same between groups based on metric invariance; (4) strict invariance: further constraining residuals to be equal between groups (57). Invariance across subgroups is depicted by significant χ^2 , chi-square should not

significantly differ between models, along with $\Delta CFI < 0.02$, $\Delta TLI < 0.02$, and $\Delta RMSEA < 0.015$ (58).

IRT analyses were conducted using the R mirt package (59). The calculation of reliability was conducted using the R ltm and coefficientalpha packages (60). CFA and MCFA were performed in AMOS 17.0. Other analyses, including descriptive analysis, t-test, and chi-square test, were conducted using SAS 9.4. The significance level was 0.05 for all statistical tests.

Results

A total of 5,829 older adults were included in this study, 50% of male, mean age of 70.29 ± 7.37 years old, about 23% from rural regions, and 28% of education level of primary school or below. In total, 1888 participants (32.39%) might be cognitive impairment screened by AD8 (scores of $AD8 \geq 2$ points). There was no significant difference in the characteristics between Sample A and Sample B ($p > 0.05$; Table 1).

The simplification of Chinese resident health literacy scale

As shown in Table 2, the final 27 items kept in the simplified version of health literacy for older adults based on IRT and DIF consisted of 11 items in the knowledge and attitudes dimension, 9 items in the behavior and lifestyle dimension, and 7 items in the health-related skills dimension. A total of 23 items were removed, of which 21 items based IRT with low information or discrimination, and/or 4 DIF-items on cognitive status or education level (Table 2).

The reliability and validation of simplified version

Due to the 23 items removed, the overall Cronbach's alpha coefficient for simplified version was 0.87 (95%CI: 0.86–0.88), slightly lower than that of original version (0.90, 95%CI: 0.89–0.90), but still considered better reliability (Table 3). The sufficient reliabilities were also found in the dimension of knowledge and attitudes (0.72, 95%CI: 0.70–0.73) and behavior and lifestyle (0.73, 95%CI: 0.72–0.75), while undesirable reliability in dimension of health-related skills (0.65, 95%CI: 0.63–0.67). There was a similar McDonald's omega between simplified version and original version (Table 3). Also, the characteristics of items, the correlation to dimension scores, and factor loading in the simplified version in Sample B were detailed in Supplementary Table 1.

The results in Table 4 indicated the goodness-of-fits of CFA. The indicators of three-factor structures in simplified version were still acceptable in CFI, TLI, GFI, and RMSEA, even improved in CFI, TLI, and GFI compared to those of original version. The test information and measurement error curves were shown in Figure 1.

Testing for measurement invariance of simplified version

After then, Sample B was stratified by gender (male or female), age (60–70 years or ≥ 70 years or over), cognitive status (normal or

impairment), or education level (illiterate or primary school, junior school, or high school or above), then fitted into the model of the simplified version separately.

As shown in Table 5, the model of the simplified version is invariant and stable across groups of gender, age, cognitive status, and education level. The indexes of ΔCFI , ΔTLI , and $\Delta RMSEA$ indicated the equivalence of the three-factor structure of the simplified version between groups. However, χ^2 might be significant in model fit across groups due to large sample size (61).

Discussion

In this study, we first simplified the CRHLS (2018 edition) based on IRT and DIF methods. Finally, the simplified version of the remaining 27 items for older adults was adapted with a good reliability and even slightly improved in construct validation. The original scale, widely used in the periodic national survey to monitor the level and dynamic changes of health literacy in Chinese adults aged 15–69 years old, was developed on the manual of “Basic Knowledge and Skills of People's Health Literacy” published by the National Health and Family Planning Commission (18), which was still an important material of health education and health communication for the public in China now. Therefore, this simplified version adopted from CRHLS focused on older adults might be more suitable for extended tracking of the dynamic changes in the whole life cycle in the future, evaluating the intervention effect of health education in chronic diseases, and investigating the relationship to health outcomes or health behaviors under the Chinese context. Moreover, it was more convenient for collection and administration of data than the original CRHLS.

Recently, IRT was more and more applied to assess and revise the scales in public health or clinical studies (43). Different from the previous simplified studies just focused on the difficulty and discrimination when using IRT model (31, 62–64), our findings also integrated the equivalence and information of items into the criteria of remaining items in order to obtain a relative accuracy of measurement of health literacy for older adults. The previous study by Shen et al. reported the overall reliability of original version was 0.95 among the population aged 15–69 years old in China (16). In contrast, the reliability of original version in our study was found to be 0.90 for the older adults, but still a good reliability. It also showed that original CRHLS, although with some shortcomings, could be used to measure the level of health literacy for the old population. Compared to the original version, simplified version still had good overall reliability ($\alpha = 0.87$) and acceptable reliability in each dimension, which was slightly lower due to the great reduction of items. Moreover, according to the Goodness-of-fits in three factors models of CFA in our study, the simplified version seemed to be better construct validation and construct invariance across groups. All of these supported that the simplified version scale might be suitable for the priority recommendation as an instrument of health literacy for older adults in Chinese social context.

We simplified the Chinese instrument for measuring health literacy by removing some redundant or DIF items to finally form the simplified version with 27 items focused on older adults in China. However, there were some limitations. Firstly, IRT model needed to meet basic assumptions of unidimensionality, monotonicity and local independence. In this study, the data was verified to satisfy these basic assumptions only except for the

TABLE 1 Characteristics of study participants.

Demographic variables	Total (N=5,829)	Simplification (Sample A, N=2,915)	Validation (Sample B, N=2,914)	p-value
<i>Gender</i>				
Male	2,891 (49.6%)	1,448 (49.67%)	1,443 (49.52%)	0.91
Female	2,938 (50.4%)	1,467 (50.33%)	1,471 (50.48%)	
<i>Subgroups of age</i>				
60–69	2,925 (50.18%)	1,473 (50.53%)	1,452 (49.83%)	0.65
70–79	2,166 (37.16%)	1,077 (36.95%)	1,089 (37.37%)	
≥80 or over	738 (12.66%)	365 (12.52%)	373 (12.80%)	
<i>Ethnicity</i>				
Han	5,614 (96.3%)	2,811 (96.43%)	2,803 (96.19%)	0.63
Other	215 (3.7%)	104 (3.57%)	111 (3.81%)	
<i>Type of residence</i>				
Urban	4,505 (77.29%)	2,252 (77.26%)	2,253 (77.32%)	0.96
Rural	1,324 (22.71%)	663 (22.74%)	661 (22.68%)	
<i>Education</i>				
Illiterate or Primary school	1,625 (27.88%)	794 (27.24%)	831 (28.52%)	0.65
Junior high school	2,129 (36.52%)	1,056 (36.23%)	1,073 (36.82%)	
Senior high school or above	2,075 (35.60%)	1,065 (36.54%)	1,010 (34.66%)	
<i>Marital status</i>				
Married	4,856 (83.30%)	2,441 (83.74%)	2,415 (82.88%)	0.38
Unmarried, Divorced, or Widow	973 (16.69%)	474 (16.26%)	499 (17.12%)	
<i>Smoking condition</i>				
Never Smoking	3,855 (66.13%)	1,951 (66.93%)	1,904 (65.34%)	0.25
Smoking	1,155 (19.81%)	562 (19.28%)	593 (20.35%)	
Smoked, but quit now	819 (14.05%)	402 (13.79%)	417 (14.31%)	
<i>Medical insurance</i>				
Yes	5,721 (98.15%)	2,852 (97.84%)	2,869 (98.46%)	0.09
No	108 (18.34%)	63 (2.16%)	45 (1.54%)	
<i>Cognitive ability</i>				
Cognitive normal	3,941 (67.60%)	1,978 (67.86%)	1,963 (67.36%)	0.69
Cognitive impairment	1,888 (32.40%)	937 (32.14%)	951 (32.64%)	
Total scores of health literacy (Mean ± SD)	39.20 ± 12.80	39.31 ± 12.69	39.08 ± 12.91	0.49
<i>Scores of three dimensions (Mean ± SD)</i>				
Knowledge and attitudes	17.24 ± 5.57	17.27 ± 5.53	17.20 ± 5.61	0.63
Behavior and lifestyle	12.79 ± 4.90	12.83 ± 4.89	12.74 ± 4.91	0.51
Health-related skills	9.17 ± 3.61	9.22 ± 3.57	9.13 ± 3.64	0.34
<i>Health literacy*</i>				
Less than 80% of total scores	4,921 (84.4%)	2,460 (84.4%)	2,461 (84.50%)	0.97
80% of total scores or above	908 (15.6%)	455 (15.6%)	453 (15.50%)	

*The scores of 80% or more was the cutoff value with health literacy in CRHLS.

eigenvalue examination in the dimension of health-related skills (the ratio of the first to second eigenvalue was 2.52, which slightly less than 3 of threshold value). Secondly, because the original version was initially developed by an expert panel from the Ministry of Health and now was widely used for community-based adults aged

15–69 years old in China, we did not further assess the content validity in this study. Moreover, the items kept in short version were selected based on IRT and DIF. The additional items recommended by researchers might be subjective, so did not be considered in this study. Thirdly, we used the data of the representative sample of older

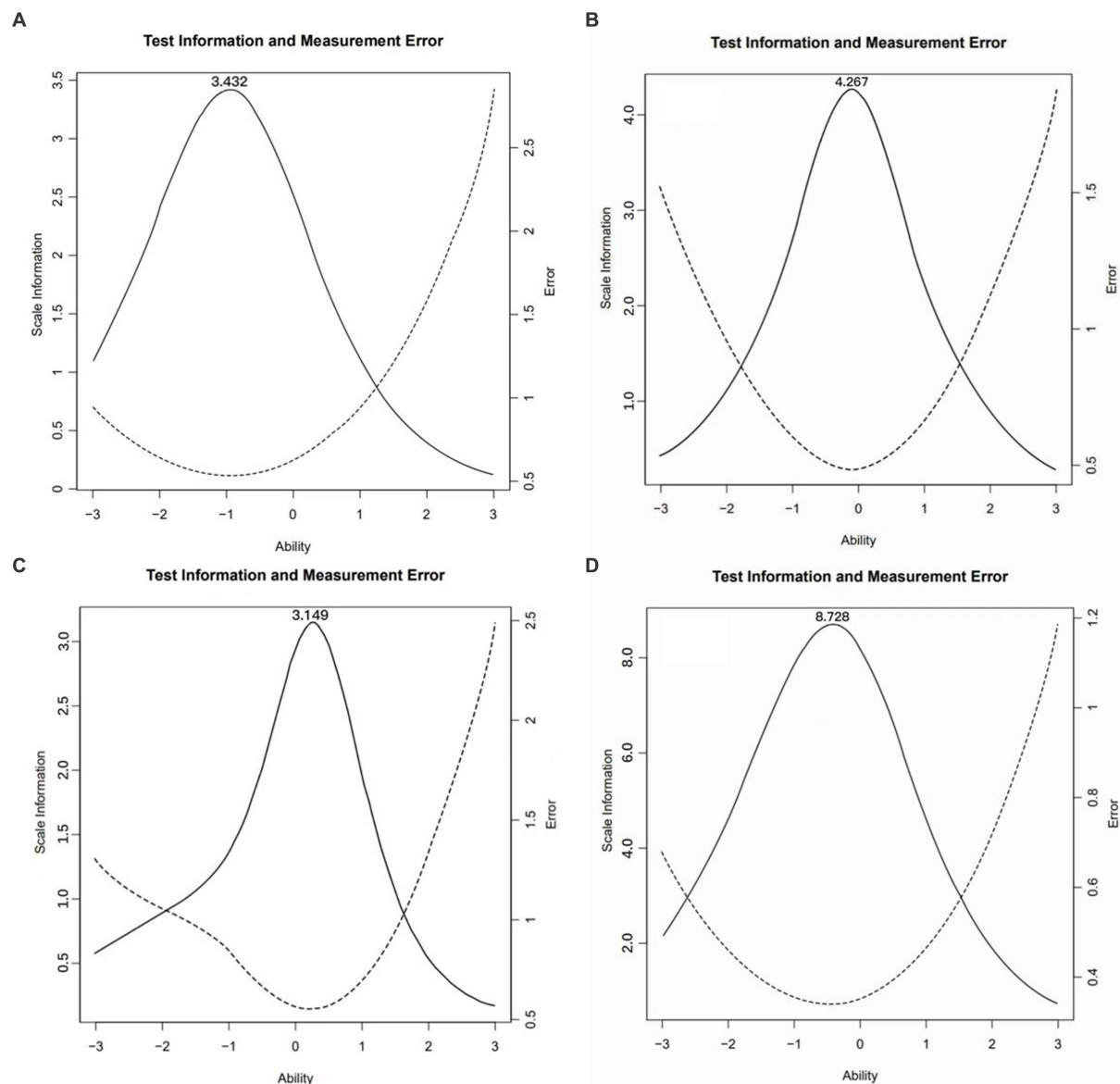


FIGURE 1
The test information and measurement error curves in simplified version. (A) Dimension of knowledge and attitudes; (B) Dimension of behavior and lifestyle; (C) Dimension of health-related skills; and (D) the total of simplified version.

adults from a cross-sectional study conducted in Beijing. There might be differences in characteristics among older adults living in different regions, which might have an impact on the reliability and validity of the scale when it was extended used in general Chinese older adults. So, this simplified version needs to be further assessed in the general populations.

Conclusion

In this study, we formed a simplified version instrument for measuring health literacy focused on older adults in China. It might be more suitable for the priority recommendation in extended tracking of the dynamic changes in the whole life cycle and assessing the level of health literacy among older adults in public health settings.

Data availability statement

The data analyzed in this study is subject to the following licenses/restrictions: The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation. Requests to access these datasets should be directed to YHW, wuhong826@pumc.edu.cn.

Ethics statement

The studies involving human participants were reviewed and approved by the institutional review board of the Institute of Basic Medical Sciences, Chinese Academy of Medical Sciences (Project No. 064–2020). The patients/participants provided their written informed consent to participate in this study.

TABLE 2 Evaluation of items based on IRT and results of DIF analysis of the original version of the Chinese Resident Health Literacy Scale in Sample A.

Item abbreviation		Correct (%)	Correlation to dimension score	IRT parameters			DIF						Kept
				Difficulty	Discrimination	Information level	Cognitive Status			Education Level			
							Statistic	<i>p</i>	<i>R</i> ²	Statistic	<i>p</i>	<i>R</i> ²	
Dimension1: Knowledge and attitudes													
A01	Prevention of the flu	79.35	0.43	-1.44	1.13	0.32	0.52	0.74	0.0001	0.65	0.52	0.0002	Yes
A02	Health care products	79.20	0.26	0.58	1.42	0.08	<0.001	0.99	<0.0001	5.20	0.07	0.0017	No
A03	Infusion	79.86	0.42	-1.43	1.08	0.29	0.22	0.76	0.0001	0.64	0.52	0.0002	Yes
A05	Health products	79.11	0.31	-2.26	0.63	0.10	<0.001	0.99	<0.0001	5.20	0.07	0.0017	No
A10	Body temperature	78.03	0.23	-3.57	0.37	0.03	7.90	0.04	0.0026	7.16	0.78	0.0002	No
B01	The definition of health	71.90	0.52	-0.92	1.43	0.51	0.20	0.76	<0.0001	4.63	0.08	0.0011	Yes
B03	Hepatitis B	56.30	0.38	-0.96	0.75	0.14	0.99	0.63	0.0003	3.87	0.05	0.0024	No
B04	Blood pressure measurement	48.75	0.38	-0.40	0.69	0.12	0.79	0.63	0.0002	0.03	0.78	0.0003	No
B06	Early symptoms of cancer	51.32	0.41	0.11	0.82	0.17	2.27	0.32	0.0006	5.71	0.90	<0.0001	No
B07	Management of gas poisoning	77.12	0.40	-1.55	0.93	0.21	4.91	0.08	0.0014	2.51	0.19	0.0007	Yes
B08	Tuberculosis treatment	61.68	0.57	-0.05	1.55	0.59	8.59	0.04	0.0019	4.38	0.10	0.0011	No
B09	Toxic and hazardous work	70.60	0.47	-0.96	1.16	0.34	6.48	0.05	0.0017	4.72	0.08	0.0013	Yes
B10	Harm of iodine deficiency	64.01	0.36	-0.79	0.63	0.10	0.09	0.80	<0.0001	0.27	0.87	<0.0001	No
B13	Vaccines for children	63.29	0.45	-0.69	0.93	0.19	0.37	0.76	0.0001	0.44	0.07	0.0016	No
B17	Meaning of warning diagram	69.47	0.51	-0.79	1.35	0.45	3.98	0.13	0.0010	5.25	0.07	0.0013	Yes
C02	Medical visits	70.31	0.45	-0.98	1.08	0.29	0.27	0.76	0.0001	6.20	0.07	0.0017	Yes
C03	Liver	14.30	0.45	-1.49	1.35	0.46	14.06	0.00	0.0039	7.69	0.08	0.0010	No
C06	Packaged food	37.29	0.28	2.75	0.73	0.13	0.89	0.63	0.0003	0.93	0.66	0.0001	No

(Continued)

TABLE 2 (Continued)

Item abbreviation		Correct (%)	Correlation to dimension score	IRT parameters			DIF						Kept
				Difficulty	Discrimination	Information level	Cognitive Status			Education Level			
							Statistic	<i>p</i>	<i>R</i> ²	Statistic	<i>p</i>	<i>R</i> ²	
C07	Treatment of sick and dead livestock	68.27	0.50	-0.70	1.22	0.37	1.04	0.64	0.0003	0.04	0.87	<0.0001	Yes
C15	Pesticide storage	58.44	0.49	-0.38	1.13	0.32	5.13	0.08	0.0013	8.44	0.06	0.0021	Yes
D03	Control weight	61.64	0.48	-0.55	1.07	0.28	0.14	0.78	<0.0001	0.77	0.52	0.0002	Yes
D04	Obesity-related disease	67.79	0.46	-0.86	1.07	0.27	6.27	0.05	0.0017	2.22	0.21	0.0006	Yes
Dimension2: Behavior and lifestyles													
A04	Fruits and vegetables	80.75	0.37	-2.17	0.70	0.12	0.35	0.74	0.0001	0.75	0.78	0.0001	No
A06	Depression	86.24	0.36	-2.33	0.88	0.19	4.80	0.23	0.0015	1.07	0.05	0.0022	No
A09	Chronic disease treatment	62.95	0.33	-1.37	0.41	0.04	1.42	0.64	0.0005	0.04	0.78	0.0001	No
B05	Dangers of smoking	52.66	0.50	-0.05	0.99	0.24	2.38	0.49	0.0006	1.46	0.73	0.0004	Yes
B11	Hydration	62.57	0.46	-0.63	0.84	0.17	0.08	0.89	<0.0001	0.49	0.78	0.0002	No
B12	National basic public health service	40.07	0.49	0.51	1.05	0.27	0.28	0.74	0.0001	2.93	0.67	0.0008	Yes
B14	Fever	77.43	0.4	-1.80	0.80	0.15	1.11	0.64	0.0003	0.23	0.90	<0.0001	No
B15	Adverse reactions	79.52	0.33	-2.30	0.62	0.09	0.03	0.89	<0.0001	0.49	0.26	0.0008	No
B19	Medical visits	71.18	0.62	-0.91	1.29	0.41	0.84	0.64	0.0002	0.15	0.80	<0.0001	Yes
B21	Opening windows for ventilation during flu season	70.05	0.45	-0.98	0.99	0.24	0.88	0.64	0.0002	0.53	0.78	0.0001	Yes
C01	Promoting mental health	52.25	0.58	0.03	1.92	0.90	1.34	0.64	<0.0001	1.86	0.69	<0.0001	Yes
C04	Fever and rash in children	56.09	0.6	-0.17	2.00	0.97	0.47	0.74	<0.0001	2.34	0.67	<0.0001	Yes
C09	Benefits of eating soy products	31.60	0.43	0.95	1.05	0.27	2.63	0.49	0.0007	0.24	0.78	0.0001	Yes

(Continued)

TABLE 2 (Continued)

Item abbreviation		Correct (%)	Correlation to dimension score	IRT parameters			DIF						Kept
				Difficulty	Discrimination	Information level	Cognitive Status			Education Level			
							Statistic	<i>p</i>	<i>R</i> ²	Statistic	<i>p</i>	<i>R</i> ²	
C10	Health benefits of exercise	45.11	0.54	0.23	1.62	0.63	0.02	0.89	<0.0001	0.01	0.96	<0.0001	Yes
C12	Coughing and sneezing	46.90	0.38	0.29	0.53	0.07	5.75	0.23	0.0018	5.89	0.26	0.0005	No
C13	Medical visits	67.75	0.52	-0.72	1.21	0.36	0.41	0.74	0.0001	0.25	0.78	0.0001	Yes
Dimension3: Health-related skills													
B16	Treatment of virulent infectious diseases	83.81	0.44	-1.63	1.24	0.36	0.03	0.93	<0.0001	0.14	0.26	0.0007	Yes
B18	Toll-free health hotline number	34.90	0.41	1.00	-0.75	0.24	2.63	0.27	0.0007	0.07	0.01	0.0033	No
B20	OTC	28.37	0.42	1.38	0.79	0.15	2.54	0.27	0.0007	0.24	0.42	0.0002	No
B22	Glass thermometer	62.74	0.44	-0.79	0.7	0.12	0.20	0.85	0.0001	0.28	0.49	0.0003	No
B25	Mild burns	73.21	0.45	-1.31	0.92	0.19	2.10	0.29	0.0006	0.05	0.10	0.0010	No
B26	Fire handling	46.14	0.4	0.41	0.55	0.07	0.01	0.93	<0.0001	0.37	0.24	0.0018	No
C08	Cardiac arrest	59.07	0.52	-0.31	1.09	0.28	4.24	0.16	0.0011	0.10	0.36	0.0004	Yes
C11	Hypoglycemic products	70.94	0.47	-1.03	0.99	0.23	7.51	0.07	0.0020	0.41	0.26	0.0006	Yes
C14	Benefits of breastfeeding for babies	44.94	0.52	0.24	1.13	0.30	0.34	0.84	0.0001	2.64	0.26	0.0006	Yes
C16	Lightning weather outdoors	82.40	0.42	-1.80	0.99	0.23	0.61	0.75	0.0002	0.04	0.42	0.0002	Yes
D01	Calculation of BMI	35.13	0.54	0.61	1.62	0.63	6.43	0.07	0.0013	0.05	0.14	0.0011	Yes
D02	Classification of BMI	42.50	0.56	0.28	1.69	0.69	0.14	0.85	<0.0001	1.22	0.26	<0.0001	Yes

*R*²: DIF magnitude, applying the cut-offs of: <0.13 = negligible; 0.13–0.26 = moderate; >0.26 = not negligible (53).

TABLE 3 Reliability of the original version and the simplified version of the Chinese Resident Health Literacy Scale in Sample B.

Dimension	Cronbach's alpha	95%CI	McDonald's omega	95%CI
Original version				
Overall	0.90	0.89–0.90	0.90	0.89–0.90
Three dimensions				
<i>Knowledge and attitudes</i>	0.78	0.77–0.80	0.79	0.78–0.80
<i>Behavior and Lifestyle</i>	0.74	0.73–0.75	0.74	0.73–0.76
<i>Health-related Skills</i>	0.70	0.68–0.71	0.70	0.68–0.72
Simplified version				
Overall	0.87	0.86–0.88	0.87	0.86–0.88
Three dimensions				
<i>Knowledge and attitudes</i>	0.72	0.70–0.73	0.72	0.70–0.73
<i>Behavior and Lifestyle</i>	0.73	0.72–0.75	0.74	0.72–0.75
<i>Health-related Skills</i>	0.65	0.63–0.67	0.65	0.62–0.67

TABLE 4 Goodness-of-fit of the CFA models of the original version and the simplified version of the Chinese resident health literacy scale in Sample B.

	χ^2	df	$p(\chi^2)$	CFI	TLI	GFI	RMSEA	RMSEA 95%CI
Original version	9420.34	1,171	<0.001	0.693	0.679	0.852	0.049	0.048–0.050
Original version (modified)*	8625.16	1,166	<0.001	0.742	0.728	0.866	0.041	0.040–0.043
Simplified version (modified)**	3099.31	319	<0.001	0.825	0.808	0.910	0.055	0.054–0.056

CFI, Comparative fit index; TLI, Tucker-Lewis index; GFI, Goodness of Fit Index; RMSEA, root mean square error of approximation.

*modified original version with five specified error covariances: A01 with A03; A09 with A10; B18 with B20; C06 with C10; D01 with D02.

**modified simplified version with two specified error covariances: A01 with A03; D01 with D02.

Author contributions

YHW and LW participated in the design of the study and organized the training of investigators. QJ, HW, YLW, and KZ participated in data collection and quality control. BY and LL participated in the literature research and collation. YLW and YHW drafted the manuscript, performed the statistical analysis, and revision of the manuscript. All authors contributed to the article and approved the submitted version.

Funding

The survey was funded by National Key R&D Program of China (2022YFC3601800), Beijing Municipal Health Commission and Beijing Health Economics Association. The funders had no role in the design of the study, analysis, and interpretation of data, or writing the manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Acknowledgments

The authors thank all the participants and staffs from primary health center involved in the survey. We appreciate the help and support from subdistrict offices and neighborhood committees.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpubh.2023.1147862/full#supplementary-material>

TABLE 5 Invariance of the simplified version of the Chinese resident health literacy scale across groups of gender, age, cognitive status, and educational levels in Sample B.

Groups	Invariance levels	χ^2	df	p-value	$\Delta\chi^2$	Δ df	$p(\Delta\chi^2)$	CFI	Δ CFI	TLI	Δ TLI	GFI	Δ GFI	RMSEA	Δ RMSEA
Gender															
	Male	1759.06	319	<0.001				0.813		0.794		0.902		0.056	
	Female	1719.63	319	<0.001				0.831		0.814		0.903		0.055	
	Configural	3478.70	638	<0.001				0.822		0.805		0.902		0.039	
	Metric	3505.46	662	<0.001	26.76	24	0.316	0.822	0.000	0.811	0.006	0.902	0.000	0.038	0.001
	Strong	3508.39	668	<0.001	2.93	6	0.818	0.822	0.000	0.813	0.002	0.902	0.000	0.038	0.000
	Strict	3539.81	697	<0.001	31.42	29	0.346	0.822	0.000	0.821	0.008	0.901	0.001	0.037	0.001
Age															
	60–69	1709.95	319	<0.001				0.824		0.806		0.902		0.055	
	≥70 or over	1719.84	319	<0.001				0.827		0.809		0.904		0.055	
	Configural	3429.78	638	<0.001				0.825		0.808		0.903		0.039	
	Metric	3452.84	662	<0.001	23.06	24	0.516	0.825	0.000	0.815	0.007	0.903	0.000	0.038	0.001
	Strong	3464.03	668	<0.001	11.19	6	0.083	0.825	0.000	0.816	0.001	0.903	0.000	0.038	0.000
	Strict	3508.25	697	<0.001	44.22	29	0.035	0.824	0.001	0.823	0.007	0.902	0.001	0.037	0.001
Cognitive status															
	Normal	2292.98	319	<0.001				0.810		0.791		0.905		0.056	
	Impairment	1285.07	319	<0.001				0.816		0.798		0.892		0.056	
	Configural	3578.11	638	<0.001				0.812		0.793		0.901		0.040	
	Metric	3631.58	662	<0.001	53.47	24	0.001	0.810	0.002	0.799	0.006	0.900	0.001	0.039	0.001
	Strong	3647.84	668	<0.001	16.26	6	0.012	0.810	0.000	0.800	0.001	0.899	0.001	0.039	0.000
	Strict	3884.22	697	<0.001	236.38	29	0.000	0.796	0.004	0.795	0.005	0.893	0.006	0.040	0.001
Education level															
	Education 1	1262.39	319	<0.001				0.820		0.802		0.882		0.060	
	Education 2	1379.87	319	<0.001				0.794		0.773		0.897		0.056	
	Education 3	1313.93	319	<0.001				0.809		0.789		0.900		0.056	
	Configural	4394.94	1,016	<0.001				0.783		0.775		0.883		0.034	
	Metric	4425.02	1,040	<0.001	30.08	24	0.182	0.783	0.000	0.780	0.005	0.882	0.001	0.033	0.001
	Strong	4430.57	1,046	<0.001	5.55	6	0.475	0.783	0.000	0.781	0.001	0.882	0.000	0.033	0.000
	Strict	4471.92	1,075	<0.001	41.35	29	0.064	0.782	0.001	0.787	0.006	0.881	0.001	0.033	0.000

χ^2 : chi-square; df: degrees of freedom; CFI, Comparative fit index; TLI: Tucker–Lewis index; GFI, Goodness of Fit Index; RMSEA, root mean square error of approximation.

Education 1: Illiterate or Primary school; Education 2: Junior school; Education 3: High school, College, or Graduate school.

References

- Wolf MJ, Gazmararian JA, Baker DW. Health literacy and functional Health status among older adults. *Arch Intern Med.* (2005) 165:1946–52. doi: 10.1001/archinte.165.17.1946
- Sheridan SL, Halpern DJ, Viera AJ, Berkman ND, Donahue KE, Crotty K. Interventions for individuals with low Health literacy: a systematic review. *J Health Commun.* (2011) 16:30–54. doi: 10.1080/10810730.2011.604391
- Baker DW, Wolf MS, Feinglass J, Thompson JA, Gazmararian JA, Huang J, et al. Health literacy and mortality among elderly persons. *Arch Intern Med.* (2007) 167:1503–09. doi: 10.1001/archinte.167.14.1503
- Baker DW, Gazmararian JA, Williams MV, Scott T, Parker RM, Green D, et al. Functional Health literacy and the risk of hospital admission among Medicare managed care enrollees. *Am J Public Health.* (2002) 92:1278–83. doi: 10.2105/ajph.92.8.1278
- Martin LR, Williams SL, Haskard KB, Dimatteo MR. The challenge of patient adherence. *Ther Clin Risk Manag.* (2005) 1:189–9.
- Liu C, Wang D, Liu C, Jiang J, Wang X, Chen H, et al. What is the meaning of Health literacy? A systematic review and qualitative synthesis. *Fam Med Community Health.* (2020) 8:e000351. doi: 10.1136/fmch-2020-000351
- Pleasant A. Advancing Health literacy measurement: a pathway to better Health and Health system performance. *J Health Commun.* (2014) 19:1481–96. doi: 10.1080/10810730.2014.954083
- O'Neill B, Gonçalves D, Ricci-Cabello I, Ziebland S, Valderas J. An overview of self-administered Health literacy instruments. *PLoS One.* (2014) 9:e109110. doi: 10.1371/journal.pone.0109110
- Haun JN, Valerio MA, McCormack LA, Sørensen K, Paasche-Orlow MK. Health literacy measurement: an inventory and descriptive summary of 51 instruments. *J Health Commun.* (2014) 19:302–3. doi: 10.1080/10810730.2014.936571
- Nguyen TH, Park H, Han H-R, Chan KS, Paasche-Orlow MK, Haun J, et al. State of the science of Health literacy measures: validity implications for minority populations. *Patient Educ Couns.* (2015) 98:1492–12. doi: 10.1016/j.pec.2015.07.013
- Nutbeam D. Health literacy as a public Health goal: a challenge for contemporary Health education and communication strategies into the 21st century. *Health Promot Int.* (2000) 15:259–7. doi: 10.1093/heapro/15.3.259
- Simon MA, Li Y, Dong X. Levels of Health literacy in a community-dwelling population of Chinese older adults. *J Gerontol A Biol Sci Med Sci.* (2014) 69:554–60. doi: 10.1093/gerona/glu179
- Xu WH, Rothman RL, Li R, Chen Y, Xia Q, Fang H, et al. Improved self-management skills in Chinese diabetes patients through a comprehensive Health literacy strategy-study protocol of a cluster randomized controlled trial. *Trials.* (2014) 15:498. doi: 10.1186/1745-6215-15-498
- Leung AY, Lou VW, Cheung MK, Chan SS, Chi I. Development and validation of Chinese Health literacy scale for diabetes. *J Clin Nurs.* (2013) 22:2099. doi: 10.1111/jocn.12018
- Leung AY, Cheung MK, Lou VW, Chan FH, Ho CK, Do TL, et al. Development and validation of the Chinese Health literacy scale for chronic care. *J Health Commun.* (2013) 18:205–2. doi: 10.1080/10810730.2013.829138
- Shen M, Hu M, Liu S, Chang Y, Sun Z. Assessment of the Chinese resident Health literacy scale in a population-based sample in South China. *BMC Public Health.* (2015) 15:637. doi: 10.1186/s12889-015-1958-0
- Health CMO. *66 tips of Health: Chinese resident Health literacy manual.* Beijing: People's Medical Publishing House (2008).
- Sun Y, Wang W, Lang Y, Zhang S, Yang F. Study on the status and influencing factors of Chinese Residents' Health literacy. *Health Health Promot.* (2022) 17:379–1. doi: 10.16117/j.cnki.31-1974/r.202204379
- Li Y, Lv X, Liang J, Dong H, Chen C. The development and Progress of Health literacy in China. *Front Public Health.* (2022) 10:10. doi: 10.3389/fpubh.2022.1034907
- Shi M, Li Y, Liu Y, Li L, Nie X, Zhang G. Study on Health literacy and its influencing factors of Chinese residents aged 60–69 years from 2012 to 2017. *Chinese journal of Health Educ.* (2019) 35:963–6–88. doi: 10.16168/j.cnki.issn.1002-9982.2019.11.001
- Chung MH, Chen LK, Peng LN, Chi MJ. Development and validation of the Health literacy assessment tool for older people in Taiwan: potential impacts of cultural differences. *Arch Gerontol Geriatr.* (2015) 61:289–5. doi: 10.1016/j.archger.2015.06.015
- He H, Lv Y, Gao L. Review on the research status of Health literacy at home and abroad. *J Med Intell.* (2017) 38:7–10. doi: 10.3969/j.issn.1673-6036.2017.01.002
- Michael LT. Advances in applications of item response theory to clinical assessment. *Psychol Assess.* (2019) 31:1442–55. doi: 10.1037/pas0000597
- Cappelleri JC, Jason Lundy J, Hays RD. Overview of classical test theory and item response theory for the quantitative assessment of items in developing patient-reported outcomes measures. *Clin Ther.* (2014) 36:648–2. doi: 10.1016/j.clinthera.2014.04.006
- Doucette A, Wolf A. Questioning the measurement precision of psychotherapy research. *Psychother Res.* (2009) 19:374–9. doi: 10.1080/10503300902894422
- McDonald RP. The dimensionality of tests and items. *Br J Math Stat Psychol.* (1981) 34:100–7. doi: 10.1111/j.2044-8317.1981.tb00621.x
- Cella D, Yount S, Rothrock N, Gershon R, Cook K, Reeve B, et al. The patient-reported outcomes measurement information system (Promis): Progress of an NIH roadmap cooperative group during its first two years. *Med Care.* (2007) 45:S3–S11. doi: 10.1097/01.mlr.0000258615.42478.55
- Wainer H, Thissen D. How is reliability related to the quality of test scores? What is the effect of local dependence on reliability? *Educ Meas Issues Pract.* (1996) 15:22–9. doi: 10.1111/j.1745-3992.1996.tb00803.x
- Yen WM. Scaling performance assessments: strategies for managing local item dependence. *J Educ Meas.* (1993) 30:187–3. doi: 10.1111/j.1745-3984.1993.tb00423.x
- Morizot J, Ainsworth AT, Reise S. Toward modern psychometrics: application of item response theory models in personality research In: RW Robins and RF Kruger, editors. *Handbook of research methods in personality psychology.* New York: Guilford Press (2007).
- Natesan Batley P, Contractor AA, Weiss NH, Compton SE, Price M. Psychometric evaluation of the Posttrauma risky behaviors questionnaire: item response theory analyses. *Assessment.* (2021) 29:1824–41. doi: 10.1177/10731911211036760
- Li Y, Xiong C, Aschenbrenner AJ, Chang CH, Weiner MW, Nosheny RL, et al. Item response theory analysis of the clinical dementia rating. *Alzheimers Dement.* (2021) 17:534–42. doi: 10.1002/alz.12210
- Liu Y-B, Liu L, Li Y-F, Chen Y-L. Relationship between Health literacy, Health-related behaviors and Health status: a survey of elderly Chinese. *Int J Environ Res Public Health.* (2015) 12:9714–25. doi: 10.3390/ijerph120809714
- Weissberger GH, Han SD, Yu L, Barnes LL, Bennett DA, Boyle PA. Financial and Health literacy discrepancies with cognition in older adults. *Neuropsychology.* (2019) 33:975–5. doi: 10.1037/neu0000565
- Goetz CG, Liu Y, Stebbins GT, Wang L, Tilley BC, Teresi JA, et al. Gender-, age-, and race/ethnicity-based differential item functioning analysis of the Movement Disorder Society-sponsored revision of the unified Parkinson's disease rating scale. *Mov Disord.* (2016) 31:1865–73. doi: 10.1002/mds.26847
- Bialosiewicz S, Murphy K, Berry T. An introduction to measurement invariance testing: resource packet for participants. *American Evaluation Association meeting.* Washington, DC (2013).
- Pattanaik S, John MT, Chung S. Assessment of differential item functioning across English and Spanish versions of the Orofacial esthetic scale. *J Oral Rehabil.* (2021) 48:73–80. doi: 10.1111/joor.13106
- Gómez-Benito J, Sireci S, Hidalgo MD, Hidalgo MD, Benítez I. Differential item functioning: beyond validity evidence based on internal structure. *Psicothema.* (2018) 30:104–9. doi: 10.7334/psicothema2017.183
- Jia Q, Wang H, Wang L, Wang Y. Association of Health Literacy with medication adherence mediated by cognitive function among the community-based elders with chronic disease in Beijing of China. *Front Public Health.* (2022) 10:824778. doi: 10.3389/fpubh.2022.824778
- Galvin JE, Roe CM, Xiong C, Morris JC. Validity and reliability of the Ad8 informant interview in dementia. *Neurology.* (2006) 67:1942–8. doi: 10.1212/01.wnl.0000247042.15547.eb
- Chen HH, Sun FJ, Yeh TL, Liu HE, Huang HL, Kuo BI, et al. The diagnostic accuracy of the ascertain dementia 8 questionnaire for detecting cognitive impairment in primary Care in the Community, clinics and hospitals: a systematic review and Meta-analysis. *Fam Pract.* (2018) 35:239–6. doi: 10.1093/fampra/cmx098
- Embretson SE, Steven P. Reise. *Item response theory.* Hove: Psychology Press (2013).
- Reise SP, Waller NG. Item response theory and clinical measurement. *Annu Rev Clin Psychol.* (2009) 5:27–48. doi: 10.1146/annurev.clinpsy.032408.153553
- Pang H, Kang X, Li Z, Zhang J, Lv R, Jiang J. An application of item response theory in item selection of Chinese self-Management of Heart Failure Instrument. *Chin J Health Stat.* (2014) 31:57–60.
- Chen G, Xiong B, Li N, Jiang Y, Chen H, Ding D. Choosing the best item types of a public recruitment examination based on item response theory. *China Exam.* (2016) 1:39–46. doi: 10.3969/j.issn.1005-8427.2016.01.006
- Biemer PP. The third chapter of Response Probability Models of Two Measurements in the book Latent Class Analysis of Survey Error, by Paul P. Biemer (2010):71–3.
- Huynh H. A clarification on the response probability criterion Rp67 for standard settings based on bookmark and item mapping. *Educ Meas Issues Pract.* (2006) 25:19–20. doi: 10.1111/j.1745-3992.2006.00053.x
- Wang X. Setting the cut score in the standards-based education examination with bookmark. *China Exam.* (2014) 7:10–18. doi: 10.3969/j.issn.1005-8427.2014.07.002
- Xiong J, Ding S, Qi S, Dai H. Use test information to analyze test quality. *Jiangxi Normal Univ.* (2002) 26:225–8. doi: 10.3969/j.issn.1000-5862.2002.03.011
- Scott NW, Fayers PM, Aaronson NK, Bottomley A, De Graeff A, Groenvold M, et al. Differential item functioning (Dif) analyses of Health-related quality of life

- instruments using logistic regression. *Health Qual Life Outcomes*. (2010) 8. doi: 10.1186/1477-7525-8-81
51. Crane PK, Cetin K, Cook KF, Johnson K, Deyo R, Amtmann D. Differential item functioning impact in a modified version of the Roland–Morris disability questionnaire. *Qual Life Res*. (2007) 16:981–0. doi: 10.1007/s11136-007-9200-x
52. Crane PK, Gibbons LE, Jolley L, van Belle G. Differential item functioning analysis with ordinal logistic regression techniques: Difdetect and Difwithpar. *Med Care*. (2006) 44:S115–23. doi: 10.1097/01.mlr.0000245183.28384.ed
53. Zumbo B, Thomas D. *A measure of effect size for a model-based approach for studying Dif*. Prince George, Canada: University of Northern British Columbia, Edgeworth Laboratory for Quantitative Behavioral Science (1997).
54. Lu X, Yeo KJ, Guo F, Zhao Z, Wu O. Psychometric property and measurement invariance of internet addiction test: the effect of socio-demographic and internet use variables. *BMC Public Health*. (2022) 22:1548. doi: 10.1186/s12889-022-13915-1
55. Ali AM, Hori H, Kim Y, Kunugi H. The depression anxiety stress scale 8-items expresses robust psychometric properties as an ideal shorter version of the depression anxiety stress scale 21 among healthy respondents from three continents. *Front Psychol*. (2022) 13:799769. doi: 10.3389/fpsyg.2022.799769
56. Yu C-Y. *Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes*. Los Angeles: University of California (2002).
57. Denovan A, Dagnall N, Dhingra K, Grogan S. Evaluating the perceived stress scale among Uk university students: implications for stress measurement and management. *Stud High Educ*. (2019) 44:120–33. doi: 10.1080/03075079.2017.1340445
58. Ali AM, Hendawy AO, Ahmad O, Al Sabbah H, Smail L, Kunugi H. The Arabic version of the Cohen perceived stress scale: factorial validity and measurement invariance. *Brain Sci*. (2021) 11:419. doi: 10.3390/brainsci11040419
59. Philip Chalmers R. Mirt-a multidimensional item response theory package for the R environment. *J Stat Softw*. (2012) 48. doi: 10.18637/jss.v048.i06
60. Zhang Z, Yuan K-H. Robust coefficients alpha and omega and confidence intervals with outlying observations and missing data. *Educ Psychol Meas*. (2016) 76:387–411. doi: 10.1177/0013164415594658
61. Chiu YH, Lu FJ, Lin JH, Nien CL, Hsu YW, Liu HY. Psychometric properties of the perceived stress scale (Pss): measurement invariance between athletes and non-athletes and construct validity. *Peer J*. (2016) 4:e2970. doi: 10.7717/peerj.2790
62. Peterson AC, Sutherland JM, Liu G, Crump RT, Karimuddin AA. Evaluation of the fecal incontinence quality of life scale (Fiql) using item response theory reveals limitations and suggests revisions. *Qual Life Res*. (2018) 27, 27:1613–23. doi: 10.1007/s11136-018-1826-3
63. Kayser L, Karnoe A, Furstrand D, Batterham R, Christensen KB, Elsworth G, et al. A multidimensional tool based on the Ehealth literacy framework: development and initial validity testing of the Ehealth literacy questionnaire (Ehlq). *J Med Internet Res*. (2018) 20:e36. doi: 10.2196/jmir.8371
64. Giacomelli SC, de Assis MAA, de Andrade DF, Schmitt J, Hinnig PF, Borgatto AF, et al. Evelopment of a food-based diet quality scale for Brazilian schoolchildren using item response theory. *Nutrients*. (2021) 13:3175. doi: 10.3390/nu13093175