tvst
Article

Psychometric Assessment of the Chinese Version of the Indian Vision Functioning Questionnaire Based on the Method of Successive Dichotomizations

Rongrong Gao[1,*], Sisi Chen[1,*], Shixiang Yan[1], Tianhao Lu[1], Haisi Chen[1], Qi Feng[1], Qinmei Wang[1], Yong Sun[2], Jinhai Huang[1], and Jyoti Khadka[1,3–5]

[1] Eye Hospital and School of Ophthalmology and Optometry, Wenzhou Medical University, Wenzhou, Zhejiang, China
[2] Shenzhen Hospital of Integrated Traditional and Western Medicine, Shenzhen, China
[3] Health and Social Care Economics Group, College of Nursing and Health Sciences, Flinders University, Adelaide, Australia
[4] Registry of Senior Australians, South Australian Health and Medical Research Institute, Adelaide, Australia
[5] Business School, University of South Australia, Adelaide, Australia

Correspondence: Yong Sun, Shenzhen Hospital of Integrated Traditional and Western Medicine, Shenzhen, China. e-mail: 383343358@qq.com
Jinhai Huang, Eye Hospital of Wenzhou Medical University, 270 West Xueyuan Road, Wenzhou, Zhejiang, 325027, China. e-mail: vip999vip@163.com

Received: June 22, 2020
Accepted: May 1, 2021
Published: June 8, 2021

Keywords: cataract; the 33-item Indian Vision Function Questionnaire; psychometric properties

Citation: Gao R, Chen S, Yan S, Lu T, Chen H, Feng Q, Wang Q, Sun Y, Huang J, Khadka J. Psychometric assessment of the Chinese version of the Indian vision functioning questionnaire based on the method of successive dichotomizations. Transl Vis Sci Technol. 2021;10(7):8, https://doi.org/10.1167/tvst.10.7.8

Purpose: The purpose of this study was to assess whether a Chinese translated version of the 33-item Indian Vision Function Questionnaire (IND-VFQ-33) forms a valid measurement scale and to evaluate its psychometric properties based on the method of successive dichotomizations (MSD).

Methods: The English version of the IND-VFQ-33 was translated, back translated, and cross-culturally adapted for use in China. It was interviewer administered to patients with cataracts. MSD, a polytomous Rasch model that estimates ordered thresholds, was used to assess and optimize psychometric properties of the overall scale and three subscales separately.

Results: One hundred and seventy-nine patients provided complete responses. After the removal of 2 misfitting items, a revised 31-item overall scale demonstrated adequate precision (person reliability [PR] = 0.92) and no misfitting items. The general functioning subscale fit the MSD model well after removing two misfitting items. The psychosocial impact subscale and the visual symptoms subscale were not considered further due to poor measurement precision. After addressing psychometric deficiencies, a 31-item overall scale (IND-VFQ-31-CN) and a 19-item general functioning subscale (IND-VFQ-GF-19-CN) were developed.

Conclusions: The original IND-VFQ-33 required re-engineering to form valid measures for use in China. The revised overall scale and general functioning subscale demonstrated adequate MSD based psychometric properties.

Translational Relevance: The revised IND-VFQ-33 is a valid patient-reported outcome assessment for Chinese patients with cataract based on MSD analysis.

Introduction

Cataract is the main cause of blindness in the world,[1,2] also one of the key eye diseases to be tackled in the prevention and control of blindness in China.[3] Epidemiological data show that there are about 3 million blind people in China because of cataract, and it is growing.[4] The number of people undergoing cataract surgery has also increased dramatically over the years in China.[5] However, due to a huge backlog of people with cataract, many people have to live with an easily treatable visual impairment and blindness.

tvst.arvojournals.org | ISSN: 2164-2591

translational vision science & technology

There are 17 patient-reported outcome measures (PROMs) that have been tested for their validity in people with cataracts. Out of the 17 PROMs, 12 are cataract-specific and the other 5 are generic vision-specific but later used on cataract populations.[6] Of them, the majority were originally designed for developed countries, by contrast, only a few were originally designed for developing countries. As a rare instrument originally developed for a developing country—India—the 33-item Indian Vision Function Questionnaire (IND-VFQ-33) has been used across different eye conditions such as uveitis,[7,8] corneal diseases,[9] cataract,[10] and primary glaucoma.[11,12] Like most ophthalmological instruments, it was developed and proved to possess good reliability and validity using the classical test theory (CTT).[13] Finger et al.[14] and Gothwal et al.[15] have verified its psychometric properties in Indian patients using the Rasch analysis, and they proved it to be an effective tool. However, they used the Andrich rating scale model, which frequently estimates disordered thresholds due to a mathematical property of the model rather than a problem with the data.[16]

Collapsing rating categories until thresholds are ordered is a common practice to address the disordered thresholds identified through the Andrich rating scale model, but that remedy changes the scale. The method of successive dichotomizations (MSD), as the only known polytomous Rasch model that always estimates ordered category thresholds, resolves these issues.[16]

Both China and India are developing countries, and we hypothesize that the IND-VFQ-33 should form a valid measure to assess cataract patients' quality of life. Therefore, the aim of this study was to assess whether the IND-VFQ-33 could be adapted to China following MSD analysis to validate its psychometric properties.

Materials and Methods

Study Population

Patients waiting for cataract surgery were recruited from the Eye Hospital of Wenzhou Medical University, Wenzhou, China. The participants were 18 years or older, had a history of cataract in one or both eyes for more than 6 months, and with normal cognitive ability to understand items and communicate their responses to items. Patients with other significant ocular comorbidities or systemic diseases that might significantly influence their quality of life other than cataract were excluded. The participants completed the IND VFQ-33 (translated into Mandarin) by face-to-face interviews. All the participants underwent a detailed clinical assessment, including habitual distant visual acuity (HDVA), slit lamp microscopy, and fundus examination. The HDVA was measured using a Snellen chart and then converted into LogMAR values for statistical analysis. Extremely poor visual acuities (hand motion and counting fingers) were converted into LogMAR equivalent, as recommended by Lange et al.[17] The study followed the Helsinki Declaration and was approved by the ethics committee of the Eye Hospital of Wenzhou Medical University. All included patients signed informed consent after agreeing with the nature and intention of the study.

The 33-Item Indian Vision Function Questionnaire

The IND-VFQ-33 has 33 items grouped into 3 subscales: general functioning subscale (item 1 to item 21), psychosocial impact subscale (item 22 to item 26), and visual symptoms subscale (item 27 to item 33). The general functioning subscale has 5 active response options: "not at all", "a little", "quite a bit", "a lot", and "cannot do this because of my sight", were coded as 1, 2, 3, 4 and 5, respectively; the response categories for the psychosocial impact subscale and the visual symptoms subscale are "not at all", "a little", "quite a bit", and "a lot", were coded as 1, 2, 3 and 4, respectively.

In this study, the original IND-VFQ-33 was translated from English into Mandarin independently by two ophthalmologists who are fluent in both languages. The two versions were revised by panel discussion to produce a second draft. The draft was then translated back by a college English teacher who was blinded to the original instrument. The panel of experts compared the back-translated version with the original version to identify any discrepancies. Finally, the cross-cultural adaptation was conducted among 20 patients to ensure the items had semantic clarity and related to their life. The Chinese version of the IND-VFQ-33 (IND-VFQ-33-CN) was consistent with the original version, except for 3 items which were revised: item 5 "Going to social functions such as weddings" was revised as "social gatherings (such as wedding, banquets, churches, temples, etc.)"; item 11 "Locking or unlocking the door" was revised as "Locking or unlocking the door with the key" (because using a door handle requires lower visual acuity than having to use a key to open a locked door); item 18 "Making out differences in coins or notes" was revised as "Identifying the different denominations of coins and notes" (to make it more relatable to Chinese people; Supplementary Table S1).

Methods of Successive Dichotomization Analysis

As a recent advancement in the psychometric methods, MSD is the only known polytomous Rasch model that always estimates ordered category thresholds. All other polytomous Rasch models, such as the Andrich rating scale model, frequently estimate disordered thresholds. Although it has been suggested that this is due to a flaw in the data, it turns out this is due to a mathematical property of the model (the multiplicative structure) that is logically inconsistent with the assumption that a rating scale, defined by ordered thresholds, is used to rate items on each trial.[16] We assessed the following MSD analysis based psychometric properties.

Measurement Precision

An instrument should have an adequate discriminative capacity with the person reliability (PR) $\geq$0.8, indicating that at least 3 strata of person abilities can be discriminated.[6,18,19]

Fit Statistics

A good instrument should be unidimensional. Unidimensionality implies that items in an instrument together measure a single underlying trait. MSD analysis uses fit statistics to assess dimensionality of an instrument. Item fit statistics (mean square [MNSQ] statistics, include infit MNSQ and outfit MNSQ) judge if items "fit" the underlying construct with the acceptable cutoff range for infit and outfit between 0.5 and 1.5. Along with the fit statistics, we also considered item content and item location before considering for item deletion.

Targeting

Targeting is defined as the match between item measures (item difficulties) and person measures (person abilities). A well targeted instrument has evenly spaced, well-fitting items with persons positioned at the same level in the person-item map.[19] The targeting is considered acceptable when there is at least one item within approximately one logit of each person, and there are multiple items within one logit of each larger cluster of persons (the more persons in a region, the more items are needed to increase discrimination ability).

Statistical Analysis

Descriptive data were analyzed using SPSS software (IBM SPSS Statistic for Windows, version 19.0.0; IBM, Armonk, NY, USA). Pearson correlation was used if both data were normally distributed, Spearman rank correlation was used otherwise. MSD analysis was performed using the "msd" package in R. $P < 0.05$ was considered as significant statistical differences in all analyses. Because there were 3 different question formats used in the IND-VFQ-33, we assessed the psychometric properties of the 3 subscales of the IND-VFQ-33-CN separately and optimized them in case of any defects.

Results

A total of 179 patients with cataract (median age = 67 years, range = 28 to 90 years) completed the IND-VFQ-33-CN, with 80 men (44.7%) and 99 women (55.3%). Forty-three percent of the participants were illiterate, 86.6% waited for the first eye operation, 36.9% had ocular comorbidities, and more than half had systemic comorbidities (Table 1).

Performance of the Overall IND-VFQ-33-CN Based on MSD Analysis

The IND-VFQ-33-CN demonstrated a high PR value of 0.91 but had 2 misfitting items (i.e. item 9 "Recognizing people from a distance" [infit = 2.00 and outfit = 2.00] and item 20 "Seeing objects fallen in the food" [infit = 2.24 and outfit = 2.19]). The misfitting items were deleted iteratively until the items fit the MSD model without compromising the PR and targeting of the scale (Table 2). The final instrument (named "IND-VFQ-31-CN") included 31 items with a high PR of 0.92. The person-item map (Fig. 1) demonstrated there was a lack of items within 1 logit for persons with abilities below −4.11 (3 persons, count for 1.7%), and that more items are needed with difficulty between −1.77 and −2.52.

Performance of the General Functioning Subscale Based on MSD Analysis

After iteratively deleting two misfitting items (i.e. item 9 "Recognizing people from a distance" [infit = 2.06 and outfit = 2.10] and item 20 "Seeing objects fallen in the food" [infit = 2.11 and outfit = 2.03]), the remaining items fit well to MSD model. PR dropped from 0.87 to 0.85, and item reliability (IR) dropped from 0.98 to 0.97 but it was still above the minimum

Table 1. Demographic Characteristics of the Participants

Characteristics	Results
Age, y (median, IQR, range)	67, 13, 28–90
Sex, n (%)	
Female	99 (55.3)
Male	80 (44.7)
First or second eye surgery *n* (%)	
First eye surgery	155 (86.6)
Second eye surgery	24 (13.4)
Visual acuity, logMAR (median, IQR, range)	
Worse eye	1.00, 1.10, 0.22–2.70
Better eye	0.52, 0.48, −0.08–2.00
Binocular	0.52, 0.48, −0.08–2.00
Ocular comorbidity,[a] n (%)	66 (36.9)
Glaucoma	5 (2.8)
DR	4 (2.2)
Pathological myopia	12 (6.7)
Corneal disorders[b]	8 (4.5)
Others[c]	49 (27.4)
Systemic comorbidity,[a] *n* (%)	95 (53.1)
Hypertension	72 (40.2)
Diabetes	28 (15.6)
Others	10 (5.6)
Education level, *n* (%)	
Illiterate	77 (43.0)
Primary school	50 (27.9)
Junior middle school	26 (14.5)
Senior middle school	18 (10.1)
University	8 (4.5)

DR, diabetic retinopathy; IQR, interquartile range.

[a]The cumulative percentage of comorbidities exceeds the total because some patients have various kinds of ocular or systemic comorbidities.

[b]Corneal macula, corneal dystrophies, etc.

[c]Pterygium, vein occlusion, uveitis, epiretinal membrane, etc.

Table 2. Properties of the Overall IND-VFQ-33-CN

	The Overall IND-VFQ-33-CN	Delete Item 9, 20
No. of items	33	31
Misfitting items	2[a]	0
PR	0.91	0.92
Targeting	−2.28	−1.72

IND-VFQ-33-CN, 33-item Indian Vision Function Questionnaire Chinese version; PR, person reliability.

[a]Misfitting items: item 9 (infit = 2.00 and outfit = 2.00), item 20 (infit = 2.24 and outfit = 2.19).

acceptable precision. Person and item mean difference improved from −2.90 to −2.23 (Fig. 2). The revised 19-item general functioning subscale was named "IND-VFQ-GF-19-CN" (Table 3).

```
MEASURE      PERSON - MAP - ITEM
                 <more>|<rare>
                       |
                     . |
                       |T 7
                       |
   3                   +  18
                     . |
                       |
                     . |  10
                       |
   2                 . +  19
                       |
                     . |S 11 21
                       |  17
                     # |  25
   1                 . T+  3 8
                     . |  12
                   .## |
                   ### |  5 14
                     . |  1 23
   0               ### +M
                    ## S|  4 13
                   .## |  24
                   .## |
                   ### |  16 22 30
  -1              #### +  2 6 32
                   ### |  15
                ###### |
                   ### M|S 28 29
                .##### |
  -2            .##### +
               ####### |
                   .## |  26
                 ##### |  27
            ########## |  33
  -3               ### S+  31
                    ## |
                  .### |T
                    .# |
                    .# |
  -4                 # +
                     . T|
                     # |
                       |
                       |
  -5                   +
                       |
                       |
                 <less>|<freq>
EACH "#" IS 2: EACH "." IS 1
```

Figure 1. Person-item map of the IND-VFQ-31-CN.

Performance of the Psychosocial Impact Subscale Based on MSD Analysis

PR was 0.46, and IR was 0.99. Item 25 "Feel you are a burden on others" was the only item that misfit-

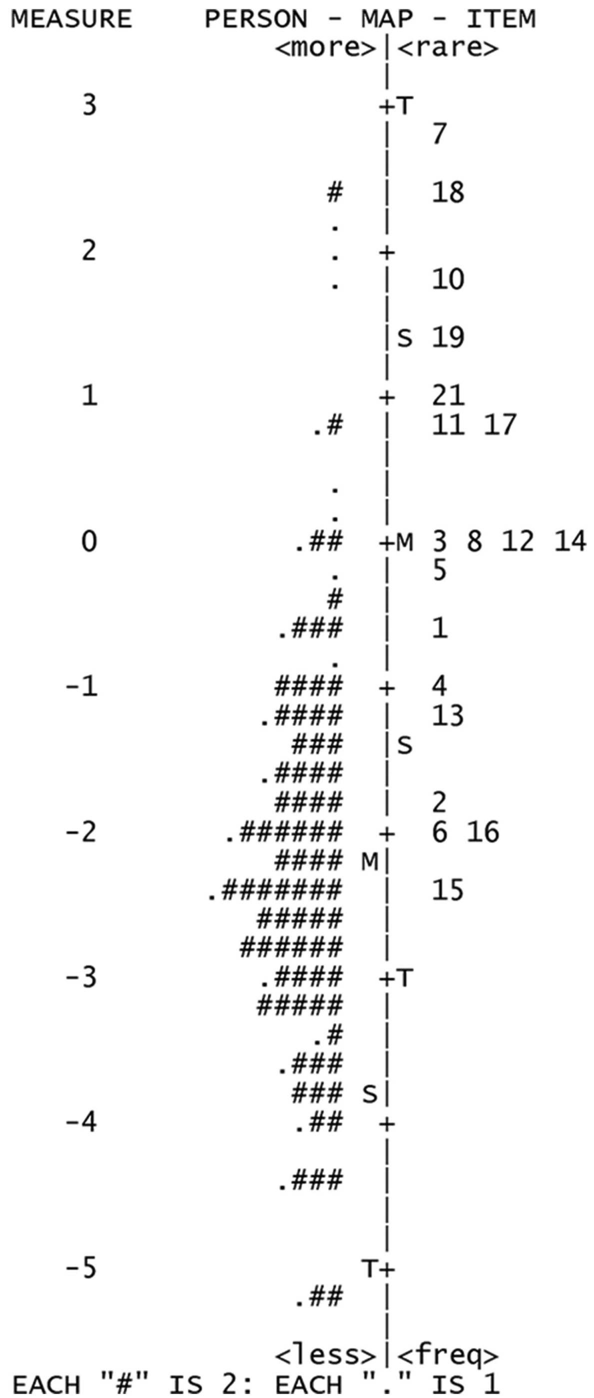

Figure 2. Person-item map of the IND-VFQ-GF-19-CN.

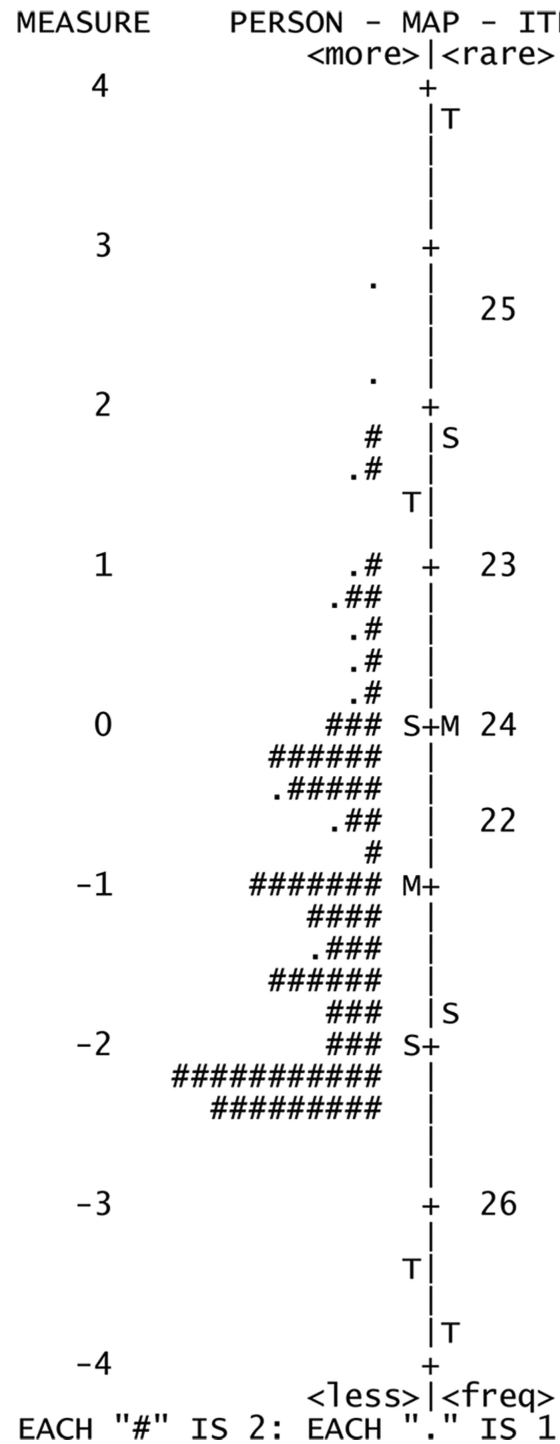

Figure 3. Person-item map of the IND-VFQ-PI-5-CN.

ted the MSD model (infit = 1.96 and outfit = 1.85), but it was retained, otherwise the PR would have worsened further. The person-item map showed most items located above the respondents, indicating that these items were generally too easy for the respondents. Of all the items, item 26 (Feel frightened to lose remaining vision) was the most impacted, and item 25 (Feel you are a burden on others) was the least impacted (Fig. 3). Table 4 shows the details of this subscale (named "IND-VFQ-PI-5-CN"). This subscale was not considered further because its PR was inadequate.

Performance of the Visual Symptoms Subscale Based on MSD Analysis

PR was 0.48, and IR was 0.97. Item 32 ("Does light seem like stars") misfit the MSD model

Table 3. Properties of the General Functioning Subscale

	Current Study		Finger et al.[14]			Gothwal et al.[15]		
Model	MSD		The Andrich model			The Andrich model		
Subscale	The general functioning subscale	the IND-VFQ-GF-19-CN	Original	Mobility	Activity limitation	Original	Mobility	Visual function
Category thresholds	Ordered	Ordered	Disordered	Ordered	Ordered	Ordered	Ordered	Ordered
No. of items	21 (1–21)	19 (1–8, 10–19, 21)	21 (1–21)	6 (1–4, 7, 8)	10 (10–14, 17–21)	21 (1–21)	7 (1–6, 8)	13 (7, 9–14, 16–21)
Misfitting items	2[a]	0	2	0	0	1[a]	0	0
PR	0.87	0.85	0.95	0.91	0.88	–	–	–
Targeting	−2.90	−2.23	−2.23	−2.94	−1.93	−0.86	−0.57	−1.13

IND-VFQ-GF-19-CN, Chinese translated version of the 33-item Indian Vision Function Questionnaire 19-item general functioning subscale; MSD, method of successive dichotomization; PR, person reliability.
[a]Misfitting items: item 9 (infit = 2.06 and outfit = 2.10), item 20 (infit = 2.11 and outfit = 2.03).

Table 4. Properties of the Psychosocial Impact Subscale

	Current Study	Finger et al.[14]	Gothwal et al.[15]
Model	MSD	The Andrich model	The Andrich model
Category thresholds	Ordered	Ordered	Ordered
No. of items	5 (22–26)	5 (22–26)	5 (22–26)
Misfitting item	1[a]	0	1[b]
PR	0.46	0.80	–
Targeting	−1.05	−1.39	−0.26

MSD, method of successive dichotomization; PR, person reliability.
[a]Misfitting item: item 25 (infit = 1.96 and outfit = 1.85).
[b]Not provided.

(infit = 1.72 and outfit = 1.63), but was retained to not decrease PR. Targeting was excellent (person mean = 0.32). Supplementary Table S2 shows the details of this subscale (named “IND-VFQ-VS-7-CN”). Item 31 (Do you close your eyes because of light from vehicles) was the most impacted and item 30 (Does bright light hurt your eyes) was the least impacted (Fig. 4). This subscale’s PR was also poor to form a valid subscale, hence it was not considered further.

Correlation Between Person Measure Score and Visual Acuity

All the IND-VFQ-31-CN and IND-VFQ-GF-19-CN scores correlated weakly with the HDVAs. The Spearman correlations between IND-VFQ-31-CN scores and better-eye, worse-eye, and binocular HDVA were $r = 0.275$, $P < 0.001$, $r = 0.211$, $P = 0.005$ and $r = 0.261$, $P < 0.001$, respectively. The correlations between IND-VFQ-GF-19-CN scores and better-eye, worse-eye, and binocular HDVA were $r = 0.259$, $P < 0.001$, $r = 0.152$, $P = 0.042$ and $r = 0.244$, $P = 0.001$, respectively. The correlations were slightly stronger between the scores and better-eye HDVA.

Comparison of the Person Measures and Item Measures Between the Translated Version and the Revised Version

The R^2 for person measures and item measures between IND-VFQ-31-CN and the overall IND-VFQ-33-CN were 0.9201 and 0.9955, respectively, the R^2 for person measures and item measures between IND-VFQ-GF-19-CN and the general functioning subscale were 0.8431 and 0.9951, respectively (Fig. 5).

Conversion From Raw Scores to MSD Equivalents

The calibrated item measures are listed in Supplementary Tables S3–S5, and the estimated rating category thresholds are listed in Supplementary Table S6. If the demographics are similar to the present study, users of the instruments can use the estimated item measures and rating category thresholds to estimate person measures using the “pms” function in the R package msd. Users should perform MSD analysis on their own sample if the demographics are considerably different from this study.

```
MEASURE     PERSON - MAP - ITEM
                    <more>|<rare>
   4                      +
                          |
                          |
                          |
                          |
   3                      +
                       .  |
                       .  |
                      .#  |
                       #  |
   2                  ## T+
                     ###  |T
                       #  |
                .#######  |
                     ###  |   30
   1                #### S+   32
                    ####  |S
              ##########  |
                      .#  |   29
             .######## M|     28
   0                ####  +M
              #########   |
                      ##  |
                      ## S|   27
           .###########   |S  33
  -1                   #  +
                      ##  |   31
                       #  |
                         T|
                      .#  |T
  -2                 ###  +
                          |
                          |
                          |
                          |
  -3                   .  +
                          |
                          |
                          |
                          |
  -4                      +
                    <less>|<freq>
EACH "#" IS 2: EACH "." IS 1
```

Figure 4. Person-item map of the IND-VFQ-VF-7-CN.

Discussion

The development and validation of a new instrument is time-consuming and laborious. At present, there is no ophthalmic instrument originally developed for Chinese people with cataracts. Most instruments used in China are translated from instruments developed in developed countries. Khadka et al.[6] reviewed 17 kinds of cataract instruments and concluded that the 3 revised subscales of the IND-VFQ (mobility, emotional well-being, and visual symptoms) were the only superior quality instruments recommended for developing countries, and they demonstrated better quality characteristics in psychological well-being as well as a relatively wide coverage of concepts being assessed.[14] However, the IND-VFQ-33 was developed and validated for the visually impaired population in India in 2004 based on CTT.[13] As China is also a developing nation, it is possible that the IND-VFQ-33 might be relevant to its settings. However, due to the cultural differences between the two countries, the original version should be revised and tested before it should be used in Chinese settings, as demonstrated by our study. From the results of this study, we further conclude that instruments developed for one developing country do not necessarily apply to another, and their applicability should be verified before use.

Most studies support a four- or five-point Likert scale.[20,21] Gothwal et al.[15] confirmed all the rating category thresholds of the original three subscales were ordered, whereas Finger et al.[14] found that certain categories of the general functioning subscale were redundant and changed the scale to a four-point by collapsing categories two (a little) and three (quite a bit). They both used the Andrich Rating model, the difference was that the former investigated patients with cataract and the latter adults with low vision. Collapsing rating categories until thresholds are ordered is an attempt to compensate for a mathematical property of the Andrich model (i.e. its multiplicative structure) rather than inherent flaws in the data. This flaw in Andrich model is addressed by using a polytomous Rasch model (MSD) that always estimates ordered thresholds.[16]

As traditionally used in Rasch analysis-based literature, an instrument should have an adequate discriminative capacity with PR ≥ 0.8.[6,18,19] We also took this PR value as our cut off. Our study found that both the psychological subscale and the visual symptoms subscale demonstrated extremely poor measurement precision (PR value of 0.46 and 0.48, separately), hence these subscales were not considered further. Gothwal et al.[15] came to a similar conclusion. One of the reasons could be attributed to the limited number of number of items these subscale had (only 5 and 7). However, it could be attributed to the fact that unlike in Western countries, Chinese people are usually emotionally conservative and might not have provided accurate responses to the items. Similar findings were reported in our previous study where the social-emotional subscale

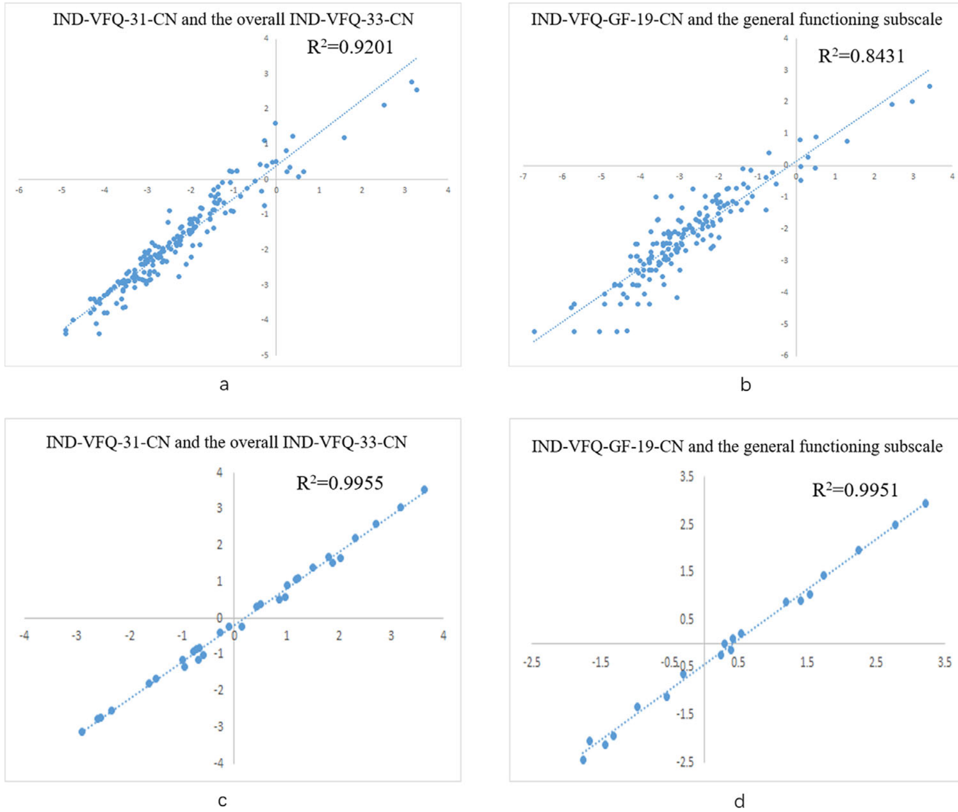

Figure 5. R^2 for person measures (**a, b**) and item measures (**c, d**) between two different Chinese versions (one with items removed).

of the NEI VFQ did not form a valid scale in our setting.[22]

Two existing Rasch analysis studies of the IND-VFQ by Finger et al.[14] and Gothwal et al.[15] both found multidimensionality of the general functioning subscale (the first contrast eigenvalues were 3.0 and 3.4, respectively), and split the subscale into two subscales (i.e. "mobility, activity limitation" and "mobility, visual function", respectively) which demonstrated unidimensionality. However, their classifications of item 7 were different, and the boundary between the definitions of the new dimensions was fuzzy, making it difficult to classify items and hard for subjects to understand. Besides, the different named dimensions may lead to confusion.[23]

The IND-VFQ-31-CN has some targeting issues. Its person-item map (Fig. 1) shows that there is a lack of items within 1 logit for persons with abilities below −4.11 (3 persons, accounting for 1.7%), and that there needs to be more items with difficulties between −1.77 and −2.52 to increase discrimination ability, because these regions have a larger cluster of persons. The revised general functioning subscale (IND-VFQ-GF-19-CN) had suboptimal targeting, as shown in its person-item map (Fig. 2). There was a lack of items within 1 logit for persons with abilities below −3.44 (32 persons, count for 17.8%), suggesting that this subscale was less sensitive for measuring ability levels of participants with a higher ability. The targeting was also worse than reported in the Finger et al.[14] and Gothwal et al.[15] studies. The difference may be attributed to sample selection. The average visual acuity of the better eye of the patients with cataract in our study was 0.55 ± 0.44 LogMAR, which is better than that of Finger et al.[14] (patients with cataract = 0.74 ± 0.45 LogMAR) and Gothwal et al.[15] (adults with low vision = 0.88 ± 0.49 LogMAR). The targeting is consistent with the results of previous Rasch validation of other instruments, such as the Quality of Life and Vision Function Questionnaire (QOL-VFQ),[24] the Activities of Daily Vision Scale (ADVS),[25] the Visual Disability Assessment (VDA),[26] the visual functioning index (VFI),[27] the visual function (VF), and quality of life questionnaires(QOL).[28] Given that sample characteristics may influence MSD model fit, a larger, more diverse sample with more severe cataracts and impairment may contribute to better targeting for the IND-VFQ.[29] On the other hand, the IND-VFQ-33 was developed in 2005, the activities referred to by its items may not be challenging enough for current patients, and items that are more relevant to modern life should be added appropriately (for example, using Facebook, WeChat, playing on a computer, and so on). The best way is to build an item bank with new items that can be updated in real time, as proposed by Pesudovs.[21,23]

Many studies have explored the relationship between patient-reported outcomes (PROs) and clinical parameters.[30–33] Although the relationship seems complex, PROs have been proved to be an effective measure of visual function in patients. The IND-VFQ-31-CN and IND-VFQ-GF-19-CN scores are both weakly correlated to visual acuity, consistent with our previous research about other instruments,[34,35] indicating that PROs and visual acuity were inter-related and mutually complementary.

All the R^2 between person and item measures from the two different Chinese versions (one with items removed) were high, indicating that the removal of items did not change the underlying latent trait being measured. However, since we were unable to find item measures from the original English version, R^2 between item measures from an English version versus Chinese version was not computed to validate the translation. Despite the meticulous translation and cultural adaptation process used, we recommend a further evaluation is required to test whether the original English version and our Chinese version measure the same underlying trait and the translation we did is valid.

Despite the encouraging findings, some limitations in the current study need to be considered. First, as the study was a single-center design, we only examined patients in Wenzhou City, most of whom had relatively mild disability, which means we are unable to generalize our result to other areas. The performance of IND-VFQ-31-CN and IND-VFQ-GF-19-CN should be further explored in different areas to enhance its validity in greater China. Besides, the cross-sectional design was notable to study the responsiveness of the instrument.

In conclusion, the current study found that the IND-VFQ-33, although suitable for India, is not optimum for Chinese settings in its original format. After MSD analysis guided revision, the IND-VFQ-31-CN and IND-VFQ-GF-19-CN demonstrate good overall functioning for Chinese population. However, further studies are warranted to test their validity in a new sample. Suboptimal targeting exists in the revised versions, which may be better suited for a more impaired population, highlighting the need to develop an item bank.

Acknowledgments

Supported in part by Zhejiang Provincial & Ministry of Health Research Fund for Medical Sciences (WKJ-ZJ-1530); Medical Scientific Research Foundation of Zhejiang Province (2017KY495); Foundation of Wenzhou City Science & Technology Bureau (Y20170195); Zhejiang Provincial High-level Talents Program (2017-102); Project of the Affiliated Eye Hospital of Wenzhou Medical University (YNCX3201906). The funding sources had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Disclosure: **R. Gao,** None; **S. Chen,** None; **S. Yan,** None; **T. Lu,** None; **H. Chen,** None; **Q. Feng,** None; **Q. Wang,** None; **Y. Sun,** None; **J. Huang,** None; **J. Khadka,** None

[*] RG and SC contributed equally to this work as first authors.

References

1. Leasher JL, Bourne RR, Flaxman SR, et al. Global estimates on the number of people blind or visually impaired by diabetic retinopathy: a meta-analysis from 1990 to 2010. *Diabetes Care*. 2016;39:1643–1649.
2. Lee CM, Afshari NA. The global state of cataract blindness. *Curr Opin Ophthalmol*. 2017;28:98.
3. Chen X, Zhou D, Shen J, et al. Prevalence and causes of visual impairment in adults in Binhu District, Wuxi, China. *Med Sci Monit.* 2018;24:317–323.
4. Gao Y, Tu CS, Xu X, Zhou F, Chen H. A new model of blindness prevention with the target of cataract elimination. *Zhonghua Yi Xue Za Zhi.* 2012;92:2055–2058.
5. Wu Y, Shang P, Che L, Ye T, Wang L, Qiu S. The necessity of strength evaluation in assessment of clinical outcome after shoulder surgery: follow-up data from patients with complex proximal humerus fractures treated by locking plate fixation. *Acta Orthop Belg*. 2016;82:189–196.
6. Khadka J, McAlinden C, Pesudovs K. Quality assessment of ophthalmic questionnaires: review and recommendations. *Optom Vis Sci*. 2013;90:720–744.
7. Venkataraman A, Rathinam SR. A pre- and post-treatment evaluation of vision-related quality of life in uveitis. *Indian J Ophthalmol*. 2008;56:307–312.
8. Niemeyer KM, Gonzales JA, Rathinam SR, et al. Quality-of-life outcomes from a randomized clinical trial comparing antimetabolites for intermedi-

ate, posterior, and panuveitis. *Am J Ophthalmol*. 2017;179:10–17.
9. Vashist P, Gupta N, Tandon R, Gupta SK, Dwivedi S, Mani K. Population-based assessment of vision-related quality of life in corneal disease: results from the CORE study. *Br J Ophthalmol.* 2016;100:588–593.
10. Finger RP, Kupitz DG, Fenwick E, et al. The impact of successful cataract surgery on quality of life, household income and social status in South India. *PLoS One.* 2012;7:e44268.
11. Arora V, Bali SJ, Gupta SK, et al. Impact of initial topical medical therapy on short-term quality of life in newly diagnosed patients with primary glaucoma. *Indian J Ophthalmol.* 2015;63:511–515.
12. Reighard CL, Pillai MR, Shroff S, et al. Glaucoma-associated visual task performance and vision-related quality of life in South India. *Ophthalmol Glaucoma*. 2019;2:357–363.
13. Gupta SK, Viswanath K, Thulasiraj RD, et al. The development of the Indian vision function questionnaire: field testing and psychometric evaluation. *Br J Ophthalmol*. 2005;89:621–627.
14. Finger RP, Kupitz DG, Holz FG, et al. The impact of the severity of vision loss on vision-related quality of life in India: an evaluation of the IND-VFQ-33. *Invest Ophthalmol Vis Sci*. 2011;52:6081–6088.
15. Gothwal VK, Bagga DK, Sumalini R. Rasch analysis of the Indian vision function questionnaire. *Br J Ophthalmol.* 2012;96:619–623.
16. Bradley C, Massof RW. Method of successive dichotomizations: An improved method for estimating measures of latent variables from rating scale data. *PLoS One*. 2018;13:e0206106.
17. Lange C, Feltgen N, Junker B, Schulze-Bonsel K, Bach M. Resolving the clinical acuity categories "hand motion" and "counting fingers" using the Freiburg Visual Acuity Test (FrACT). *Graefes Arch Clin Exp Ophthalmol.* 2009;247:137–142.
18. Khadka J, Pesudovs K, McAlinden C, Vogel M, Kernt M, Hirneiss C. Reengineering the glaucoma quality of life-15 questionnaire with Rasch analysis. *Invest Ophthalmol Vis Sci.* 2011;52:6971–6977.
19. Mallinson T, Stelmack J, Velozo C. A comparison of the separation ratio and coefficient alpha in the creation of minimum item sets. *Med Care*. 2004;42(1 Suppl):I17–I24.
20. Gothwal VK, Wright TA, Lamoureux EL, Pesudovs K. Measuring outcomes of cataract surgery using the Visual Function Index-14. *J Cataract Refract Surg*. 2010;36: 1181–1188.
21. Pesudovs K, Gothwal VK, Wright T, Lamoureux EL. Remediating serious flaws in the National Eye Institute Visual Function Questionnaire. *J Cataract Refract Surg.* 2010;36:718–732.
22. Mollazadegan K, Huang J, Khadka J, et al. Cross-cultural validation of the National Eye Institute Visual Function Questionnaire. *J Cataract Refract Surg.* 2014;40:774–784.
23. Braithwaite T, Calvert M, Gray A, Pesudovs K, Denniston AK. The use of patient-reported outcome research in modern ophthalmology: impact on clinical trials and routine clinical practice. *Patient Relat Outcome Meas*. 2019;10:9–24.
24. Gothwal VK, Wright TA, Lamoureux EL, Pesudovs K. Rasch analysis of the quality of life and vision function questionnaire. *Optom Vis Sci.* 2009;86:E836–E844.
25. Gothwal VK, Wright TA, Lamoureux EL, Pesudovs K. Activities of daily vision scale: what do the subscales measure? *Invest Ophthalmol Vis Sci.* 2018;51:694–700.
26. Pesudovs K, Wright TA, Gothwal VK. Visual disability assessment: valid measurement of activity limitation and mobility in cataract patients. *Br J Ophthalmol.* 2010;94:777–781.
27. Gothwal VK, Wright T, Lamoureux EL, Pesudovs K. Psychometric properties of visual functioning index using Rasch analysis. *Acta Ophthalmol*. 2010;88:797–803.
28. Gothwal VK, Wright TA, Lamoureux EL, Pesudovs K. Rasch analysis of visual function and quality of life questionnaires. *Optom Vis Sci*. 2009;86:1160–1168.
29. Ningrum E, Evans S, Soh SE. Validation of the Indonesian version of the Safety Attitudes Questionnaire: A Rasch analysis. *PLoS One*. 2019;14:e0215128.
30. McKean-Cowdin R, Varma R, Hays RD, et al. Longitudinal changes in visual acuity and health-related quality of life: the Los Angeles Latino Eye study. *Ophthalmology.* 2010;117:1900–1907, 1907.e1.
31. Pondorfer SG, Terheyden JH, Heinemann M, Wintergerst MWM, Holz FG, Finger RP. Association of vision-related quality of life with visual function in age-related macular degeneration. *Sci Rep*. 2019;9:15326.
32. Kandel H, Pesudovs K, Watson SL. Measurement of quality of life in keratoconus. *Cornea*. 2020;39:386–393.
33. Schmelter V, Dirisamer M, Siedlecki J, et al. Determinants of subjective patient-reported quality of vision after small-incision lenticule extraction. *J Cataract Refract Surg*. 2019;45:1575–1583.

translational vision science & technology

34. Khadka J, Huang J, Chen H, et al. Assessment of cataract surgery outcome using the modified Catquest Short-Form Instrument in China. *PLoS One*. 2016;11:e0164182.

35. Khadka J, Huang J, Mollazadegan K, et al. Translation, cultural adaptation, and Rasch analysis of the visual function (VF-14) questionnaire. *Invest Ophthalmol Vis Sci*. 2014;55:4413–4420.