



Archived at the Flinders Academic Commons:

<http://dspace.flinders.edu.au/dspace/>

'This is the peer reviewed version of the following article:
Jyoti Khadka, Colm McAlinden, Vijaya K. Gothwal, Ecosse L.
Lamoureux, Konrad Pesudovs; The Importance of Rating
Scale Design in the Measurement of Patient-Reported
Outcomes Using Questionnaires or Item Banks. Invest.
Ophthalmol. Vis. Sci. 2012;53(7):4042-4054. doi: 10.1167/
iovs.12-9728.

which has been published in final form at

<http://dx.doi.org/10.1186/1477-7525-10-80>

This article is made available with the permission of the
publisher, Association for Research in Vision and
Ophthalmology

Copyright © Association for Research in Vision and
Ophthalmology

The Importance of Rating Scale Design in the Measurement of Patient-Reported Outcomes Using Questionnaires or Item Banks

Jyoti Khadka,¹ Colm McAlinden,¹ Vijaya K. Gothwal,² Ecosse L. Lamoureux,^{3,4} and Konrad Pesudovs¹

PURPOSE. To investigate the effect of rating scale designs (question formats and response categories) on item difficulty calibrations and assess the impact that rating scale differences have on overall vision-related activity limitation (VRAL) scores.

METHODS. Sixteen existing patient-reported outcome instruments (PROs) suitable for cataract assessment, with different rating scales, were self-administered by patients on a cataract surgery waiting list. A total of 226 VRAL items from these PROs in their native rating scales were included in an item bank and calibrated using Rasch analysis. Fifteen item/content areas (e.g., reading newspapers) appearing in at least three different PROs were identified. Within each content area, item calibrations were compared and their range calculated. Similarly, five PROs having at least three items in common with the Visual Function (VF-14) were compared in terms of average item measures.

RESULTS. A total of 614 patients (mean age \pm SD, 74.1 \pm 9.4 years) participated. Items with the same content varied in their calibration by as much as two logits; “reading the small print” had the largest range (1.99 logits) followed by “watching TV” (1.60). Compared with the VF-14 (0.00 logits), the rating scale of the Visual Disability Assessment (1.13 logits) produced the most difficult items and the Cataract Symptom Scale (0.24 logits) produced the least difficult items. The VRAL item bank was suboptimally targeted to the ability level of the participants (2.00 logits).

CONCLUSIONS. Rating scale designs have a significant effect on item calibrations. Therefore, constructing item banks from existing items in their native formats carries risks to face validity and transmission of problems inherent in existing instruments, such as poor targeting. (*Invest Ophthalmol Vis Sci.* 2012;53:4042–4054) DOI:10.1167/iovs.12-9728

From the ¹National Health and Medical Research Council Centre for Clinical Eye Research, Discipline of Optometry and Vision Science, Flinders Medical Centre and Flinders University of South Australia, Adelaide, Australia; ²Meera and L B Deshpande Centre for Sight Enhancement, Vision Rehabilitation Centres, L V Prasad Eye Institute, Hyderabad, India; ³Centre for Eye Research Australia, Department of Ophthalmology, University of Melbourne, Victoria, Australia; and the ⁴Singapore Eye Research Institute, Singapore National Eye Centre, Singapore.

Supported by NHMRC Grant 1031838.

Submitted for publication February 19, 2012; revised April 30, 2012; accepted May 18, 2012.

Disclosure: **J. Khadka**, None; **C. McAlinden**, None; **V.K. Gothwal**, None; **E.L. Lamoureux**, None; **K. Pesudovs**, None

Corresponding author: Konrad Pesudovs, National Health and Medical Research Council Centre for Clinical Eye Research, Discipline of Optometry and Vision Science, Flinders University, Bedford Park, South Australia, 5042, Australia; Konrad.Pesudovs@flinders.edu.au.

Patient-reported outcomes (PROs) are widely used in health care settings and recently the use of PROs in clinical trials has become a requirement of many funding bodies.^{1–4} This requirement has prompted researchers to develop better, higher quality, and more comprehensive PROs compared with previously developed instruments (legacy instruments). Therefore, the field of patient-centered measurement is rapidly evolving from traditional labor-intensive paper-pencil-based PROs to comprehensive item banks.⁵ Item banks are further benefited by the use of Computer Adaptive Testing (CAT) systems for implementation. CAT offers more flexible, precise, and reliable measurements while requiring fewer items and, consequently, a shorter administration time.^{6,7} Item banks and CAT systems are being developed and implemented in many disciplines in health care.^{8,9}

An item bank is simply a large collection of calibrated items that measures a defined latent trait (e.g., physical disability, symptoms, psycho-social effects of a disease on patients).¹⁰ Items in an item bank are calibrated by using item response theory, such as Rasch analysis.¹¹ However, the first and the most important phase of item bank development is content (item) identification. A plethora of traditional health-specific legacy PROs have been developed in the past 2 decades with existing items from these PROs serving as a rich source of item content; however, item bank developers can take two approaches to using existing content.^{12,13} Either the existing items from available PROs can be adopted in their native rating scale formats (original question format and response categories) or the existing items are modified (i.e., revising and refining item wordings), usually to make all questions in a bank use a consistent format. Existing content can also be supplemented by new content from patient/expert focus groups.^{13,14}

Since 2006, our group has been developing comprehensive item banks in ophthalmology. Initially, we adopted the native rating scale format approach to develop a vision-related activity limitation (VRAL) item bank for cataract. The VRAL item bank had items drawn from 16 different PRO instruments that were either developed or used to assess cataract surgery outcomes.¹⁰ Several of these items had the same content (e.g., reading newspapers, driving by night, recognizing faces at distance) but had different rating scales in their native formats (exactly as in the source PROs). When Rasch analysis was used to calibrate these items, we found that the items with the same content had significantly different item difficulty calibrations. As the content was the same, the only variable that differed across these items was the rating scale design. Therefore, we hypothesized that the differences in item calibrations were attributable to rating scales. The effect of rating scales on item calibration was also noticed in other studies^{15,16}; however, to our knowledge, this issue has not been investigated in detail. Therefore, we used our VRAL item bank data to explore the

actual effect of rating scales on item calibrations. Further, we investigated the impact of rating scales on the overall difference in VRAL scores measured by different PROs.

METHODS

Participants and PROs

Participants were patients on the waiting list for cataract surgery at Flinders Medical Centre, Adelaide, South Australia. All participants were 18 years or older, English speaking, and cognitively able to self-administer PROs. A pack containing 10 PROs randomly selected from the 16 PROs (Table 1) were mailed to the participants for self-administration. The 16 PROs were the Visual Disability Assessment (VDA),¹⁷ the National Eye Institute Visual Function Questionnaire (NEIVFQ),¹⁸ the Cataract Symptom Scale (CatScale),¹⁹ the Impact of Visual Impairment (IVI),²⁰ the Visual Activities Questionnaire (VAQ),²¹ the Visual Symptoms and Quality of Life Questionnaire (VSQ),²² the Technology of Patient Experience (TyPE),²³ the Activities of Daily Vision Scale (ADVS),²⁴ the Quality of Life-Visual Function Questionnaire (QOLVFQ),²⁵ the Houston Vision Assessment Test (HVAT),²⁶ the CatQuest (CATQ),²⁷ the Visual Function and Quality of Life (VFQOL),²⁸ The Impact of Cataract Surgery (ICS),²⁹ The Visual Functioning Index (VFI),³⁰ the Visual Function-14 (VF-14),³¹ and the Distance Visual Impairment Questionnaire (DVI).³² These PROs were chosen because they were either developed to, or have been used to assess the impact of cataract and/or outcomes of cataract surgery.

The VRAL item bank was developed using items from 16 different PROs (Table 1). It is likely to add extra respondent burden and could compromise the quality of data, if the participants had to complete all the 16 PROs. Therefore, each participant was sent a pack of 10 PRO instruments. However, it is required to calibrate all the items from the 16 PROs on a common metric scale to feed into the item bank. This requires a common set of linking items to link with all other items, which we refer as “anchoring.” The common set of items we used in this study was the items of the VDA. Therefore, each pack contained the VDA and 9 PROs randomly selected from other 15 PROs. To ensure data that could be linked, all the participants completed the VDA first followed by as many of the other instruments as they could complete. All the participants completed at least the VDA and it was the instrument that was completed before the other PRO instruments.

The study adhered to the ethical standards laid by the tenants of the 2008 Declaration of Helsinki and was approved by the Flinders Clinical Ethics Committee. All participants provided written informed consent before their inclusion in the study.

Vision-Related Activity Limitation Item Bank

A total of 226 items from the 16 legacy PROs measuring VRAL in their native rating scale format were included in the item bank. Data collected for these items were pooled together for Rasch analysis to calibrate the 226 items.

Rasch Analysis

Rasch analysis is a probabilistic mathematical model that estimates a person's ability and item difficulty on a common measurement continuum defined by an interval scale, expressed in log-odd units (logits).³³ A logit is a unit of measurement to report relative differences between a person's ability and item difficulty in an equal interval scale. At a detailed level, Rasch analysis estimates threshold values for response categories in logits. The threshold is the point at which the ability required to endorse either of the adjacent categories is equal; that is, the transition point from one category to the next. The number of thresholds is always one less than the number of categories; for example, for five-response categories the number of thresholds is four. Each threshold has its own difficulty estimate. Difficulty estimates of the thresholds are expected to increase monotonically, that is, the response categories should represent increasing levels of underlying trait being measured from lower to higher category. For a threshold “*k*” with the difficulty value “*F_k*” the probability of choosing category “*k*” is calculated by the Rasch model by the following equation³³:

$$P_{nik} = \frac{e^{(B - [D_i + F_k])}}{1 + e^{(B - [D_i + F_k])}}$$

where, P_{nik} = probability of person “*n*” choosing category “*k*” over the adjacent category on any item “*I*,” B = person ability, F_k = difficulty of threshold k , D_i = item difficulty, and $D_i + F_k$ is the item difficulty added to the threshold difficulty indicating the effect of threshold difficulty on item i . The difficulty estimates of the threshold k are estimated across the entire set of items once in the Andrich Rating scale model. The same logic is followed to calculate difficulty estimates for other thresholds. If the assumption is met that all items have equal

TABLE 1. Sixteen PRO Instruments, Number of Items, and Response Categories Used to Develop a Vision-Related Activity Limitation Item Bank

PROs	Number of Items	Response Categories
Visual Disability Assessment, VDA (Pesudovs and Coster, 1998) ¹⁷	18	4
Visual Function-14, VF-14 (Steinberg et al., 1994) ³¹	14	5
Visual Functioning index, VFI (Bernth-Petersen, 1981) ³⁰	8	2 and 3*
Technology of Patient Experiences, TyPE (Javitt et al., 1997) ²³	12	5
Distance Visual Impairment Questionnaire, DVI (Haase and Bryant, 1973) ³²	5	2
Activities of Daily Vision Scale, ADVS (Mangione et al., 1992) ²⁴	21	5
Cataract Symptom Scale, CatScale (Crabtree et al., 1999) ¹⁹	12	5
Visual Activities Questionnaire, VAQ (Sloane et al., 1992) ²¹	27	5
Quality of Life and Visual Function Questionnaire, QOLVFQ (Carta et al., 1998) ²⁵	11	3
Visual Function and Quality of Life, VF&QOL (Fletcher et al., 1997) ²⁸	17	4
Impact of Vision Impairment, IVI (Hassell et al., 2000) ²⁰	21	6
Catquest, CATQ (Lundstrom et al., 1997) ²⁷	10	4
Visual Symptoms and Quality of Life Questionnaire, VSQ (Donovan et al., 2003) ²²	14	2, 3, 4, 5, and 7†
National Eye Institute Visual Function Questionnaire, NEIVFQ (Mangione et al., 1998) ¹⁸	21	5
Houston Vision Assessment Test, HVAT (Prager et al., 2000) ²⁶	10	5
Impact of Cataract Surgery, ICS (Monestam and Wachtmeister, 1999) ²⁹	5	4
Total items	226	

The following PROs had a different number of response categories across their items:

* VFI: items 3, 4, 5, 6, 7, and 8 = 2-category option; items 1 and 2 = 3-category option.

† VSQ: items 17, 21, and 23 = 2-category option; item 16 = 3-category option; items 2 and 15 = 4-category option; items 1, 3, 5, 6, 7, 8, and 18 = 5-category-option; item VI = 7-category option.

estimates (i.e., item calibration) and the person's ability. For this map, more difficult items and more able participants are located at the bottom of the map. As can be observed on the map, items with a similar level of difficulty are expected to lie on the same line on the map and have similar item calibration (e.g., item 16, "move in unfamiliar surroundings" and item 8, "driving a car by the day" were of the same difficulty). The 18-item VDA was slightly mistargeted to the study population (difference between person and item means = 1.4). Within the VRAL item bank, items with the same content are also expected to demonstrate the same or nearly equal calibrations and, therefore, the same locations on the person-item map. Whether items with the same content calibrate at similar or different levels of item difficulty can, therefore, be demonstrated on person/item maps.

Analysis

To further investigate the impact of rating scales on item calibration, we specifically searched among the 226 items and grouped items with the same content areas. Items would qualify for further analysis only if the content appeared in at least three items across three different PROs. We identified 57 items across 15 different content areas from 13 PROs (Appendix 1). Item difficulty calibrations on a common metric scale were assessed.

Rasch analysis was performed according to the Andrich Rating scale model using joint likelihood estimation with Winsteps (version 3.67.0, Chicago, IL) software. Combined Rasch analysis was carried out on the 226 items measuring VRAL (as a single large questionnaire). The number of response categories ranged between two and seven across 226 items (Table 1). To calibrate items on a single common metric scale, the polarity of the response categories was modified such that all the items were scaled in the same direction. This was done by arranging the first response as the most positive and the last response as the most negative response for all the items. For this analysis, a positive item indicates that the item requires a lower level of ability than the average, or the item is less difficult and vice versa. Item calibration (difficulty estimates in logits) and locations of the items were assessed. The range of the item calibrations within each content area was also calculated.

Similarly, we also searched among the 13 PROs for those that had at least three common item/content pairs. We identified five PROs (CatScale, TyPE, ADVS, NEIVFQ, and VDA) having at least three item/content pairs in common with the VF-14 (Appendix 2).

For the five PROs with at least three common item/content pairs, the difference between the item calibration with each PRO and with the VF-14 was calculated. Using the common items, a simple average estimate of item difficulty (VRAL score) for each PRO was calculated. The difference in VRAL scores (i.e., relative item difficulty in reference to the VF-14) between each PRO and the VF-14 was calculated and the values were compared.

RESULTS

A total of 614 patients participated. The mean age of the participants was 74.1 years (SD \pm 9.4) and 56% were female. Among the 614 patients, 59% had bilateral cataracts, 41% were awaiting second eye surgery, and 51% had ocular comorbidities (glaucoma, 16%; age-related macular degeneration, 9%; and diabetic retinopathy, 4%). The participants had been diagnosed with cataract for an average of 3.2 ± 8.7 years. The mean visual acuity was 0.22 ± 0.20 logarithm of the minimum angle of resolution (LogMAR) ($\sim 6/9.5^{-1}$) in the better eyes and 0.55 ± 0.36 LogMAR ($\sim 6/24^{+2}$) in the worse eyes. The systemic comorbidities of the study population were representative of an elderly cataract population in Australia.⁴⁵

Figure 2 shows the person-item map of the VRAL item bank. Items at the top of the map (DVI1 and DVI2) were least difficult

items, whereas items near the bottom the map (VAQ15, DIVI5) were the most difficult items. The highlighted items were those with the content "reading small print" but with varying question formats and response categories; these were widely spread out on the map. The VRAL item bank also demonstrated suboptimal targeting to the study population (difference between person and item means was 2 logits).

Table 2 shows the item measure estimates of the 57 items across 15 content areas derived from the 13 PRO instruments (Appendix 1); on average, item calibrations ranged over 1.07 logits for each content area. The widest spread of item calibration was observed for the items with the content "reading small print" (1.99 logits), followed by "watching TV" (1.6 logits). The narrowest range was observed for items with the content "reading newspapers and books" (0.17). For 9 (60%) of 15 content areas, the range was more than 1 logit. Among the remaining six content areas, only half had a range less than 0.5 logit. When compared, most of the items of the VDA had greater difficulty (more negative item difficulty estimates) than the items of other PROs.

Content contamination (items with mixed content versus a single content) was identified as having an additional confounding effect on item calibration. For example, in Figure 2, the multibarreled hobby item NEIVFQ6 ("Doing work or hobbies that require you to see well up close, such as cooking, sewing, fixing things around the house, or using hand tools," -0.80 logits) was noticeably more difficult than the single-barreled hobby item IVI2 ("Favorite pastime hobbies," -0.33 logits).

We also investigated the impact of rating scales on an overall difference in VRAL score, measured by six different PROs. The rating scales of the CatScale, TyPE, ADVS, NEIVFQ, and VDA produced more difficult items than the rating scale of the VF-14 (Fig. 3). Among these six PROs, the rating scale of the VDA (question format: "to what extent, if at all, does your vision interfere with your ability to...?") and response categories: "not at all, a little, quite a bit, and a lot") produced the most difficult items (average item measure was 1.13 logits more difficult than the VF-14). In terms of difficulty, the VDA was followed by the rating scales of the NEIVFQ (0.66 logits), ADVS (0.55 logits), TyPE (0.43), and CatScale (0.24) (Fig. 3).

DISCUSSION

This study has provided empirical evidence that differences in rating scale formats have a significant effect on item calibrations beyond item content. A large number of PROs and rating scales were used that enabled a head-to-head comparison at item-level. We found that the effect of rating scales on item calibrations could be marked, as demonstrated by the spread of item difficulty estimates between the items with the same content. This implies that using a common rating scale across all items measuring an underlying trait provides an advantage of allowing item calibration based on item content only. An item bank, or questionnaire, where all items use the same rating scale design, and items vary in difficulty only according to their content, is an elegant design that provides a level of face validity for the potential user. An item bank comprising legacy items in their native rating scale format can still be used to measure the underlying latent trait. The measurement properties of the bank may not be degraded per se, but it may be more difficult to see where items are required to further optimize measurement.

The influence of rating scales on item calibration was also briefly reported by Massof.^{15,16} He observed this effect when he merged data from four different PROs (ADVS, NEIVFQ, VF-



EACH "#" IS 5. EACH "." IS 1 TO 4

FIGURE 2. Person-item map of the 226 legacy items measuring vision-related activity limitation. Items are located on the *right* and participants (represented by “#” and “.”) are located on the *left* of the *dashed line*. Items are denoted by the abbreviation of the source PROs followed by the original item number. Difficult items and more able participants are located at the *bottom* of the map. Items highlighted by *gray shades* are the items having the similar content “reading small print,” note these are spread over a wide range. M = mean; S = 1 SD from the mean; T = 2 SDs from the mean. The M, S, and T are shown for both the participants and the items on either side of the *dashed line*. The difference between participant and item means is 2.00 logits. VDA, Visual Disability Assessment; NEIVFQ, National Eye Institute Visual Function Questionnaire; CatScale, Cataract Symptom Scale; IVI, Impact of Visual Impairment; VAQ, Visual Activities Questionnaire; VSQ, Visual Symptoms and Quality of Life Questionnaire; TyPE, Technology of Patient Experience; ADVS, Activities of Daily Vision Scale; QOLVFQ, Quality of Life-Visual Function Questionnaire; HVAT, Houston Vision Assessment Test; Catquest, CATQ; VFQOL, Visual Function and Quality of Life; ICS, Impact of Cataract Surgery; VFI, Visual Functioning Index; VF-14, Visual Function-14; DVI, Distance Visual Impairment Questionnaire.

14, and VAQ). He argued that the effect was probably due to the subtle differences in rating scales between the PROs.^{15,16} Our findings also support his argument. As defined in the methods section, the difference in item difficulty estimates can be attributed to the difference in rating scale designs between the items with the same content. We observed that question formats (i.e., how an item was asked) probably has more effect on item calibrations than wording and number of response categories. For example, the items that had similar question formats (CatScale item 9: “Do you have difficulty reading books or newspaper?” VSQ item 1: “Do you have difficulty reading normal print in books or newspapers because of trouble with your eyesight?” and VF-14 item 2: “Do you have difficulty, even with glasses, reading a newspaper or a book?”) shared similar item calibrations (Table 2). The differences in item calibrations were found to be higher when items with the same content

were stated differently. For example, the VDA item 4 (“To what extent, if at all, does your vision interfere with your ability to watch television?”) had much higher item calibration (−0.85 logits) than the CatScale item 2 (0.25 logits) (“Do you have difficulty watching television?”), despite having only subtle differences in response categories. However, the item calibration of the VDA item 4 (−0.85 logits) was closer to the HVAT item 9a/b (−0.34 logits) (“To what extent is your watching television impaired?”) probably due to the similar question format, despite having completely different response categories in terms of number and wording (Table 2 and Appendix 1). This observation was consistent among most of the items across the 15 content areas. In our study, we found a much higher effect of rating scales on item calibration than in Massof’s study.¹⁶ It was possibly because we compared the ratings scales of 13 PROs compared with 4 PROs in the study

TABLE 2. Comparison of Item Measures and Their Range Values (in logits) of the Items Distributed across 15 Different Contents, Identified from the 13 PRO Instruments

PROs	Street Signs/														
	Small Print	News Papers	Newspapers and Books	TV	Driving Day	Driving Night	Faces Distance	Faces Near	Store Names	Pouring Liquids	Peripheral Vision	Color Vision	Steps	Crossing the Road	Public Transport
VDA	—	—	—	-0.8	0.06	-1.88	-1.19	—	—	—	—	—	—	0.11	0.2
NEIVFQ	-1.57	-0.97	—	—	—	-1.34	—	-0.34	—	—	0.15	—	—	—	—
CatScale	-1.73	—	-0.7	0.25	—	0	—	—	1.53	—	—	—	0.02	0.89	—
IVI	-0.51	-0.78	—	—	—	—	—	—	—	—	—	—	-0.27	—	0.6
VAQ	-2.5	—	—	—	—	—	—	—	-0.01	—	-0.63	0.65	—	—	—
VF-14	-1.12	—	-0.53	0.45	1.07	-0.75	—	0.4	—	—	—	—	—	—	—
VSQ	—	—	-0.61	—	—	—	—	0.07	1.15	—	—	—	—	—	—
TYPE	—	—	—	0.04	0.56	-1.12	-0.02	—	—	—	—	—	—	—	—
ADVS	—	-0.82	—	0.8	0.01	-1.57	—	-0.26	—	—	—	—	-1.21	—	0.48
QOIVQ	—	—	—	—	—	—	0.2	—	—	—	—	1.9	—	1.08	—
HVAI	—	—	—	-0.34	0.29	—	—	—	—	—	—	—	—	—	—
CATQ	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
VFQOL	—	—	—	—	—	-0.41	—	—	—	—	0.27	1.36	—	—	—
Range (logit)	1.99	0.19	0.17	1.60	1.06	1.13	1.39	0.74	1.54	1.54	0.9	1.25	1.23	0.97	0.4

Higher range indicates a wider spread of item calibration.

by Massof.¹⁶ Our study had also included the four PROs used in Massof's study.¹⁶ As expected, we found in our study that items from these four PROs had the same hierarchy of item calibrations as reported by Massof.^{15,16}

Our VRAL item bank also demonstrated suboptimal targeting to the study population, despite having a large volume of items. Interestingly, all of these PROs, except the VDA, demonstrated suboptimal targeting to the study population in previous studies.^{43,44,46-50} This implies that our VRAL item bank had inherited the targeting problems associated with these PROs. This was probably because the content of these PROs was developed when cataract surgery was performed at a later stage of cataract maturation and disability than occurs today.⁵¹ Therefore, these item contents were outdated and were not able to capture the important issues that matter to people with cataracts at the present time. This was reflected by poor targeting demonstrated by our VRAL item bank. This highlights a key disadvantage of using legacy items alone to form an item bank. To improve the targeting, more difficult items that match the population should be added to the item bank. A better overall strategy for developing an item bank would be to extract content from the existing items, supplement with content derived from patient focus groups, and rewrite all as items with a common rating scale. This argument is also supported by previous studies that have reported that rewritten items provided better information, targeting, and increased measurement precision when compared with the existing items.^{1,52}

Besides the rating scales, we also observed that items with single and mixed contents demonstrated different item calibration. For example, items with the content "reading newspapers" were found to be more difficult than "reading newspapers and books." This was probably because books have a larger print size than newspapers. Therefore, when "reading newspapers" was combined with books, the task was perceived easier than just "reading newspapers." Even though the differences in item measures were not very high, they were more pronounced when item calibration was compared between single- and multibarreled items.

When compared with the VF-14, we found that the rating scales of the five other PROs produced more difficult item calibrations for the same items (Fig. 3). The VF-14 provided the easiest ratings. This was probably because the VF-14 had two-part questions for each item (Part 1: "Do you have difficulty, even with glasses...?" Yes; No; Not applicable), where respondents could answer that they had no difficulty before they knew other options were available (Part 2: "if yes, how much difficulty do you currently have"? A little = 1; A moderate amount = 2; A great deal = 3, Unable to do the activity = 4). Therefore someone with "a little difficulty" may choose "no difficulty" rather than "yes difficulty" because it is only a small problem. In this way, the VF-14 may underestimate VRAL; whereas, in other PROs, rating scales were laid out in such a way that directly allowed respondents to choose options appropriate to them. Among all the PROs, the VDA gave the most difficult items, probably because the VDA items really push respondents to admit any disability at all (e.g., "To what extent, if at all, does your vision interfere with your ability to...?" Not at all = 1; A little = 2; Quite a bit = 3; A lot = 4). This was probably the reason why the VDA was a better-targeted PRO to a cataract population than other PROs.⁴⁴ For this analysis, the VF-14 was selected as the reference PRO because it shared a maximum number of common items with other PROs. Although this approach neatly estimates the relative difficulty of the rating scale used in each PRO, this approach is limited in that it does not use the same items for all PROs and is restricted to a small number of items. Therefore, the logit

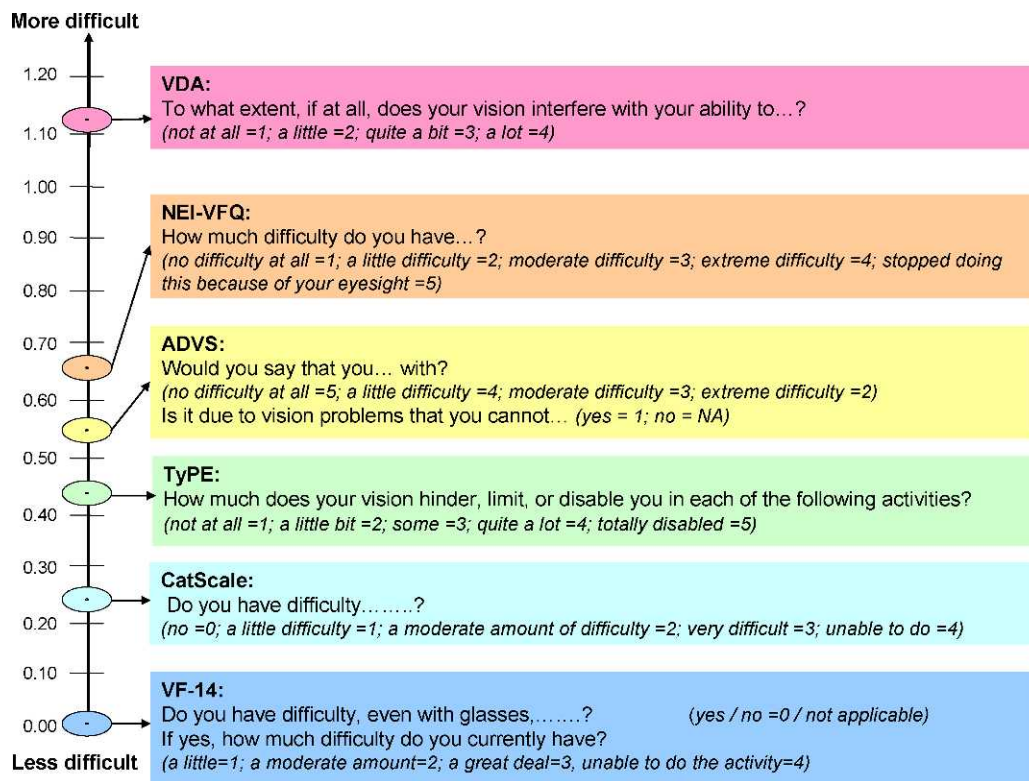


FIGURE 3. The position of the five PROs (VDA, NEIVFQ, ADVS, TyPE, and CatScale) on a linear difficulty scale in logits, in reference to the VF-14. All instruments are relative to the VF-14, which is rated at 0 logits.

differences calculated herein are a guide only and do not represent conversion factors that could be used to equate different PROs.

We acknowledge that this study used items that measure vision-related disability only; consequently, the generalizability of the results to other latent traits (symptoms/emotional well-being) and other disciplines in health care may not be appropriate. However, the concept of rating scales and measuring patient reports crosses the boundary of underlying traits being measured and the study population used. Therefore, we expect that the results of this study would likely apply to other latent traits and disciplines in health care. The other potential limitation of the study was that the participants self-administered multiple PROs. The VDA was used as the anchoring PRO instrument; therefore, it was always completed first. This might have introduced a fatigue effect in their responses, particularly to the PROs that were administered after a few had been completed. However, the ordering of other PROs was randomly varied to control for the fatigue effect.

In conclusion, we have demonstrated that differences in rating scale design could have a significant effect on item calibrations beyond item content. Question format and response category labels appeared to influence item calibrations and, ultimately, overall measurement by PROs. Therefore, it would be difficult to directly compare research findings using different PROs. Moreover, it would be inelegant to use items from different PROs in their native rating scale formats to develop an item bank in which it is desirable that item calibration reflects item content only. A preferred strategy would be to extract content from the existing items, refine the wording, add new content from other sources (e.g., patient focus groups) and fit all the items to a simple, uniform, and common rating scale.

Acknowledgments

The authors acknowledge Thomas A. Wright for his contributions in recruiting patients, posting the questionnaires, and collating data for analysis.

References

1. Fenwick E, Rees G, Pesudovs K, et al. Social and emotional impact of diabetic retinopathy: a review. *Clin Experiment Ophthalmol*. 2012;40:27-38.
2. Hobart JC, Cano SJ, Zajicek JP, Thompson AJ. Rating scales as outcome measures for clinical trials in neurology: problems, solutions, and recommendations. *Lancet Neurol*. 2007;6:1094-1105.
3. Pesudovs K. Patient-centred measurement in ophthalmology—a paradigm shift. *BMC Ophthalmol*. 2006;6:25.
4. Weinreb RN, Kaufman PL. The glaucoma research community and FDA look to the future: a report from the NEI/FDA CDER Glaucoma Clinical Trial Design and Endpoints Symposium. *Invest Ophthalmol Vis Sci*. 2009;50:1497-1505.
5. Cella D, Gershon R, Lai JS, Choi S. The future of outcomes measurement: item banking, tailored short-forms, and computerized adaptive assessment. *Qual Life Res*. 2007;16:133-141.
6. Wainer H, Dorans N, Flaugher R. *Computerized Adaptive Testing: A Primer*. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates; 2000.
7. Bjorner JB, Chang CH, Thissen D, Reeve BB. Developing tailored instruments: item banking and computerized adaptive assessment. *Qual Life Res*. 2007;16:95-108.
8. Cella D, Riley W, Stone A, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health

- outcome item banks: 2005-2008. *J Clin Epidemiol.* 2010;63:1179-1194.
9. Fenwick EK, Pesudovs K, Khadka J, et al. The impact of diabetic retinopathy on quality of life: qualitative findings from an item bank development project [published online ahead of print January 31, 2012]. *Qual Life Res.* 2012. doi:10.1007/s11136-012-0110-1.
 10. Pesudovs K. Item banking: a generational change in patient-reported outcome measurement. *Optom Vis Sci.* 2010;87:285-293.
 11. Irwin DE, Stucky BD, Langer MM, et al. PROMIS Pediatric Anger Scale: an item response theory analysis. *Qual Life Res.* 2012;21:697-706.
 12. Hays RD, Liu H, Spritzer K, Cella D. Item response theory analyses of physical functioning items in the medical outcomes study. *Med Care.* 2007;45:S32-38.
 13. DeWalt DA, Rothrock N, Yount S, Stone AA. Evaluation of item candidates: the PROMIS qualitative item review. *Med Care.* 2007;45:S12-21.
 14. Vogels AG, Jacobusse GW, Reijneveld SA. An accurate and efficient identification of children with psychosocial problems by means of computerized adaptive testing. *BMC Med Res Methodol.* 2011;11:111.
 15. Massof RW. Application of stochastic measurement models to visual function rating scale questionnaires. *Ophthalmic Epidemiol.* 2005;12:103-124.
 16. Massof RW. An interval-scaled scoring algorithm for visual function questionnaires. *Optom Vis Sci.* 2007;84:689-704.
 17. Pesudovs K, Coster DJ. An instrument for assessment of subjective visual disability in cataract patients. *Br J Ophthalmol.* 1998;82:617-624.
 18. Mangione CM, Lee PP, Pitts J, Gutierrez P, Berry S, Hays RD. Psychometric properties of the National Eye Institute Visual Function Questionnaire (NEI-VFQ). NEI-VFQ Field Test Investigators. *Arch Ophthalmol.* 1998;116:1496-1504.
 19. Crabtree HL, Hildreth AJ, O'Connell JE, Phelan PS, Allen D, Gray CS. Measuring visual symptoms in British cataract patients: the cataract symptom scale. *Br J Ophthalmol.* 1999;83:519-523.
 20. Hassell JB, Weih LM, Keeffe JE. A measure of handicap for low vision rehabilitation: the impact of vision impairment profile. *Clin Experiment Ophthalmol.* 2000;28:156-161.
 21. Sloane ME, Ball K, Owsley C, Bruni JR, Roenker DL. The Visual Activities Questionnaire: Developing an instrument for assessing problems in everyday visual tasks. *Technical Digest, Noninvasive Assessment of the Visual System, Topical Meeting of the Optical Society of America.* January 26-28, 1992. Santa Fe, NM.
 22. Donovan JL, Brookes ST, Laidlaw DA, Hopper CD, Sparrow JM, Peters TJ. The development and validation of a questionnaire to assess visual symptoms/dysfunction and impact on quality of life in cataract patients: the Visual Symptoms and Quality of Life (VSQ) Questionnaire. *Ophthalmic Epidemiol.* 2003;10:49-65.
 23. Javitt JC, Wang F, Trentacost DJ, Rowe M, Tarantino N. Outcomes of cataract extraction with multifocal intraocular lens implantation: functional status and quality of life. *Ophthalmology.* 1997;104:589-599.
 24. Mangione CM, Phillips RS, Seddon JM, et al. Development of the 'Activities of Daily Vision Scale'. A measure of visual functional status. *Med Care.* 1992;30:1111-1126.
 25. Carta A, Braccio L, Belpoliti M, et al. Self-assessment of the quality of vision: association of questionnaire score with objective clinical tests. *Curr Eye Res.* 1998;17:506-511.
 26. Prager TC, Chuang AZ, Slater CH, Glasser JH, Ruiz RS. The Houston Vision Assessment Test (HVAT): an assessment of validity. The Cataract Outcome Study Group. *Ophthalmic Epidemiol.* 2000;7:87-102.
 27. Lundstrom M, Roos P, Jensen S, Fregell G. Catquest questionnaire for use in cataract surgery care: description, validity, and reliability. *J Cataract Refract Surg.* 1997;23:1226-1236.
 28. Fletcher AE, Ellwein LB, Selvaraj S, Vijaykumar V, Rahmathullah R, Thulasiraj RD. Measurements of vision function and quality of life in patients with cataracts in southern India. Report of instrument development. *Arch Ophthalmol.* 1997;115:767-774.
 29. Monestam E, Wachtmeister L. Impact of cataract surgery on visual acuity and subjective functional outcomes: a population-based study in Sweden. *Eye (Lond).* 1999;13:711-719.
 30. Bernth-Petersen P. The effectiveness of cataract surgery. A retrospective study. *Acta Ophthalmol (Copenh).* 1981;59:50-56.
 31. Steinberg EP, Tielsch JM, Schein OD, et al. The VF-14. An index of functional impairment in patients with cataract. *Arch Ophthalmol.* 1994;112:630-638.
 32. Haase KW, Bryant EE. Development of a scale designed to measure functional vision loss using an interview technique. *Proc Am Stat Assoc.* 1973;SS:274-279.
 33. Bond TG, Fox CM. *Applying the Rasch Model: Fundamental Measurement in the Human Science.* 2nd ed. London: Lawrence Erlbaum Associates; 2007.
 34. Muraki E. A generalized partial credit model: application of an EM algorithm. *Appl Psych Meas.* 1992;16:159-176.
 35. Massof RW. The measurement of vision disability. *Optom Vis Sci.* 2002;79:516-552.
 36. Mallinson T. Why measurement matters for measuring patient vision outcomes. *Optom Vis Sci.* 2007;84:675-682.
 37. Linacre JM. Optimizing rating scale category effectiveness. *J Appl Meas.* 2002;3:85-106.
 38. Khadka J, Ryan B, Margrain TH, Court H, Woodhouse JM. Development of the 25-item Cardiff Visual Ability Questionnaire for Children (CVAQC). *Br J Ophthalmol.* 2010;94:730-735.
 39. McAlinden C, Pesudovs K, Moore JE. The development of an instrument to measure quality of vision: the Quality of Vision (QoV) questionnaire. *Invest Ophthalmol Vis Sci.* 2010;51:5537-5545.
 40. McAlinden C, Gothwal VK, Khadka J, Wright TA, Lamoureux EL, Pesudovs KA. Head-to-head comparison of 16 cataract surgery outcome questionnaires. *Ophthalmology.* 2011;118:2374-2381.
 41. 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.
 42. Haley SM, Ni P, Jette AM, et al. Replenishing a computerized adaptive test of patient-reported daily activity functioning. *Qual Life Res.* 2009;18:461-471.
 43. Pesudovs K, Gothwal VK, Wright T, Lamoureux EL. Remediate serious flaws in the National Eye Institute Visual Function Questionnaire. *J Cataract Refract Surg.* 2010;36:718-732.
 44. 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.
 45. Kirkwood BJ, Pesudovs K, Latimer P, Coster DJ. The efficacy of a nurse-led preoperative cataract assessment and postoperative care clinic. *Med J Aust.* 2006;184:278-281.
 46. Lamoureux EL, Pesudovs K, Thumboo J, Saw SM, Wong TY. An evaluation of the reliability and validity of the visual functioning questionnaire (VF-11) using Rasch analysis in an Asian population. *Invest Ophthalmol Vis Sci.* 2009;50:2607-2613.
 47. 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.

48. Pesudovs K, Garamendi E, Keeves JP, Elliott DB. The Activities of Daily Vision Scale for cataract surgery outcomes: re-evaluating validity with Rasch analysis. *Invest Ophthalmol Vis Sci.* 2003;44:2892-2899.
49. Gothwal VK, Wright TA, Lamoureux EL, Pesudovs K. Using Rasch analysis to revisit the validity of the Cataract TyPE Spec instrument for measuring cataract surgery outcomes. *J Cataract Refract Surg.* 2009;35:1509-1517.
50. Gothwal VK, Wright TA, Lamoureux EL, Pesudovs K. Cataract Symptom Scale: clarifying measurement. *Br J Ophthalmol.* 2009;93:1652-1656.
51. Lundstrom M, Pesudovs K. Questionnaires for measuring cataract surgery outcomes. *J Cataract Refract Surg.* 2011;37:945-959.
52. Bruce B, Fries JF, Ambrosini D, et al. Better assessment of physical function: item improvement is neglected but essential. *Arthritis Res Ther.* 2009;11:R191.

APPENDIX 1. The 15 Different Content Areas and the Items That Were Obtained from the 13 PRO Instruments (ADVS, Activities of Daily Vision Scale; CatScale, Cataract Symptom Scale; Catquest= CATQ; HVAT, Houston Vision Assessment Test; IVI, Impact of Visual Impairment; NEIVFQ, National Eye Institute Visual Function Questionnaire; QOLVFQ, Quality of Life-Visual Function Questionnaire; TyPE, Technology of Patient Experience; VAQ, Visual Activities Questionnaire; VDA, Visual Disability Assessment; VFQOL, Visual Function and Quality of Life; VF-14, Visual Function-14; VSQ, Visual Symptoms and Quality of Life Questionnaire). Items are Denoted by the Abbreviation of the Source PRO Followed by the Original Item Number.

Content 1. Reading Small Print

S.No	PRO_item Number	Item/Response Option
1	CatScale_1	Do you have difficulty reading small print? (No; A little difficulty; A moderate amount of difficulty; Very difficult; Unable to do)
2	VAQ_7	I have problems reading small print (for example, phone book, newspapers). (Never; Rarely; Sometimes; Often; Always)
3	NEIVFQ_A3	Wearing glasses, how much difficulty do you have reading the small print in a telephone book, on a medicine bottle, or on legal forms? Would you say: (No difficulty at all; A little difficulty; Moderate difficulty; Extreme difficulty; Stopped doing this because of your eyesight)
4	IVI_13	In the past month, how much has your eyesight interfered with reading labels or instructions on medicines? (Not at all; Hardly at all; A little; A fair amount; A lot; Can't do because of eyesight)
5	VF-14_1	Do you have difficulty, even with glasses, reading small print, such as labels on medicine bottles, a telephone book, and food labels? (Yes/No/Not applicable) If yes, how much difficulty do you currently have? (A little; A moderate amount; A great deal)

Content 2. Reading newspapers

1	NEIVFQ_5	How much difficulty do you have reading ordinary print in newspapers? Would you say you have: (No difficulty at all; A little difficulty; Moderate difficulty; Extreme difficulty; Stopped doing this because of your eyesight)
2	IVI_7	In the past month, how much has your eyesight interfered with reading ordinary size print? (for example newspapers) (Not at all; Hardly at all; A little; A fair amount; A lot; Can't do because of eyesight)
3	ADVS_11bc	Would you say that you read the ordinary print in newspapers with: (PLEASE CHECK ONLY 1 ANSWER) (No difficulty at all; A little difficulty; Moderate difficulty; Extreme difficulty) Is it because of visual problems that you cannot read the ordinary print in newspapers? (Yes; No; Not applicable)

Content 3. Reading books or newspapers

1	CatScale_9	Do you have difficulty reading books or newspapers? (No; A little difficulty; A moderate amount of difficulty; Very difficult; Unable to do)
2	VSQ_1	Do you have difficulty reading normal print in books or newspapers because of trouble with your eyesight? (No difficulty; Yes, a little difficulty; Yes, some difficulty; Yes, a great deal of difficulty; I cannot read because of my eyesight)
3	VF-14_2	Do you have difficulty, even with glasses, reading a newspaper or a book? (Yes; No; Not applicable) If yes, how much difficulty do you currently have? (A little; A moderate amount; A great deal, Are you unable to do the activity)

Content 4. Watching television

1	VDA_4	To what extent, if at all, does your vision interfere with your ability to watch television? (Not at all; A little; Quite a bit; A lot)
2	CatScale_2	Do you have difficulty watching television? (No; A little difficulty; A moderate amount of difficulty; Very difficult; Unable to do)

APPENDIX 1. Continued

S.No	PRO_item Number	Item/Response Option
3	ADVS_9b/c	Would you say that you are able to see television with: (PLEASE CHECK ONLY 1 ANSWER) <i>(No difficulty at all; A little difficulty; Moderate difficulty; Extreme difficulty)</i> Is it because of visual problems that you are unable to watch television? <i>(Yes; No; Not applicable)</i>
4	HVAT_9a/b	To what extent is your watching television impaired? <i>(Not at all limited; Slightly limited; Somewhat limited; Moderately limited; Severely limited)</i> If there are limitations, how much is because of eyesight? <i>(I have no visual or other physical limitations; None due to eyesight; Some due to eyesight; Half due to eyesight; Most due to eyesight; All due to eyesight)</i>
5	VF-14_12	Do you have difficulty, even with glasses, watching television ? <i>(Yes; No; Not applicable)</i> If yes, how much difficulty do you currently have? <i>(A little; A moderate amount; A great deal; Are unable to do the activity)</i>
6	TYPE_7	How much does your vision hinder, limit, or disable you in each of the following activities? Watching television <i>(Not at all; A little bit; Some; Quite a lot; Totally disabled)</i>
Content 5. Driving by day		
1	VDA_8	To what extent, if at all, does your vision interfere with your ability to drive a car by day ? <i>(Not at all; A little; Quite a bit; A lot)</i>
2	ADVS_2b	Would you say that you drive during the day with: (PLEASE CHECK ONLY 1 ANSWER) <i>(No visual difficulty at all; A little difficulty because of vision; Moderate difficulty because of vision; Extreme difficulty because of vision)</i> Is it because of visual problems that you are unable to drive during the day? <i>(Yes; No)</i>
3	HVAT_3a/b	To what extent is your driving during the day impaired? <i>(Not at all limited; Slightly limited; Somewhat limited; Moderately limited; Severely limited)</i> If there are limitations, how much is because of eyesight? <i>(I have no visual or other physical limitations; None due to eyesight; Some due to eyesight; Half due to eyesight; Most due to eyesight; All due to eyesight)</i>
4	VF-14_13	How much difficulty do you have driving during the day because of your vision? Do you have: <i>(No difficulty; A little difficulty; A moderate amount of difficulty; A great deal of difficulty)</i>
5	TYPE_8	How much does your vision hinder, limit, or disable you in each of the following activities? Daytime driving <i>(Not at all; A little bit; Some; Quite a lot; Totally disabled)</i>
Content 6. Driving by night		
1	VDA_9	To what extent, if at all, does your vision interfere with your ability to drive a car by night ? <i>(Not at all; A little; Quite a bit; A lot)</i>
2	NEIVFQ_16	How much difficulty do you have driving at night ? Would you say you have: <i>(No difficulty at all; A little difficulty; Moderate difficulty; Extreme difficulty; Have you stopped doing this because of your eyesight)</i>
3	ADVS_1b	Would you say you drive at night with: <i>(No difficulty at all; 4. A little difficulty [go to 1d]; 3. Moderate difficulty [go to 1d]; 2. Extreme difficulty [go to 1d])</i> Is it because of your visual problems that you are unable to drive at night? <i>(1. Yes [go to 2a]; 2. No [go to 2a])</i>
4	VF-14_14	How much difficulty do you have driving at night because of your vision? Do you have: <i>(No difficulty = 1; A little difficulty = 2; A moderate amount of difficulty = 3; A great deal of difficulty = 4)</i>
5	TYPE_9	How much does your vision hinder, limit, or disable you in each of the following activities? Nighttime driving <i>(Not at all = 1; A little bit = 2; Some = 3; Quite a lot = 4; Totally disabled = 5; Don't do for other reasons = 0)</i>
Content 7. Face recognition at distance		
1	VDA_3	To what extent, if at all, does your vision interfere with your ability to recognize faces across the street ? <i>(Not at all = 1; A little = 2; Quite a bit = 3; A lot = 4)</i>
2	VFQOL_3	How much problem do you have recognizing people across the street ? <i>(Not at all; A little; Quite a lot; A lot)</i>
3	CatScale_3	Do you have difficulty recognizing someone across the street ? <i>(No; A little difficulty; A moderate amount of difficulty; Very difficult; Unable to do)</i>
4	QOLVFQ_10	Because of your vision, do you have problems in recognizing people across the street ? <i>(Not at all; Quite a lot; Very much)</i>
5	TYPE_3	How much does your vision hinder, limit, or disable you in each of the following activities? Recognizing people or objects across the street. <i>(Not at all; A little bit; Some; Quite a lot; Totally disabled)</i>

APPENDIX 1. Continued

S.No	PRO_item Number	Item/Response Option
Content 8. Face recognition at near		
1	VFQOL_4	How much problem do you have recognizing the face of a person standing near you? (Not at all; A little; Quite a lot; A lot)
2	VF-14_4	Do you have difficulty, even with glasses, recognizing people when they are close to you? (Yes; No; Not applicable) If yes, how much difficulty do you currently have? (A little; A moderate amount; A great deal; Are unable to do the activity)
3	Catquest_8	Recognize the faces of people you meet. . . . (Yes, extreme difficulty; Yes, much difficulty; Yes, some difficulty; No, no difficulty; Cannot say)
Content 9. Reading street signs/name of the stores		
1	NEIVFQ_8	How much difficulty do you have reading street signs or the name of stores? (No difficulty at all; A little difficulty; Moderate difficulty; Extreme difficulty; Stopped doing this because of your eyesight)
2	VF-14_6	Do you have difficulty, even with glasses, reading traffic signs, street signs, or store signs? (Yes; No; Not applicable) If yes, how much difficulty do you currently have? (A little; A moderate amount; A great deal; Are unable to do the activity)
3	VSQ_5	Is it difficult for you to read signs in the streets, shops or other public places? (No, not difficult; Yes, a little difficult; Yes, quite difficult; Yes, very difficult; I cannot see signs now because of my eyesight)
4	ADVS_4b	Would you say that you read street signs in daylight with: (PLEASE CHECK ONLY 1 ANSWER) (No difficulty at all; A little difficulty; Moderate difficulty; Extreme difficulty)
Content 10. Pouring liquid		
1	VAQ_9	When pouring liquid , I have trouble judging the level of the liquid in a container, such as the level of coffee in a cup. (Never; Rarely; Sometimes; Often; Always)
2	VSQ_8	How much difficulty do you have pouring liquids (water, tea, coffee from a jug or pot into a cup) because of trouble with your eyesight? (No difficulty; Yes, a little difficulty; Yes, some difficulty; Yes, a great deal of difficulty; I cannot pour the liquid)
3	CatScale_10	Because of your eyesight, do you have difficulty pouring a hot drink? (No; A little difficulty; A moderate amount of difficulty; Very difficult; Unable to do)
Content 11. Peripheral vision		
1	VFQOL_6	When you are walking along, how much problem do you have noticing objects off to the side? (Not at all; A little; Quite a lot; A lot)
2	NEIVFQ_10	Because of your eyesight, how much difficulty do you have noticing objects off to the side while you are walking along? (No difficulty at all; A little difficulty; Moderate difficulty; Extreme difficulty; Stopped doing this because of your eyesight)
3	VAQ_2	I have trouble noticing things in my peripheral vision. (Never; Rarely; Sometimes; Often; Always)
Content 12. Color vision		
1	VFQOL_9	How much problem do you have in recognizing colors? (Not at all; A little; Quite a lot; A lot)
2	QOLVFQ_17	How much problem do you have in recognizing colors? (Not at all; Quite a lot; Very much)
3	VAQ_25	I have difficulty distinguishing between colors. (Never; Rarely; Sometimes; Often; Always)
Content 13. Steps		
1	IVI_9	In the past month, how much has your eyesight interfered with going down steps, stairs, or curbs? (Not at all; Hardly at all; A little; A fair amount; A lot; Can't do because of eyesight)
2	CatScale_5	Because of your eyesight, do you have difficulty going down stairs? (No; A little difficulty; A moderate amount of difficulty; Very difficult; Unable to do)
3	ADVS_6b	Would you say that you walk down steps with: (PLEASE CHECK ONLY 1 ANSWER) (No apprehension [or fear] at all; A little apprehension [or fear]; Moderate apprehension [or fear]; Extreme apprehension [or fear])
Content 14. Crossing the road		
1	VDA_13	To what extent, if at all, does your vision interfere with your ability to cross the road? (Not at all; A little; Quite a bit; A lot)

APPENDIX 1. Continued

S.No	PRO_item Number	Item/Response Option
2	CatScale_6	Because of your eyesight, do you have difficulty crossing the road ? (No; A little difficulty; A moderate amount of difficulty; Very difficult; Unable to do)
3	ADVS_5b/c	Because of your vision, do you have problems crossing the street ? (Not at all; Quite a lot; Very much)
Content 15. Public transport		
1	VDA_14	To what extent, if at all, does your vision interfere with your ability to use public transport ? (Not at all; A little; Quite a bit; A lot)
2	IVI_18	In general, how much has your eyesight interfered with traveling or using transport ? (bus and train) (Not at all; Hardly at all; A little; A fair amount; A lot; Can't do because of eyesight)
3	ADVS_5b/c	Would you say that you use public transportation with: (PLEASE CHECK ONLY 1 ANSWER) (No visual difficulty at all; A little visual difficulty; Moderate difficulty because of vision; Extreme difficulty because of vision) Is it because of visual problems that you do not use public transportation? (Yes; No)

APPENDIX 2. Items of the Five PROs (ADVS, Activities of Daily Living Scale; CatScale, Cataract Symptom Scale; NEIVFQ, National Eye Institute Visual Function Questionnaire; TYPE, Technology of Patient Experience; VDA, Visual Disability Assessment) That Are Common to at Least Three Item Contents of the Visual Function-14 (VF-14).

S.No	PRO_item Number	Item/Response Option
1	VF-14_1	Do you have difficulty, even with glasses, reading small print, such as labels on medicine bottles, a telephone book, and food labels ? (Yes; No; Not applicable) If yes, how much difficulty do you currently have? (A little; A moderate amount; A great deal)
2	VF-14_2	Do you have difficulty, even with glasses, reading a newspaper or a book ? (Yes; No; Not applicable) If yes, how much difficulty do you currently have? (A little; A moderate amount; A great deal; Are unable to do the activity)
3	VF-14_6	Do you have difficulty, even with glasses, reading traffic signs, street signs, or store signs ? (Yes; No; Not applicable) If yes, how much difficulty do you currently have? (A little; A moderate amount; A great deal; Are unable to do the activity)
4	VF-14_12	Do you have difficulty, even with glasses, watching television ? (Yes; No; Not applicable) If yes, how much difficulty do you currently have? (A little; A moderate amount; A great deal; Are unable to do the activity)
5	VF-14_13	How much difficulty do you have driving during the day because of your vision? Do you have: (No difficulty; A little difficulty; A moderate amount of difficulty; A great deal of difficulty)
6	VF-14_14	How much difficulty do you have driving at night because of your vision? Do you have: (No difficulty = 1; A little difficulty = 2; A moderate amount of difficulty = 3; A great deal of difficulty? = 4)
7	VDA_4	To what extent, if at all, does your vision interfere with your ability to watch television ? (Not at all; A little; Quite a bit; A lot)
8	VDA_8	To what extent, if at all, does your vision interfere with your ability to drive a car by day ? (Not at all; A little; Quite a bit; A lot)
9	VDA_9	To what extent, if at all, does your vision interfere with your ability to drive a car by night ? (Not at all; A little; Quite a bit; A lot)
10	NEIVFQ_8	How much difficulty do you have reading street signs or the name of stores ? (No difficulty at all; A little difficulty; Moderate difficulty; Extreme difficulty; Stopped doing this because of your eyesight)
11	NEIVFQ_16	How much difficulty do you have driving at night ? Would you say you have: (No difficulty at all; A little difficulty; Moderate difficulty; Extreme difficulty; Have stopped doing this because of your eyesight)
12	NEIVFQ_A3	Wearing glasses, how much difficulty do you have reading the small print in a telephone book, on a medicine bottle, or on legal forms ? Would you say: (No difficulty at all; A little difficulty; Moderate difficulty; Extreme difficulty; Stopped doing this because of your eyesight)

APPENDIX 2. Continued

S.No	PRO_item Number	Item/Response Option
13	ADVS_1b	Would you say you drive at night with: (No difficulty at all; 4. A little difficulty [go to 1d]; 3. Moderate difficulty [go to 1d]; 2. Extreme difficulty [go to 1d]) Is it because of your visual problems that you are unable to drive at night? (1. Yes [go to 2a]; 2. No [go to 2a])
14	ADVS_2b	Would you say that you drive during the day with: (PLEASE CHECK ONLY 1 ANSWER) (No visual difficulty at all; A little difficulty because of vision; Moderate difficulty because of vision; Extreme difficulty because of vision) It is because of visual problems that you are unable to drive during the day? (Yes; No)
15	ADVS_9b/c	Would you say that you are able to see television with: (PLEASE CHECK ONLY 1 ANSWER) (No difficulty at all ; A little difficulty ; Moderate difficulty; Extreme difficulty) Is it because of visual problems that you are unable to watch television? (Yes; No; Not applicable)
16	TyPE_7	How much does your vision hinder, limit, or disable you in each of the following activities? Watching television (Not at all; A little bit; Some; Quite a lot; Totally disabled)
17	TyPE_8	How much does your vision hinder, limit, or disable you in each of the following activities? Daytime driving (Not at all; A little bit; Some; Quite a lot; Totally disabled)
18	TyPE_9	How much does your vision hinder, limit, or disable you in each of the following activities? Night-time driving (Not at all = 1; A little bit = 2; Some = 3; Quite a lot = 4; Totally disabled = 5; Don't do for other reasons = 0)
19	CatScale_1	Do you have difficulty reading small print ? (No; A little difficulty; A moderate amount of difficulty; Very difficult; Unable to do)
20	CatScale_2	Do you have difficulty watching television ? (No; A little difficulty; A moderate amount of difficulty; Very difficult; Unable to do)
21	CatScale_9	Do you have difficulty reading books or newspapers ? (No; A little difficulty; A moderate amount of difficulty; Very difficult; Unable to do)