Validity of a visual impairment questionnaire in measuring cataract surgery outcomes

Konrad Pesudovs, PhD, Lynda E. Caudle, GCertPH, Gwyneth Rees, PhD, Ecosse L. Lamoureux, PhD

PURPOSE: To test the validity of the Impact of Visual Impairment (IVI) questionnaire in a cataract population.

SETTING: Flinders Eye Centre, Flinders Medical Centre, Flinders University, Adelaide, South Australia.

METHODS: Cataract patients recruited from a hospital waiting list completed the IVI questionnaire. The scale was assessed for fit to the Rasch model. Unidimensionality, item and person fit to the model, response category performance, differential item functioning (whether different subgroups responded differently), and targeting of item difficulty to patient ability were assessed.

RESULTS: Overall, the IVI questionnaire performed well; there were ordered thresholds, person separation reliability was 0.97, and it was free from differential item functioning. One item (worry about eyesight getting worse) misfit the model and was removed. There was evidence of multidimensionality, indicating that the overall IVI score should be discarded; however, the 3 subscales (reading and accessing information, mobility and independence, and emotional well-being) functioned well. Several items calibrated differently in cataract patients compared with low-vision patients, indicating different issues are important to each population and that there is a need for population-specific conversion algorithms. Targeting of the IVI items was biased toward more impaired patients.

CONCLUSIONS: The 3 subscales of the IVI questionnaire functioned well in a cataract population. However, additional items targeting the less impaired patients, especially second-eye cataract patients, would improve measurement.

J Cataract Refract Surg 2008; 34:925–933 © 2008 ASCRS and ESCRS

Cataract is the leading cause of blindness worldwide[1] and is the most frequent eye condition in the elderly.[1,2] Cataract surgery has a high level of efficacy, has minimal complications, and is convenient for patients. Although the patient's visual acuity remains the most important clinical outcome of cataract surgery to the surgeon, the ability to perform routine daily activities is critically important to the patient.[3] Visual acuity may underestimate the value of surgery because it does not necessarily reflect postoperative functional improvement, changes in activities of daily living, and satisfaction with vision.[4] Thus, the quality-of-life, economic, and social benefits of improved vision often remain implicit.

Several questionnaires have been validated in patients with cataract. They include the Activities of Daily Vision Scale (ADVS),[5,6] the Visual Disability Assessment,[7] and the VF-14.[8] They focus on functional status related to vision (visual ability) and have been shown to be sensitive in the detection of clinically meaningful changes after cataract surgery and to provide information predictive of the outcomes of cataract surgery that is as powerful as that provided by the traditional predictors of age and ocular comorbidity.[6,9] These questionnaires, however, are limited to the domain of visual disability and thus do not address other areas of quality of life that are potentially important to

Accepted for publication December 21, 2007.

From the NH&MRC Centre for Clinical Eye Research (Pesudovs, Caudle), Flinders University and Flinders Medical Centre-Australia, Bedford Park, the Centre for Eye Research Australia (Rees, Lamoureux), Department of Ophthalmology, University of Melbourne, Melbourne, and Vision CRC (Lamoureux), Sydney, Australia.

No author has a financial or proprietary interest in any material or method mentioned.

Supported in part by NHMRC Centre of Clinical Research Excellence Grant 264620, NHMRC (Canberra, Australia) Career Development Award 426765 (Dr. Pesudovs), and NHMRC Public Health Fellowship 359277 (Dr. Lamoureux).

Corresponding author: Konrad Pesudovs, NH&MRC Centre for Clinical Eye Research, Department of Ophthalmology, Flinders Medical Centre, Bedford Park, South Australia, 5042, Australia. E-mail: konrad.pesudovs@flinders.edu.au.

Published by Elsevier Inc.
0886-3350/08/$—see front matter

doi:10.1016/j.jcrs.2007.12.052

the cataract patient. They are also limited in their development and validation as shown using classical test theory and have been shown not to hold up under testing with item response theory, Rasch analysis in particular.[10–12]

The Impact of Visual Impairment instrument (IVI) is a demonstrated valid scale to assess participation in daily activities by visually impaired individuals.[13–15] It is also a sensitive measure to assess the impact of cataract surgery on daily functioning in patients with early age-related macular degeneration (ARMD).[16] Compared with other instruments that have been used to assess vision-specific quality of life in cataract patients, the IVI has undergone substantial validation by Rasch analysis, a modern and sophisticated technique to aid questionnaire validation. A Rasch-calibrated instrument estimates linear interval measures from ordinal raw scores, facilitating the use of parametric statistical techniques. This improves the accuracy of scoring and removes measurement noise, which in turn improves sensitivity to intervention-induced changes.[17–20] Rasch analysis also assesses the instrument's validity, particularly if the scale items fit with the measurement of a single underlying latent trait (unidimensionality), and whether the items target the spectrum of the overall trait being measured; that is, whether they cover the range of participation in daily life or visual disability to suit the patient population (targeting).[10] Items not fitting a single dimension of visual ability and poor targeting of items to patients have been problems of the ADVS and the VF-14, and both lack items targeted at the more able cataract patient.[10,11]

The original IVI was validated on people with low vision, and the analysis included only a small proportion of individuals with cataract.[21] Instruments should only be applied to populations that they have been validated on as validity across different conditions cannot be assumed. Therefore, the validity of the IVI in cataract patients is unclear. The aim of this study was to empirically determine whether the IVI provides a valid assessment of perceived restriction of participation in patients with cataract.

PATIENTS AND METHODS

Patients

Patients with cataract were drawn from the public surgery waiting list of the ophthalmology service at Flinders Medical Centre, Adelaide, South Australia. All patients on the list had been previously assessed in the eye clinic and were deemed to have cataract causing visual disability that required surgical intervention. Consecutive patients on the list were invited to participate. This included those with unilateral and bilateral cataract, those awaiting second-eye surgery, and those with ocular comorbidity (eg, ARMD). It is important to be inclusive of all these characteristics to test the validity of the IVI for all types of cataract patients. Confining the analysis to bilateral cataract patients without ocular comorbidity, for example, would skew the assessment of the instrument and leave it untested in the other groups. Other inclusion criteria were age 18 years or older, no severe cognitive impairment, and ability to converse in English without the need for an interpreter. Ethical approval was obtained, and all patients who agreed to participate signed a consent form. This research adhered to the tenets of the Declaration of Helsinki.

Impact of the Visual Impairment Instrument

The IVI was developed to assess the restriction of participation in daily activities in people with low vision. It can be self-administered or administered by an interviewer. Recently, the IVI was further validated to examine its response scale and internal consistency as well as to provide the true linear scoring benefits of Rasch analysis.[13] This resulted in a 28-item questionnaire with a 4-category response scale for 26 items (0 = not at all; 1 = a little; 2 = a fair amount; 3 = a lot) and a 3-category response scale for 2 items (0 = not at all; 1 = a fair amount; 2 = a lot).[13] A 3-subscale structure possessing interval level measurement characteristics was subsequently confirmed using confirmatory factor and Rasch analyses.[15] The subscales are emotional well-being, reading and accessing information, and mobility and independence. The revised 28-item IVI was used in this study.

Rasch Analysis

The IVI data were fitted to the Rasch model[22] using the RUMM2020 software (RUMM Laboratory Pty. Ltd.).[23] When the scale data meet the Rasch model expectations, the ordinal raw score is transformed into an interval (linear) scale.[24,25] Among a number of advantages, normally distributed interval-level measurement allows the use of parametric analysis of data. The use of Rasch analysis to validate the IVI has been described extensively.[13,15,26] Briefly, 3 overall fit statistics are considered. The first is an item–trait interaction score, reported as chi square, which reflects the property of invariance across the trait and therefore indicates whether the data fit the model. An item–trait interaction probability value greater than 0.002 (Bonferroni-adjusted P value) was used to indicate no substantial deviation from the Rasch model. Two other fit statistics represent the residuals between the expected estimate and actual values for each person–item, summed over all items for each person and over all persons for each item. The residuals are transformed to approximate a z score and represent a standardized normal distribution in which a perfect fit to the model would have a mean of approximately 0 and a standard deviation (SD) of 1. This allows identification of which persons and items do and do not fit the model. A person-separation reliability score ranging between 0 and 1 indicates how well the items of the instrument separate the respondents. Larger values indicate a greater ability to distinguish between strata of person ability. A value of 0.9, for example, represents an ability to distinguish 4 strata of person ability. Individual item or person statistics with fit residual values greater than 2.5 or probability values below the Bonferroni adjusted α value were used to indicate misfit of the data to the model. Item removal was also considered if items had fit residual values greater than 2.5 or less than the Bonferroni-adjusted probability scores ($P = .002$).

The ordering of thresholds (ie, how the patents interpret the transition from 1 category to the next) was evaluated.

The presence of disordered thresholds shows that the categories are not working as intended. This can occur when there are too many response options or when the labeling of 2 or more options is similar, which is potentially confusing or open to misinterpretation. The collapsing of adjacent categories was considered in the event of disordered thresholds. Similarly, the occurrence of differential item functioning was statistically tested to ascertain whether subgroups within the sample (eg, sex), despite equal levels of the underlying trait, responded differently to an individual item. It is an important aspect of validity that items behave consistently across groups; that is, questionnaires should be free from differential item functioning. Targeting was also assessed as it was important to determine whether the IVI items were particularly suitable to assess visual disability associated with cataract. Poorly targeted measures are limited by floor or ceiling effects, display an uneven spread of items across the full range of respondent's scores, and show insufficient items to assess the full range of the sample trait.

Dimensionality testing determines whether the instrument is purely measuring the underlying trait (participation in daily living) that it purports to measure. The unidimensionality of the IVI was assessed using principal components analysis of the residuals. Unidimensionality was formally tested in RUMM2020[23] by allowing the pattern of factor loadings on the first component to determine subsets of items (positive and negative loadings subsets). If person estimates derived from these 2 subsets of items statistically differ (using independent *t* test provided in RUMM2020) from the estimates derived from the full scale, a breach of the assumption of unidimensionality is indicated.[27] While person estimates for each of these 2 sets of items should not be significantly different, less than or equal to 5% of cases being dissimilar is tolerated.[27,28]

The subscale structure of the IVI was retested in this population using confirmatory factor analysis. Valid subscales were then assessed within the Rasch model as described above for the entire IVI. Overall subscale performance was reported in terms of the item–trait interaction chi square, mean person and item fit residuals, person separation reliability, differential item functioning, unidimensionality, and targeting of items to persons.

The relationship between raw scores and Rasch person measures was determined by double-asymptotic nonlinear regression.[29] Documenting this relationship allows other investigators wishing to use the IVI subscales to use these validation data to convert raw scores into Rasch person measures without performing Rasch analysis.

RESULTS

Table 1 shows the characteristics of the 181 cataract patients who completed the IVI. Most were elderly, female, and reported some general medical comorbidity. The majority had bilateral cataract and did not have ocular comorbidity. Two thirds had visual acuity better than 6/12.

Table 1. Patient characteristics (N = 181).

Characteristic	Result
Mean age (y) ± SD	72.2 ±11.9
Sex, n (%)	
Male	71 (39)
Female	110 (61)
Binocular visual acuity	
Mean ± SD	
LogMAR	0.23 ± 0.19
Snellen	6/10
Range	
LogMAR	−0.20 to 0.70
Snellen	6/3.8 to 6/30
Awaiting second-eye surgery, n (%)	74 (41)
Ocular comorbidity*, n (%)	
Yes	46 (25)
No	135 (75)
Duration of cataract (y)	
Median	2
Range	0 to 31
Systemic comorbidity†, n (%)	
Yes	117 (65)
No	64 (35)

*For example, ARMD, glaucoma
†For example, hypertension, diabetes

Fit of Impact of Visual Impairment Instrument Data to Rasch Model

Initially, the IVI scores were reversed for Rasch analysis, giving participants with higher levels of participation higher scores. The partial credit approach (which allows each item to have its own threshold parameters) was used because the likelihood-ratio test in RUMM2020 was statistically significant ($P<.001$), indicating that the rating scale model (which requires equivalent thresholds across all items) was not appropriate.

The initial fit of the IVI data to the Rasch model showed a significant (less than the Bonferroni adjusted value of 0.002) item–trait interaction probability value (chi-square = 84, degrees of freedom = 56, P = .0015). This suggests that the data do not fit the Rasch model. There was no evidence of disordered thresholds (Figure 1), suggesting that the patients correctly discriminated the response options of the IVI.

The mean person fit residual value, however, was −0.61 ± 1.79 (SD). Ideally, the mean and SD values are expected to be closer to 0 and 1, respectively, suggesting misfit to the model by respondents. Individual person fit statistics showed that 7 patients (3.9%) had fit residual values outside the acceptable range (>2.5). Further analysis of the misfit patients showed inconsistent patterns in the items where extreme responses were observed. Upon removal of these misfit persons, the mean person fit residual value improved to −0.52 ± 1.53. The mean item fit residual ± SD value (−0.41 ± 1.33) did not show serious misfit. There was, however, no change in the overall item–trait interaction probability value (P = .0015).

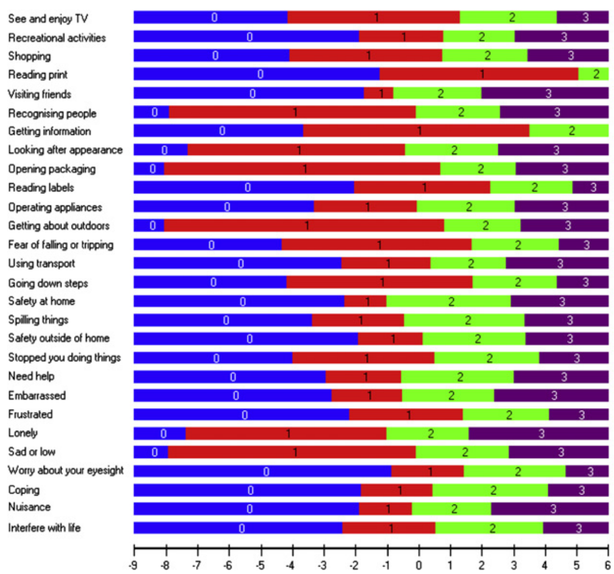

Figure 1. Threshold map of the IVI showing ordered thresholds.

Further examination of the items however showed that the IVI item 25 (worry about your eyesight getting worse) had an extreme fit residual value (3.49) and a probability value of 0.0005, which is substantially less than the Bonferroni-adjusted α value of 0.002. This item was subsequently removed. The item–trait interaction statistics substantially improved thereafter, showing a nonsignificant probability value (chi square = 73.2, P = .056) and thus demonstrating that IVI data fit the expectations of the measurement model. The item fit residual also improved for the mean (−0.41 to −0.35) and SD (1.33 to 1.21) values. The person separation reliability score for the IVI was 0.97, which indicates that the scale can discriminate between groups of respondents with 4 or more different levels of restriction of participation in daily living.

Examination of the rating scale categories, however, showed that the participants essentially endorsed only 3 responses: not at all, a little, and a fair amount (Table 2). Twenty-six (96.3%) of the 27 items had less than 4% of patients selecting the response "a lot," which suggests that overall the participants did not have a lot of difficulty with the IVI items. At the item level, this pattern was evident for several items such as operating household appliances, fear of falling or tripping, going down steps, safety at home, and feeling sad or low (Table 2).

Overall, the 5 most difficult items for patients on the IVI were reading ordinary size print, reading labels or instructions on medicine, feeling frustrated, unable to cope, and going down steps, with logit scores of 2.18, 2.04, 1.42, 1.13, and 0.97, respectively (Table 2). Conversely, the 5 least difficult items were feeling lonely, recognizing people, looking after appearance, safety at home, and feeling sad or low, with logit scores of −2.22, −1.78, −1.72, −1.69, and −1.63, respectively (Table 2). The logit calibrations in this cataract population are compared with those for the 28-item IVI in a low-vision population in Table 2.

Differential Item Functioning

Sex, ocular comorbidity, systemic comorbidity, and the impact of comorbidity on daily living were assessed and found to be free from differential item functioning, with probability values exceeding the adjusted α value for each factor assessed. This finding indicates that the IVI performs with similar accuracy regardless of subgroups within these person factors.

Targeting

The participant's range of ability (−2.39 to 7.15 logits) was found to have a normal distribution (Kolmogorov-Smirnov Z test score = 0.63; P = .82). The person–item threshold map (Figure 2) shows the person and item thresholds on the same calibrated scale (upper and lower sections of the graph, respectively). The map shows an uneven spread of items across the full range of respondents' scores, suggesting less than optimal targeting of the cataract patients (*top*) to the IVI items and thresholds (*bottom*). For example, many participants on the right of the graph had no difficulty, even with the most difficult items of the questionnaire. Furthermore, the mean person location logit value (3.36) substantiates that overall the questionnaire was not appropriately targeted and that overall the participants had a substantially higher level of participation than the average of the scale items (which would be 0 logit).

Unidimensionality

The principal components analysis of the residuals identified 2 subsets of items for the IVI consisting of the highest positive and negative loading items. Person estimates (location values) generated for the subsets in each case were subjected to independent t tests to compare each person's estimates. Ideally, person estimates for each of the 2 sets of items should not be significantly different from one another, although less than or equal to 5% of cases being dissimilar is tolerated.

The negative subset (principal component loadings less than −0.3) comprised 7 items, and the positive subset (principal component loadings more than 0.3) comprised 5 items (Table 3). As 16.98% (95% confidence interval, 14%-20%) of the person estimates in

Table 2. Category response proportions (categories reversed) and fit indices of the 27 IVI items to the Rasch model.

	Category Response Proportion (%)									
IVI Item	1	2	3	4	Location	SE	Fit Resid	χ^2	Prob	Location in an LV Population
See and enjoy TV	1	27	49	23	0.72	0.17	1.93	5.06	0.08	0.38
Recreational activities	3	17	38	42	0.76	0.19	0.99	7.74	0.02	0.26
Shopping	1	19	43	37	0.19	0.17	−1.88	3.47	0.18	0.66
Reading print	7	75	18	—	2.18	0.21	0.05	2.31	0.32	2.12
Visiting friends	1	7	35	58	−0.67	0.19	−1.04	3.98	0.14	−0.87
Recognizing people	0	12	39	49	−1.78	0.18	−0.33	1.13	0.57	0.38
Getting information	1	61	38	—	0.14	0.22	−1.12	2.97	0.23	0.03
Looking after appearance	0	10	41	49	−1.72	0.18	1.30	0.77	0.68	−0.78
Opening packaging	0	19	42	40	−1.31	0.17	−1.13	1.02	0.60	−0.54
Reading labels	3	39	42	15	2.04	0.16	1.52	0.66	0.72	1.20
Operating appliances	1	13	45	41	0.02	0.17	−1.01	1.91	0.38	−0.24
Getting about outdoors	0	20	40	40	−1.25	0.17	−0.98	1.05	0.59	0.37
Fear of falling or tripping	1	33	44	23	0.84	0.16	0.92	5.30	0.07	−0.07
Traveling or using transport	2	16	36	46	0.43	0.18	−0.21	1.31	0.52	0.42
Going down steps	1	34	43	23	0.97	0.16	−0.05	1.15	0.56	0.41
Safety at home	0	9	48	43	−1.69	0.19	−2.60	6.91	0.03	−1.47
Spilling things	1	11	51	37	0.05	0.18	−0.60	2.39	0.30	−1.39
Safety outside of home	1	15	48	36	0.29	0.17	−2.36	5.30	0.07	−0.31
Stopped you doing things	1	19	48	32	0.32	0.17	−1.95	1.22	0.54	0.53
Need help	1	10	47	42	−0.01	0.17	−0.94	0.18	0.92	0.06
Embarrassed	1	10	39	50	−0.16	0.17	−0.39	0.34	0.84	0.75
Frustrated	3	27	46	24	1.42	0.15	1.26	5.51	0.06	−1.08
Lonely	0	8	30	61	−2.22	0.19	−0.75	0.88	0.64	−0.60
Sad or low	0	13	43	44	−1.63	0.17	1.45	5.13	0.08	0.74
Coping	3	16	53	28	1.13	0.16	−0.36	0.79	0.67	−0.04
Nuisance	2	10	36	52	0.06	0.17	−0.31	1.15	0.56	0.51
Interfere with life	2	18	50	31	0.90	0.16	−0.84	3.59	0.17	−0.50

Fit Resid = fit residual value; LV = low vision; Prob = probability value; SE = standard error

the positive subset were found to be significantly different from person estimates in the negative subset, evidence of multidimensionality was detected for the IVI. This suggests that the IVI in this population was measuring more than 1 construct and that the scale could operate optimally if these constructs were assessed individually.

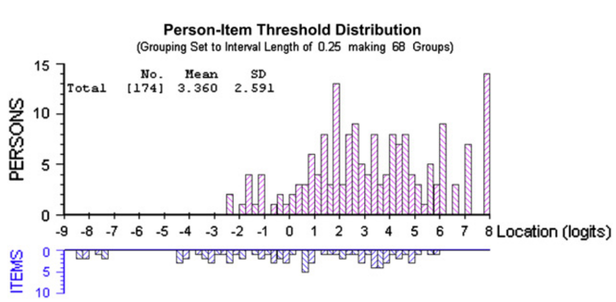

Figure 2. The targeting map shows an uneven spread of items across the full range of respondents' scores suggesting less than optimal targeting of the cataract patients (*top*) to the IVI items and thresholds (*bottom*).

Confirmatory Factor Analysis

Table 4 shows the goodness-of-fit statistics for the 3-factor model. All indices showed a reasonable fit between the IVI data and the 3-factor model. The β coefficients of all items were statistically significant ($P < .001$) and ranged between 0.68 and 0.84, 0.79 and 0.83, and 0.65 and 0.84 for the mobility and independence, emotional well-being, and reading and accessing information subscales, respectively. These findings provide evidence of the 3-subscale structure of the IVI.

Performance of Subscales Within Rasch Model

Each subscale was tested for fit to the Rasch model as per the approach taken to the entire IVI above. Table 5 shows the results. Each subscale showed good overall performance with a nonsignificant item–trait interaction chi square, acceptable person and item fit residuals, and good person separation reliability (>0.90). There was no evidence of differential item

Table 3. Principal component analysis of the residuals showing the first factor loading with 2 subsets (positive and negative loading items). Only items with factor loadings greater than ± 0.3 (followed by *) from each subset were used.

IVI Item	PC Loading
Getting about outdoors	0.627*
Going down steps	0.511*
Safety outside of home	0.441*
Fear of falling or tripping	0.424*
Using transport	0.337*
Spilling things	0.299
Stopped you doing things	0.283
Safety at home	0.238
Recreational activities	0.221
Operating appliances	0.153
Getting information	0.142
Visiting friends	0.120
Looking after appearance	0.119
Shopping	0.073
Recognizing people	0.054
Reading labels	0.018
Reading print	−0.043
Need help	−0.075
See and enjoy TV	−0.120
Opening packaging	−0.160
Unable to cope	−0.391*
Vision interfere with life	−0.419*
Feeling embarrassed	−0.510*
Feeling like a nuisance	−0.510*
Feeling lonely	−0.524*
Feeling frustrated	−0.538*
Feeling sad or low	−0.581*

PC = principal component

functioning or multidimensionality in any other subscale. As with the overall IVI, each subscale was targeted toward the less able end of the population (mean person location more than 3.0), with the more able participants having little or no difficulty with the more difficult items.

Table 4. Goodness-of-fit statistics for the 3-factor model.

Fit Index	Recommended Value	Value
χ^2	NA	587.3
df	NA	320
χ^2/df	≤2.00	1.78
Root-mean-square error of approximation	≤0.08	0.067
Goodness of fit index	≥0.9	0.812
Comparative fit index	≥0.9	0.902
Tucker-Lewis index	≥0.9	0.901

NA = not applicable

Raw Score to Rasch Scale Conversion

Other investigators wishing to use the IVI subscales can use these validation data to convert raw scores into Rasch person measures without having to perform Rasch analysis. Raw scores are calculated by first reversing scores (0, 1, 2, 3, 4, 5) (5, 4, 3, 2, 1, 0) to give better IVI scores to those experiencing less restriction of participation. The categories are then collapsed to 4 (3, 2, 2, 1, 1, 0) or 3 (2, 1, 1, 1, 1, 0) as described previously. Then, for each subscale the average of the items gives the IVI raw score. This is related to the IVI Rasch person measure, as shown in Figure 3. The relationship is double asymptotic because the average raw rating has a floor and a ceiling (at 0 and 3). The relationship can be described as a double-asymptotic nonlinear regression. The equations shown in Table 6 can be used to convert raw scores to Rasch person measures for each subscale.

DISCUSSION

Overall, the IVI questionnaire performed well in a cataract population. The response categories were used appropriately, as illustrated by ordered thresholds, and there was good person separation, indicating that the IVI can discriminate between 4 strata of respondents. The IVI in our cataract population was free from differential item functioning, indicating it is consistent across subpopulations. In these ways, the IVI performed comparably to previously published performance in a low-vision population.[13–15] However, there were differences in the IVI's performance specific to the cataract population.

Overall fit to the Rasch model suggested a problem with item fit. One item (worry about eyesight getting worse) did not fit the Rasch model. This suggests that this item behaves differently in cataract patients than in low-vision patients. This may partly be explained by the fact that cataract patients are awaiting an operation that will, in most cases, remove the eye problem and thus patients' concern about the progression of cataract. In contrast, most low-vision patients commonly have an eye disease that cannot be treated; thus, these patients are likely to be concerned about progression of the disease. A quarter of the cataract patients had ocular comorbidity, indicating that the misfit arose because some patients were worried about progression of their comorbid eye disease, with the remainder not worried at all. Removing this item confirmed the remaining 27 items fit the Rasch model.

Principal components analysis found evidence of multidimensionality. The IVI was previously shown to contain 3 viable subscales. When assessed with confirmatory factor analysis, it was again shown that 3

Table 5. Results of testing of subscale fit to the Rasch model.

	Subscale		
Parameter	Emotional Well-Being	Mobility and Independence	Reading and Accessing Information
Total chi square (df)	14.4 (14)	34 (26.3)	25.2 (18)
Bonferroni-adjusted χ^2 probability	0.42	0.09	0.12
Mean items fit residual	−0.32	−0.56	−0.48
Mean persons fit residual	−0.41	−0.52	−0.51
Person-separation reliability	0.93	0.95	0.92
Unidimensionality (%)	3.66	3.57	3.38
Mean item location	0	0	0
Mean person location	3.91	3.53	3.23

df = degrees of freedom

subscales exist. The high level of difference in person estimates from the different subsets tested in the principal components analysis suggests that the calculation of an overall score for the IVI in cataract patients should be abandoned. It would be more appropriate to report only the 3 subscale scores for this population. The subscales were tested with Rasch analysis and found to be valid. Raw score to Rasch scale conversion algorithms have been provided for the 3 subscales.

Although the response scale categories were ordered, the more impaired choice was underused. This suggests a shorter response scale could be used in cataract patients, which is consistent with Rasch analysis of the ADVS.[10] This response category use belies a problem with targeting. The items were, on the whole, too easy for the patients; this is best illustrated by a 3.36 logit difference between person and item mean values for the overall IVI, with a similar disparity for each subscale. Although not a fatal flaw, as illustrated by the retention of good person separation, the IVI would benefit from items that better target less impaired patients. Possible items could cover very difficult tasks such as driving in the rain or doing very fine needlework; could be more specific to second-eye cataract patients (who are less impaired), including tasks such as judging depth (eg, pouring drinks or putting a key into a keyhole); or could be specific to unilateral visual impairment (eg, Do you have trouble seeing on one side?). This suggests there are patient-centered issues that lead patients to desire cataract surgery and that are not tapped by the IVI. All questionnaires that have been used to assess cataract patients, including the VF-14 and the ADVS, have the same problem.[10,11,30] This problem can be avoided by using Rasch analysis in the development of questionnaires and using item targeting as a reason for retaining items.[20,30,31]

It is important to determine not simply whether the IVI questionnaire performs the functions required of a vision-related instrument but also whether it performs differently under different conditions. The IVI was previously developed and validated in a low-vision population.[13–15] One cannot assume this confers validity in cataract or any other eye disease; this must be tested and examined for differences. It is worth comparing the item calibrations between a low-vision population[15] and the cataract population in this study (Table 2). Of the 27 items retained in the analysis, 15 varied by 0.5 logits and 8 varied by 1.0 logit or more when compared to the results in the low-vision population. The largest disparities were in recognizing people (−2.2), getting about outdoors (−1.6), Lonely (−1.6), and sad or low (−2.4)—items easier for cataract patients—and spilling things (1.4), frustrated (2.5), coping (1.2), and interfering with life (1.4)—items more difficult for cataract patients. This suggests that different issues are important to people with cataract. For instance, cataract patients seem less troubled than low-vision patients by depression type of emotional issues such as sadness or loneliness. On the other hand, cataract patients were more troubled by emotional issues such as difficulty coping, experiencing frustration, and vision interfering with life overall. These differences emphasize the need to revalidate a questionnaire using different disease populations. Different calibrations for items do not make a major difference to the overall functioning of a questionnaire, nor do they interfere with Rasch analysis of data in this population. However, different calibrations across conditions preclude the use of simple conversions of raw scores to Rasch scores when the conversion is calibrated according to data obtained for a different disease. Thus, investigators must use calibrations derived from the same condition they are studying or perform an individual Rasch analysis on their data. Therefore, cataract-specific algorithms have been provided for converting raw scores to Rasch scores for investigators wishing to take advantage of

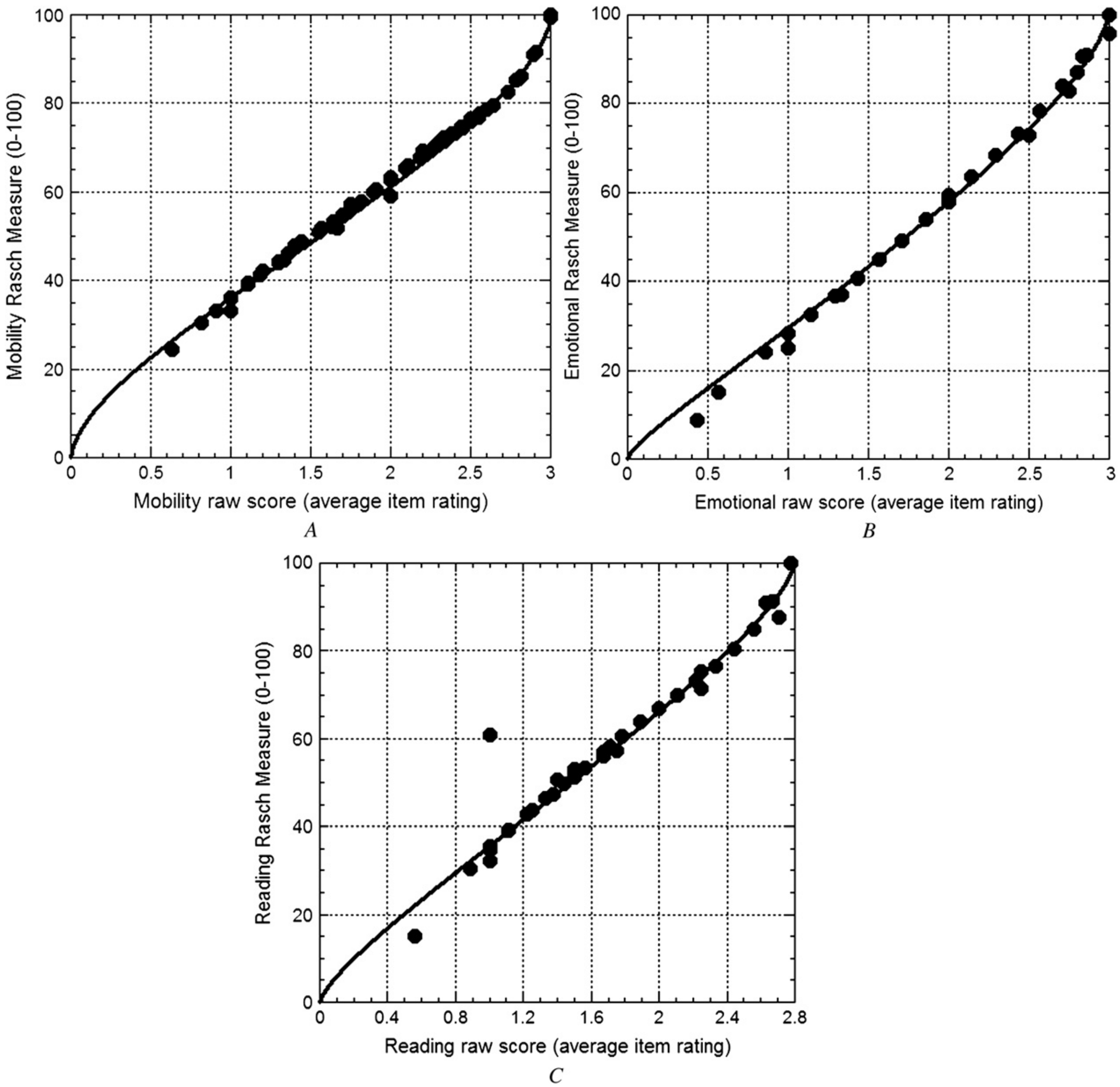

Figure 3. Scatterplot of the person measure estimated from Rasch analysis versus the average rating for each person (multiple cases overlap) across items (raw subscale score). The fit lines are generated with double asymptotic nonlinear regression. *A*: Mobility and independence. *B*: Emotional well-being. *C*: Reading and accessing information.

the scoring benefits of Rasch analysis without performing the analysis.

This study suggests that the IVI questionnaire is suitable for use as a cataract surgery outcome measure. However, for optimum performance, 1 item (worry about eyesight getting worse) should be removed from the analysis. Also, if simple calculation of Rasch scaling from raw scores is used, a cataract-specific conversion algorithm is required. The IVI has advantages over other cataract surgery outcome measures that simply score visual disability in that it reports results across 3 subscales of participation in activities of daily living; that is,

Table 6. Equations that can be used to convert raw scores to Rasch person measures for each subscale.

Subscale	Equation Converting Raw Score to Rasch Person Measure
Mobility and independence	$IVI_{person\ measure} = 35.00\log(IVI_{raw\ score}/3 - IVI_{raw\ score}) + 48.71$
Emotional well-being	$IVI_{person\ measure} = 43.23\log(IVI_{raw\ score}/3 - IVI_{raw\ score}) + 42.90$
Reading and accessing information	$IVI_{person\ measure} = 37.37\log(IVI_{raw\ score}/2.8 - IVI_{raw\ score}) + 44.92$

reading and accessing information, mobility and independence, and emotional well-being. However, like other cataract surgery outcome measures, the IVI lacks items to target more able patients, especially second-eye cataract patients. The ideal cataract surgery outcomes instrument should include such items.

REFERENCES

1. Thylefors B, Negrel AD, Pararajasegaram R, Dadzie KY. Global data on blindness. Bull World Health Organ 1995; 73:115–121
2. Hyman L. Epidemiology of eye disease in the elderly. Eye 1987; 1:330–341
3. Pager CK. Expectations and outcomes in cataract surgery; a prospective test of 2 models of satisfaction. Arch Ophthalmol 2004; 122:1788–1792
4. Javitt JC, Brenner MH, Curbow B, Legro MW, Street DA. Outcomes of cataract surgery; improvement in visual acuity and subjective visual function after surgery in the first, second, and both eyes. Arch Ophthalmol 1993; 111:686–691
5. Mangione CM, Phillips RS, Seddon JM, Lawrence MG, Cook EF, Dailey R, Goldman L. Development of the "Activities of Daily Vision Scale." A measure of visual functional status. Med Care 1992; 30:1111–1126
6. Mangione CM, Orav EJ, Lawrence MG, Phillips RS, Seddon JM, Goldman L. Prediction of visual function after cataract surgery; a prospectively validated model. Arch Ophthalmol 1995; 113:1305–1311
7. Pesudovs K, Coster DJ. An instrument for assessment of subjective visual disability in cataract patients. Br J Ophthalmol 1998; 82:617–624
8. Steinberg EP, Tielsch JM, Schein OD, Javitt JC, Sharkey P, Cassard SD, Legro MW, Diener-WestM Bass EB, Damiano AM. The VF-14; an index of functional impairment in patients with cataract. Arch Ophthalmol 1994; 112:630–638
9. Schein OD, Steinberg EP, Cassard SD, Tielsch JM, Javitt JC, Sommer A. Predictors of outcome in patients who underwent cataract surgery. Ophthalmology 1995; 102:817–823
10. 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. Available at: http://www.iovs.org/cgi/reprint/47/7/2789. Accessed February 5, 2008
11. Velozo CA, Lai JS, Mallinson T, Hauselman E. Maintaining instrument quality while reducing items: application of Rasch analysis to a self-report of visual function. J Outcome Meas 2000–2001; 4:667–680
12. Mallinson T, Stelmack J, Velozo C. A comparison of the separation ratio and coefficient α in the creation of minimum item sets. Med Care 2004; 42(suppl 1):I17–I24
13. Lamoureux EL, Pallant JF, Pesudovs K, Hassell JB, Keeffe JE. The Impact of Vision Impairment Questionnaire: an evaluation of its measurement properties using Rasch analysis. Invest Ophthalmol Vis Sci 2006; 47:4732–4741. Available at: http://iovs.org/cgi/reprint/47/11/4732.pdf. Accessed February 5, 2008
14. Lamoureux EL, Pallant JF, Pesudovs K, Rees G, Hassell JB, Keeffe JE. The effectiveness of low-vision rehabilitation on participation in daily living and quality of life. Invest Ophthalmol Vis Sci 2007; 48:1476–1482
15. Lamoureux EL, Pallant JF, Pesudovs K, Rees G, Hassell JB, Keeffe JE. The Impact of Vision Impairment Questionnaire: an assessment of its domain structure using confirmatory factor analysis and Rasch analysis. Invest Ophthalmol Vis Sci 2007; 48:1001–1006
16. Lamoureux EL, Hooper CY, Lim L, Pallant JF, Hunt N, Keeffe JE, Guymer RH. Impact of cataract surgery on quality of life in patients with early age-related macular degeneration. Optom Vis Sci 2007; 84:683–688
17. Norquist JM, Fitzpatrick R, Dawson J, Jenkinson C. Comparing alternative Rasch-based methods vs raw scores in measuring change in health. Med Care 2004; 42(suppl):I25–I36
18. Garamendi E, Pesudovs K, Stevens MJ, Elliott DB. The Refractive Status and Vision Profile: evaluation of psychometric properties and comparison of Rasch and summated Likert-scaling. Vision Res 2006; 46:1375–1383
19. Pesudovs K. Patient-centered measurement in ophthalmology—a paradigm shift. BMC Ophthalmol 2006; 6:25. Available at: http://www.biomedcentral.com/content/pdf/1471-2415-6-25.pdf. Accessed February 5, 2008
20. Pesudovs K, Garamendi E, Elliott DB. The Quality of Life Impact of Refractive Correction (QIRC) Questionnaire: development and validation. Optom Vis Sci 2004; 81:769–777
21. Weih LM, Hassell JB, Keeffe J. Assessment of the Impact of Vision Impairment. Invest Ophthalmol Vis Sci 2002; 43:927–935. Available at: http://www.iovs.org/cgi/reprint/43/4/927.pdf. Accessed February 5, 2008
22. Rasch G. Probabilistic Models for Some Intelligence and Attainment Tests. Chicago, IL, University of Chicago Press, 1960
23. Andrich D, Lyne A, Sheridan B, Luo G. RUMM 2020 [computer program]. Perth, Australia, RUMM Laboratory, 2003
24. van Alphen A, Halfens R, Hasman A, Imbos T. Likert or Rasch? Nothing is more applicable than good theory. J Adv Nurs 1994; 20:196–201
25. Tennant A, McKenna SP, Hagell P. Application of Rasch analysis in the development and application of quality of life instruments. Value Health 2004; 7(suppl 1):S22–S26
26. Lamoureux EL, Ferraro JG, Pallant JE, Pesudovs K, Rees G, Keeffe JE. Are standard instruments valid for the assessment of quality of life and symptoms in glaucoma? Optom Vis Sci 2007; 84:789–796
27. Smith EV Jr. Understanding Rasch measurement: detecting and evaluating the impact of multidimensionality using item fit statistics and principal component analysis of residuals. J Appl Measure 2002; 3:205–231
28. Tennant A, Pallant JF. Unidimensionality matters! (A tale of two Smiths?). Rasch Measurement Transactions 2006; 20(1):1048–1051. Available at: http://www.rasch.org/rmt/rmt201c.htm. Accessed February 5, 2008
29. Massof RW. Application of stochastic measurement models to visual function rating scale questionnaires. Ophthalmic Epidemiol 2005; 12:103–124
30. Pesudovs K, Garamendi E, Elliott DB. The Contact Lens Impact on Quality of Life (CLIQ) Questionnaire: development and validation. Invest Ophthalmol Vis Sci(47):2789–2796. Available at: http://www.iovs.org/cgi/reprint/47/7/2789.pdf. Accessed February 5, 2008
31. Lamoureux EL, Pallant JF, Pesudovs K, et al. An evaluation of the VCM1 (the core module of the Vision-Related Quality of Life scale) using Rasch analysis. In press, Am J Ophthalmol 2008

First author:
Konrad Pesudovs, PhD

NH;MRC Centre for Clinical Eye Research, Department of Ophthalmology, Flinders Medical Centre, Bedford Park