ARTICLE

Measuring outcomes of cataract surgery using the Visual Function Index-14

Vijaya K. Gothwal, PhD, Thomas A. Wright, Ecosse L. Lamoureux, PhD, Konrad Pesudovs, PhD

PURPOSE: To determine which version of the Visual Function Index-14 (VF-14) most precisely measured cataract surgery outcomes, rescale the VF-14 using Rasch analysis, and create a short-form version for comparison.

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

METHODS: In this cohort study incorporating questionnaire development, participants were drawn from the cataract surgery waiting list at Flinders Medical Centre. There were 2 cohorts: a preoperative cohort used for questionnaire development and an outcomes cohort. All patients had cataract surgery by phacoemulsification with intraocular lens implantation. Rasch analysis was used to refine the VF-14 into valid long-form (VF-11R) and short-form (VF-8R) versions. The ability of 8 versions (original; 2 proposed versions; 5 previously proposed versions) of the VF-14 to discriminate cataract surgery outcomes was compared with that of the standard VF-14 using the relative precision method.

RESULTS: The preoperative cohort comprised 210 patients and the outcomes cohort, 51 patients. Large gains in visual functioning occurred with cataract surgery, and these were detectable with all versions of the VF-14. The largest gain in precision, 125% (relative precision. 2.25), occurred for VF-8R. Short forms that were not Rasch scaled showed gains in precision, from 23% to 80%. The VF-8R also showed the largest gains in precision in 2 subgroups: with ocular comorbidity (relative precision, 2.14) and without ocular comorbidity (relative precision, 2.48).

CONCLUSIONS: Results show an unequivocal advantage to using Rasch-scaled scores for assessing cataract surgery outcomes. The 8-item, Rasch-scaled VF-8R appears ideally suited for measuring cataract surgery outcomes given its high precision and short test time.

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

J Cataract Refract Surg 2010; 36:1181–1188 © 2010 ASCRS and ESCRS

A patient's perspective is critical in evaluating the need for, and outcomes of, cataract surgery.[1–3] Questionnaires are increasingly being required for these evaluations. One such questionnaire is the Visual Function Index-14 (VF-14), which was developed to assess functional impairment in cataract patients.[4] The VF-14 is a popular questionnaire. It possesses adequate traditional psychometric properties,[5,6] has a concise format, is easy to administer, and has been validated internationally.[3,7] However, researchers have suggested it is too time consuming for routine use and therefore have proposed shortened versions.[8–10] Uusitalo et al.[8] proposed a VF-7, derived by selecting items that best correlated with patient satisfaction. Pager[9] also advocated a VF-7, which included items (different from Uusitalo et al.) that closely correlated with the overall preoperative VF-14 score. Moghimi et al.[10] advocated a VF-9 for use in specific conditions, including cataract surgery outcomes in traumatic aniridia.

The most recent short-form of the VF-14 is the VF-9, a Rasch-scaled version proposed by Lamoureux et al.[11] for use in a population-based study. Before this, Mallinson et al.[12] had used the VF-14 as an illustrative example to show the benefits of using Rasch analysis to shorten questionnaires. In contrast, Friedman et al.[13] proposed a shortened VF-11 but questioned the advantages of shortening the original VF-14.

Given there are many short forms of the VF-14, each varying in item content and number, which version best measures cataract surgery outcomes is unclear. To bring clarity to this problem, we aimed to compare the precision (ie, usefulness in making comparisons between preoperative and postoperative participants)[14] of current short-form versions of the VF-14

0886-3350/$—see front matter
doi:10.1016/j.jcrs.2010.01.029

in assessing cataract surgery outcomes to determine the preferred version for future use.

Furthermore, questionnaires reexamined using Rasch analysis have shown more sensitivity to change postoperatively[2]; therefore, we hypothesized that Rasch-scaled versions of the VF-14 may improve the precision of outcomes measurement. Although this has been done in a population-based setting, the high rate of normal visual functioning may make such a population unsuitable for refining the instrument. Therefore, we evaluated a cataract population to revise the VF-14 using Rasch analysis and included this version in our comparison.

PATIENTS AND METHODS

Study Group and Protocol

Since 2005, as part of a long-term Cataract Outcomes Assessment Study, data on a number of cataract-specific questionnaires (including the VF-14) were collected. This assessment was implemented by routinely mailing packs of questionnaires (10) to consecutive patients on the waiting list for cataract extraction surgery at Flinders Medical Centre, Adelaide, South Australia. Inclusion criteria were English speaking, aged 18 years or older, and ability to provide written informed consent. Patients self-administered the questionnaires and returned them in a prepaid envelope. Patients chose to complete as many questionnaires as they wished. A demographic data form was included in the pack to obtain information regarding ocular and systemic status, which was subsequently confirmed from the patient's medical record at the time of data entry.

During a single 6-month data-collection window, the same pack was mailed 6 months after cataract surgery. Patients had coexisting systemic and ocular conditions, which is typical of an elderly cataract patient cohort in Australia.[15] Ethics approval for this research was obtained from the Flinders Clinical Ethics Committee. This research adhered to the tenets of the Declaration of Helsinki.

Submitted: November 23, 2009.
Final revision submitted: January 2, 2010.
Accepted: January 21, 2010.

From NH&MRC Centre for Clinical Eye Research (Gothwal, Wright, Pesudovs), Department of Optometry and Vision Science, Flinders Medical Centre and Flinders University of South Australia, Bedford Park, South Australia; Centre for Eye Research Australia (Lamoureux), Department of Ophthalmology, University of Melbourne, Victoria; Vision CRC (Lamoureux), Sydney, New South Wales, Australia; Meera and LB Deshpande Centre for Sight Enhancement (Gothwal), Vision Rehabilitation Centres, LV Prasad Eye Institute, Hyderabad, India; Singapore Eye Research Institute (Lamoureux), Singapore National Eye Centre, Singapore.

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

There were 2 patient populations. The first cohort comprised preoperative cataract patients whose data were used to refine the VF-14 with Rasch analysis (development group). The second cohort comprised patients whose data were used to measure cataract surgery outcomes (outcomes group).

Standardized eye examinations were performed before and after (minimum 21 days) cataract surgery. Habitual monocular and binocular visual acuity assessments were performed using computerized testing based on the logMAR principles with screen illumination of 150 candelas/m^2.[16] The visual acuity in the operated and fellow eyes of patients who had cataract surgery is presented here.

Questionnaires

Visual Function Index-14 The VF-14 contains questions (items) related to the degree of difficulty in performing 14 vision-dependent activities (eg, reading, watching television).[4] Table 1 shows the activities the VF-14 addresses and the response categories. Responses were coded as recommended by the developers. "Not applicable" responses were treated as missing data in the analysis. Higher scores represent better visual functioning (ie, less difficulty) and, therefore, greater ability in performing the activity.

Short-Form Versions of the Visual Function Index-14 Five studies that proposed short-form versions of the VF-14 were identified. They were Friedman et al.'s VF-11 (11 items),[13] Uusitalo et al.'s VF-7 (VF-7U; 7 items),[8] Pager's VF-7 (VF-7P; 7 items),[9] Moghimi et al.'s VF-9 for women (VF-9MF; 9 items) and for men (VF-9MM; 10 items),[10] and Lamoureux et al.'s Rasch-analyzed VF-9 (VF-9L).[11] Each shortened version contains a different set of the original VF-14 items.

The response options used in all short-form versions were similar to the original VF-14. Although Lamoureux et al.[11] proposed a reduction in categories from 5 to 4 for their VF-9L, in this study the original 5 categories for data collection were retained as that was how Lamoureux et al. collected their data.

Outcome Measures

Change in overall visual functioning with cataract surgery was the primary outcome measure. This outcome was also tested for 2 subgroups: with ocular comorbidity and without ocular comorbidity. Change in visual acuity was the secondary outcome measure.

Assessment of the Psychometric Properties of Visual Function Index-14 by Rasch Analysis

The native scoring system of the VF-14 is an ordinal (Likert) scale (ie, numerical values in an increasing order are assigned to categories of increasing difficulty) that uses summary scoring. This approach falsely assumes the spacing between response categories is equal and that all the items have the same level of difficulty. Ordinal scores are not a measurement; thus, they are inappropriate for measuring the degree of difference between patients or between preoperative and postoperative periods.[17] Therefore, before using the VF-14, it was imperative to assess its psychometric properties using Rasch analysis. A series of analyses was performed that included assessment of the following: (1) behavior of response categories (ie, whether higher

Table 1. Item content for VF-14 and the 2 Rasch-scaled versions of the VF-14 (VF-11R and VF-8R).*

Item	Item Description in VF-14	Items in VF-11R	Items in VF-8R
1	Reading small print, such as labels on medicine bottles, a telephone book, food labels	Retained	Retained
2	Reading a newspaper or a book	Retained	Retained
3	Reading a large-print book or large-print newspaper or numbers on a telephone	Retained	Eliminated
4	Recognizing people when they are close to you	Retained	Eliminated
5	Seeing steps, stairs, or curbs	Retained	Retained
6	Reading traffic signs, street signs, or store signs	Retained	Retained
7	Doing fine handwork, such as sewing, knitting, crocheting, carpentry	Retained	Retained
8	Writing checks or filling out forms	Retained	Retained
9	Playing games, such as bingo, dominos, card games, mahjong	Retained	Retained
10	Taking part in sports, such as bowling, handball, tennis, golf	Eliminated	Eliminated
11	Cooking	Retained	Eliminated
12	Watching television	Retained	Retained

*For items 1 through 12, the frame question was, "Do you have any difficulty, even with glasses?"; there were 5 scoring response options (no = 4; a little = 3; a moderate amount = 2; a great deal = 1; unable to do the activity = 0). Items 13 through 18 are driving items. Two are scoring items with 5 response options, and there are different frame question for these items; these were eliminated from the Rasch-scaled versions (VF-11R and VF-8R).

categories represented better visual functioning), (2) measurement precision (represented by person separation; minimum acceptable value of 2.0[18]), (3) unidimensionality (ie, whether all the items contribute and measure a single underlying latent trait of visual functioning measured by infit mean square statistic with acceptable range of 0.7 to 1.3 and also by principal components analysis,), and (4) whether items match the patient's visual functioning (represented by targeting; ideal <0.5 logits). If all the items did not measure visual functioning (representing lack of unidimensionality), the goal was to provide remedial measures. As in other studies, this one considered shortening the VF-14 without compromising its original properties. Details about applying Rasch analysis to the questionnaires for this purpose have been described[2,19,20] and are reported in brief here. In the context of Rasch analysis, an item (activity) is considered difficult if a high level of visual functioning is required to complete it. In Rasch analysis, item difficulty and patient ability are calibrated on the same scale and are expressed in logit units.[18,21]

Using the data from all preoperative cataract patients, Rasch analysis was performed using the Andrich rating scale model for polytomous data (ie, multiple response options for an item) in the Winsteps software (version 3.68).[22,23] In contrast to the need to combine categories, as reported by Lamoureux et al.[11] for the VF-9L, the patients in this study used the response options as they were intended to and, therefore, the original 5 response categories were retained. The VF-14 showed adequate stratification of visual functioning evidenced by a person separation of 2.45 (minimum acceptable value, 2.0) indicating that it was able to discriminate between 3 strata of patient's visual functioning (Table 2). Targeting was suboptimum (1.86 logits), indicating that the items were mismatched to the patient's visual functioning. This result indicated that, overall, the items were too easy for patients.[12]

Two items did not fit. This indicated a lack of unidimensionality (ie, these 2 items measured a construct different than the remaining 12 items [not visual functioning]). Principal component analysis further confirmed the lack of unidimensionality by revealing the presence of a secondary dimension, which could be described as relating to driving.

Taken together, the above findings suggested that the VF-14 required revision. Specifically, unidimensionality had to be restored and item misfit minimized. Unidimensionality was restored by deleting the 2 driving items. However, after deletion of the items, a further item (playing games) showed misfit and therefore was also deleted. The remaining 11 items then fit the Rasch model. That is, these items formed a unidimensional measure of visual functioning that could be used in the comparisons along with previously proposed short-form versions. This new version is referred to here as the VF-11R (R for Rasch) (Table 2).

In the VF-11R, certain items possessed the same difficulty level as others. This suggested redundancy in the measure and that further items in the VF-11R could be removed. The following criteria were used to drive the selection of items to be retained in the short-form: (1) maintain a minimum person separation value of 2.0 and (2) maintain targeting.

Two further items were removed from the VF-11R. In this process, an additional item also misfit and was deleted. Thus, 8 items remained in this unidimensional short-form version, which is referred to here as the VF-8R (Table 1). In terms of being a unidimensional measure of visual functioning, the VF-8R was superior to the VF-14, although person separation and targeting were marginally lower than for the VF-11R (Table 2). Nevertheless the VF-8R was shorter than the original scale by 6 items. The reliability of these short-form versions was not tested.

To fulfill the study's main aim of determining the best version of VF-14 for assessing the change in visual functioning after cataract surgery, the VF-11R and VF-8R were appended to the existing list of the 5 shortened versions of the VF-14.[8–11, 13]

Statistical Analysis

For the Rasch analysis of the outcomes, the data obtained from the preoperative patients and postoperative patients were combined; that is, all data were assembled in a single data set, with the postoperative data treated as "new patients".[24] Preoperative and postoperative visual functioning scores (in logits) were then estimated for each patient. This

Table 2. Overall performance of the VF-14 and the included short-form versions of the VF-14.

Parameter	VF-14	VF-11R*	VF-8R*	VF-11†	VF-7U†	VF-7P†	VF-9MF†	VF-9MM†	VF-9L*
Misfitting items (n)	2	0	0	1	1	1	0	2	1
Person separation	2.45	2.46	2.29	2.29	1.86	2.07	2.31	2.18	2.73
Mean item location	0	0	0	0	0	0	0	0	0
Mean person location	1.86	2.57	1.97	1.39	1.53	1.75	2.64	1.67	2.26
Principal components analysis (eigenvalue)	2.3	1.6	1.6	2.2	1.6	1.6	1.7	2.3	1.7

VF-14 = Visual Functioning Index 14 (14 items[4]; VF-11R = 11 items (Rasch scaled version from present study); VF-8R = 8 items (Rasch scaled version from present study); VF-11 = 11 items (Friedman et al.[13]); VF-7U = 7 items (Uusitalo et al.[8]); VF-7P = 7 items (Pager[9]); VF-9MF = 9 items for females (Moghimi et al.[10]); VF-9MM = 9 items for males (Moghimi et al.[10]); VF-9L = 9 items (Lamoureux et al.[11])
*Rasch-scaled versions
†Non Rasch–scaled versions

was done so that the preoperative and postoperative scores were derived on the same scale and would therefore provide an accurate measure of outcomes.

A 1-way analysis of variance (ANOVA) was used to determine whether the change in preoperative to postoperative score for the original VF-14 and each shortened version differed significantly from zero. The F statistic with a P value less than 0.05 was considered significant. Relative precision was then used to examine how well each version of the VF-14 distinguished visual functioning between preoperative and postoperative periods, relative to the Likert scoring of the original VF-14.[25] Relative precision is a ratio of pairwise F statistics (F for each version versus F for the Likert scoring of VF-14). The extent to which the relative precision ratio differed from 1.0 indicated the degree to which the 2 scoring methods differed in their ability to detect the change in scores; values greater than 1.0 indicated increased precision.

To maximize comparability, the ordinal raw scores (from VF-14, VF-11, VF-7U, VF-7P, VF-9MF, and VF-9MM) and Rasch measures (from VF-11R, VF-8R, and VF-9L) were transformed from their original scale to a 0 to 100 metric; minimum visual functioning (maximum difficulty) was set at 0 and maximum visual functioning (minimum difficulty), at 100.[26]

SPSS for Windows software (version 15.0, SPSS, Inc.) was used for all general descriptive statistics. A paired t test was used to compare improvements in visual acuity within the group for those with ocular comorbidity and without ocular comorbidity. Independent-samples t tests were used to compare the improvement in visual acuity between these groups. A P value less than 0.05 was considered statistically significant.

RESULTS

Response and Patient Characteristics

The VF-14 was mailed to 414 patients, of whom 210 (50.7% response rate) returned the completed questionnaire. Postoperatively, 51 of the 81 patients who were mailed the VF-14 returned it (62.9% response rate). Table 3 shows the baseline characteristics of the patients by group.

Clinical Outcomes

Combining the data of the preoperative patients and postoperative patients for Rasch analysis of the outcomes yielded 102 patient records. Table 4 shows the mean preoperative and postoperative visual acuity values in the operated eyes and fellow eyes. Visual acuity improved significantly from preoperatively to postoperatively overall ($P < .0001$) and in the comorbidity subgroup ($P < .0001$) and no-comorbidity subgroup ($P = .02$). The final postoperative visual acuity was not significantly different between the 3 groups (F = 2.69 and $P = .08$, ANOVA).

Relative Precision: Clinical Discrimination

Tables 5, 6, and 7 show the mean preoperative and postoperative scores (and mean change) for the VF-14 and the various short-form versions in the overall group, the ocular comorbidity subgroup, and the no-ocular comorbidity subgroup, respectively. Overall, regardless of the scoring method used, the mean postoperative scores were consistently higher than the preoperative scores across all versions (Table 5). The

Table 3. Baseline sociodemographic and clinical characteristics of the cataract patients who completed the VF-14.

	Group	
Characteristic	Development	Outcomes
Patients (n)	210	51
Mean age (y) ± SD	74.3 ± 9.3	73.0 ± 7.5
Sex, n (%)		
Male	88 (42)	29 (57)
Female	122 (58)	22 (43)
Ocular comorbidity,* n (%)		
Present	98 (48)	30 (59)
Absent	106 (52)	21 (41)
Systemic comorbidity,† n (%)		
Present	142 (84)	40 (78)
Absent	27 (16)	11 (22)

*Includes age-related macular degeneration, glaucoma, diabetic retinopathy, etc. Data were missing for 6 cases in the development group.
†Includes hypertension, diabetes, angina, etc. Data were missing for 41 cases in the development group.

Table 4. Preoperative and postoperative visual acuity of 51 patients who had cataract surgery and completed the VF-14.

Group/Exam Time	Visual Acuity
Operated eyes	
All (n = 51)	
Preoperative*	
Mean logMAR ± SD	0.52 ± 0.40
Range	0.00 to 2.00
Snellen	$6/19^{-1}$
Postoperative*	
Mean logMAR ± SD	0.18 ± 0.21
Range	−0.12 to 0.80
Snellen	$6/7.5^{-1}$
With comorbidity (n = 30)	
Preoperative*	
Mean logMAR ± SD	0.41 ± 0.32
Range	0.00 to 1.30
Snellen	6/15
Postoperative*	
Mean logMAR ± SD	0.23 ± 0.21
Range	−0.10 to 0.80
Snellen	$6/9.5^{-1}$
With no comorbidity (n = 20)	
Preoperative (better eye)	
Mean logMAR ± SD	0.69 ± 0.45
Range	0.10 to 2.00
Snellen	$6/30^{-1}$
Postoperative (better eye) †	
Mean logMAR ± SD	0.07 ± 0.17
Range	−0.12 to 0.44
Snellen	$6/7.5^{+1}$
Fellow eyes	
All	
Mean logMAR ± SD	0.20 ± 0.20
Range	−0.30 to 0.80
Snellen	6/9.5
With comorbidity	
Mean logMAR ± SD	0.18 ± 0.18
Range	−0.30 to 0.50
Snellen	$6/9.5^{+1}$
Without comorbidity	
Mean logMAR ± SD	0.22 ± 0.24
Range	−0.1 to 0.80
Snellen	$6/9.5^{-1}$

Notes on logMAR values: 1.3 represents visual acuity of 3/60 or 6/120; 2.00 represents light perception, 0 represents 6/6, negative logMAR values indicate visual acuity of better than 6/6.
Snellen notation: Minus sign in the superscript indicates patient could not read the line completely and missed letters, for example, $6/19^{-2}$ indicates patient missed 2 letters from this line. Plus sign indicates patient read this line completely correctly and read 2 more letters correctly in the subsequent smaller line
*$P < 0.0001$ (Paired t test)
†$P = 0.02$ (Paired t test)

largest improvement in scores occurred for the VF-8R. Figure 1 shows the relative distribution of the VF-14 and VF-8R scores preoperatively and postoperatively in the overall group.

As hypothesized, all Rasch-scaled versions achieved significantly greater gains in precision in discriminating between visual functioning of preoperative and postoperative patients (Table 5). The gain in precision compared with the original Likert scored VF-14 was 98% for the VF-11R, 125% for the VF-8R, and 98% for the VF-9L.

Similar to the overall group, the mean postoperative scores were significantly higher than the preoperative scores in both subgroups (Tables 6 and 7). The gain in precision was consistently the largest for VF-8R with ocular comorbidity (114%) and without ocular comorbidity (148%). In the subgroup without ocular comorbidity, 2 Likert-scored versions (VF-11 and the VF-9MM) had less precision (12% and 22%, respectively) than the VF-14.

DISCUSSION

After cataract surgery, visual acuity improved significantly overall (by a mean of 3.4 lines) and in both subgroups, with the largest gains in eyes without ocular comorbidity (6.2 lines). Visual acuity is, of course, a surrogate for visual functioning, albeit limited to the high contrast acuity spectrum of function. More important, visual functioning also improved significantly overall and in both subgroups. For example, postoperatively, patients in the overall group had a mean VF-8R Rasch-score of 83.15 logits (15.39-logit improvement from preoperative assessment), while the ocular comorbidity subgroup had a mean VF-8R Rasch score gain of 13.87 logits, and the no comorbidity subgroup gained 17.35 logits. Similar improvements, albeit smaller in magnitude by comparison, were observed for the VF-14 and the other 7 short-form versions.

The main objective of our study was to determine the best short-form version of the VF-14 by comparing the relative precision of 8 short-form versions against the original VF-14 in measuring the outcomes of cataract surgery. We found larger gains in precision for Rasch-scoring (range of relative precision 98% to 125% increase) in discriminating the visual functioning in the overall group; the largest gain of 125% was for the VF-8R (relative precision, 2.25). Similar large gains were observed for Rasch-scoring across both subgroups. In fact, the largest gain in precision (relative precision = 2.48) was for the VF-8R in discriminating the visual functioning for those who did not have ocular comorbidity. That is, the precision of VF-8R in this subgroup was 2.48 times better than that of the original VF-14. Thus, the results in our study provide strong evidence of the benefits of Rasch-scaling questionnaires. These results are consistent with those of other researchers, who have also showed the benefits

Table 5. Mean preoperative and postoperative scores for cataract surgery patients (overall, n = 51) and relative precision for the VF-14 and the included short-form versions of the VF-14.

	Mean ± SE				
Version	Preoperative	Postoperative	Mean Differences* ± SE: Preop Vs Postop	F Statistic†	Relative Precision‡
VF-14/Likert	82.49 ± 1.99	90.61 ± 1.79	8.12 ± 1.87	9.18	1.00
VF-11R/Rasch	79.59 ± 1.50	88.92 ± 1.59	9.33 ± 1.61	18.14	1.98
VF-8R/Rasch	67.75 ± 2.36	83.15 ± 2.43	15.39 ± 2.66	20.67	2.25
VF-11/Likert	78.68 ± 2.20	89.43 ± 2.07	10.75 ± 2.31	12.66	1.38
VF-7U/Likert	78.17 ± 2.10	88.37 ± 1.96	10.20 ± 2.00	12.57	1.37
VF-7P/Likert	77.26 ± 2.38	90.17 ± 2.10	12.91 ± 2.43	16.53	1.80
VF-9MF/Likert	83.18 ± 1.95	92.14 ± 1.73	8.95 ± 1.68	11.77	1.28
VF-9MM/Likert	81.34 ± 2.03	90.50 ± 1.82	9.16 ± 1.85	11.27	1.23
VF-9L/Rasch	79.49 ± 1.55	89.17 ± 1.66	9.68 ± 1.68	18.14	1.98

SE = standard error

*The follow-up time for self-administration of the VF-14 postoperatively was a minimum of 6 months from the date of surgery. The mean difference was calculated by subtracting the postoperative score from the preoperative score, with a positive result indicating a gain postoperatively.

†$P < .05$

‡Relative precision was calculated by dividing the F statistic for each version by that of the VF-14 (as baseline).

of Rasch-scaled versions over Likert scores for ophthalmic and nonophthalmic questionnaires.[25–28]

The main reason the Rasch-scaled versions had relatively greater precision in measuring outcomes is the reduction in error in estimating the measurement of visual disability, as evidenced by reduced standard errors of the measures.[25,29] Smaller standard errors, typical of Rasch scaling, were noted in the present study for the VF-11R and VF-9L, but not for the VF-8R.[28] Second, as a result of logistic transformation, Rasch-scaling increases measurement precision by expanding the range of measurement. It is the larger range of measurement for the VF-8R that probably caused its increased standard errors, although further reliability testing of this version could be informative. In contrast, Likert-scaled scores are constrained at each end of the scale. The larger range of measurement in the Rasch-scaled versions implies reduced ceiling and floor effects (ie, patients with extreme scores), as was evidenced with the use of VF-8R. Patients with

Table 6. Mean preoperative and postoperative scores for cataract surgery patients who had ocular comorbidity (n = 30) and relative precision for the VF-14 and the included short-form versions of the VF-14.

	Mean ± SE				
Version	Preoperative	Postoperative	Mean Differences* ± SE: Preop Vs Postop	F Statistic	Relative Precision¶
VF-14/Likert	81.54 ± 2.78	89.21 ± 2.76	7.66 ± 2.36	3.81†	1.00
VF-11R/Rasch	79.45 ± 2.13	87.69 ± 2.35	8.24 ± 2.26	6.73‡	1.77
VF-8R/Rasch	67.71 ± 3.29	81.58 ± 3.57	13.87 ± 3.81	8.15‡	2.14
VF-11/Likert	77.87 ± 2.94	88.60 ± 2.99	10.73 ± 2.77	6.54‡	1.72
VF-7U/Likert	76.11 ± 2.91	86.73 ± 2.99	10.62 ± 2.83	6.48‡	1.70
VF- 7P/Likert	77.54 ± 3.17	88.21 ± 3.33	10.66 ± 3.33	5.37‡	1.41
VF-9MF/Likert	82.51 ± 2.79	90.18 ± 2.74	7.67 ± 2.22	3.84‡	1.01
VF-9MM/Likert	80.01 ± 2.74	89.39 ± 2.74	9.31 ± 2.14	5.77‡	1.51
VF-9L/Rasch	79.59 ± 2.18	88.30 ± 2.46	8.71 ± 2.35	7.02‡	1.84

SE = standard error

*The follow-up time for self-administration of the VF-14 postoperatively was a minimum of 6 months from the date of surgery. The mean difference was calculated by subtracting the postoperative score from the preoperative score, with a positive result indicating a gain postoperatively.

†$P > .05$ for VF-14 only

‡$P < .05$

¶Relative precision was calculated by dividing the F statistic for each version by that of the VF-14 (as baseline).

Table 7. Mean preoperative and postoperative scores for cataract surgery patients who did not have ocular comorbidity (n = 20) and relative precision for the VF-14 and the included short-form versions of the VF-14.

	Mean ± SE				
Version	Preoperative	Postoperative	Mean Differences* ± SE: Preop Vs Postop	F statistic†	Relative Precision‡
VF-14/Likert	84.97 ± 2.71	92.97 ± 1.86	8.00 ± 3.20	5.92	1.00
VF-11R/Rasch	80.50 ± 2.03	91.20 ± 1.94	10.70 ± 2.31	14.44	2.44
VF-8R/Rasch	68.87 ± 3.36	86.22 ± 3.03	17.35 ± 3.72	14.70	2.48
VF-11/Likert	81.09 ± 3.29	91.02 ± 2.82	9.93 ± 4.17	5.26	0.89
VF-7U/Likert	82.49 ± 2.69	91.70 ± 1.97	9.21 ± 2.92	7.62	1.29
VF- 7P/Likert	78.21 ± 3.59	93.46 ± 1.79	15.25 ± 3.54	14.42	2.43
VF-9MF/Likert	85.29 ± 2.49	95.15 ± 1.46	9.86 ± 2.53	11.63	1.96
VF-9MM/Likert	84.43 ± 2.92	92.39 ± 2.14	7.96 ± 3.39	4.84	0.82
VF-9L/Rasch	80.04 ± 2.16	91.04 ± 2.02	11.03 ± 2.47	13.90	2.35

SE = standard error

*The follow-up time for self-administration of the VF-14 postoperatively was a minimum of 6 months from the date of surgery. The mean difference was calculated by subtracting the postoperative score from the preoperative score, with a positive result indicating a gain postoperatively.

†$P < .05$

‡Relative precision was calculated by dividing the F statistic for each version by that of the VF-14 (as baseline).

high visual functioning scored at the upper end of the VF-8R, while those with low visual functioning scored at the lower end. Although it appears as though there was some truncation of measurement in the postoperative samples, the truncation seemed to be less with the VF-8R (Figure 1).

Nevertheless, the overarching question is which version(s) of VF-14 should be used for assessing outcomes of cataract surgery? Our results clearly indicate that the Rasch-scaled VF-8R is the most appropriate. There are many potential benefits to using it. First, it provides interval-level measurement, making comparison between patients meaningful. Second, all items measure a single construct of visual functioning (implying unidimensionality, which is an essential measurement property); this is unlike the original VF-14, which is confounded by more than 1 construct. Third, it has better measurement precision for discriminating outcomes, indicating a smaller sample size will be required to find significant differences. Finally, with only 8 items, respondent burden and administration time are minimal.

The proposed VF-8R version is not without limitations. It has suboptimum targeting, marginally lower than the original VF-14. Except for the Catquest-9SF,[2] problems with targeting (ie, items being too easy) have been evident for cataract patients with all other questionnaires.[1,19,20] There may also be marginal differences in patient response if the questionnaire were administered in an 8-item format instead of a 14-item format[30]; however, this has not been tested.

In conclusion, our results show that Rasch-scaled versions of VF-14 perform better than Likert-scored versions. In particular, the VF-8R measures cataract surgery outcomes with high precision, possesses psychometric properties comparable to those of the original VF-14, and performs even better than VF-14 in terms of measuring a single construct. Given these benefits, we believe the VF-8R would prove to be a superior tool in cataract outcomes assessment.

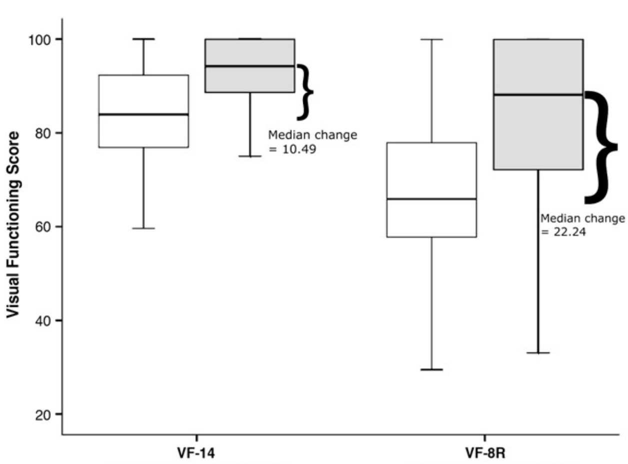

Figure 1. Box-and-whisker plot of preoperative scores (*empty boxes*) and postoperative scores (*solid boxes*) of visual functioning in the overall group of patients (n = 51) using the VF-14 and the Rasch-scaled version, the VF-8R. The boxes contain the interquartile range, and the line running across the center of each box represents the median. The change in the median score was statistically significantly larger for the VF-8R than for the VF-14 (both $P < .0001$, paired *t* test).

REFERENCES

1. 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/44/7/2892. Accessed March 29, 2010
2. Lundström M, Pesudovs K. Catquest-9SF patient outcomes questionnaire; nine-item short-form Rasch-scaled revision of the Catquest questionnaire. J Cataract Refract Surg 2009; 35:504–513
3. Steinberg EP, Tielsch JM, Schein OD, Javitt JC, Sharkey P, Cassard SD, Legro MW, Diener-West M, Bass EB, Damiano AM, Steinwachs DM, Sommer A. National study of cataract surgery outcomes; variation in 4-month postoperative outcomes as reflected in multiple outcome measures. Ophthalmology 1994; 101:1131–1140; discussion by DM O'Day, 1140–1131
4. Steinberg EP, Tielsch JM, Schein OD, Javitt JC, Sharkey P, Cassard SD, Legro MW, Diener-West M, Bass EB, Damiano AM, Steinwachs DM, Sommer A. The VF-14. An index of functional impairment in patients with cataract. Arch Ophthalmol 1994; 112:630–638. Available at: http://archopht.ama-assn.org/cgi/reprint/112/5/630. Accessed March 29, 2010
5. Cassard SD, Patrick DL, Damiano AM, Legro MW, Tielsch JM, Diener-West M, Schein OD, Javitt JC, Bass EB, Steinberg EP. Reproducibility and responsiveness of the VF-14. An index of functional impairment in patients with cataracts. Arch Ophthalmol 1995; 113:1508–1513. Available at: http://archopht.ama-assn.org/cgi/reprint/113/12/1508. Accessed March 29, 2010
6. Bilbao A, Quintana JM, Escobar A, García S, Andradas E, Baré M, Elizalde B. Responsiveness and clinically important differences for the VF-14 index, SF-36, and visual acuity in patients undergoing cataract surgery; for the IRYSS-Cataract Group. Ophthalmology 2009; 116:418–424; e1
7. Alonso J, Espallargues M, Andersen TF, Cassard SD, Dunn E, Bernth-Petersen P, Norregaard JC, Black C, Steinberg EP, Anderson GF. International applicability of the VF-14; an index of visual function in patients with cataracts. Ophthalmology 1997; 104:799–807
8. Uusitalo RJ, Brans T, Pessi T, Tarkkanen A. Evaluating cataract surgery gains by assessing patients' quality of life using the VF-7. J Cataract Refract Surg 1999; 25:989–994
9. Pager CK. Assessment of visual satisfaction and function after cataract surgery. J Cataract Refract Surg 2004; 30:2510–2516
10. Moghimi S, Riazi Esfahani M, Maghsoudipour M. Visual function after implantation of aniridia intraocular lens for traumatic aniridia in vitrectomized eye. Eur J Ophthalmol 2007; 17:660–665
11. Lamoureux EL, Pesudovs K, Thumboo J, Saw S-M, 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
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):I-17–I-24
13. Friedman DS, Tielsch JM, Vitale S, Bass EB, Schein OD, Steinberg EP. VF-14 item specific responses in patients undergoing first eye cataract surgery: can the length of the VF-14 be reduced? Br J Ophthalmol 2002; 86:885–891. Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1771229/pdf/bjo08600885.pdf. Accessed March 29, 2010
14. McHorney CA, Ware JE Jr, Rogers W, Raczek AE, Lu JFR. The validity and relative precision of MOS short- and long-form health status scales and Dartmouth COOP charts; results from the Medical Outcomes Study. Med Care 1992; 30(suppl 5):MS253–MS265
15. 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. Available at: http://www.mja.com.au/public/issues/184_06_200306/kir10534_fm.pdf. Accessed March 29, 2010
16. Bailey IL, Bullimore MA, Raasch TW, Taylor HR. Clinical grading and the effects of scaling. Invest Ophthalmol Vis Sci 1991; 32:422–432. Available at: http://www.iovs.org/cgi/reprint/32/2/422.pdf. Accessed March 29, 2010
17. Mallinson T. Why measurement matters for measuring patient vision outcomes. Optom Vis Sci 2007; 84:E675–E682. Available at: http://journals.lww.com/optvissci/Fulltext/2007/08000/Why_Measurement_Matters_for_Measuring_Patient.6.aspx. Accessed March 29, 2010
18. Wright BD, Masters GN. Rating Scale Analysis (Rasch Measurement Book). Chicago, IL, MESA Press, 1982
19. Gothwal VK, Wright TA, Lamoureux EL, Pesudovs K. Rasch Analysis of Visual Function and Quality of Life Questionnaires. Optom Vis Sci 2009; 86:1160–1168
20. Gothwal VK, Wright TA, Lamoureux EL, Pesudovs K. Rasch analysis of the Quality of Life and Vision Function Questionnaire. Optom Vis Sci 2009; 86:E836–E844
21. Wright BD, Stone MH. Best Test Design; Rausch Measurement. Chicago, IL, MESA Press, 1979
22. Linacre JM. WINSTEPS® Rasch measurement computer program. Beaverton, OR, Winsteps.com, 2009
23. Andrich D. A rating scale formulation for ordered response categories. Psychometrika 1978; 43:561–573
24. Wolfe EW, Chiu CW. Measuring change across multiple occasions using the Rasch Rating Scale Model. J Outcome Meas 1999; 3:360–381
25. McHorney CA, Haley SM, Ware JE Jr. Evaluation of the MOS SF-36 Physical Functioning Scale (PF-10): II. Comparison of relative precision using Likert and Rasch scoring methods. J Clin Epidemiol 1997; 50:451–461
26. 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
27. 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 1):I-25–I-36
28. Fitzpatrick R, Norquist JM, Dawson J, Jenkinson C. Rasch scoring of outcomes of total hip replacement. J Clin Epidemiol 2003; 56:68–74
29. Raczek AE, Ware JE, Bjorner JB, Gandek B, Haley SM, Aaronson NK, Apolone G, Bech P, Brazier JE, Bullinger M, Sullivan M. Comparison of Rasch and summated rating scales constructed from SF-36 physical functioning items in seven countries: results from the IQOLA Project. J Clin Epidemiol 1998; 51:1203–1214
30. Pesudovs K, Burr JM, Harley C, Elliott DB. The development, assessment, and selection of questionnaires. Optom Vis Sci 2007; 84:663–674. Available at: http://www.optvissci.com/pt/re/ovs/pdfhandler.00006324-200708000-00005.pdf;jsessionid=KZcJrBxMKHzzcx4VGhCDgGwV2vDSD5HknjcG729LKxDlX1XPvfm1!-701738752!181195629!8091!-1. Accessed March 29, 2010

First author
Vijaya K Gothwal, PhD

Department of Ophthalmology, Flinders Medical Centre and Flinders University of South Australia, Bedford Park, South Australia