Validity of a visual impairment questionnaire in measuring cataract surgery outcomes

Konrad Pesudovs (PhD), Lynda E Caudle (Grad Cert-Public Health), Gwyneth Rees PhD, Ecosse L Lamoureux, PhD

Accepted for publication December 21, 2007.

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

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

This research was supported in part by NHMRC Centre of Clinical Research Excellence Grant 264620. Konrad Pesudovs is supported by National Health and Medical Research Council (NHMRC, Canberra, Australia) Career Development Award 426765. Ecosse Lamoureux is supported by an NHMRC Public Health Fellowship 359277.

Konrad Pesudovs had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

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.

**<BEGIN ABSTRACT>**

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

**SETTING:** The Flinders Eye Centre of the Flinders Medical Centre, Flinders University, 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 (DIF, whether different subgroups respond 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 DIF. One item misfitted the model *(worry about eyesight getting worse*) and was removed. There was evidence of multidimensionality indicating that the overall IVI score should be discarded, but the 3 subscales (*reading and accessing information*, *mobility and independence* and *emotional well-being*) functioned well. A number of items calibrated differently in cataract compared to low vision indicating different issues are important to each population, and the need for population-specific conversion algorithms. Targeting of the IVI items was biased towards the more impaired patients.

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

*J Cataract Refract Surg 2008; 34:•••–••• © 2008 ASCRS and ESCRS*

**<END ABSTRACT>**

**<BEGIN ARTICLE>**

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 patients' visual acuity remains the most important clinical outcome of cataract surgery for the surgeon, the ability to perform routine daily activities is critically important for the patient.[3] Visual acuity may underestimate the value of surgery because it doesn't necessarily reflect postoperative functional improvement, changes in activities of daily living, and satisfaction with vision.[4] As a consequence, the quality-of-life, economic, and social benefits of improved vision often remain implicit.

Several questionnaires have been validated in patients with cataract such as 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 following cataract surgery and to provide information predictive of the outcome 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 so do not tap other areas of quality of life potentially important to the cataract patient. They are also limited in their development and validation by 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 in 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 AMD.[16] Compared to other instruments which have assessed vision-specific quality of life in patients with cataract, the IVI has undergone substantial validation using 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 they target the spectrum of the overall trait being measured i.e. 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 for both the ADVS and the VF-14; with a lack of items targeted at the more able cataract patient.[10,11]

The original IVI was validated with people with low vision but 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 is unclear. The aim of this study was therefore to empirically determine if the IVI provides a valid assessment of perceived restriction of participation in patients with cataract.

## SUBJECTS AND MATERIALS

### Participants

Participants with cataract were drawn from the public surgery waiting list of the ophthalmology service at Flinders Medical Centre, Adelaide (South Australia). All patients on this list had been previously assessed in the eye clinic, were deemed to have cataract causing visual disability that required surgical intervention. Consecutive patients on the list were invited to participate. This included patients with unilateral and bilateral cataract, those awaiting second eye surgery and those with ocular comorbidity (eg, ARMD). It is important to be inclusive if all these types of cataract patients, in order 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 on the other groups. Other inclusion criteria were aged 18 years or older, no severe cognitive impairment and ability to converse in English without the need for an interpreter. Ethical approval was obtained for both populations and a consent form was signed by each patient who agreed to participate. This research adhered to the tenets of the Declaration of Helsinki.

### The Impact of 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-or-interviewer-administered. 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-*"not at all"* (0), *"a little"* (1), *"a fair amount"* (2), *"a lot"* (3) and a 3-category response scale for 2 items-*"not at all"* (0),*"a fair amount"* (1), *"a lot"* (2).[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.[23] Where 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 for the use of parametric analysis of data. The use of Rasch analysis to validate the IVI has been described extensively previously.[13,15,26] Briefly, 3 overall fit statistics are considered. An item-trait interaction score reported as a Chi-Square ($\chi^2$), which reflects the property of invariance across the trait and therefore indicates whether the data fit the model. An item-trait interaction probability value more 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 where perfect fit to the model would have a mean of approximately 0 and a standard deviation 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 four strata of person ability. Individual item or person statistics with Fit Residuals values more than 2.5 or probability values below the Bonferroni adjusted alpha value were used to indicate misfit of the data to the model. Item removal was also considered if items demonstrated Fit residual values more than 2.5 or less than Bonferroni-adjusted probability scores (*P* = .002).

The ordering of thresholds (ie, how the patents interpret the transition from one category to the next) was investigated. The presence of disordered thresholds demonstrates that the categories are not working as intended. This can occur when there are too many response options, or when the labeling of options is similar to each other, 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

(DIF) was statistically tested to ascertain if 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 DIF. Targeting was also assessed as it was important to determine if 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 (PCA) of the residuals. Unidimensionality is formally tested in RUMM2020 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 two subsets of items statistically differ (using independent t-test provided in RUMM) 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 two sets of items should not be significantly different, ≤5% of cases being dissimilar is tolerated.[27,28]

The subscale structure of the IVI was re-tested for this population using confirmatory factor analysis (CFA). 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 $\chi^2$, mean person and item fit residuals, person separation reliability, DIF, unidimensionality and targeting of items to persons.

The relationship between raw scores and Rasch person measures was determined by a double-asymptotic non-linear regression.[29] Documenting this relationship allows 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.

## RESULTS

The characteristics of the cataract participants (n = 181) who completed the IVI are shown in Table 1. Most participants were elderly, female and reported some general medical comorbidity. The majority of cataract participants had bilateral cataract and did not have ocular comorbidity. Two thirds of the participants had visual acuity better than 6/12.

### Fit of the Impact of Visual Impairment Instrument Data to the 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

RUMM 2020 was statistically significant ($p<0.001$) 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 ($\chi^2$ (df) = 84 (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 response options of the IVI were correctly discriminated by cataract participants.

The mean (SD) person fit residual value however was -0.61 (1.79). 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 participants (3.9%) had Fit Residuals values outside the acceptable range (more than 2.5). Further analysis of the misfitting participants showed inconsistent patterns in the items where extreme responses were observed. Upon removal of these misfitting persons, the mean (SD) person fit Residual value improved to −0.52 (1.53). The mean (SD) item fit residual value did not show any serious misfit −0.41 (1.33). There was however no change in the overall Item-Trait Interaction probability value (*P* = .0015).

Further examination of the items however showed that the item IVI25 *'Worry about your eyesight getting worse'* had an extreme fit residual value (3.49) and probability value of 0.0005 which is substantially less than the Bonferroni-adjusted alpha value of 0.002. This item was subsequently removed. The item-trait interaction statistics substantially improved thereafter showing a non significant probability value ($\chi^2$ =73.2; *P* = .056) demonstrating that IVI data fit the expectations of the measurement model. The item Fit Residual also improved for 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 three responses namely "Not at all"; "A little", "A fair amount" (Table 2). Twenty-six of the twenty-seven items (96.3%) had less than 4% of the participants selecting the response 'A lot' which suggests that overall the participants did not have 'a lot of difficult with the IVI items. At the item level, this pattern was evident for a number of items such as *'operating household appliances'; 'fear of falling or tripping'; 'going down step'; 'safety at home'*; and *'feeling sad or low'* (Table 2).

Overall, the 5 most difficult items in the IVI for our cataract patients 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 (Table 2). Conversely, the five least difficult items were *'Feeling lonely'; 'Recognising 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 to those seen 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 were found to be free from DIF, with probability values exceeding the adjusted alpha value for each of the factors assessed. This finding indicates that the IVI performs with similar accuracy regardless of subgroups within these person factors.

**Targeting**

The participants' 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, there are a number of participants on the right of the graph that have 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 PCA of the residuals identified two 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.  The negative subset (PC loadings < −0.3) comprised 7 items and the positive subset (PC loadings >0.3) comprised 5 items (Table 3). 16.98% (95% CI: 14%-20%) of the person estimates were found to be significantly different and therefore evidence of multidimensionality was detected for the IVI in this population of cataract patients. This finding suggests that the IVI in this population was measuring more than one construct and that the scale could operate optimally if these constructs were assessed individually.

**Confirmatory Factor Analysis**

The goodness of fit statistics for the 3-factor model are shown in Table 4. All the indices showed a reasonable fit between the IVI data and the 3-factor model. The beta coefficients of all items were all 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 the evidence of the 3-subscale structure of the IVI as previously demonstrated with a low vision population.

### Performance of the Subscales Within the Rasch Model

Each subscale was tested for fit to the Rasch model as per the approach taken to the entire IVI above. The results are summarized in Table 5. Each subscale showed good overall performance with a non-significant item-trait interaction $\chi^2$, acceptable person and item fit residuals and good person separation reliability (more than 0.90). There was no evidence of DIF or multidimensionality in any other the subscales. 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.

### 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 firstly 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 illustrated in Figure 3. The relationship is double-asymptotic because the average raw rating has a floor and a ceiling (at 0 and 3). The double-asymptotic non-linear regression equations listed in Table 6 can be used to convert raw scores to Rasch person measures for each subscale.

## DISCUSSION

Overall, the IVI questionnaire performs well in a cataract population. The response categories are used appropriately illustrated by ordered thresholds and there is good person separation indicating that the IVI can discriminate between 4 strata of respondents. The IVI in this cataract population is free from differential item functioning indicating it is consistent across sub populations. 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. The item *"Worry about eyesight getting worse"* did not fit the Rasch model. This suggests that this item behaves differently in cataract patients as compared with low vision patients. This may partly be explained by the fact that cataract patients are awaiting an operation that will in most cases remove their eye problem, together with their concern about its progression. In contrast, low vision patients, commonly have an eye disease which cannot be treated, so progression of which is likely to concern the individual. A quarter of the cataract patients had ocular comorbidity, so perhaps the misfit arose in some being worried about progression of their comorbid eye disease, with the remainder not worried at all. Removing this item confirmed the remaining 27 items performed as a unidimensional item set fitting the Rasch model.

Principal Components Analysis revealed evidence of multi-dimensionality. The IVI has previously been shown to contain 3 viable subscales. When assessed with confirmatory factor analysis, again it was shown that 3 subscales were existent. The high level of difference in person estimates from the different subsets tested in the PCA, suggest that the calculation of an overall score for the IVI in cataract patients should be abandoned. It would be more appropriate to only report 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.

While the response scale categories were ordered, the more impaired choice was underutilized. This suggests a shorter response scale could be utilized in cataract patients which is consistent with Rasch analysis of the ADVS.[10] This response category usage 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. While not a fatal flaw as illustrated by retaining good person separation, the IVI would benefit from items which better targeted the less impaired patients. Possible items could cover very difficult tasks eg, driving in the rain, doing very fine needlework, or possible could be more specific to second eye cataract patients (who are less impaired) such as judging depth tasks like pouring drinks or putting a key into a keyhole, or specific for unilateral visual impairment like do you have trouble seeing on one side. This suggests there are patient-centered issues which lead patients to desire cataract surgery which are not tapped by the IVI. Although, it is worth noting that all questionnaires which have been assessed in cataract patients, including the VF-14 and the ADVS, suffer the same problem.[10,11] 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 whether it performs differently for different conditions. The IVI was previously developed and validated in a low vision population.[13–15] One cannot assume that 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 and this one (Table 2).[15] Of the 27 items retained in the analysis, 15 of the items varied by 0.5 logits and 8 varied by 1.0 logit or more when compared to a low vision population. The largest disparities were *recognizing people (−2.2)*, *getting about outdoors (−1.6)*, *Lonely (−1.6)*, *Sad or low (−2.4)* (items easier for cataract patients) and *spilling things (1.4)*, *Frustrated (2.5)*, *Coping (1.2)*, *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 by depression type of emotional issues such as sadness or loneliness compared to low vision patients. On the other hand cataract patients were more troubled by emotional issues such as difficulty coping, experiencing frustration and their vision interfering with their life overall. These differences emphasize the need to revalidate a questionnaire on different disease populations. Different calibrations for items don't make a major difference to the overall functioning of a questionnaire, nor do they interfere with Rasch analyzing data from this population. However, different calibrations across conditions preclude the use of simple conversions of raw scores to Rasch scores where the conversion is calibrated according to data obtained for a different disease. Thus investigators need to use calibrations derived from the same condition that they are studying or they need to 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 the scoring benefits of Rasch analysis without needing to do the analysis.

This study suggests that the IVI questionnaire is suitable for use as a cataract surgery outcome measure. However, for optimal performance one 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 which simply score visual disability in that it reports results across 3 subscales of participation in activities of daily living namely *reading and accessing information*, *mobility and independence* and *emotional well being*. However, like other cataract surgery outcome measure, the IVI lacks items to target the more able, 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, et al. 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, et al. 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, et al. 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, et al. 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, et al. 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

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; 4:667–680

12. Mallinson T, Stelmack J, Velozo C. A comparison of the separation ratio and coefficient alphas in the creation of minimum data sets. Med Care 2004; 42:1–17

13. Lamoureux E, Pallant J, Pesudovs K, et al. The Impact of Vision Impairment Questionnaire: an evaluation of its measurement properties using Rasch analysis. Invest Ophthalmol Vis Sci 2006; 47:4732–4741

14. Lamoureux EL, Pallant JF, Pesudovs K, et al. 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, et al. 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 E, Hooper C, Lim L, et al. 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 versus raw scores in measuring change in health. Med Care 2004; 42:I25–36
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
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 JE. Assessment of the impact of vision impairment. Invest Ophthalmol Vis Sci 2002; 43:927–935
22. Rasch G. Probabilistic Models for Some Intelligence and Attainment Tests. Chicago: University of Chicago Press 1960
23. Andrich D, Lyne A, Sheridan B, Luo G. RUMM 2020. Perth: 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 E, Ferraro J, Pallant J, et al. Are standard instruments valid for the assessment of quality of life and symptoms in glaucoma? Optom Vis Sci 2007; 84:689–696

27. Smith EV. Detecting and evaluation 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 J. Unidimensionality matters (A tale of two Smiths?). Rasch Measurement Transactions 2006; 20:1

29. Massof RW. Application of stochastic measurement models to visual function rating scale questionnaires. Ophthalmic Epidemiol 2005; 12:103–124

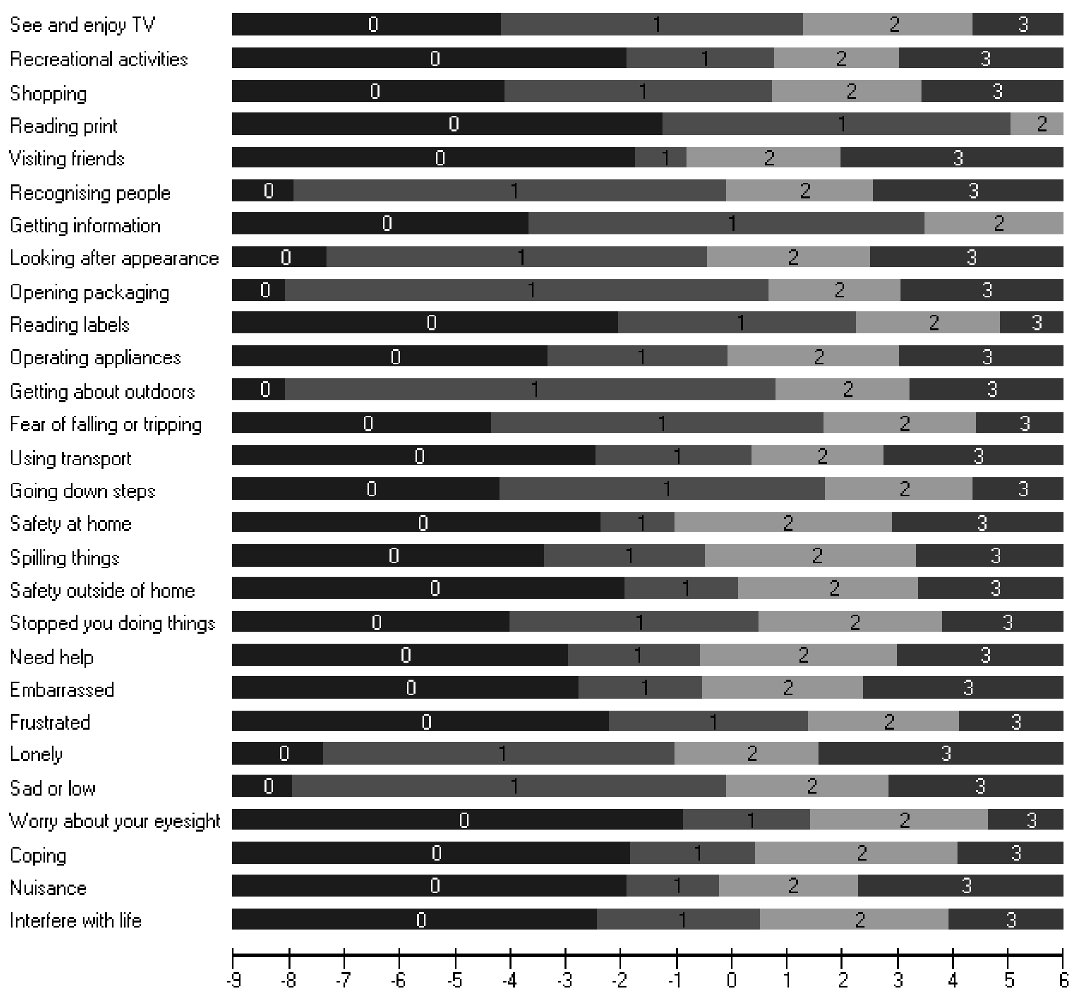
30. Pesudovs K, Garamendi E, Elliott D. The Contact Lens Impact on Quality of Life (CLIQ) Questionnaire: Development and Validation. Invest Ophthalmol Vis Sci 2006; 47:2789–2796

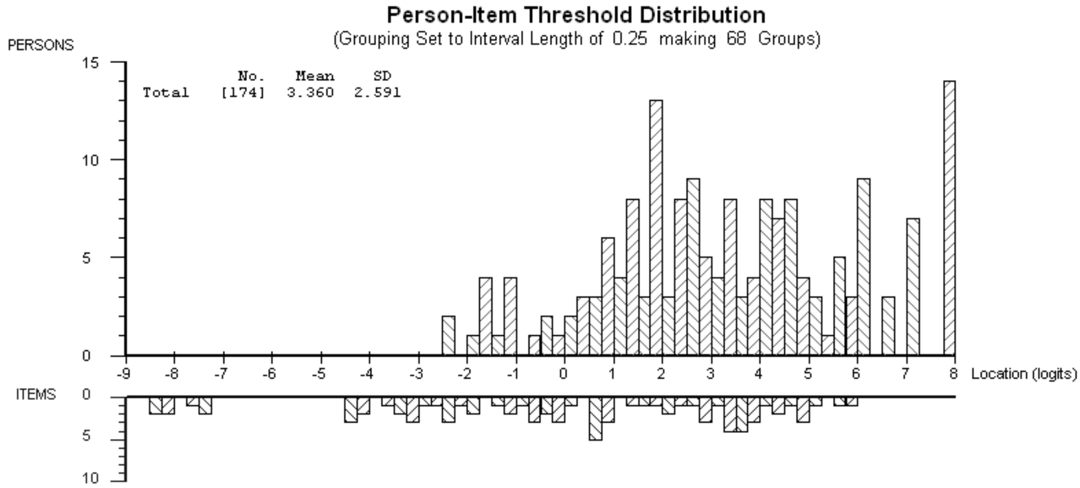
31. Pesudovs K, Burr JM, Elliott DB. How to develop, validate and assess a questionnaire. Optom Vis Sci 2007;84:664-675

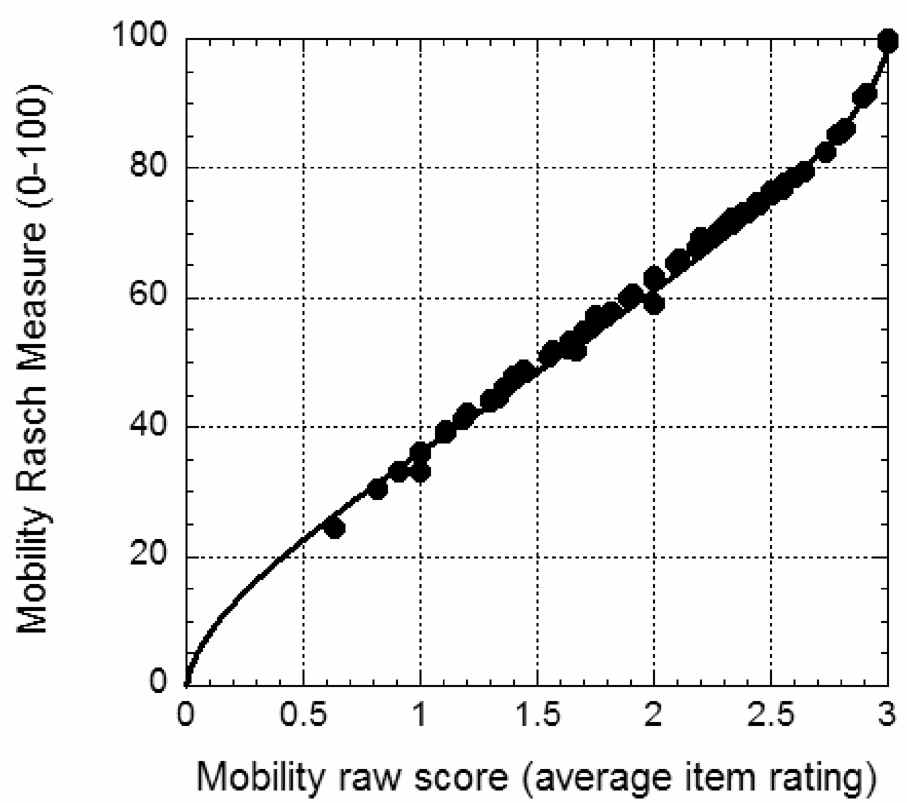
**Figure 1.** Threshold map of the IVI showing ‘ordered thresholds’ which indicates that the participants could reliably discriminate between the categories of difficulty of the IVI.

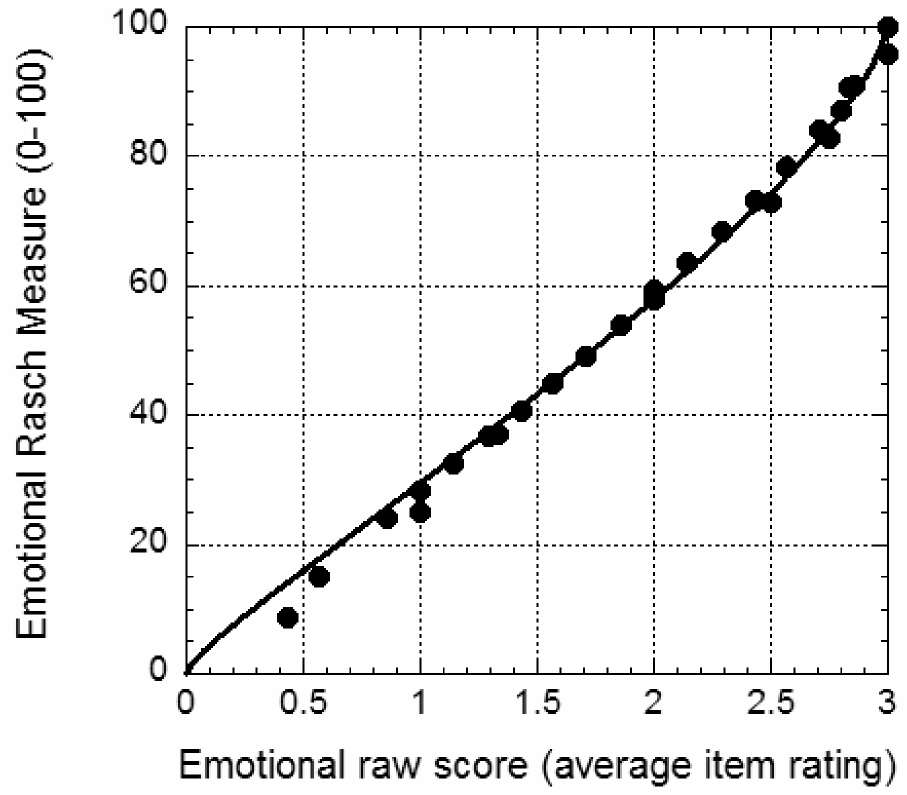
**Figure 2.** The targeting map shows an uneven spread of items across the full range of respondents' scores suggesting less than optimal targeting of the cataract patients (*top*) to the IVI items and thresholds (*bottom*). There are number of participants on the right of the graph that have no difficulty even with the most difficult items of the questionnaire.

**Figure 3.** Scatter plots 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.

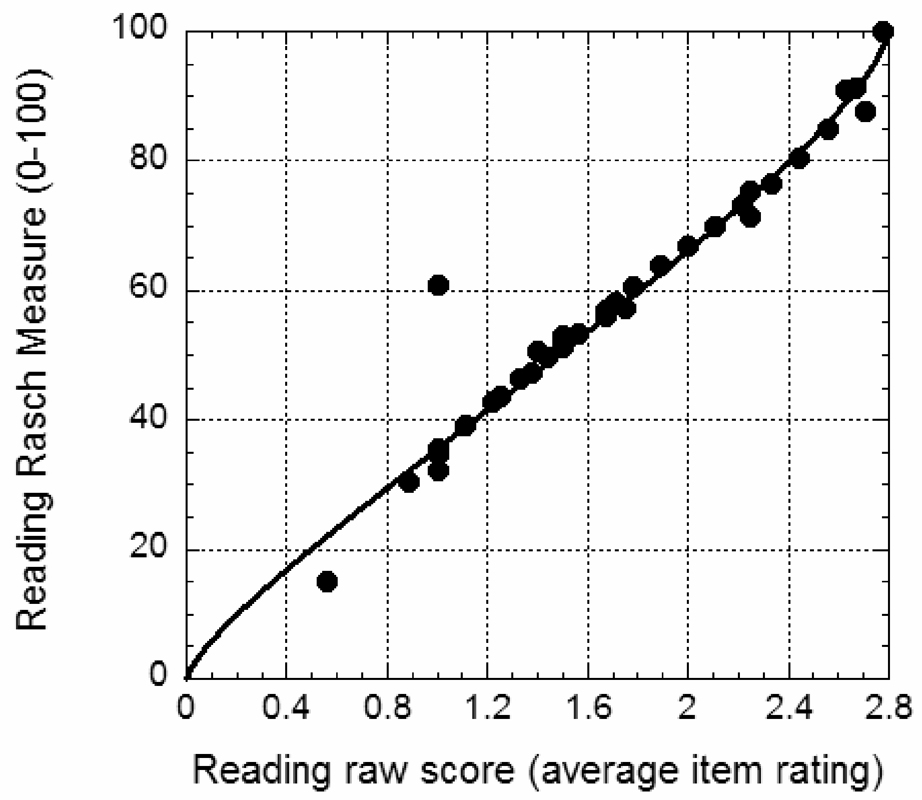

See and enjoy TV
Recreational activities
Shopping
Reading print
Visiting friends
Recognising people
Getting information
Looking after appearance
Opening packaging
Reading labels
Operating appliances
Getting about outdoors
Fear of falling or tripping
Using transport
Going down steps
Safety at home
Spilling things
Safety outside of home
Stopped you doing things
Need help
Embarrassed
Frustrated
Lonely
Sad or low
Worry about your eyesight
Coping
Nuisance
Interfere with life
-9 -8 -7 -6 -5 -4 -3 -2 -1 0 1 2 3 4 5 6


Person-Item Threshold Distribution
(Grouping Set to Interval Length of 0.25 making 68 Groups)
PERSONS
No. Mean SD
Total [174] 3.360 2.591
Location (logits)
ITEMS


Mobility Rasch Measure (0-100)
Mobility raw score (average item rating)



Emotional Rasch Measure (0-100)
100
80
60
40
20
0
0
0.5
1
1.5
2
2.5
3
Emotional raw score (average item rating)


Reading Rasch Measure (0-100)
Reading raw score (average item rating)
0
20
40
60
80
100
0
0.4
0.8
1.2
1.6
2
2.4
2.8