ORIGINAL ARTICLE

Are Standard Instruments Valid for the Assessment of Quality of Life and Symptoms in Glaucoma?

ECOSSE L. LAMOUREUX, PhD, JOHN G. FERRARO, MHA, JULIE F. PALLANT, PhD, KONRAD PESUDOVS, PhD, FAAO, GWYN REES, PhD, and JILL E. KEEFFE, PhD

Centre for Eye Research Australia, The University of Melbourne, Melbourne, Victoria, Australia (ELL, JGF, GR, JEK), School of Rural Health, University of Melbourne, Victoria, Australia (JFP), NH&MRC Centre for Clinical Eye Research-Flinders University and Flinders Medical Centre, Bedford Park South Australia, Australia (KP), Vision CRC, Sydney, Queensland, Australia (ELL, JEK), and Royal Victorian Eye and Ear Hospital, Melbourne, Victoria, Australia (JGF)

ABSTRACT

Purpose. To determine if the impact of Visual Impairment Instrument (IVI) and Glaucoma Symptom Scale (GSS) are valid instruments to assess participation in daily living and ophthalmic complaints, respectively, in a glaucoma population.
Methods. Patients with glaucoma were recruited from private and public clinics and completed the IVI and GSS questionnaires. The two scales were assessed for fit to the Rasch model. Unidimensionality, individual item and person fit to the model, response category performance (how respondents differentiate between the response options), differential item functioning (how subgroups, despite equal levels of the underlying trait, respond differently to an individual item), and targeting of items to patients (good spread of items across the full range of patients' scores) were assessed.
Results. One hundred seventy-five participants (mean age = 71 year) were recruited. The majority (65%) had primary open angle glaucoma and good presenting visual acuity ≥6/9 in the better eye (87%). Only one-third of the participants had severe visual field loss in both eyes. Disordered thresholds were evident across all GSS items, indicating that the categories were difficult to discriminate and required category collapsing (5 to 3 categories). There was no evidence of person and item misfit, differential item functioning, and multidimensionality. However, both scales displayed ineffective person-item targeting as a large number of participants demonstrated little difficulty with the most difficult items.
Conclusions. Because of unsatisfactory targeting, The IVI and GSS are suboptimal scales to assess patients with glaucoma but relatively good vision. It is likely that items could be added to optimize the performance of both instruments. There may however be a need to develop a glaucoma-specific instrument to assess Quality of Life in this population.
(Optom Vis Sci 2007;84:789–796)

Key Words: Quality of Life, glaucoma, Rasch analysis, ocular complaints, IVI

G laucoma is the third most common eye disease in the world and has been projected to become the second leading cause of blindness globally as well as in most regions.[1,2] In Australia alone, it is estimated that there are approximately 200,000 people with glaucoma,[3] and a further understanding of the impact of this eye condition on participation in daily living and Quality of Life (QOL) is needed.

However, most vision-specific QOL assessments of patients with glaucoma have not been performed using a Rasch-scaled measure.[4,5] A Rasch-calibrated instrument provides a transformation of the ordinal raw scores to a linear interval scale permitting the use

of parametric statistical techniques. This process improves the accuracy of scoring and removes noise from the measurement which in turn improves sensitivity to change and correlations with other variables.[6–9] Rasch analysis also assesses the instrument's validity, particularly the fit of items to the overall construct, and the effectiveness of targeting the range of item difficulty to patients' ability.[10] In an earlier study,[11] we demonstrated that the Impact of Visual Impairment Instrument (IVI) was a valid scale to assess participation in daily activities in visually impaired people. Using Rasch analysis, we derived from the ordinal difficulty ratings, interval measures of perceived visual ability for restriction of participation for each pa-

tient (person measure) and the required ability for each item (item measures).

To assess the impact of the vision loss on participation in daily living and QOL, it would be beneficial to have a single and valid instrument which is generalizable and stable across different causes of visual impairment. Considering that the original IVI[12] was validated with almost a quarter of its sample diagnosed with glaucoma and that almost one-third of its items are related to orientation and mobility [activities found to be particularly difficult for glaucoma patients with visual field loss (VF)], the main aim of this study was to determine if the IVI was a valid instrument to assess perceived restriction of participation and QOL specifically in a glaucoma population.

A secondary aim of this study was to assess the Glaucoma Symptom Scale (GSS)[13] using Rasch analysis to determine if it is a valid measure of ophthalmic complaints associated with glaucoma. In the original validation of the GSS,[13] the sample was deliberately selected to represent a full range of the severity of the condition as well as to optimize the process of scale development and testing. However, as argued by the initial authors, their study population may not accurately represent the milder levels of glaucoma generally seen in community-living individuals. Whether the GSS is a valid patient-centered scale for the assessment of symptoms experienced by glaucoma patients living in an Australian community has also not been demonstrated and was explored in this study.

METHODS
Participants and Setting

Participants were recruited from a tertiary public hospital glaucoma clinic at the Royal Victorian Eye and Ear Hospital (RVEEH) and a private specialist glaucoma practice between February 2002 and April 2003. Medical records were reviewed before each clinical session to identify eligible participants who were then approached at the time of their appointment. Most consented to participate (81%). Eligibility criteria included an ability to converse in English, visual acuity (VA) $\geq$6/60 in both eyes, glaucomatous VF in either eye, and a reliable automated visual field test (defined as fixation loss, false negative and false positive errors <33%). Participants who had a history of optic neuropathy other than glaucoma, any form of retinal pathology, corneal opacity, pupil miosis, or had ocular surgery in the previous 3 months were excluded from the study. Informed consent was obtained from each participant. All participants had a clinical assessment and completed the 28-item IV1 and GSS questionnaires which were interviewer-administered. Ethical approval was obtained from the RVEEH Human Research and Ethics Committee. This research adhered to the tenets of the Declaration of Helsinki.

Clinical Assessment

Diagnosis of the type of glaucoma was determined by review of medical record (i.e., as diagnosed by the ophthalmologist). Measurement of presenting VA was performed using a Snellen chart at 6 m. Visual acuity was measured in each eye with the participant using their habitual distance correction. Monocular visual fields were tested using 24–2 threshold program on the Humphrey Visual Field Analyzer (Humphrey Instruments, San Leandro, CA). Participants who had unreliable automated field

results (e.g., fixation loss, false-negative, and false-positive errors >33%) on the day of testing were excluded from the study. Our visual field assessment procedures followed those of Advanced Glaucoma Intervention Study (AGIS) for appropriate age-related plus power lens, test stimulus, room lighting, pupil size, testing right eye first, and test instructions and encouragement throughout testing.[14] Visual field results are reported as mean defect and by AGIS scores.

The Impact of Vision Impairment

The IVI profile was developed specifically to assess patient's vision rehabilitation needs in the context of limitation of participation resulting from impaired vision. A detailed description of the IVI questionnaire has been fully published elsewhere.[12] Responses to the IVI items were rated (numerical score in brackets) as "not at all" (0), "hardly at all," (1) "a little," (2) "a fair amount," (3) "a lot," (4) "can't do because of eyesight," (5) or "don't do because of other reasons" (8). The IVI was either administered by a trained interviewer or self-administered by the participant. Recently, the IVI was further validated using Rasch analysis.[11] This resulted in a 28-item questionnaire with a four-category response scale for 26 items and a three-category response scale for two items.[11] The revised 28-item IVI was used in this study.

The Glaucoma Symptom Scale

The GSS is a questionnaire used to quantify complaints or symptoms experienced by patients with glaucoma.[13] The questionnaire comprises 10 ocular complaints, 6 of which are nonvisual and 4 of which are visual. The nonvisual complaints include "burning/smarting/stinging," "tearing," "dryness," "itching," "soreness/tiredness," and "feeling of something in the eye." The visual complaints include "blurry/dim vision," "hard to see in daylight," "hard to see in darkness," and "halos around lights." Initially, the patient was asked the following question: "Have you experienced the following problem in the last four weeks?" and if so, they were asked the following question: "How bothersome has it been?" For each eye, a five-level score is recorded ranging from 0 (complaint present and very bothersome) to 4 (no complaint).

Statistical Analyses

Descriptive analyses were performed to characterize the demographic and clinical characteristics of the participants (SPSS statistical software, Version 14.1, SPSS Science, Chicago, IL).

Rasch Analysis.   The IVI and GSS data were assessed for fit to the Rasch model[15] using the RUMM 2020 software (RUMM Laboratory Pty Ltd., Perth, Australia).[16] The Rasch model assumes that the probability of a patient affirming an item is a logistic function of the relative distance between the item and patient locations on a linear scale. Hence, it is anticipated that the probability of endorsing a particular rating category will increase monotonically with the difference between the person's level of difficulty in performing daily activities and the level of difficulty required for the task. Where the IVI and GSS data meet the Rasch model expectations, the ordinal raw score is transformed into a true Rasch scale (logit).[17,18]

The logit or log-odds unit is the mathematical unit of Rasch measurement. A positive logit item indicates that the item requires a higher level of ability than the mean of the items, whereas a negative item logit suggests that the item requires a lower level of ability than the average. For ease of interpretation, the IVI rating scale scoring was reversed for the Rasch analysis, in which the participants with the high visual ability (low level of restriction of participation) were given the high scores. The rating scale for the GSS was not reversed, as the most able participants (or participants with the least problem) were given a high score.

Parameters Determining Fit of the GSS or IVI to the Rasch Model

The partial credit approach[19] was used as the likelihood-ratio test of RUMM 2020 was found to be statistically significant (p < 0.001) indicating that the rating scale model (which requires all items to have the same response options) was inappropriate. An item-trait interaction statistic reported as a χ^2 value is used and a p value >0.05 indicates no substantial deviation between the IVI or GSS data and the Rasch model. Individual item or person statistics with Fit Residuals values >2.5 or p values below the Bonferroni-adjusted alpha value are used to indicate misfit of the data to the model. Fit to the model was evaluated using person and item fit residual statistics, which are transformed residuals approximating a z-score and representing a standardized normal distribution, in which an optimal fit to the model would have an expected mean value of 0 and variance of 1.

Item removal is also considered if items demonstrate fit residual values >2.5 or less than Bonferroni-adjusted probability scores [p = 0.001 (0.05/32) and = 0.005 (0.05/10), for the IVI and GSS, respectively]. 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.

Disordered Threshold and Differential Item Functioning

Disordered thresholds occur when participants have difficulty discriminating between the response options. This means literally that a category expected to be "harder" than an adjacent category was actually "easier," but often represents interchangeability of categories. Category collapsing is often the solution to disordered thresholds, which can improve overall fit to the model.

Misfit of the data to the Rasch model could also be linked with differential item functioning (DIF) where different groups within the sample (e.g., gender), despite equal levels of the underlying trait, respond differently to an individual item. DIF can be detected both graphically and statistically using analysis of variance comparing scores across each level of the person factor and across different levels of trait (referred to as class intervals).

Targeting

Targeting was also assessed as it was important to determine if the IVI and GSS items were particularly suitable to assess participation in daily living and complaints associated with glaucoma in this specific population. 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. Finally, the person-item deviation residuals are examined by principal components analysis for associations which may be indicative of the breach of the assumption of local independence. The absence of such associations, in addition with adequate fit of the data to the Rasch model, support unidimensionality of the construct.[20]

RESULTS

The mean age of the 175 participants (95 males, 54%) was 71.1 years and the majority of them (113, 65%) had primary open angle glaucoma (Table 1). One hundred fifty-three (87%) participants had presenting VA ≥6/9 in the better eye. The mean deviation average scores (dB) for the better and worse eyes were −10.4 ± 8.3 and −10.2 ± 8.6, respectively. The mean VF, as assessed by the AGIS score, was considered moderate for the better and worse eyes (AGIS = 8.3 and 8.6, respectively). Only a third of the participants were categorized as having severe VF (AGIS ≥12) in both eyes.

TABLE 1.
The characteristics of the 175 study participants

|	
Age (yr), mean ± SD (range)	71.1 ± 11.8 (23–93)
Gender	
Men	95 (54%)
Women	80 (46%)
Main cause of vision loss	
Primary open angle glaucoma	113 (65%)
Other types of glaucoma (e.g., angle-closure, secondary glaucoma, etc.)	62 (35%)
Visual acuity	
Better eye	
6/9 or better	153 (87.4%)
6/12–6/18	19 (10.9%)
<6/18–6/60	3 (1.7%)
Worse eye	
6/9 or better	106 (60.6%)
6/12–6/18	54 (30.9%)
<6/18–6/60	15 (8.5%)
HVF 24–2: mean deviation (dB)	
Better eye, mean ± SD (range)	−10.4 ± 8.3 (−32.7–2.4)
Worse eye, mean ± SD (range)	−10.2 ± 8.6 (−33.7–19.9)
AGIS score	
Better eye	
Mean SD (range)	8.3 ± 5.9 (0−20)
Mild field loss (AGIS = 0–5)	73 (42%)
Moderate field loss (AGIS = 6–11)	46 (26%)
Severe field loss (AGIS = ≥12)	56 (32%)
Worse eye	
Mean± SD (range)	8.6 ± 6.6 (0−20)
Mild field loss (AGIS = 0–5)	75 (43%)
Moderate field loss (AGIS = 6–11)	35 (20%)
Severe field loss (AGIS = ≥12)	65 (37%)
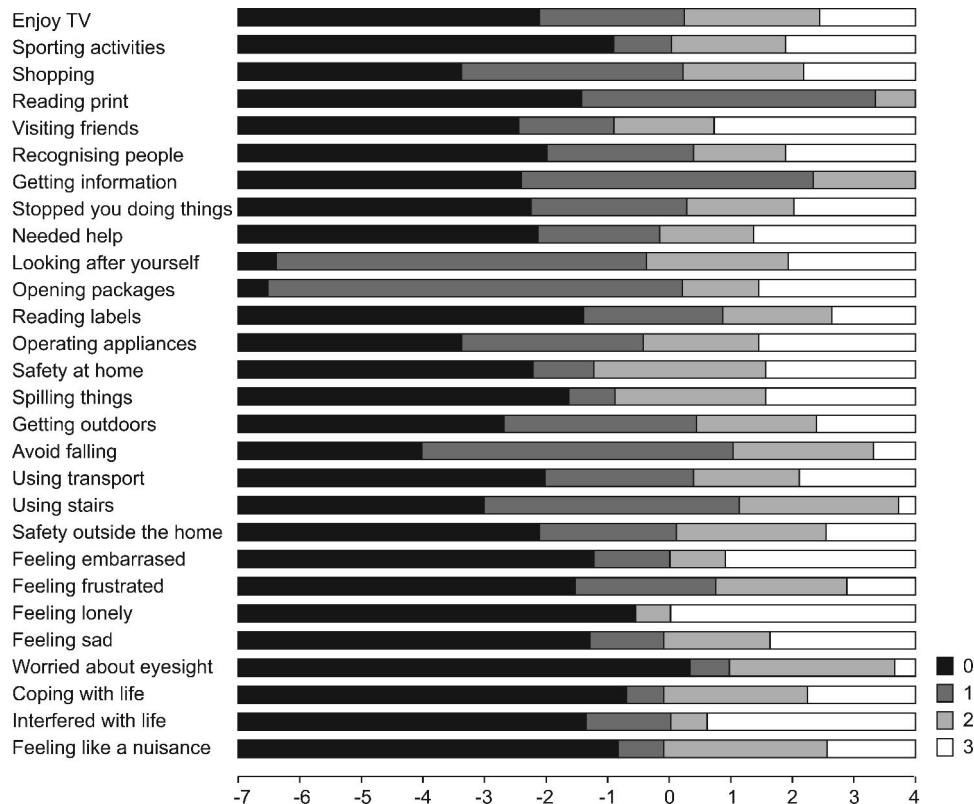
FIGURE 1.
Threshold map of the IVI showing "ordered thresholds," which indicates that the participants could reliably discriminate between the categories of difficulty of the IVI.

Overall Fit of Data to the Rasch Model

The IVI.   Rasch analysis of the 28-item IVI showed fit to the Rasch model with a nonsignificant item-trait interaction total χ^2 probability value (total $\chi^2 = 72.2$, df = 56, p = 0.07). The pattern of item thresholds showed all items having ordered thresholds (Fig. 1).

The GSS.   The GSS items for the left and right eyes were Rasch-analyzed separately and were found to be almost identical. Consequently, only the results for the left eye are presented here. As opposed to the IVI, a pattern of disordered threshold was evident across all GSS items. For the item "itching," for example (Fig. 2), there is evidence of two response categories not being used consistently. Categories 1 and 3 are used interchangeably with adjacent categories. The five categories were therefore collapsed to three categories and recoded as 0,1,1,2,2. Recoding produced ordered thresholds for all items (Fig. 3).

Estimates of Person and Item Measures and Overall Fit Values

The IVI.   All items showed Fit Residuals values <2.5 with Bonferroni-adjusted probability scores >0.001 (0.05/32 items) indicating no significant deviation from the model. The mean (SD) item and person Fit Residuals were -0.39 (1.41) and -0.23 (1.24), respectively. When the items and persons fit the Rasch model, the mean and standard deviation values tend to approximate 0 and 1, respectively. The Person Separation Reliability was 0.94 indicating a capacity for the IVI instrument to

separate four or more strata of person ability. These values, together with a nonsignificant item-trait interaction (p = 0.14), substantiated that the IVI data fitted the Rasch model and the hierarchical ordering of the items is consistent across all levels of the trait "restriction of participation in daily living."

The GSS.   Subsequent to recoding, all the GSS items displayed Fit Residual values <2.5 and probability values >0.005 (Bonferroni-adjusted -0.05/10 items) indicating no significant deviation from the model (Table 2). The mean (SD) item and person Fit Residuals were -0.21 (0.88) and -0.21 (0.94), respectively, suggesting that items and persons tended to fit the Rasch model as the mean and SD values approximated 0 and 1, respectively. Because there were no missing values, RUMM indicated that the GSS data had substantial internal reliability (Cronbach's alpha value = 0.86) and person separation reliability (0.86). A nonsignificant item-trait interaction (total $\chi^2 = 20.5$, df = 20, p = 0.42) supported that the GSS data fitted the Rasch model.

Differential Item Functioning

The DIF method was used to determine whether different subgroups, despite equal levels of participation in daily living, respond in a different manner to an individual item of IVI and GSS. All items were found to be free from DIF, with probability values exceeding the adjusted alpha value for each of the person factors assessed (gender, level of VF and VA, and type of glaucoma).

litch Itching   Locn = .0.029   Unit = 0.116   FitRes = 0.126   ChiSq[Pr] = 0.434   SampleN = 159
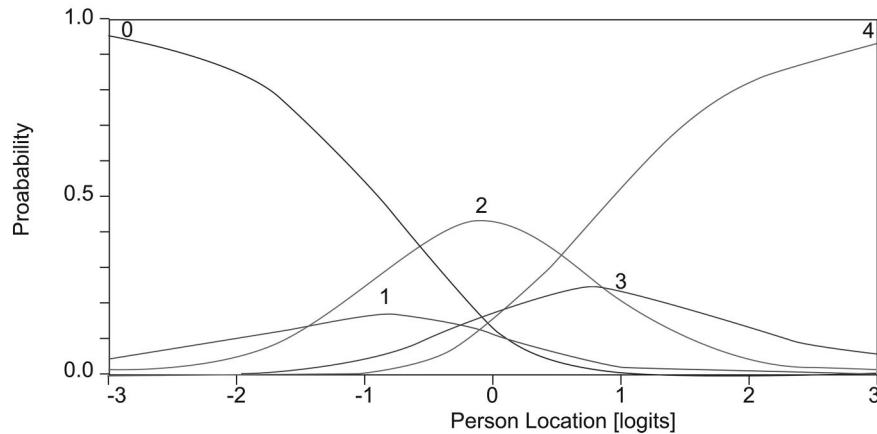

FIGURE 2.

Category probability curve showing disordered thresholds for item litch-"itching." Categories 1 and 3 do not have a range along the ability scale suggesting that they are the least likely categories to be selected. These two categories are underutilized and used interchangeably with categories 2 and 4, respectively. Collapsing categories 1 and 3 with adjacent ones may help produce "ordered threshold."
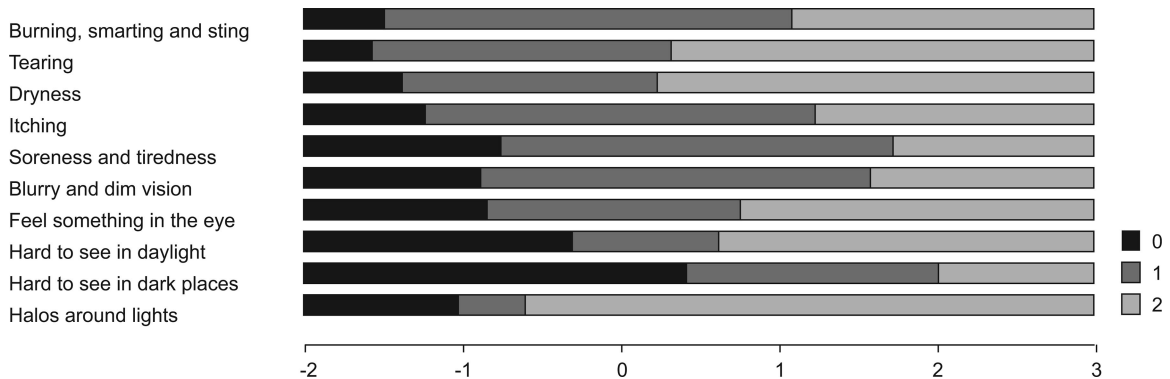

FIGURE 3.

Threshold map of the GSS showing ordered thresholds following item rescoring.

TABLE 2.

Fit of the 10 GSS items (left eye) to the Rasch model after rescoring

Item	Location	Fit residuals	χ^2	p
Burning/stinging	−0.19	−1.67	4.31	0.12
Tearing	−0.62	0.56	1.01	0.60
Dryness	−0.57	−0.15	1.60	0.45
Itching	0.00	0.17	1.56	0.46
Soreness/tiredness	0.48	−0.67	2.08	0.35
Foreign-body sensation	−0.04	−0.67	1.16	0.56
Blurry/dim vision	0.35	−1.28	5.83	0.05
Difficulty seeing in daylight	0.16	−0.20	0.59	0.74
Difficulty seeing in darkness	1.22	1.12	1.09	0.58
Halos around lights	−0.81	0.65	1.32	0.52

All items showed Fit Residuals values <2.5, degrees of freedom = 27.7, and Bonferroni adjusted probability scores >0.005.

Targeting

The IVI. The participants' range of ability (−2.2 to 5.3 logits) was found to have a normal distribution (Kolmogorov-Smirnov Z test score = 0.92; p = 0.36). The three most difficult items were "Worried about your eyesight getting worse?"; "Reading a sign across the street?" and "Feeling frustrated or annoyed?" with logit scores of 1.67, 0.96 and 0.72, respectively.

Conversely, the three least difficult items were "Opening packaging?"; "Generally looking after your appearance?" and "Visiting family and friends?" recording logit values of −1.61, −1.60, and −0.87, respectively. The left panel of the targeting map (Fig. 4) shows the frequency distribution of participants along the Rasch calibrated scale for "participation in daily living" and the right-hand panel shows the position of each item
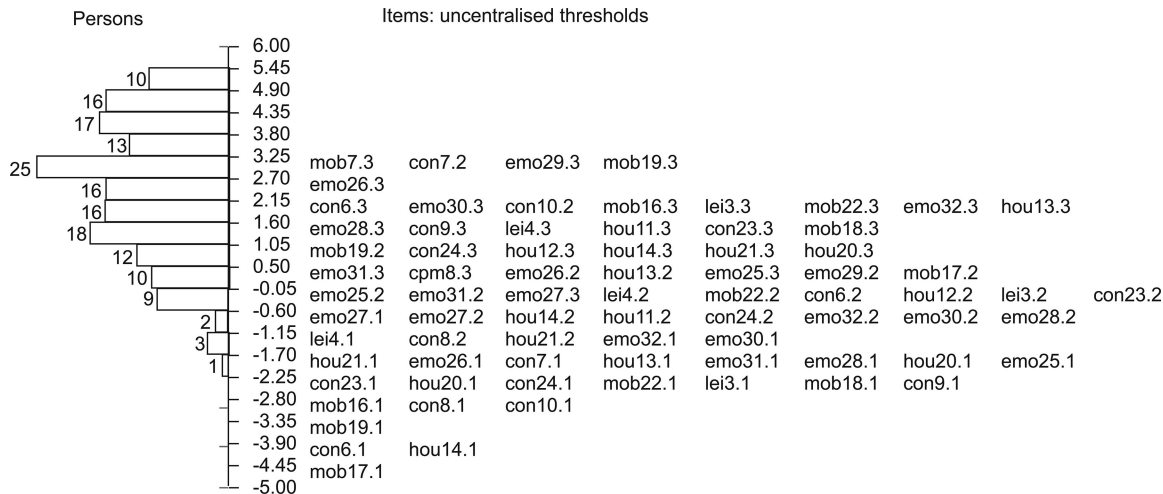
FIGURE 4.

The Person-Item threshold map of the 28-item IVI for the glaucoma participants which shows an uneven spread of items across the full range of respondent's scores. Most items are located in the bottom half of the graph and there were more participants who did not have difficulty endorsing even the most difficult items (top half), compared to those who had substantial difficulty performing even the easiest activities (bottom half).

and their thresholds. Values after the decimal point indicate the specific threshold. The mean person location logit score was 2.6 (SD, ±1.8), which indicates that overall the participants possessed a substantially higher level of participation in daily living than the average of the scale items (0 logit). A large number of participants demonstrated little or no difficulty with the most difficult items (clustered at the top left) and most of the items were found at the bottom of the graph (clustered at the bottom right). The uneven spread of items across the full range of the participant's scores and the relatively high mean person score (2.6) indicate poor item-person targeting for the glaucoma patients on the IVI.

The GSS.   Similar to the IVI, the participants' range of ability (−2.31 to 3.0 logits) was found to be normally distributed (Kolmogorov-Smirnov Z test score = 1.21; p = 0.11). The three most bothersome items were "difficulty seeing in darkness," "soreness/tiredness" and "blurry/dim vision" (1.22, 0.48, and 0.36, respectively). Conversely, "seeing halos around lights," "tearing" and "dryness" were the three least experienced complaints recording logit vales of −0.81, −0.62, and −0.57, respectively. The mean person (SD) logit score was 1.76 (1.5), which indicated that overall the participants experienced a substantially lower level of complaint than the average of the scale items (0 logit). Overall, there was an uneven spread of items across the full range of respondent's scores with a number of patients finding it easy to endorse the most bothersome items (Fig. 5). Most of the GSS items tended to cluster in the bottom half of the graph suggesting an ineffective targeting of the GSS items to the participant's level of complaint.

Unidimensionality

Principal Components Analysis of the residuals identified two subsets of items consisting of the highest positive and negative loading items. Person estimates generated for these two subsets were subjected to a series of independent t-tests to compare the estimates for each person. Only 2.5% and 3.16% of the IVI and GSS estimates, respectively, were found to be significantly differ-
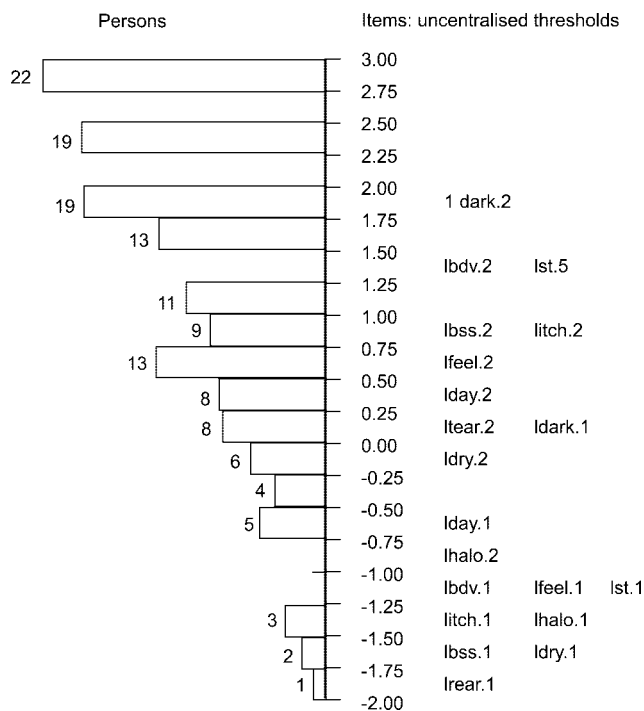

FIGURE 5.

The Person-Item threshold map of the 10-item GSS for the glaucoma participants which also shows an uneven spread and lack of items across the full range of respondent's scores. Most items are located in the bottom half of the graph and there were a number of participants (top half) who did not have difficulty endorsing even the most difficult items.

ent. These values are less than the recommended cut point of 5% and therefore no evidence of multidimensionality was detected.

Criterion Validity

The criterion validity of the Rasch-scaled IVI was assessed by its ability to discriminate between participants of different levels of glaucoma namely mild (AGIS = 0 to 5), moderate (AGIS = 6 to

11) and severe (AGIS = ≥12). There was a significant difference between the three categories for the better [ANOVA; $F(2, 172) = 3.48$; $p = 0.03$] and worse eyes [ANOVA; $F(2, 172) = 4.66$; $p = 0.01$]. Poorer visual field was associated with greater restriction of participation in the better (2.9. 2.6 and 2.0 mean logit values for mild, moderate and severe, respectively) and worse eyes (3.0. 2.6 and 2.1 mean logit values for mild, moderate, and severe, respectively). In contrast, no significant differences were found between the three visual field categories on the GSS person measures for the better [ANOVA; $F(2, 172) = 0.22$; $p = 0.80$] and worse eyes [ANOVA; $F(2, 172) = 0.46$; $p = 0.63$].

DISCUSSION

The IVI and the GSS scales underwent Rasch analysis to determine their validity in assessing restriction of participation and ocular complaints, respectively, in patients with glaucoma. Our results demonstrate that both scales in their current forms do not meet the standards of measurement defined by the Rasch model. The evidence of disordered thresholds was apparent in the GSS and category collapsing was required. The GSS scale also showed poor criterion validity and both instruments demonstrated substantial ineffective item-person targeting, in particular the IVI. This finding suggests that these two instruments are not effectively assessing the measured traits and could be optimized if other items were added to assess the participation and problems in patients with glaucoma but relatively good vision. However, the performance of both measures in people at the more severe spectrum of the disease has not been comprehensively evaluated in this study and further work is needed.

One of the goals of this study was to determine whether the patient-based assessment of participation in daily life developed for patients with low vision was equally valid for patients with relatively good VA found in glaucoma. Our overall conclusion does not support the hypothesis despite a good person separation reliability value, no evidence of multidimensionality and DIF. The main problem is poor targeting of item difficulty to patient ability. Poor targeting is probably due to 90% of the glaucoma patients having good central vision (6/9 or better). Considering that several studies using the IVI have shown that restriction of participation is positively related with VA,[12,21–23] our finding that the items were relatively easy to endorse and not effectively targeting this specific population is therefore not surprising. It appears that many of the causes of visual impairment that characterize patients with low vision are not similar to those with glaucoma where visual fields may be impaired in the presence of good VA. This finding indicates that there may be a need for the development of eye disease-specific quality of life questionnaires; notably for glaucoma patients with visual impairment.

The GSS had similar problems as the IVI, with ineffective targeting. Poor targeting could be partly due to the nature of the population we employed. The GSS was initially validated with a convenience sample designed to balance participants by age groups, gender and age as well as glaucoma patients representing the full range of vision impairment recruited from university-based ophthalmology practices.[13] As opposed, our sample included glaucoma patients attending a public tertiary clinic and a private practice. It is also possible that our sample did not possess the range of

the severity of VF to experience the ocular complaints associated with glaucoma. This argument appears to have some validity as two-thirds of our patients were considered to have mild to moderate VF in the worse eye (AGIS scores = 0 to 11) and that our mean worse eye AGIS score was lower than that recorded in the initial study (8.6 vs. 10.5). Nevertheless, this is an important population for testing the GSS on as it represents the two main modes of practice for glaucoma patients in Australia. Overall, it appears that the GSS scale, in spite of possessing a number of good psychometric characteristics (i.e., unidimensionality, internal validity, person separation reliability and valid measurement characteristics) may not be an optimal instrument to assess fully the range of symptoms and complaints experienced by this sample of glaucoma patients attending private or public clinics.

The problem of poor targeting of relatively minimally impaired patients is a problem in other areas of ophthalmology. Questionnaires which measure visual disability are frequently used to measure the outcome of cataract surgery.[24] Commonly, after cataract surgery, particularly after both eyes have been operated on, patients have no visual disability.[10,25] This leads to a ceiling effect on measurement, by definition, which is difficult to avoid. Perhaps the same problem is difficult to eliminate in a glaucoma symptom scale as one of the goals of management is to minimize symptoms, so if this is done effectively in a majority of cases, a ceiling effect on symptom measurement may be inevitable.

In summary, the GSS demonstrated poor criterion validity and both measures showed evidence of unsatisfactory item targeting for people with glaucoma and relatively good vision. It is likely that items could be added to optimize the performance of both instruments. Future research should assess the performance of other existing glaucoma specific instruments[26,27] and there may be a need to develop a new one.

REFERENCES

1. Resnikoff S, Pascolini D, Etya'ale D, Kocur I, Pararajasegaram R, Pokharel GP, Mariotti SP. Global data on visual impairment in the year 2002. Bull World Health Organ 2004;82:844–51.
2. Quigley HA. Number of people with glaucoma worldwide. Br J Ophthalmol 1996;80:389–93.
3. Taylor HR. Eye care for the future: the Weisenfeld lecture. Invest Ophthalmol Vis Sci 2003;44:1413–8.
4. Noë G, Ferraro J, Lamoureux E, Rait J, Keeffe JE. Associations between glaucomatous visual field loss and participation in activities of daily living. Clin Experiment Ophthalmol 2003;31:482–6.
5. Spaeth G, Walt J, Keener J. Evaluation of quality of life for patients with glaucoma. Am J Ophthalmol 2006;141:S3–14.
6. 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:I25–36.
7. 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;46:1375–83.
8. Pesudovs K. Patient-centred measurement in ophthalmology–a paradigm shift. BMC Ophthalmol 2006;6:25.
9. Pesudovs K, Garamendi E, Elliott DB. The Quality of Life Impact of Refractive Correction (QIRC) Questionnaire: development and validation. Optom Vis Sci 2004;81:769–77.

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;44:2892–9.

11. 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;47:4732–41.

12. Weih LM, Hassell JB, Keeffe J. Assessment of the impact of vision impairment. Invest Ophthalmol Vis Sci 2002;43:927–35.

13. Lee BL, Gutierrez P, Gordon M, Wilson MR, Cioffi GA, Ritch R, Sherwood M, Mangione CM. The Glaucoma Symptom Scale. A brief index of glaucoma-specific symptoms. Arch Ophthalmol 1998;116:861–6.

14. Advanced Glaucoma Intervention Study. 2. Visual field test scoring and reliability. Ophthalmology 1994;101:1445–55.

15. Rasch G. Probabilistic Models for Some Intelligence and Attainment Tests. Chicago: University of Chicago Press; 1980.

16. Andrich D, Lyne A, Sheridan B, Luo G. RUMM 2020 [Computer program]. Perth: RUMM Laboratory; 2003.

17. 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.

18. 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):S22–6.

19. Masters GN. A Rasch model for partial credit scoring. Psychometrika 1982;47:149–74.

20. Lundgren-Nilsson A, Grimby G, Ring H, Tesio L, Lawton G, Slade A, Penta M, Tripolski M, Biering-Sorensen F, Carter J, Marincek C, Phillips S, Simone A, Tennant A. Cross-cultural validity of functional independence measure items in stroke: a study using Rasch analysis. J Rehabil Med 2005;37:23–31.

21. Hassell JB, Lamoureux EL, Keeffe JE. Impact of age related macular degeneration on quality of life. Br J Ophthalmol 2006;90:593–6.

22. Lamoureux EL, Hassell JB, Keeffe JE. The determinants of participation in activities of daily living in people with impaired vision. Am J Ophthalmol 2004;137:265–70.

23. Lamoureux EL, Hassell JB, Keeffe JE. The impact of diabetic retinopathy on participation in daily living. Arch Ophthalmol 2004;122:84–8.

24. de Boer MR, Moll AC, de Vet HC, Terwee CB, Volker-Dieben HJ, van Rens GH. Psychometric properties of vision-related quality of life questionnaires: a systematic review. Ophthal Physiol Opt 2004;24:257–73.

25. Pesudovs K, Elliott DB, Coster DJ. The Cataract Outcomes Questionnaire: a Rasch-scaled measure of visual disability. Invest Ophthalmol Vis Sci 2005;46:E-abstract 3844.

26. Nelson P, Aspinall P, Papasouliotis O, Worton B, O'Brien C. Quality of life in glaucoma and its relationship with visual function. J Glaucoma 2003;12:139–50.

27. Janz NK, Wren PA, Lichter PR, Musch DC, Gillespie BW, Guire KE. Quality of life in newly diagnosed glaucoma patients: The Collaborative Initial Glaucoma Treatment Study. Ophthalmology 2001;108:887–97.

Ecosse Lamoureux
Centre for Eye Research Australia
Department of Ophthalmology
University of Melbourne
32 Gisborne St
East Melbourne, Victoria 3002, Australia
e-mail: ecosse@unimelb.edu.au