1040-5488/10/8704-0285/0 VOL. 87, NO. 4, PP. 285–293
OPTOMETRY AND VISION SCIENCE

REVIEW

Item Banking: A Generational Change in Patient-Reported Outcome Measurement

Konrad Pesudovs*

ABSTRACT

Purpose. Patient-reported outcomes are traditionally measured with questionnaires and many have been developed to measure Vision-Related Activity Limitation (VRAL; visual disability or visual functioning), Symptoms, and Quality Of Life (QOL). These vary in quality and can be classified as First or Second Generation instruments. First generation instruments are characterized by simple summary scoring of ordinal responses, which precludes interval measurement. This problem is solved in second generation instruments where Rasch analysis is used to optimize psychometric properties. However, second generation instruments retain limitations; difficulties in comparing scores across instruments, limited applicability to populations and inability to adapt to change. A third generation approach to patient-reported outcomes measurement, item banking, can solve these problems. The aim of this project was to use Rasch analysis to calibrate all items from all instruments to form VRAL, Symptoms, and QOL Item Banks.

Methods. Six hundred twenty-four people on the waiting list for cataract surgery were recruited. Each participant completed, by self-administration, a number of the 19 instruments. A total of 353 items were calibrated using Rasch analysis (Winsteps v3.67). The psychometric properties of each item bank were optimized; items fitting the Rasch model were retained (Infit and Outfit range, 0.50 to 1.50).

Results. Items were sorted into the three traits; 226 tapped VRAL, 22 symptoms, and 60 QOL. Satisfactory measurement of each latent trait occurred with person separation of 8.11 for VRAL, 2.33 for Symptoms, and 3.20 for QOL. Rasch estimates of item difficulty were highly stable with an average standard error of 0.11 logits.

Conclusions. Item banks for the measurement of the latent traits of VRAL, symptoms, and QOL have been formed. New items can be added to enable evolution of measurement. Item banking facilitates accurate and precise measurement through computer adaptive testing. This approach provides common measurement scales, facilitating worldwide comparison of results.

(Optom Vis Sci 2010;87:285–293)

Key Words: cataract, Rasch analysis, visual function, Quality of Life, questionnaire

Today, the practice of Optometry is accompanied by the obligation to provide treatments supported by an evidence base. This has been stressed for the treatment of glaucoma, for example, where large multi-center randomized controlled trials have been performed,[1] but it is also true across the spectrum of optometric interventions including the provision of spectacles, contact lenses, and visual training. The source of the evidence for the benefit of interventions is outcomes research. A number of end points may be used to study the effect of interventions and these are called outcome measures. In Optometry and Ophthalmology, outcome measures include survival (e.g., corneal graft), anatomic measures, physiological measures (e.g., intraocular pressure), optical quality, and visual performance. Like many others, I've long been interested in the impact of eye disease and its treatments on optical quality and visual performance.[2–4] Although this field is endlessly fascinating, we need to face the harsh reality that measures of optical quality of visual performance are only surrogate measures. Of course, what really matters is the outcome from the patient's point of view (after all it doesn't matter how much further down the visual acuity chart a patient can read with their new glasses, if they find them intolerable to wear).

*PhD, FAAO
Department of Optometry and Vision Science, NH&MRC Centre for Clinical Eye Research, Flinders Medical Centre and Flinders University of South Australia, Bedford Park, South Australia, Australia.

Patient-Reported Outcomes

Any report coming directly from a patient concerning the outcome of an intervention is a patient-reported outcome

(PRO). Usually these take the form of questionnaires and can be as simple as asking a patient whether they are satisfied with the outcome of the intervention up to large instruments with complicated scoring systems. The need for PROs should be self-evident given that outcomes research simply formally tests what occurs in clinical practice where a good practitioner would naturally ask their patients about the benefits of interventions. This is widely recognized across medicine, where the inclusion of PROs in outcomes research is often mandated by research funding bodies, ethics committees, third party payers, and regulatory agencies like the Food and Drug Administration.[5–8] Although the importance of measuring PROs has moved beyond question, the key to performing good outcomes research is to have good outcome measures, including PROs. Optometrists are experts in the measurement of visual performance and therefore understand the advantages logMAR visual acuity testing has over Snellen visual acuity testing and why the former is preferred as an outcome measure, but we may not be so expert in the measurement properties of questionnaires. Of course, questionnaires vary in their measurement properties just as other outcome measures do. The conduct of high quality outcomes research depends on an understanding of these measurement issues.

What is Being Measured?

A PRO can be designed to measure innumerable traits (concept represented by the questions in the PRO), so it is critical to determine what needs to be measured. In Optometry and Ophthalmology, the most commonly measured trait is visual disability. Visual disability is a person's reduced ability to perform tasks due to impairment of visual performance. This is also known as visual functioning (VF) or vision-related activity limitation (VRAL), which is the appropriate nomenclature to be in line with the World Health Organization International Classification of Functioning, Disability, and Health.[9] The reason this trait is so commonly measured is because visual-related activity limitation is the traditional indication for cataract surgery and the majority of ophthalmic outcomes research has occurred for this procedure. Other traits that can be measured include satisfaction, ocular surface symptoms,[10] pain,[11] optical/visual symptoms,[12] well being,[13] depression[14] and perhaps most importantly, quality of life (QOL).[15] QOL is a holistic assessment of the effects of disease on the person and includes many dimensions such as a patient's physical, social and emotional wellbeing, spiritual, vocational, economic issues, convenience et cetera. Therefore, QOL instruments are essentially complex. Confusingly, investigators often mis-represent their instruments as a measure of QOL, when it only measures one or two dimensions; often VRAL only. Nevertheless, the trait under measurement will be evident from inspection of the questions asked. An activity limitation instrument will include questions like: how much difficulty do you have reading? A symptoms instrument may ask: how troubled are you by gritty eyes? When choosing a questionnaire, determining the trait that is sought to be measured and identifying an instrument that measures that trait are critical decisions. However, the quality of the questionnaire is also critical.

Questionnaire Quality

The development and validation of questionnaires, or instruments as they are called in the literature, is a long and complex process. The extent to which instruments have been developed and validated can be used to define their quality. We have previously published questionnaire quality assessment criteria in this journal.[16] These quality assessment criteria build on previous authors' work and incorporate the latest advances in questionnaire development methodology. The field of questionnaire science is a rapidly evolving field just as Optometry is. One way to make sense of the progression in questionnaire technology is to consider instruments as belonging to different generations.

First Generation—Conventional Questionnaires

First Generation instruments dominate the questionnaires available to Optometry and Ophthalmology. These are characterized by conventional instrument development and validation (true score theory). These instruments use Likert or summary scoring of ordinal values applied to response categories. For example, "How much trouble do you have…… not at all, (1) a little, (2) quite a bit, (3) a lot (4)." This scoring approach involves two assumptions regarding the response scale: it is assumed that the spacing between response categories is equidistant, and it is assumed that all questions have the same "value." For example, questions on difficulty driving during the day and driving at night are afforded the same value yet driving at night is a more difficult task so should be weighted accordingly.[17] Neither assumption is valid which makes the scoring non-linear. As a result, first generation instruments should not be considered as measures. This is important because it limits the ability of first generation instruments to detect differences or measure outcomes.[18, 19]

First generation instruments are common and still widely used. Table 1 lists a number of VRAL instruments that are best defined

TABLE 1.
Vision-related activity limitation instruments included in the analysis

Distance Vision Scale,[60] 1973	Cataract Symptom Scale, 1999
Visual Function Index,[61] 1981	TyPE, 1999
Activities of Daily Vision Scale,[62] 1992	Visual Functioning & Quality of Life (Fletcher),[63] 1999
Visual Activities Questionnaire,[64] 1992	Quality of Life & Visual Functioning (Carta),[65] 1999
VF-14,[21] 1994	Impact of Cataract Surgery (Monestam),[66] 1999
Cataract Symptom Score,[21] 1994	Houston Vision Assessment Test,[67] 2000
Catquest,[68] 1997	NEI-VFQ,[20] 2000
Visual Disability Assessment,[69] 1998	Impact of Vision Impairment,[31] 2000
Vision Core Measure,[70] 1998	Visual Symptoms and Quality of life,[71] 2003
Adaptation to Vision Loss,[72] 1998	

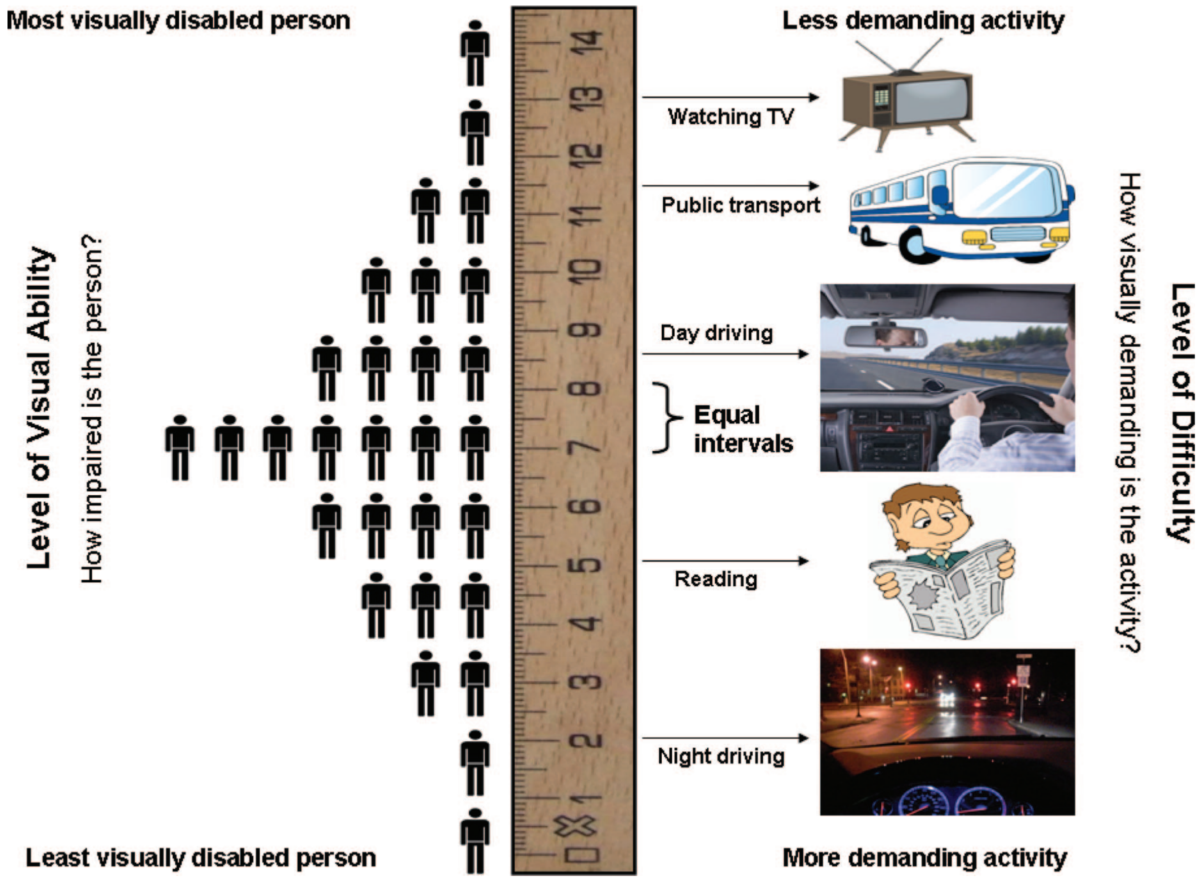

FIGURE 1.
Schematic person/item map showing persons and items calibrated onto the same linear scale. The scale is represented by a ruler and the relative difficulty of each item is apparent by its position along the ruler.

as first generation instruments. These include many of the most commonly used instruments like the National Eye Institute Visual Functioning Questionnaire (NEI-VFQ)[20] and the VF-14.[21] Other popular first generation instruments include the Ocular Surface Disease Index,[22] the NEI Refractive QOL Questionnaire among others.[23]

Second Generation—Questionnaires Using Rasch Analysis

Second generation instruments use Rasch analysis to deal with the assumptions that corrupt the scoring method used in first generation instruments. Rasch analysis provides interval scaling with both persons and questions (called items in the PRO field) calibrated to the same scale (Fig. 1). Rasch analysis also provides greater insight into the psychometric properties of the instrument. Specifically, how well items fit to the latent trait being measured; how well the items discriminate the people (separation); and how well item difficulty targets person ability. These two benefits of Rasch analysis (scoring and insight into psychometric properties) have popularized its application in ophthalmic questionnaires. Several articles in this journal have described in detail the science behind Rasch analysis and its advantages.[24, 25]

Second generation instruments fall into two categories: new instruments and legacy instruments. Rasch analysis can be applied in the development of new instruments. Examples of this include the QOL Impact of Refractive Correction,[26] the Contact Lens Impact on QOL,[27] the Ocular Comfort Index,[28] and the Veterans Affairs Low Vision VFQ-48.[29] Legacy instruments are first generation instruments, which have been re-engineered using Rasch analysis. However, the re-engineering process can be as simple as scoring using Rasch analysis[30] or may involve using Rasch analysis to optimize the psychometric properties of the instrument. Examples of second generation legacy instruments are the Impact of Visual Impairment,[31–33] the Catquest-9SF,[34] and the McMonnies Questionnaire.[35]

Why Rasch Analysis?

Empirical evidence for the benefits afforded by Rasch analysis can be drawn from a number of recent studies. Garamendi et al.[18] used the Refractive Status and Vision Profile instrument to try to discriminate between spectacle wearers seeking refractive surgery and those happily wearing spectacles on the basis of Refractive Status and Vision Profile QOL scores. They found that simply Rasch scaling all 42 items provided a 10% gain in precision, but also removing items that poorly fitted the underlying latent trait (QOL) resulted in an instrument with a 197% gain in precision for discriminating the groups. Similarly, Gothwal et al.[19] used the VF-14 to measure cataract surgery outcomes and found that a Rasch scaled and revised version of the VF-14 could measure surgical outcome with a gain in precision of 148% compared with the original version. The studies demonstrate that by optimizing the psychometric properties of an instrument using Rasch analysis large gains in signal (reduction in noise) are possible facilitating more efficient measurement.

Recently, our group has embarked on a project to create second generation legacy instruments out of all existing VRAL instruments. The strategy used was to rescale the instruments using Rasch analysis and to optimize the psychometric properties of the instruments. Second generation instruments were successfully developed for the Visual Disability Assessment,[36] the Cataract TyPE Spec,[37] the Cataract Symptom Scale,[38] Catquest,[39] the VF-14,[19] and the NEI-VFQ.[40] For a number of other instruments, the revision process was unsuccessful or the revised instrument was suboptimal. Examples included the Activities of Daily Vision Scale,[41] the Vision Core Measure-1,[42] Visual Function Index,[43] Visual Activities Questionnaire,[44] the Impact of Cataract Surgery,[44] Cataract Symptom Score,[45] Visual Function and QOL,[46] QOL and VF,[47] and the Distance Vision Scale.[48]

This line of research illustrated that second generation (legacy) instruments have some limitations. Almost all have poor targeting of item difficulty to person ability; items are designed to target people less able than the population. Therefore, the addition of items to suit the more able is required. Poor targeting may not be a problem with *de novo* second generation instruments, at least initially. Nevertheless, a fundamental problem of questionnaires is that they are not adaptable. The origin of the targeting problem is in the shifting indication for cataract surgery; surgery is now performed at lower levels of disability that a decade ago.[34] Therefore, these questionnaires have become outdated. Second, although all of these instruments are calibrated onto an interval scale using logits as the unit of measurement, the scale is not identical for all instruments. The range of the latent trait over which each measures varies also, some suit the third world, others a more Western population, but no instrument suits all populations. Finally, in refining these instruments, items which did not tap the latent trait under measurement were discarded to remove noise from the measure. However, these items may be able to be used in a different instrument to tap another latent trait; but this possibility was lost in a Second Generation instrument.

Third Generation—Item Banking

All these problems can be addressed in the Third Generation approach to PRO measurement: item banking. An item bank is simply a list of items, far more than would be included in a single questionnaire, which are all calibrated for difficulty using Rasch analysis. Item banks are generated by drawing items from multiple instruments and including them in one large Rasch analysis. Once all items are calibrated onto a single scale they can be drawn on to make measurements. Items can be selected, either manually or by a computer algorithm to target the ability of the patients under test. Because all items are calibrated to a single scale, whether "easy" items are used to measure in a third world setting, or "difficult" items are used to measure an urban Western population, the scores are meaningful in the same terms. Item banks are adaptable, in that new items can be added simply by pilot testing the new items with already calibrated items and performing a Rasch analysis to calibrate the new items.[49] As with second generation instruments, item banks benefit from optimized psychometric properties, such as all items tapping the same latent trait. Therefore, item banks need to be developed for each latent trait. In the context of pooling items from multiple instruments, this may have an advantage whereby items discarded from a VRAL item bank may be able to be pooled to form another item bank (e.g., QOL).

Item banking of ability or QOL is occurring for other medical outcomes, e.g., spinal cord injury,[50] chronic pain,[51] and depression[52] all of which are current National Institutes of Health funded projects. Similarly, Massof et al.[53–55] has pioneered item banking of ability items in low vision. However, it is time item banking occurred for all PROs in Optometry and Ophthalmology, which is the aim of our current line of research.

The Vision-Related Item Banking Project—Aim

The aim was to calibrate all items from all existing VRAL instruments onto a single measurement scale to create a VRAL item bank. Our secondary aim was to identify items belonging to other latent traits, and develop item banks for these traits also.

METHODS

Instruments

Nineteen instruments were identified as including VRAL content. They are listed in Table 1. These instruments contained a total of 353 items.

Study Population

Because the majority of these instruments were developed for cataract surgery outcomes assessment, we confined our initial data collection to a cataract population. Participants were invited from the waiting list for cataract surgery at the Flinders Medical Centre, Adelaide, South Australia (average waiting period 3 to 4 months). Six hundred twenty-four participants were enrolled. Their demographic details are displayed in Table 2.

TABLE 2.
Participant demographics

Characteristic	Value
Participants (N)	624
Gender (%)	
Female	56
Male	44
Age in years (mean ± SD)	74.1 ± 9.4
Cataract status (%)	
Bilateral	59
Awaiting second surgery	41
Ocular co-morbidity (%)	
Present	48
Absent	52
LogMAR visual acuity (mean ± SD)	
Better eye	0.22 ± 0.20
Worse eye	0.55 ± 0.36

Enrolled participants were mailed a pack containing a subset of the 19 questionnaires for self-administration. Questionnaires were presented in their native format using their native response scales. Completed questionnaires were returned via a self-addressed and prepaid envelope.

Rasch Analysis

The data were analyzed using the Andrich rating scale model[56] with Winsteps software (version 3.68).[57] The Rasch model is a probabilistic mathematical model, which provides estimates of person ability and item difficulty along a common measurement continuum, expressed in log-odd units (logits). For the VRAL, a positive item logit indicates that the item requires a lower level of ability than the average (i.e., the item is relatively easier). A positive logit for participants (person ability) suggests that the participant's visual disability is greater than the mean required level of ability for the items (i.e., the overall ability required for the task is greater than the ability that the participant possesses).

Rasch analysis was also used to optimize the psychometric properties of the item bank. First, the suitability of each item to be included in the item bank was assessed using fit statistics. A range for Infit and Outfit of 0.50 to 1.50 was considered acceptable. This is more lenient than has been proposed for questionnaires,[16] but this position was deliberately taken to make the item bank inclusive of as much content as possible. Fit of all items to the construct was confirmed using principal components analysis of the residuals (observed minus expected scores).[58] The performance of the item bank was assessed in terms of measurement precision using standard error on the calibrations and overall person separation (and reliability) statistics as well as targeting of item difficulty to person ability (difference between person and item means in logits).

RESULTS

Rasch estimates of item difficulty were generated with an average standard error on the measure of 0.11 logits indicating a high level of stability across items. However, 21.2% of items mis-fit the model (measured something other than VRAL). There was contamination with other latent traits including symptoms and QOL. As a remedy, the items were organized into 3 item banks representing three latent traits using principal components analysis of the residuals, item fit, and face validity: VRAL; Symptoms (The impact of disease or interventions on sensation e.g., visual; blur, glare . . . or comfort; gritty eyes, pain . . .); and QOL (Holistic assessment of the effects of disease on the person, this includes

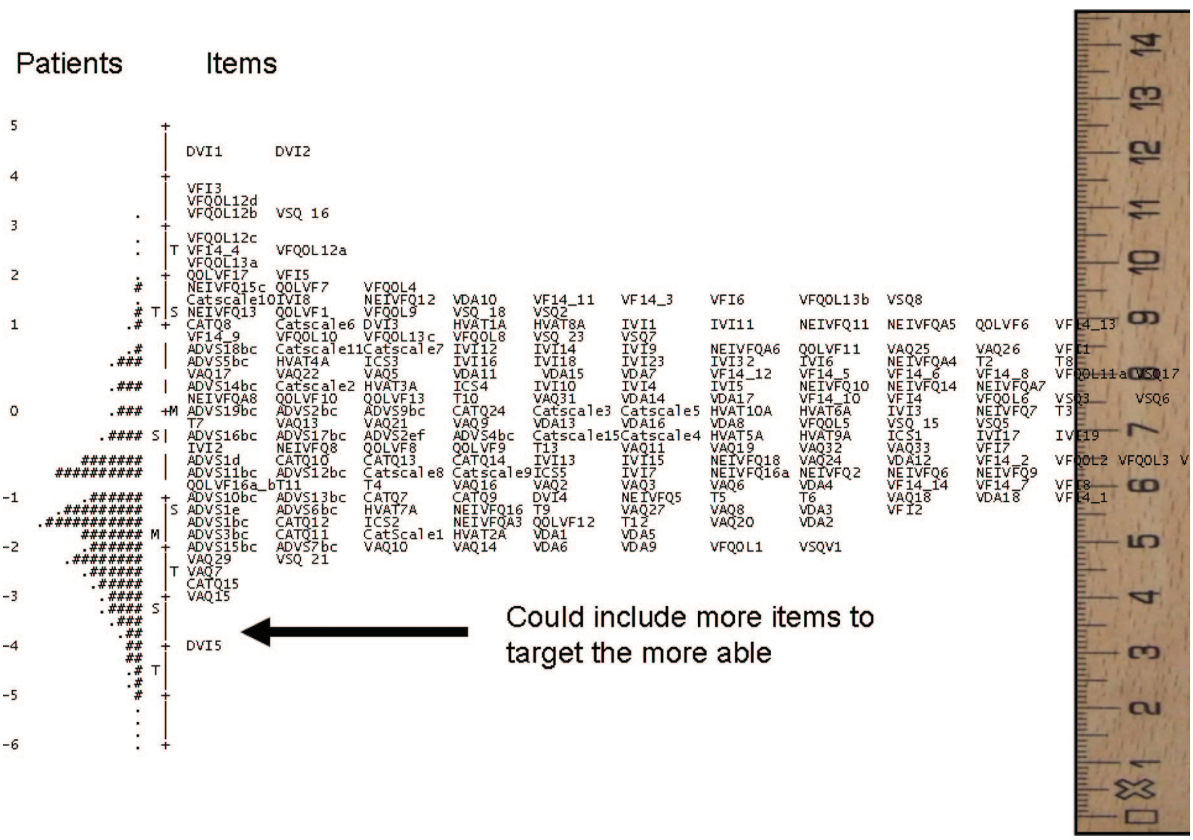

FIGURE 2.

Actual person/item map for all the VRAL items drawn from all 19 instruments. Persons and items are calibrated along the same linear scale with the ruler from Fig. 1 included to illustrate the parallel between the two figures. Each item is identified with the acronym of the instrument's title, and its item number. The distribution of persons and items mismatches slightly indicating that more "difficult" items need to be added.

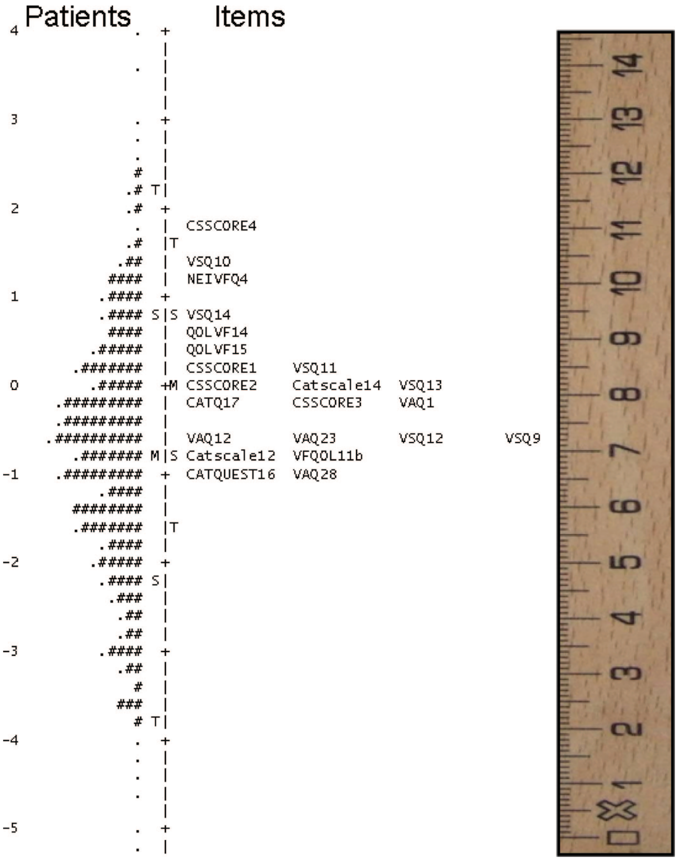

FIGURE 3.

Actual person/item map for all the symptoms items drawn from 8 of the instruments. Persons and items are calibrated along the same linear scale with the ruler from Fig. 1 included to illustrate the parallel between the two figures. Each item is identified with the acronym of the instrument's title, and its item number. The distribution of persons and items mismatches slightly indicating that more "difficult" items need to be added.

multiple dimensions; physical, social, emotional wellbeing, concerns, convenience, independence, economic). These 3 item banks were then re-analyzed separately.

The VRAL item bank contained 231 items from 16 questionnaires. This was trimmed to 226 items to meet the fit statistic criteria of 0.50 to 1.50. The real person separation was 8.11 (reliability 0.99) and the standard error on the measure was 0.11 logits. Therefore, this was a highly precise measurement of VRAL in cataract. However, there was poor targeting of item difficulty to person ability (person mean −1.64 logits). This is illustrated in Fig. 2. Targeting could be improved with the addition of items, which target the more able.

The symptoms item bank contained 24 items from 8 questionnaires. This was trimmed to 22 items to meet the fit statistic criteria. The real person separation was 2.33 (reliability, 0.84). Again this represented precise measurement of symptoms. There was excellent targeting of item difficulty to person ability (person mean −0.56). This is illustrated in Fig. 3.

The QOL item bank contained 63 items from 7 questionnaires. This was trimmed to 60 items to meet the fit statistic criteria. The real person separation was 3.20 (reliability, 0.91). Again this represented precise measurement of QOL. There was poor targeting of items to persons (person mean −2.11). This is illustrated in Fig. 4.

DISCUSSION

The research presented herein represents "proof of concept" of item banking for vision-related latent traits. Pooling the items enables effective measurement of all three traits. The vision-related item banking project has identified a family of three latent traits for which items can be banked: VRAL, symptoms, and QOL. Each of these item banks contained sufficient items to measure each latent trait. However, each could benefit from the addition of content. The VRAL item bank could benefit from additional items to target the more able cataract patients. Both the symptoms and QOL item banks effectively contain residual items, which did not fit the VRAL item bank. Therefore both require the addition of content to become comprehensive measures. This is especially true of the QOL item bank for which the items were chiefly drawn from the Impact of Vision Impairment (emotional well being subscale) and NEI-VFQ (social-emotional items) instruments. However, the expansion of item banks is simple, by adding uncalibrated items to the bank and determining their calibration with Rasch analysis against the calibrated items already in the bank.[49]

The calibration and banking of items enable more flexible implementation strategies like computer adaptive testing (Fig. 5). An "adaptive test" presents items that will most accurately target the individuals' visual ability. The test commences with the median item and every time the candidate answers a question, the computer re-estimates his or her ability. Based on the most recent, revised ability estimate, the computer selects the next item to be presented, such that the candidate will find the item challenging. This process provides highly accurate measurement of ability and the test can be continued until a desired level of precision is achieved. By tailoring the test to the individual, the problem of poor targeting of item difficulty to patient ability is eliminated.

Computer adaptive testing creates the possibility for electronic implementation for direct data capture by an internet portal with the potential for various platforms, e.g., iPhone app.[59] Our intention is to create online testing providing real-time scoring and recording of data. The platform could be designed to allow for more latent traits and other variables such as different eye diseases. Of course, item banking and computer adaptive testing need not be confined to cataract patients. Vision-related PRO measures should be developed to suit all eye diseases, all latent traits and patients from all countries of the world.

Our next step from the research presented herein was to reformat all the items on to a single response format for each latent trait; thus moving away from native questionnaire format. This removes redundancy involved in items with similar content and eliminates variance related to question format. We have done this for each of these three latent traits. We are now collecting data in different eye diseases across international settings. We are also expanding the content of the item banks by identifying new content at interviews and in focus groups. Over time, we hope this will lead us to the ultimate goal of comprehensive measurement of all vision-related PROs.

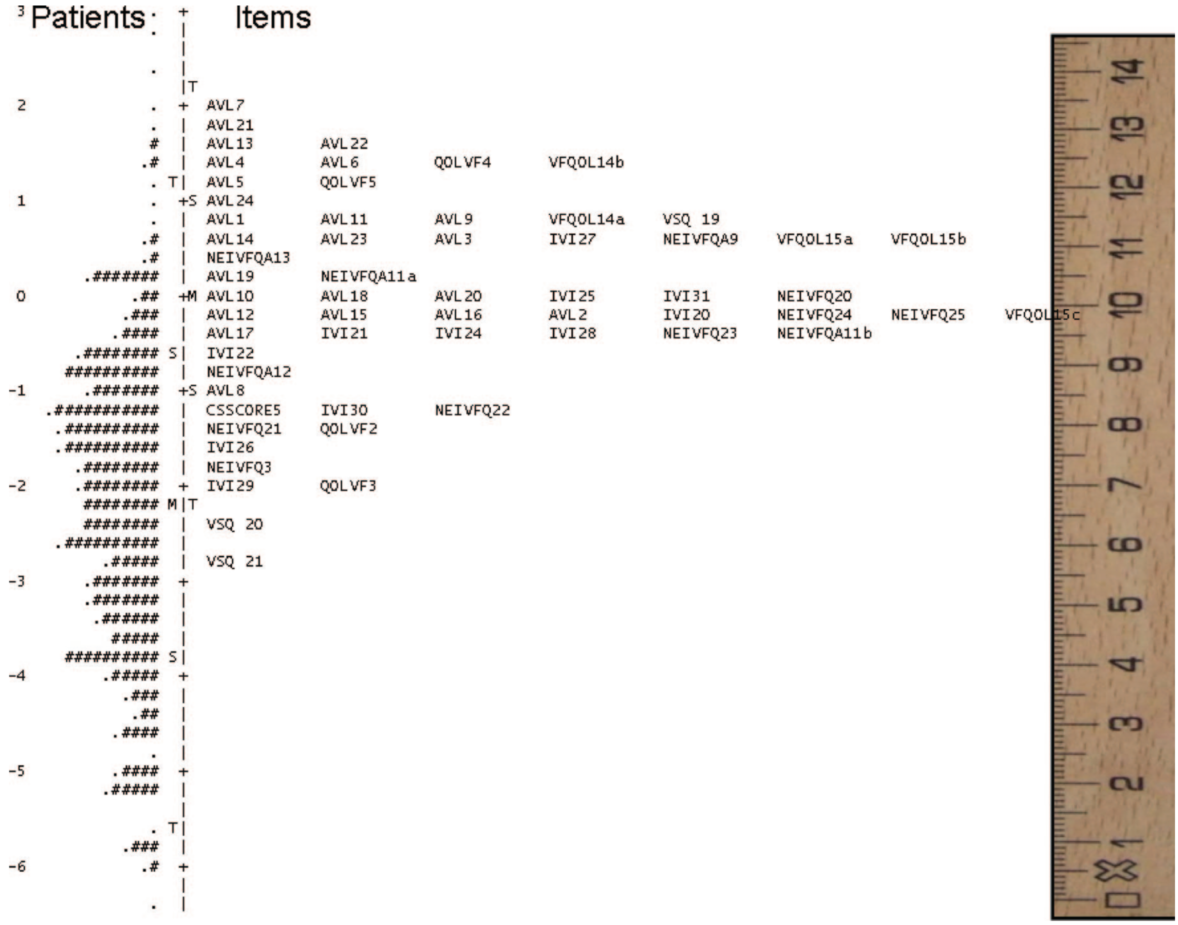

FIGURE 4.

Actual person/item map for all the QOL items drawn from 7 of the instruments. Persons and items are calibrated along the same linear scale with the ruler from Fig. 1 included to illustrate the parallel between the two figures. Each item is identified with the acronym of the instrument's title and its item number. The distribution of persons and items mismatches slightly indicating that more "difficult" items need to be added.

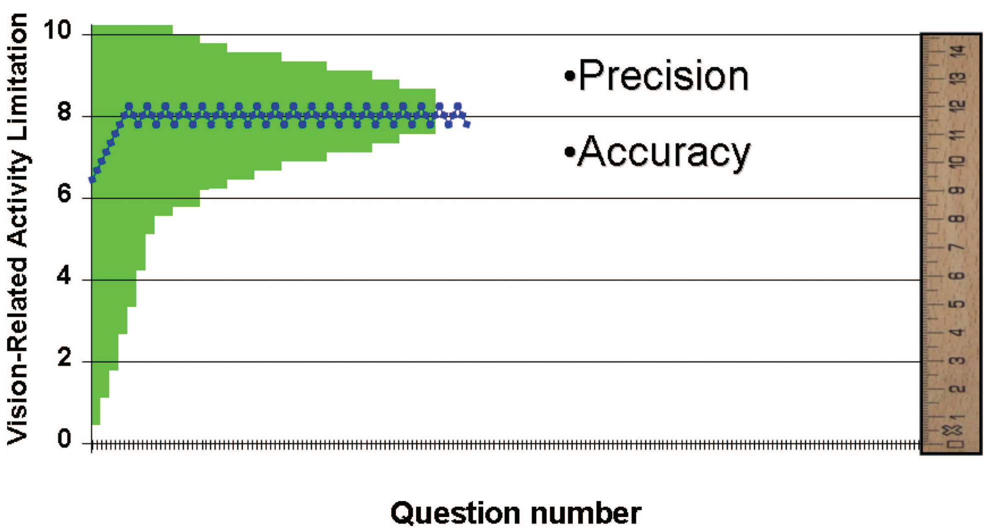

FIGURE 5.

Model of computer adaptive testing for VRAL. The difficulty of items presented is varied depending on the response to the previous item. In this example, items are selected starting with the median item with progressively easier items (increasing VRAL) offered after each task is reported as difficult. After the 9th item is reported as easy a more difficult item is offered. The algorithm then operates a staircase about the estimated threshold. The error on the measurement is denoted by the green background. The staircase is continued until a desired level of precision is achieved.

ACKNOWLEDGMENTS

My collaborators Vijaya Gothwal, Thomas Wright, and Ecosse Lamoureux contributed significantly to the work described herein. All the ophthalmology staff at Flinders Medical Centre consented to their patients being available to this project.

The author has no personal financial interest in the development, production, or sale of any device discussed herein.

Received December 12, 2009; accepted December 15, 2009.

REFERENCES

1. Chauhan BC. Truths and lesser truths: an interpretation of clinical trials in glaucoma. Optom Vis Sci 2008;85:376–9.
2. Pesudovs K, Coster DJ. Assessment of visual function in cataract patients with a mean visual acuity of 6/9. Aust N Z J Ophthalmol 1996;24:5–9.
3. Vingrys AJ, Pesudovs K. Localized scotomata detected with temporal modulation perimetry in central serous chorioretinopathy. Aust N Z J Ophthalmol 1999;27:109–16.
4. Lim L, Pesudovs K, Coster DJ. Penetrating keratoplasty for keratoconus: visual outcome and success. Ophthalmology 2000;107: 1125–31.
5. Patrick DL, Burke LB, Powers JH, Scott JA, Rock EP, Dawisha S, O'Neill R, Kennedy DL. Patient-reported outcomes to support medical product labeling claims: FDA perspective. Value Health 2007; 10(Suppl 2):S125–37.
6. Csaky KG, Richman EA, Ferris FL III. Report from the NEI/FDA Ophthalmic Clinical Trial Design and Endpoints Symposium. Invest Ophthalmol Vis Sci 2008;49:479–89.
7. Pham HH, Schrag D, O'Malley AS, Wu B, Bach PB. Care patterns in Medicare and their implications for pay for performance. N Engl J Med 2007;356:1130–9.
8. Sharma S, Brown GC, Brown MM, Hollands H, Shah GK. The cost-effectiveness of photodynamic therapy for fellow eyes with subfoveal choroidal neovascularization secondary to age-related macular degeneration. Ophthalmology 2001;108:2051–9.
9. World Health Organization. The International Classification of Functioning, Disability and Health (ICF). Geneva: World Health Organization; 2001.
10. Pult H, Murphy PJ, Purslow C. A novel method to predict the dry eye symptoms in new contact lens wearers. Optom Vis Sci 2009;86: 1042–50.
11. Yu JO, Gundel RE. Use of acular LS in the pain management of keratoconus: a pilot study. Optom Vis Sci 2009. (e-pub December 4, 2009). doi:10.1097/OPX. 0b013e3181c75170.
12. Chase C, Tosha C, Borsting E, Ridder WH III. Visual discomfort and objective measures of static accommodation. Optom Vis Sci 2009; 86:883–9.
13. Court H, Greenland K, Margrain TH. Predicting state anxiety in optometric practice. Optom Vis Sci 2009;86:1295–302.
14. Lamoureux EL, Tee HW, Pesudovs K, Pallant JF, Keeffe JE, Rees G. Can clinicians use the PHQ-9 to assess depression in people with vision loss? Optom Vis Sci 2009;86:139–45.
15. Cui Y, Stapleton F, Suttle C, Bundy A. Health- and vision-related quality of life in intellectually disabled children. Optom Vis Sci 2010; 87:37–44.
16. Pesudovs K, Burr JM, Harley C, Elliott DB. The development, assessment, and selection of questionnaires. Optom Vis Sci 2007;84: 663–74.
17. 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.
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; 46:1375–83.
19. Gothwal VK, Wright TA, Lamoureux EL, Pesudovs K. Measuring outcomes of cataract surgery using Visual Function Index-14 (VF-14). J Cataract Refract Surg 2010;36:in press.
20. Mangione CM, Lee PP, Gutierrez PR, Spritzer K, Berry S, Hays RD. Development of the 25-item National Eye Institute Visual Function Questionnaire. Arch Ophthalmol 2001;119:1050–8.
21. 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;112:630–8.
22. Schiffman RM, Christianson MD, Jacobsen G, Hirsch JD, Reis BL. Reliability and validity of the Ocular Surface Disease Index. Arch Ophthalmol 2000;118:615–21.
23. McDonnell PJ, Mangione C, Lee P, Lindblad AS, Spritzer KL, Berry S, Hays RD. Responsiveness of the National Eye Institute Refractive Error Quality of Life instrument to surgical correction of refractive error. Ophthalmology 2003;110:2302–9.
24. Mallinson T. Why measurement matters for measuring patient vision outcomes. Optom Vis Sci 2007;84:675–82.
25. Massof RW. The measurement of vision disability. Optom Vis Sci 2002;79:516–52.
26. 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.
27. Pesudovs K, Garamendi E, Elliott DB. The Contact Lens Impact on Quality of Life (CLIQ) Questionnaire: development and validation. Invest Ophthalmol Vis Sci 2006;47:2789–96.
28. Johnson ME, Murphy PJ. Measurement of ocular surface irritation on a linear interval scale with the ocular comfort index. Invest Ophthalmol Vis Sci 2007;48:4451–8.
29. Stelmack JA, Massof RW. Using the VA LV VFQ-48 and LV VFQ-20 in low vision rehabilitation. Optom Vis Sci 2007;84:705–9.
30. Massof RW. An interval-scaled scoring algorithm for visual function questionnaires. Optom Vis Sci 2007;84:689–704.
31. 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.
32. Lamoureux EL, Pallant JF, Pesudovs K, Rees G, Hassell JB, Keeffe JE. The effectiveness of low-vision rehabilitation on participation in daily living and quality of life. Invest Ophthalmol Vis Sci 2007;48: 1476–82.
33. Lamoureux EL, Pallant JF, Pesudovs K, Rees G, Hassell JB, Keeffe JE. The impact of vision impairment questionnaire: an assessment of its domain structure using confirmatory factor analysis and Rasch analysis. Invest Ophthalmol Vis Sci 2007;48:1001–6.
34. 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;35:504–13.
35. Gothwal VK, Pesudovs K, Wright T, McMonnies C. McMonnies Questionnaire: enhancing screening for dry eye syndromes using Rasch Analysis. Invest Ophthalmol Vis Sci 2009. (e-pub November 5, 2009).
36. Pesudovs K, Gothwal VK, Wright TA. Visual disability assessment—valid measurement of activity limitation and mobility in cataract patients. Br J Ophthalmol 2009. (e-pub December 3, 2009).
37. Gothwal VK, Wright TA, Lamoureux EL, Pesudovs K. Using Rasch analysis to revisit the validity of the Cataract TyPE Spec instrument for measuring cataract surgery outcomes. J Cataract Refract Surg 2009;35:1509–17.
38. Gothwal VK, Wright TA, Lamoureux EL, Pesudovs K. Cataract

Symptom Scale: clarifying measurement. Br J Ophthalmol 2009;93: 1652–6.
39. Gothwal VK, Wright TA, Lamoureux EL, Lundstrom M, Pesudovs K. Catquest questionnaire: re-validation in an Australian cataract population. Clin Experiment Ophthalmol 2009;37:785–94.
40. Pesudovs K, Gothwal VK, Wright TA, Lamoureux EL. National Eye Institute-Visual Function Questionnaire: What does it measure? J Cataract Refract Surg 2010;36:in press.
41. Gothwal VK, Wright T, Lamoureux EL, Pesudovs K. Activities of Daily Vision Scale: what do the subscales measure? Invest Ophthalmol Vis Sci 2010;51:694–700.
42. Lamoureux EL, Pesudovs K, Pallant JF, Rees G, Hassell JB, Caudle LE, Keeffe JE. An evaluation of the 10-item vision core measure 1 (VCM1) scale (the Core Module of the Vision-Related Quality of Life scale) using Rasch analysis. Ophthalmic Epidemiol 2008;15: 224–33.
43. Gothwal VK, Wright T, Lamoureux EL, Pesudovs K. Psychometric properties of visual functioning index using Rasch analysis. Acta Ophthalmol 2009. (e-pub June 26 2009).
44. Gothwal VK, Wright TA, Lamoureux EL, Pesudovs K. Visual Activities Questionnaire: assessment of subscale validity for cataract surgery outcomes. J Cataract Refract Surg 2009;35:1961–9.
45. Gothwal VK, Wright TA, Lamoureux EL, Pesudovs K. Cataract symptom score questionnaire: Rasch revalidation. Ophthalmic Epidemiol 2009;16:296–303.
46. Gothwal VK, Wright TA, Lamoureux EL, Pesudovs K. Rasch analysis of visual function and quality of life questionnaires. Optom Vis Sci 2009;86:1160–8.
47. Gothwal VK, Wright TA, Lamoureux EL, Pesudovs K. Rasch analysis of the quality of life and vision function questionnaire. Optom Vis Sci 2009;86:E836–44.
48. Gothwal VK, Wright TA, Lamoureux EL, Pesudovs K. Guttman scale analysis of the distance vision scale. Invest Ophthalmol Vis Sci 2009;50:4496–501.
49. Haley SM, Ni P, Jette AM, Tao W, Moed R, Meyers D, Ludlow LH. Replenishing a computerized adaptive test of patient-reported daily activity functioning. Qual Life Res 2009;18:461–71.
50. Slavin MD, Kisala PA, Jette AM, Tulsky DS. Developing a contemporary functional outcome measure for spinal cord injury research. Spinal Cord 2009. (e-pub October 20, 2009).
51. Anatchkova MD, Saris-Baglama RN, Kosinski M, Bjorner JB. Development and preliminary testing of a computerized adaptive assessment of chronic pain. J Pain 2009;10:932–43.
52. Choi SW, Reise SP, Pilkonis PA, Hays RD, Cella D. Efficiency of static and computer adaptive short forms compared to full-length measures of depressive symptoms. Qual Life Res 2009. (e-pub November 26. 2009).
53. Massof RW, Ahmadian L, Grover LL, Deremeik JT, Goldstein JE, Rainey C, Epstein C, Barnett GD. The Activity Inventory: an adaptive visual function questionnaire. Optom Vis Sci 2007;84:763–74.
54. Massof RW, Hsu CT, Baker FH, Barnett GD, Park WL, Deremeik JT, Rainey C, Epstein C. Visual disability variables. II: the difficulty of tasks for a sample of low-vision patients. Arch Phys Med Rehabil 2005;86:954–67.
55. Massof RW, Hsu CT, Baker FH, Barnett GD, Park WL, Deremeik JT, Rainey C, Epstein C. Visual disability variables. I: the importance and difficulty of activity goals for a sample of low-vision patients. Arch Phys Med Rehabil 2005;86:946–53.
56. Andrich D. A rating scale formulation for ordered response categories. Psychometrika 1978;43:561–73.
57. Linacre JM. A user's guide to WINSTEPS [computer program]. Chicago: Winsteps.com; 2009.
58. Wright BD. Local dependency, correlations and principal components. Rasch Meas Trans 1996;10:509–11. Available at: http://www.rasch.org/rmt/rmt103b.htm. Accessed December 30 2009.
59. Gothwal VK, Pesudovs K. Interactive computer-based self-reported, visual function questionnaire: the Palmpilot-VFQ. Eye (Lond) 2009. (e-pub November 13).
60. Haase KW, Byrant EE. Development of a scale designed to measure functional distance vision loss using an interview technique. Proc Am Stat Assoc 1973;SS:274–9.
61. Bernth-Petersen P. Visual functioning in cataract patients. Methods of measuring and results. Acta Ophthalmol (Copenh) 1981;59: 198–205.
62. Mangione CM, Phillips RS, Seddon JM, Lawrence MG, Cook EF, Dailey R, Goldman L. Development of the 'Activities of Daily Vision Scale.' A measure of visual functional status. Med Care 1992;30: 1111–26.
63. Fletcher A, Vijaykumar V, Selvaraj S, Thulasiraj RD, Ellwein LB. The Madurai Intraocular Lens Study. III: Visual functioning and quality of life outcomes. Am J Ophthalmol 1998;125:26–35.
64. Sloane ME, Ball K, Owsley C, Bruni JR, Roenker DL. The Visual Activities Questionnaire: developing an instrument for assessing problems in everyday visual tasks. In: Nonivasive Assessment of the Visual System, 1992 OSA Technical Digest Series. Washington, DC: Optical Society of America; 1992:26–9.
65. Carta A, Braccio L, Belpoliti M, Soliani L, Sartore F, Gandolfi SA, Maraini G. Self-assessment of the quality of vision: association of questionnaire score with objective clinical tests. Curr Eye Res 1998; 17:506–11.
66. Monestam E, Wachtmeister L. Impact of cataract surgery on visual acuity and subjective functional outcomes: a population-based study in Sweden. Eye (Lond) 1999;13(Pt 6):711–9.
67. Prager TC, Chuang AZ, Slater CH, Glasser JH, Ruiz RS. The Houston Vision Assessment Test (HVAT): an assessment of validity. The Cataract Outcome Study Group. Ophthalmic Epidemiol 2000;7: 87–102.
68. Lundström M, Roos P, Jensen S, Fregell G. Catquest questionnaire for use in cataract surgery care: description, validity, and reliability. J Cataract Refract Surg 1997;23:1226–36.
69. Pesudovs K, Coster DJ. An instrument for assessment of subjective visual disability in cataract patients. Br J Ophthalmol 1998;82: 617–24.
70. Frost NA, Sparrow JM, Hopper CD, Peters TJ. Reliability of the VCM1 Questionnaire when administered by post and by telephone. Ophthalmic Epidemiol 2001;8:1–11.
71. Donovan JL, Brookes ST, Laidlaw DA, Hopper CD, Sparrow JM, Peters TJ. The development and validation of a questionnaire to assess visual symptoms/dysfunction and impact on quality of life in cataract patients: the Visual Symptoms and Quality of life (VSQ) Questionnaire. Ophthalmic Epidemiol 2003;10:49–65.
72. Horowitz A, Reinhardt JP. Development of the adaptation to age-related vision loss scale. J Vis Impair Blind 1998;92:30–41.

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