Acta Ophthalmologica

The Impact of Cataract Surgery questionnaire: re-evaluation using Rasch analysis

Vijaya K. Gothwal,[1,2] Thomas A. Wright,[1]
Ecosse L. Lamoureux[3,4,5] and Konrad Pesudovs[1]

[1]National Health and Medical Research Council, Centre for Clinical Eye Research, Department of Ophthalmology, Flinders Medical Centre and Flinders University of South Australia, Bedford Park, South Australia, Australia
[2]Meera and LB Deshpande Centre for Sight Enhancement, Vision Rehabilitation Centres, LV Prasad Eye Institute, Hyderabad, India
[3]Centre for Eye Research Australia, The University of Melbourne, Melbourne, Victoria, Australia
[4]Vision Cooperative Research Centre, Sydney, New South Wales, Australia
[5]Singapore Eye Research Institute, Singapore National Eye Centre, Singapore

ABSTRACT.
Purpose: **The Impact of Cataract Surgery (ICS) questionnaire was designed to assess cataract surgery outcomes. The aim of this study was to describe the psychometric properties of the ICS questionnaire using the Rasch model in a cataract population.**
Methods: **Ninety-one patients waiting to undergo cataract surgery in the first or second eye at the Flinders Medical Centre, Adelaide, South Australia self-administered the four-item ICS questionnaire. Rasch analysis was performed to assess behaviour of response categories, ability to differentiate between participants' visual abilities (person separation; minimum acceptable 2.0), if items measure a single underlying construct [i.e. unidimensionality assessed by fit statistics and further by principal components analysis (PCA)] and matching of item difficulty to participant ability (targeting; ideal < 0.5 logits). Adequate person separation was defined as basic requirement for a measure, failing which further assessment such as PCA was not performed.**
Results: **The four-item ICS questionnaire did not meet the required measurement properties (person separation zero). Response categories did not behave as intended, requiring the collapsing of categories for one item (read ordinary newspaper-size print). One item misfit (estimating distance) indicating that it was not measuring the same construct as other items. However, person separation failed to improve following the deletion of this item. Targeting was –0.46 logits, indicating that the item difficulty was well suited to the visual abilities of the participants.**
Conclusion: **In its present form, the ICS is unsuitable for visual disability assessment in patients awaiting cataract surgery. Other, better visual function questionnaires are available and preferred.**

Key words: cataract – impact of cataract surgery – questionnaire – Rasch analysis – surgery

Acta Ophthalmol.

doi: 10.1111/j.1755-3768.2009.01733.x

Introduction

Cataract surgery is the most performed ophthalmic surgical procedure worldwide and it is increasingly being offered at lower disease thresholds in Australia, Scandinavia and other developed nations (Keeffe & Taylor 1996; Taylor 2000; Leinonen & Laatikainen 2002). Patients' reported outcomes, including measurement of visual functioning and quality of life, can play an important role in justifying cataract surgery and measuring its outcome (Lundström et al. 2001; Mozaffarieh et al. 2005; Lundqvist & Monestam 2008). In recent decades several generic and disease-specific (including cataract) visual function questionnaires (VFQs) have been developed. The majority of these VFQs were developed using classical test theory (CTT), the limitations of which are now acknowledged by most investigators (Massof 2002). The major limitations of CTT relate to the assumptions associated with instrument scoring and the limited ability to assess item characteristics, including whether the items form a unidimensional contruct (Hambleton 2000).

The availability of item response theory-based (IRT) models, specifically the Rasch model, has encouraged psychometric research to address fundamental measurement issues associated with CTT (McHorney et al. 1997; Hambleton 2000; Tennant et al. 2004). Rasch analysis is the most commonly used IRT model to create new VFQs (Gothwal et al. 2003; Pesudovs et al. 2004; Pesudovs et al. 2006) or test existing VFQs (Velozo et al. 2000; Massof & Fletcher 2001; Pesudovs et al. 2003; Lamoureux et al. 2008). The Rasch measurement models checks two important assumptions: (i) the probability of endorsing one question does not increase the probability of endorsing another one identically (local independence); and (ii) all questions in the questionnaire measure a single underlying construct (unidimensionality). A number of cataract-specific questionnaires [Visual Function-14 (VF-14), Activities of Daily Vision Scale (ADVS), Visual Disability Assessment (VDA)] have been re-examined using Rasch analysis (Velozo et al. 2000; Pesudovs et al. 2003; Pesudovs et al. 2005); the Impact of Cataract Surgery questionnaire (ICS) (Monestam & Wachtmeister 1999) has not.

Therefore, the overall goal of the present study was to use the Rasch model to examine the responses of cataract patients in Australia to the ICS questionnaire.

Materials and Methods

Patients of the Flinders Medical Centre (Adelaide, South Australia) currently on the public waiting list to have cataract surgery participated in this study. Patients were mailed the ICS questionnaire for self-completion, which they returned via a self-addressed envelope. Included patients were aged 18 years or older, English-speaking and had no severe cognitive impairment. Co-existing ocular and systemic comorbidities representative of a typical cataract population in Australia were apparent in the current patient group (Kirkwood et al. 2006). Ethical approval was obtained from the Flinders Clinical Research Ethics Committee and all participants signed a consent form. The study was conducted in accordance with the tenets of the Declaration of Helsinki.

Clinical assessment

Routine clinical assessments occurred prior to listing for cataract extraction. Visual acuity assessments were performed using computerized testing based on LogMAR principles with a screen illumination of 150 cd/m^2.

The ICS questionnaire

The ICS questionnaire consists of four items (Table 1). The items relate, respectively, to ability to read, watch television, orientate in unfamiliar surroundings and estimate distance (near and far). As can be seen from the table, each item uses a different rating scale.

Rasch analysis

The data were analysed with Winsteps software (Linacre 2008) (Winsteps 3.66; Chicago, Illinois, USA) using the Andrich rating scale model for polytomous data (Andrich 1978). The nature of Rasch analysis has been detailed elsewhere (Massof 2002; Mallinson 2007; Pesudovs et al. 2007). A brief description follows.

The first step was an assessment of the behaviour of the response categories. Because each item had varying number and labelling of the response categories (Table 1), an individual Andrich rating scale was applied for each item format (four) to examine the performance of the response categories. Once the response categories were found to show the intended hierarchy, other characteristics were investigated as follows.

The Winsteps software was used to obtain the estimates of person ability and item difficulties, together with an assessment of measurement precision (using person separation statistics, an indicator of the number of statistically different levels (or strata) of participant ability distinguished by the items; minimum acceptable value of 2.0), unidimensionality [i.e. the extent to which all items in the ICS measure a single underlying construct, reflected in the information-weighted or infit mean square statistics; acceptable fit criterion of 0.7–1.3 and further confirmation of unidimensionality by principal components analysis (PCA) of residuals], differential item functioning (DIF) (i.e. if items perform equally between subgroups; for example, men and women), targeting (the extent to which the difficulty of the items match the abilities of the participants represented by a difference between person and item mean values in the person–item map; a perfectly targeted instrument would have targeting of 0; a difference between means of more than one logit indicates notable mistargetting) and item hierarchy (i.e. items should form a hierarchy of difficulty, ranging from least to most difficult to perform, also visualized in the person–item map).

The minimum acceptable measurement property for the ICS to be termed as a measure was a person separation > 2.0. In case the questionnaire failed this fundamental requirement, further assessments such

Table 1. Contents of the Impact of Cataract Surgery Questionnaire.

Item number	Item description	Number of categories
1	Can you read ordinary newspaper-size print? If YES, what visual aids do you need to be able to read?	4*
2	Do you experience any problems while watching TV caused by your cataractous eye?	2†
3	Do you experience difficulties when orientating in unfamiliar surroundings?	3‡
4	Do you experience difficulties in estimating distance (nearby/far away)?	3§

* Response options: none or spectacles (1), hand-held or stand magnifiers (2), others (please specify) (3) and no (4, i.e. cannot read).
† Response options: yes, difficulties (2) and no (1).
‡ Response options: no problems (1), some problems (2) and severe problems (3).
§ Response options: No (1), Yes, difficulties to nearby/No to far away (2) or Yes, difficulties to far away/No to near by (2), and Yes, difficulties to nearby and far away (3).

as PCA and DIF were not carried out.

Results

Participant characteristics

Ninety-one participants completed the ICS questionnaire. The mean age of the patients was 74.9 years (range 50–91 years) and 50 (54.9%) were female. Table 2 summarizes the participant characteristics.

Rasch analysis of the ICS data

Assessment of response categories
Figure 1 illustrates the category probability curves (CPCs) of item 1 in its four-point original form showing disordered thresholds necessitating collapsing of categories. The CPC plots visual disability on the x-axis against the probability of endorsing each response category on the y-axis. Threshold refers to the point between two adjacent response categories, for example 1 and 2, where either response (1 or 2) has equal probability of being selected. For a given item, the number of thresholds is always one less than the number of categories. Consequently, one would expect to see three thresholds, but category 3 (others, please specify) was never used and so was not observed. Furthermore, one can see disordered thresholds (i.e. category 2 does not have a range along the scale where it is the most likely category to be selected). Threshold disordering suggests that the response scale is not working adequately to order participants with distinct levels of ability. Therefore, categories were reorganized by collapsing the first two categories to generate dichotomous response categories: 1 represented 'none/some visual aids (spectacles or hand-held or stand magnifiers)' and the remaining response option ('no') was recoded as 2. The remainder of the item groups demonstrated ordered thresholds.

Person separation and item fit
The person separation for the four-item ICS was unacceptably low (Table 3). Misfit was observed for only for one item, 'Do you experience difficulties in estimating distance near-by/far away?' (Table 4). However, the person separation did not improve following the deletion of this misfitting item (Table 3).

Item hierarchy and targeting
The person–item map (Fig. 2) illustrates that 'orientating in unfamiliar surroundings' was the least difficult activity to perform whereas 'estimating distance' was the most difficult. The mean person measure was −0.46 logits (ranging between +3.0 and −3.0 logits), indicating that the item difficulty mostly matched the visual abilities of the participants. In comparison, the distribution of items covered a very narrow range (−1.50 to 1.17 logits). Furthermore, the items were sparsely spread: two items (orientation, reading) were located above the mean item difficulty and the other two (watching TV, estimating distance) were located below the mean item difficulty.

Table 2. Characteristics of respondents to the Impact of Cataract Surgery Questionnaire (n = 91).

Characteristic		n (%) or mean ± standard deviation
Age (year)		74.9 ± 9.0
Gender		
Male		41 (45.1)
Female		50 (54.9)
Visual acuity		
Eye to be operated on (habitual)	LogMAR (Snellen)	0.53 ± 0.35 (6/19^{-1}*)
	Range	0.04–1.60 (6/6^{-2} to 6/240)
Fellow eye (habitual)	LogMAR (Snellen)	0.30 ± 0.31 (6/12)
	Range	−0.26 to 1.30 (6/3^{-2} to 6/120)
Binocular (habitual)	LogMAR	0.24 ± 0.21 (6/9.5^{-2})
	Range	−0.26 to 0.92 (6/3^{-2} to 6/48^{-1})
Awaiting second-eye surgery		40 (44.0)
Ocular comorbidity†		
Present		49 (53.8)
Absent		42 (46.2)
Systemic comorbidity‡		
Present		10 (11.0)
Absent		81 (89.0)

* Minus symbol in superscript indicates that a participant missed some letters from that particular line. For example, 6/19^{-1} indicates one missed letter from this line.
† Includes glaucoma, diabetic retinopathy, age-related macular degeneration etc.
‡ Includes diabetes, hypertension, angina etc.

Discussion

Rasch analysis revealed the overall performance of the ICS questionnaire to be poor. The fundamental limitation was inadequate person separation, indicating that the questionnaire was unable to differentiate between the visual abilities of participants awaiting cataract surgery; therefore, it does not supplement the results of a clinical evaluation for cataract surgery. It is ineffective to retain items of no discriminatory value in a questionnaire. However the items of the ICS do constitute important day-to-day activities of elderly patients. Given this, the simplest way to increase the level of discrimination would be to add more items – specifically, those items that cover a wider range of activities such as self-care, face recognition etc., as seen in other VFQs that have good person separation (Velozo et al. 2000; Pesudovs et al. 2003; Pesudovs et al. 2005). Poor person separation can occur for several reasons, but is a common problem with questionnaires containing a smaller number of items. Monestam & Wachtmeister (1999) reported that the ICS was intentionally kept short to ensure simplicity of completion for the elderly population. Although respondent burden is reduced with fewer items (Mallinson et al. 2004), the undue shortening or inclusion of a smaller number of items can disrupt the psychometric properties of the questionnaire and thereby limit its use.

Misfitting item was the other limitation of the ICS. One item (estimating distance) misfit, indicating that this item is not in tandem with the other three items in the measurement of the underlying construct. In other words participants responded to this item erratically, perhaps because the item is

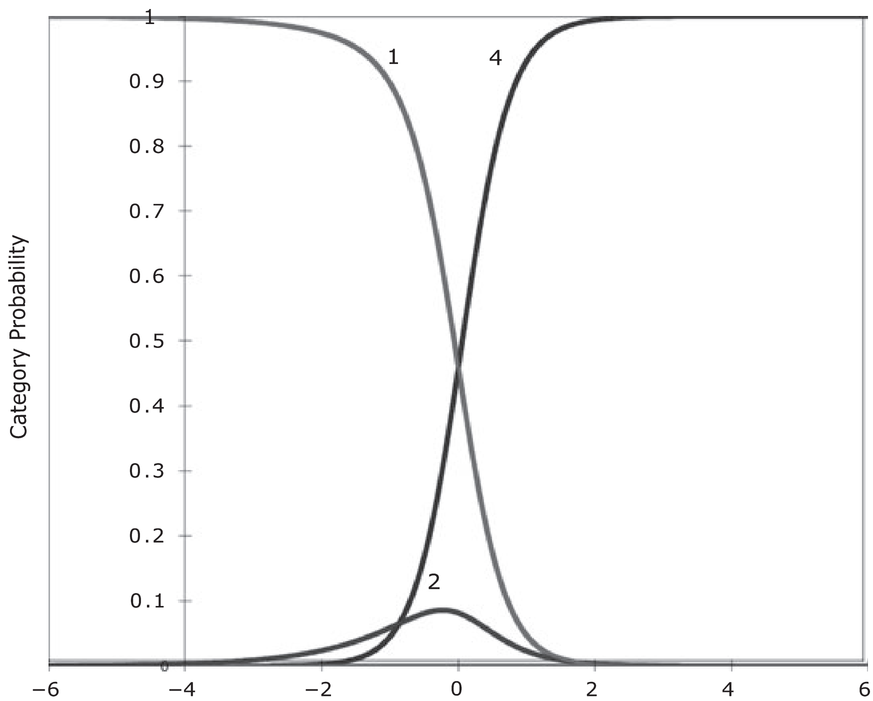

Fig. 1. Category probability curves for item 1 of the Impact of Cataract Surgery questionnaire showing disordered thresholds for category 2. This item has four response categories but category 3 was never used and so was not observed. Response category 2 ('hand-held or stand magnifiers') does not have a range along the visual disability scale, where it is the most likely category to be selected. Therefore, it is less likely to be endorsed by the cataract participants.

Table 3. Overall performance of all versions of the Impact of Cataract Surgery (ICS) questionnaire.

	Version of the ICS questionnaire tested	
Parameter	Four-item ICS	Three-item ICS
Number of misfitting items	1	0
Person separation	0	0
Mean item location	0	0
Mean person location	−0.46	−0.47

Table 4. Item fit statistics (ordered from least to most difficult) for the Impact of Cataract Surgery Questionnaire.

Item no.	Item description	Item calibration (logits)*	Standard error	Infit mean square
3	Orientating in unfamiliar surroundings	1.17	0.21	0.92
1	Read ordinary newspaper-size print	0.90	0.27	0.95
2	Watching TV	−0.57	0.23	0.76
4	Estimating distance (nearby/far away)	−1.50	0.17	1.34

* Logit is the natural logarithm of the odds of a participant being successful at a specific task or an item being carried out successfully; positive item logit indicates that the item requires a lower visual ability than the mean of the items and is an easier item, while a negative item logit indicates that the item requires a higher visual ability than the mean of the items and is a harder item.

not understood well, is ambiguous or measures a second dimension (Pesudovs et al. 2007). Of these, ambiguity appears the most likely in the present case because the item pertains to estimating distances either nearby or far away, which is rather a vague description. Removal of misfitting items usually improves the fit of the model (Pallant et al. 2006). However, this was not the case with the ICS questionnaire and the discriminatory ability continued to remain suboptimal.

The limitations of the performance of VFQs such as the ICS questionnaire could be avoided if Rasch analysis was used in their development, wherein the focus on the 'item' gives a direct connection between the item and the location of the item on the latent variable (Pesudovs et al. 2007). The person–item map of the ICS questionnaire highlights the inadequacies that can be associated with CCT-developed questionnaires, specifically the presence of significant gaps in the locations of the items. This limitation could explain the lack of subjective visual improvement in a proportion of patients (aged 90 years or older) in a cataract-outcomes study that used the modified version of the ICS questionnaire (Monestam & Wachmeister 2004).

Despite these shortcomings, the ICS questionnaire demonstrated one adequate property: targeting. With the exception of the Catquest-9SF (Lundström & Pesudovs 2009), most VFQs that have recently been examined using Rasch analysis have demonstrated poor targeting (Velozo et al. 2000; Pesudovs et al. 2003; Lamoureux et al. 2008; Pesudovs et al. 2008). A visual acuity measurement of 6/15 or 6/12 is considered sufficient for watching television and performing most everyday activities (Bergman & Sjostrand 1992). The mean binocular visual acuity of the participants in the present study was $6/9.5^{-2}$, and the good targeting perhaps suggests that this vision was just sufficient for the participants to perform the activities included in the ICS questionnaire.

Thus, the fundamental limitation of the ICS questionnaire remains its poor discriminatory ability, which could be improved with the addition of appropriately targeted items. Items can be added to static questionnaires; however, an even better strategy would be the creation of item banks for computer adaptive testing (Cook et al. 2005; Fayers 2007; Hays & Lipscomb 2007). Item banks, where items from different questionnaires are pooled, have been created and used in other areas of health assessment (Haley et al. 2004, 2006). It is now time for ophthalmic research to develop such an item bank.

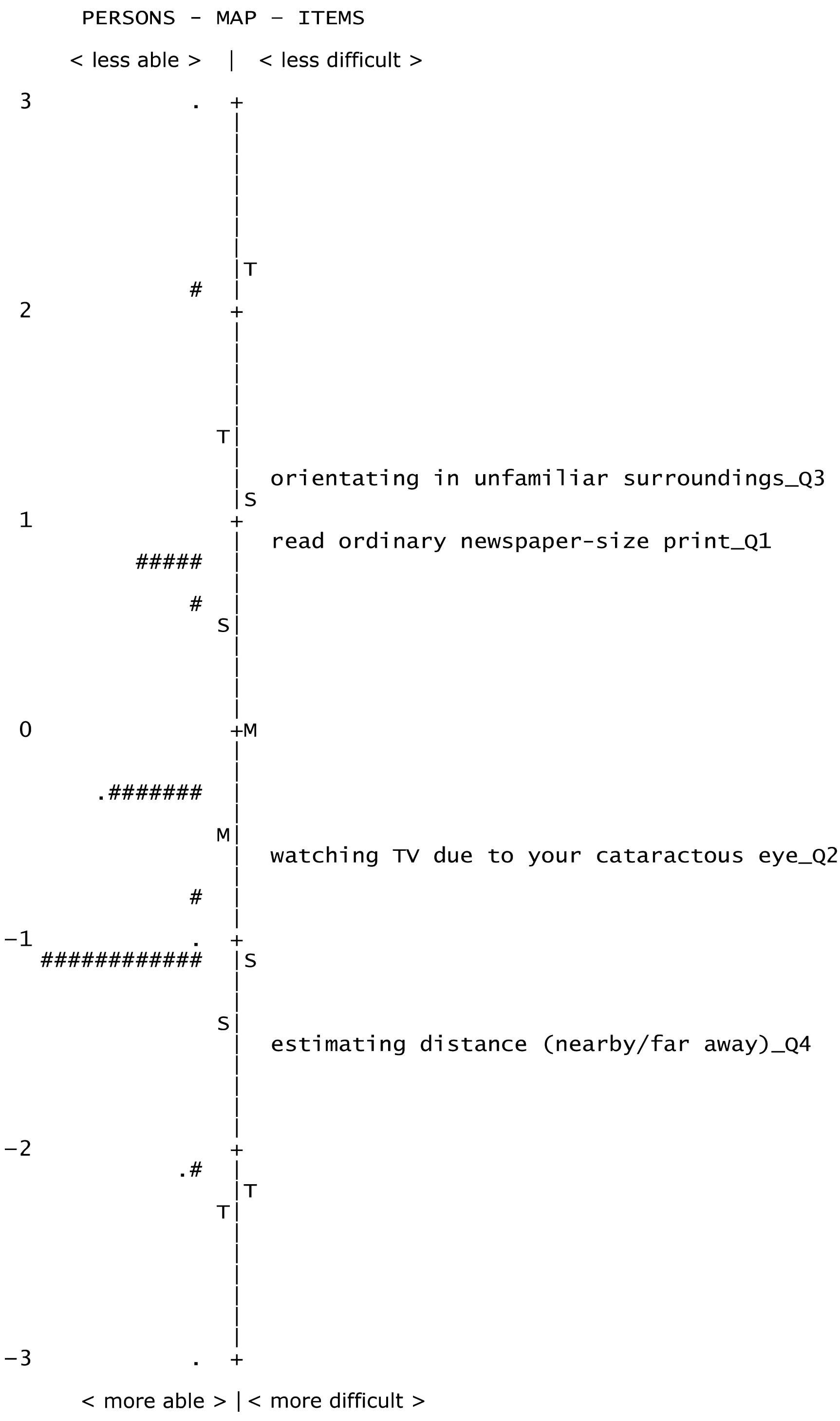

Fig. 2. Person–item map for the four-item Impact of Cataract Surgery questionnaire. The participants are located on the left of the dashed line, with more able participants located at the bottom of the map. Items are located on the right of the dashed line and more difficult items are located at the bottom of the map. Each # represents three participants and each . represents one to three participants. M, mean; S, 1 standard deviation (SD) from the mean; T, 2 SD from the mean.

In the meantime, clinicians and researchers can use other VFQs such as the impact of vision impairment (Pesudovs et al. 2008), vision core measure 1 (Lamoureux et al. 2008) or the Catquest-9SF (Lundström & Pesudovs 2009), which have a relatively larger number of items covering a wider range of the construct and have been demonstrated to fulfil the stringent requirements of the Rasch model for the assessment of cataract outcomes.

In conclusion, the ICS questionnaire does not meet the requirements of the Rasch model and thus, in its present form, it appears unsuitable for measuring visual disability in patients awaiting cataract surgery. The fundamental limitation of the ICS questionnaire, its inability to distinguish between the visual abilities of patients with cataract, makes it limited as an outcome measure. Other, better VFQs can be used for the assessment of cataract outcomes.

Acknowledgements

The authors thank all the participants who volunteered to participate in this study. They acknowledge the support of the National Health and Medical Research Council (Canberra, Australia) (Centre of Clinical Research Excellence Grant 264620). Konrad Pesudovs is supported by National Health and Medical Research Council (Canberra, Australia) Career Development Award 426765.

References

Andrich DA (1978): A rating scale formulation for ordered response categories. Psychometrika **43**: 561–573.

Bergman B & Sjostrand J (1992): Vision and visual disability in the daily life of a representative population sample aged 82 years. Acta Ophthalmol (Copenh) **70**: 33–43.

Cook KF, O'Malley KJ & Roddey TS (2005): Dynamic assessment of health outcomes: time to let the CAT out of the bag? Health Serv Res **40**: 1694–1711.

Fayers PM (2007): Applying item response theory and computer adaptive testing: the challenges for health outcomes assessment. Qual Life Res **16**(Suppl. 1): 187–194.

Gothwal VK, Lovie-Kitchin JE & Nutheti R (2003): The development of the LV Prasad-Functional Vision Questionnaire: a measure of functional vision performance of visually impaired children. Invest Ophthalmol Vis Sci **44**: 4131–4139.

Haley SM, Coster WJ, Andres PL, Kosinski M & Ni P (2004): Score comparability of short forms and computerized adaptive testing: simulation study with the activity measure for post-acute care. Arch Phys Med Rehabil **85**: 661–666.

Haley SM, Ni P, Hambleton RK, Slavin MD & Jette AM (2006): Computer adaptive testing improved accuracy and precision of scores over random item selection in a physical functioning item bank. J Clin Epidemiol **59**: 1174–1182.

Hambleton RK (2000): Emergence of item response modeling in instrument development and data analysis. Med Care **38**: 60–65.

Hays RD & Lipscomb J (2007): Next steps for use of item response theory in the assessment of health outcomes. Qual Life Res **16**(Suppl. 1): 195–199.

Keeffe JE & Taylor HR (1996): Cataract surgery in Australia 1985–94. Aust NZ J Ophthalmol **24**: 313–317.

Kirkwood BJ, Pesudovs K, Latimer P & Coster DJ (2006): The efficacy of a nurse-led preoperative cataract assessment and postoperative care clinic. Med J Aust **184**: 278–281.

Lamoureux EL, Pesudovs K, Pallant JF, Rees G, Hassell JB, Caudle LE & Keeffe JE (2008): 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 **15**: 224–233.

Leinonen J & Laatikainen L (2002): Changes in visual acuity of patients undergoing cataract surgery during the last two decades. Acta Ophthalmol Scand **80**: 506–511.

Linacre JM (2008): WINSTEPS Rasch measurement computer program. Chicago: Winsteps.com.

Lundqvist B & Monestam E (2008): Gender-related differences in cataract surgery outcome: a 5-year follow-up. Acta Ophthalmol **86**: 543–548.

Lundström M & Pesudovs K (2009): Catquest-9SF patient outcomes questionnaire nine-item short-form Rasch-scaled revision of the Catquest questionnaire. J Cataract Refract Surg **35**: 504–513.

Lundström M, Roos P, Brege KG, Floren I, Stenevi U & Thorburn W (2001): Cataract surgery and effectiveness. 2. An index approach for the measurement of output and efficiency of cataract surgery at different surgery departments. Acta Ophthalmol Scand **79**: 147–153.

Mallinson T (2007): Why measurement matters for measuring patient vision outcomes. Optom Vis Sci **84**: 675–682.

Mallinson T, Stelmack J & Velozo C (2004): A comparison of the separation ratio and coefficient alpha in the creation of minimum item sets. Med Care **42**: 17–24.

Massof RW (2002): The measurement of vision disability. Optom Vis Sci **79**: 516–552.

Massof RW & Fletcher DC (2001): Evaluation of the NEI visual functioning questionnaire as an interval measure of visual ability in low vision. Vision Res **41**: 397–413.

McHorney CA, Haley SM & Ware JE Jr (1997): Evaluation of the MOS SF-36 Physical Functioning Scale (PF-10): II. Comparison of relative precision using Likert and Rasch scoring methods. J Clin Epidemiol **50**: 451–461.

Monestam E & Wachmeister L (2004): Impact of cataract surgery on the visual ability of the very old. Am J Ophthalmol **137**: 145–155.

Monestam E & Wachtmeister L (1999): Impact of cataract surgery on visual acuity and subjective functional outcomes: a population-based study in Sweden. Eye **13**: 711–719.

Mozaffarieh M, Heinzl H, Sacu S & Wedrich A (2005): Clinical outcomes of phacoemulsification cataract surgery in diabetes patients: visual function (VF-14), visual acuity and patient satisfaction. Acta Ophthalmol Scand **83**: 176–183.

Pallant JF, Miller RL & Tennant A (2006): Evaluation of the Edinburgh Post Natal Depression Scale using Rasch analysis. BMC Psychiatry **6**: 28.

Pesudovs K, Burr JM, Harley C & Elliott DB (2007): The development, assessment, and selection of questionnaires. Optom Vis Sci **84**: 663–674.

Pesudovs K, Caudle LE, Rees G & Lamoureux EL (2008): Validity of a visual impairment questionnaire in measuring cataract surgery outcomes. J Cataract Refract Surg **34**: 925–933.

Pesudovs K, Elliott DB & Coster DJ (2005): The cataract outcomes questionnaire – a Rasch scaled measure of visual disability. Invest Ophthalmol Vis Sci **45**: ARVO e-abstract 3844.

Pesudovs K, Garamendi E & Elliott DB (2004): The quality of life impact of refractive correction (QIRC) questionnaire: development and validation. Optom Vis Sci **81**: 769–777.

Pesudovs K, Garamendi E & Elliott DB (2006): The contact lens impact on quality of life (CLIQ) questionnaire: development and validation. Invest Ophthalmol Vis Sci **47**: 2789–2796.

Pesudovs K, Garamendi E, Keeves JP & Elliott DB (2003): The Activities of Daily Vision Scale for cataract surgery outcomes: re-evaluating validity with Rasch analysis. Invest Ophthalmol Vis Sci **44**: 2892–2899.

Taylor HR (2000): Cataract: how much surgery do we have to do? Br J Ophthalmol **84**: 1–2.

Tennant A, McKenna SP & Hagell P (2004): Application of Rasch analysis in the development and application of quality of life instruments. Value Health **7**(Suppl. 1): 22–26.

Velozo CA, Lai JS, Mallinson T & Hauselman E (2000): Maintaining instrument quality while reducing items: application of Rasch analysis to a self-report of visual function. J Outcome Meas **4**: 667–680.

Received on March 9th, 2009.
Accepted on July 27th, 2009.

Correspondence:
Konrad Pesudovs
National Health and Medical Research Council Centre for Clinical Eye Research
Department of Ophthalmology
Flinders Medical Centre
Bedford Park
5042 South Australia
Australia
Tel: + 618 8204 4899
Fax: + 618 8277 0899
Email: konrad.pesudovs@flinders.edu.au