

An instrument for assessment of subjective visual disability in cataract patients

Konrad Pesudovs, Douglas J Coster

Abstract

Aims/background—The construction and validation of an instrument for the assessment of subjective visual disability in the cataract patient is described. This instrument is specifically designed for measuring the outcome of cataract surgery with respect to visual disability.

Methods—Visually related activities thought to be affected by cataract were considered for the questionnaire. These were reduced by pilot study and principal components analysis to 18 items. A patient's assessment of his/her ability to perform each task was scored on a four point scale. Scores were averaged to create an overall index of visual disability, as well as subscale indices for mobility related disability, distance/lighting/reading related disability, and near and related tasks visual disability. The questionnaire, administered verbally is entitled "The Visual Disability Assessment (VDA)". Reliability testing included test-retest reliability, interobserver reliability (ρ , the intraclass correlation coefficient), and internal consistency reliability (Cronbach's α). Construct validation, the process for proving that a test measures what it is supposed to measure, included consideration of content validity, comparison with the established Activities of Daily Vision Scale (ADVS) and empirical support with factor analysis.

Results—For the four indices, interobserver reliability varied from 0.92 to 0.94, test-retest reliability varied from 0.96 to 0.98, and internal consistency reliability varied from 0.80 to 0.93. The VDA compared favourably with the ADVS by correlation, but Bland-Altman analysis demonstrated that the two instruments were not clinically interchangeable. Factor analysis suggests that all test items measure a common theme, and the subgroupings reflect common themes.

Conclusions—The VDA is easy to administer because it has a short test time and scoring is straightforward. It has excellent interobserver, test-retest, and internal consistency reliability, and compares favourably with the ADVS, another test of visual disability. Factor analysis demonstrated that the 18 items measure a related theme, which can be assumed to be visual disability. The VDA is a valid instrument which provides a comprehensive assessment of visual disability in cataract

patients and is designed to detect changes within a patient over time.

(Br J Ophthalmol 1998;82:617-624)

Cataract is the leading cause of blindness worldwide and the leading cause of reversible blindness in most developed countries.¹⁻³ In developed countries, the prevalence of cataract and the success of surgery are reflected in the high expenditure on cataract surgery by government health services.³⁻⁵ However, in many countries resources are limited and expenditure may need to be justified in terms of patient benefit.²⁻⁶ The traditional measure of clinical progress, visual acuity, can both be insensitive to the presence of eye disease⁷⁻¹² and fail to completely capture visual disability.¹³⁻¹⁹ Other tests of visual function have been proposed to meet this shortfall, such as contrast sensitivity and glare loss.¹⁴⁻²¹ However, these have yet to be proved to relate to preoperative visual disability and postoperative changes in visual disability.² Some even propose that an assessment of subjective visual disability is sufficient to evaluate the cataract patient and that measures of visual functions, such as contrast sensitivity and glare loss, are unnecessary.²² Since the goal of cataract surgery is to reduce visual disability,² it is necessary first to define and then measure visual disability reliably in order to quantify patient benefit.

Disability is the restriction or lack of ability to perform an activity in a manner or within a range considered normal for a human being.²³ Visual disability is disability caused by impairment of vision. Assessment of visual disability can be used as part of any ophthalmological appraisal of patients with sight threatening disease to assess the impact of the disease on the patient as well as the impact of any treatment. The need to rate the status of patients using an index of functional disability has been recognised increasingly in many medical fields for both clinical research and clinical practice.²⁴⁻²⁶

We have constructed and validated an instrument for the assessment of subjective visual disability in cataract patients. This instrument is designed to look at the outcome of cataract surgery, including the relation between subjective visual disability and objective measures of visual function, how these change with cataract surgery, and which variables influence patient satisfaction. The construction and validation process described demonstrates how to create such an instrument. This may be of value to other investigators who need to create a tool for use in different communities. Since disability connects the

Department of
Ophthalmology,
Flinders Medical
Centre and Flinders
University of South
Australia, Australia
K Pesudovs
D J Coster

Correspondence to:
Konrad Pesudovs,
Department of
Ophthalmology, Flinders
Medical Centre, Bedford
Park, South Australia, 5042,
Australia.

Accepted for publication
9 December 1997

individual's ability to function with the demands of their environment, the system of assessment must reflect these demands. For example, aboriginal hunter gatherers from central Australia will need a different questionnaire from urban dwellers in Adelaide, South Australia. This has recently been considered in a study based in India.²⁷ Similarly, more subtle differences in communities still require differences in assessment. Thus, scoring systems are affected by locality.

In May 1993, when construction and validation of this tool for the assessment of subjective visual disability began, other questionnaires had been described and validated,^{17 28-33} but only one of these, the Activities of Daily Vision Scale (ADVS), seemed suitable for the assessment of visual disability in cataract patients in Western urban society in the 1990s. The earlier tools focused too much on gross levels of disability to be relevant to a cataract surgery population in the 1990s^{3 34-36}; this was proved in the item reduction phase. However, even the ADVS was considered to have several disadvantages that could have been overcome with the creation of a new instrument. Other groups thought similarly. In the United States, a similar instrument to ours, the VF-14, was being developed at the same time.³⁷

The major disadvantage of the ADVS for use in a study of the outcome of cataract surgery carried out in Adelaide, was the relevance of the items to the local population. There was also concern over the time taken to complete the ADVS and to calculate the scores. The paper describing the ADVS also failed to demonstrate the internal consistency of the subscales.¹⁷ One planned purpose for the VDA was to compare subscales of visual disability with various measures of visual function, such as contrast sensitivity, glare loss, and colour vision, so robust subscales were essential. Furthermore, the content of the ADVS was different from some areas we were keen to study. Specifically, mobility related disability was not sufficiently probed, there was extensive examination of driving ability for a population in which many do not drive and there was concern about the specific nature of the near task questions; these two latter content issues may waste time in data collection. Finally, there was a concern as to whether the ADVS was designed to discriminate between cataract patients along a continuum of visual disability or to evaluate change after cataract surgery. Mangione and others' original paper stressed the former and clearly stated that more research was required to establish the latter.³⁸ Subsequently, the ADVS has been successfully used to measure change in disability after cataract surgery.³⁸

An instrument for quantifying the visual disability of cataract patients in an outcome study needed not only to measure disability but also to be particularly sensitive to clinical progress. An instrument can both discriminate between individuals along a continuum of visual disability and evaluate change in an individual over time, but optimising for the former may impede the latter.³⁹ The responsiveness of an

instrument to change is a ratio of the change in subjects after intervention to the variability of stable subjects.⁴⁰ The intention was to create an instrument which was sensitive to the impact of intervention but was stable to subtle changes in clinical state.⁴¹ The obvious step was to consider the gradations of the scale; the ADVS uses a five point scale of patient response, but intuitively a four point scale may be more robust. Fewer points may be more stable again, but this would sacrifice the discrimination of individuals along a continuum of disability.⁴² This discrimination along a continuum is assisted for the subscales if they include a large number of items.³⁹ Thus, reducing the response choices to four may improve responsiveness (by increasing stability) without a great sacrifice in discriminating between individuals if factor analysis suggests that subscales should include multiple items.

The instrument also needed to be quick and easy to use because the patients in the main study were subjected to a long series of visual function testing, including measures of low contrast visual acuity, contrast sensitivity, glare loss, and colour vision, in addition to verbal questioning. Hence, further lengthy questioning would have been too tiring, especially considering the age distribution of a cataract population. Similarly, a short testing time would also make the instrument more suitable for other studies implemented in busy clinical situations where time limitations preclude the use of more lengthy tests.

Methods

SELECTING THE ITEM POOL

To ensure content validity, it was necessary to identify as many visually related tasks as possible that were potentially affected in the cataract patient. All such activities were included in the pool of items considered for the questionnaire. A retrospective examination of 50 cases of cataract yielded a list of many activities cited as impaired by cataract but which were expected to be improved by cataract surgery. Activities listed by other authors in attempts to quantify visual disability were also included,^{14 17 22 28-32 43} as were those cited in the specific literature on visual impairment and cataract.⁴⁴⁻⁵²

ITEM REDUCTION

Items that were too specific to be relevant to the majority of patients (for example, oil painting or cross stitch) were eliminated or grouped into a general item (for example, hobbies). This left 37 items which were all included in a pilot questionnaire administered to 15 cataract patients. Several items, comprising mainly self care activities such as eating, personal grooming, and use of the telephone, were eliminated after the pilot study because this cataract patient population was not sufficiently impaired to have difficulty with these activities. Such items had come chiefly from early studies of cataract and visual disability which used patient populations more impaired than that typically operated on in the 1990s.^{28 29} This reflects the increased tendency to operate at an earlier stage of cataract than 15 years ago.^{3 34-36}

Table 1 Activities listed in the visual disability assessment. The patient is asked: "To what extent, if at all, does your vision interfere with your ability to carry out the following activities?" The patient is asked to take into account both the degree to which they can perform each task as well as the extra effort involved. Assessment of visual disability is done for both eyes open and habitual spectacle correction is worn. The scoring system is included. All are counted for the total score, those marked with ^m are included in the mobility score, those marked with ^d are included in the distance/lighting/reading score, and those marked with ⁿ are included in the near and related tasks score

Activity	Not at all (1)	A little (2)	Quite a bit (3)	A lot (4)
Reading ^{d,n}	1	2	3	4
Seeing in the distance ^d	1	2	3	4
Recognising faces across the street ^d	1	2	3	4
Watching TV ^d	1	2	3	4
Seeing in bright light/glare ^d	1	2	3	4
Seeing in poor or dim light ^{d, n}	1	2	3	4
Appreciating colours ⁿ	1	2	3	4
Driving a car, by day ^d	1	2	3	4
Driving a car, by night ^d	1	2	3	4
Walking inside ^m	1	2	3	4
Walking outside ^m	1	2	3	4
Using steps ^m	1	2	3	4
Crossing the road ^m	1	2	3	4
Using public transport ^m	1	2	3	4
Travelling independently ^m	1	2	3	4
Moving in unfamiliar surroundings ⁿ	1	2	3	4
Employment/housework activities ⁿ	1	2	3	4
Hobbies/leisure activities ⁿ	1	2	3	4

Principal components analysis is a multivariate statistical method which can demonstrate how well items relate to each other and how much each item contributes to the overall variance. While it is important that all items measure a common theme, items which are too closely related are redundant. Items contributing to less than 0.4% of the variance of all items were eliminated. Redundant items included reading normal print books and reading newspapers when a more general reading question was already included. Reduction to a minimum number of items improves instrument efficiency, shortens test time, and reduces user and subject burden.³⁹ This reduced the list to 18 items (Table 1). Each remaining item was thought to contribute to the pool of information about visual disability.

RESPONSE SCALE

Categorical scales are amenable to statistical analysis if they have enough categories; a seven point scale is sufficient.³⁹ Continuous scales such as Rosser lines or visual analogue scales are also excellent but they require the questionnaire to be administered in a written format. The main hurdle for applying this questionnaire to cataract patients is that some will not be able to see well enough to complete a written format, so the questionnaire needs to be administered verbally. This leads to the problem, especially relevant to a geriatric population, that a patient can only remember a limited number of categories on a scale. Pilot work suggested that the maximum number of scaled responses that could be remembered by most subjects was four, so in response to the question: "To what extent, if at all, does your vision interfere with your ability to carry out the following activities?", the responses of choice were: not at all, a little, quite a bit, and a lot. The patient responses were recorded with the numerical values 1, 2, 3, 4, respectively. Hence, all items are scored in the same direction and in the same units. The use of a short scale also assists uniformity of interpretation.³⁹

In assessing visual disability, patients are also asked to take into account both the degree to which they can perform each task and the extra effort involved. This is because ratings of the magnitude of performance on tasks are misleading if the patient's effort is not considered.²⁴ Patients are instructed to assess their disability with both eyes open and their habitual spectacle correction worn. The patients receive no further explanation.

The questionnaire is scored by adding all the numerical scores together and dividing by the number of questions answered. Missing data are dealt with by including only answered questions in the index. For example, if all 18 questions are answered then the sum of the answers is divided by 18. If the individual has never driven a car, two questions cannot be answered and the sum of the answers is divided by 16 (18-2). Several subscales are also scored (Table 1); the choice of items in each subscale is justified by factor analysis. All questions relating to the mobility index (seven items), distance/lighting/reading index (eight items), and the near and related tasks index (five items) are aggregated in the same manner. Missing data are treated the same way as for total visual disability index. Thus all scales give a score in a range of 1 (no disability) to 4 (severe disability). The resulting instrument is called the "Visual Disability Assessment" (VDA).

Short integer scales, such as four point scales used here, lack responsiveness to subtle changes in clinical state.⁴² However, this is an advantage for detecting major changes in clinical state, such as the impact of surgery.⁴⁰ If the instrument is stable to subtle variations in clinical state, then it is more likely to be responsive to larger changes, which enhances instrument utility in outcome studies.³⁹ The creation of indices by combining several four point scale items has the effect of improving the sensitivity of the instrument to clinical change.⁴² For example, the total visual disability index includes 18 items, so the number of steps becomes $4 \times 18 = 72$. Similarly, the number of steps on the other subscales is extended through item combination. This has the advantage of creating an instrument that is extremely sensitive to major events without completely sacrificing the ability of the tool to discriminate between patients along a continuum of disability.

RELIABILITY AND VALIDITY TESTING

Reliability is the proportion of the total variance which is attributable to true differences among subjects. The remainder of the variance is noise which is considered to result from test-retest variation, interobserver variation, and internal inconsistency.⁵³ The reliability of an instrument is determined by measuring test-retest reliability, interobserver reliability, and internal consistency reliability, usually with ρ , the intraclass correlation coefficient, and Cronbach's α .⁵³⁻⁵⁵

Validity is the extent to which the instrument measures what it is intended to measure.⁵³ This is assessed by comparison with a universally

Table 2 Reliability results for the visual disability assessment

Visual disability assessment index	Interobserver reliability	Test-retest reliability	Internal consistency reliability
	ρ	ρ	Cronbach's α
Total visual disability score	0.94*	0.98*	0.93
Mobility visual disability score	0.93*	0.97*	0.92
Distance/lighting/reading visual disability score	0.92*	0.98*	0.89
Near and related tasks visual disability score	0.92*	0.96*	0.80

* $p < 0.001$ in all cases.

accepted standard (criterion validity), if one exists. However, for disability measurements there are no universally acknowledged standards.^{24 41 53 56} For this reason, validity is established by ensuring that all relevant aspects of visual disability are included in the instrument (content validity), by comparison with instruments which purport to measure the same thing, and by factorial validity which helps establish how the items in the instrument can be grouped into scales measuring the same thing.^{53 57 58} This method of validation using these three techniques together is called construct validity.

In order to examine the reliability and validity of the VDA, it was administered to 438 cataract patients attending the outpatients ophthalmology clinic at Flinders Medical Centre, Adelaide, South Australia. All patients consented to participate and had sufficient language skills and cognitive function to complete the questionnaire.

Reliability

Test-retest reliability was conducted by the same observer administering the questionnaire twice at an interval of one week. This should be sufficient separation time to negate the effects of memory without allowing the condition to change.⁵⁹ Interobserver variation testing was also conducted at an interval of one week using one other trained observer. These results were compared using ρ , the intraclass correlation coefficient. The VDA was also assessed for internal consistency reliability using the standard Cronbach's α ,⁵⁵ that tests whether multiple items in an instrument measure the same thing; this is assumed if they are summed to create a single index.

Validity

Content validity is the extent to which the items chosen reflect all visually related activities that are potentially affected in the cataract patient. Content validity cannot be formally assessed because it is difficult to prove conclusively that the items chosen were representative of all possible items.⁶⁰ However, the methods outlined above in item selection are important steps for establishment of content validity.⁵³

Criterion validation of the VDA is not possible as there is no universally accepted standard in visual disability scoring,^{24 41 53 56} but established instruments can be used to see how close they come to measuring the same thing. However, the error cannot be ascribed to either instrument to declare its departure from reality. Despite these difficulties, an important part of construct validity is to compare the

Table 3 Loadings for each item with the single factor* identified in factor analysis

Reading	0.70
Seeing in the distance	0.67
Recognising faces across the street	0.72
Watching TV	0.72
Seeing in bright light/glare	0.61
Seeing in poor (dim) light	0.69
Appreciating colours	0.30
Driving a car, by day	0.80
Driving a car, by night	0.60
Walking inside	0.66
Walking outside	0.75
Using steps	0.69
Crossing the road	0.84
Using public transport	0.83
Travelling independently	0.81
Moving in unfamiliar surroundings	0.80
Employment/housework activities	0.68
Hobbies/leisure activities	0.62

* Kaiser–Caffrey reliability coefficient of 0.94.

VDA with a surrogate standard. In order to do so, the VDA was compared with the ADVS, which is a validated instrument for measuring visual disability, and although different in construct with the VDA, both purport to measure visual disability.¹⁷ The ADVS was administered as described by its authors.¹⁷ The correlation of the two instruments was measured with Spearman correlation coefficients and the agreement or interchangeability was assessed with Bland–Altman analysis.⁶¹

The final method considered for construct validity was factor analysis which was used to provide empirical support for the instrument's scales (factorial validity).^{53 57} Factor analysis is a multivariate statistical method which when applied to a matrix of variables reduces those variables to a number of factors.⁶² The grouping of variables into factors depends on how well each variable relates to each factor. The proportion of the variance described by the principal factor indicates whether the instrument tests in one or more content areas.⁵⁸ However, factor analysis does not provide a unique solution. The analysis can be "rotated" by various techniques such as varimax or oblimin to find items which can have high communality and thus form additional factors.⁶² This grouping of items into additional factors can be used to justify the creation of subscale indices as it is proof that the items sample the same content area specified by the factor to which they contribute.⁵⁸

ANALYSIS

Item reduction utilised principal components analysis. Internal consistency reliability was estimated using Cronbach's α .⁵⁵ Test-retest reliability and interobserver reliability were tested with ρ , the intraclass correlation coefficient. Criterion validity was estimated by comparison of the VDA with the ADVS using Spearman correlation coefficients and examination with Bland–Altman analysis. Finally, factor analysis was used to support the construct validity of the VDA, including a Kaiser–Caffrey reliability coefficient.⁶³ All statistical analyses were performed on SPSS for Windows (SPSS Inc) or done manually.

Table 4 Community for the three factors identified in rotated factor analysis with the items included in the subscales. These items are the most strongly related items for each factor

Factor 1: Mobility		Factor 2: Distance/lighting/reading		Factor 3: Near and related tasks	
Walking inside	0.79	Reading	0.72	Reading	0.42
Walking outside	0.85	Seeing in the distance	0.80	Seeing in poor light	0.40
Using steps	0.75	Recognising faces	0.77	Appreciating colours	0.74
Crossing the road	0.90	Watching TV	0.71	Employment/housework	0.82
Using public transport	0.91	Seeing in bright light/glare	0.79	Hobbies	0.61
Travelling independently	0.85	Seeing in poor/dim light	0.75		
Moving in unfamiliar surroundings	0.87	Driving during the day	0.73		
		Driving at night	0.79		

Results

The VDA was administered to 438 cataract patients. Ages ranged from 40 to 91 years with a mean of 74.1 (SD 7.7) years. Sixty six per cent of subjects were female. Factorial validity testing used data from all 438 patients. Eighty six patients had repeat administration of the VDA for reliability testing and 40 were administered the ADVS for criterion validity testing.

For the purposes of instrument comparison, the time taken to administer the VDA and the ADVS was recorded on 20 patients who received both questionnaires. The VDA took on average 5 minutes to complete (range 3–9 minutes) and the ADVS took 8 minutes to complete (range 5–14 minutes). The VF-14 has been reported to take 4 minutes, on average, to administer (range 2–14 minutes).⁶⁴ The VF-14 range is not comparable with the other results because the patient group was different. In patients with poor communication skills or poor cognitive ability the VDA may also take longer. The time taken to calculate scores on the VDA varied from 2 to 3 minutes, but for the ADVS score calculation required between 5 and 7 minutes.

The data for all 438 patients gave results across the full range of each index. The total visual disability score ranged from 1.00 to 4.00 (mean 1.66 (SD 0.68)). The near visual disability score ranged from 1.00 to 4.00 (1.73 (0.71)). The distance visual disability score ranged from 1.00 to 4.00 (1.98 (0.85)). The mobility visual disability score ranged from 1.00 to 4.00 (1.40 (0.67)).

RELIABILITY

Interobserver reliability was estimated with ρ to be 0.94 for the total visual disability score and ranged from 0.92 to 0.93 for the three subscales (Table 2). Test-retest reliability was also estimated with ρ and found to be 0.98 for total visual disability score and to vary from 0.96 to 0.98 for the three subscales (Table 2). Internal consistency reliability as estimated by Cronbach's α was 0.93 for the total visual disability score and ranged from 0.80 to 0.92 for the three subscales (Table 2).

VALIDITY

Construct validity consists of a surrogate criterion validity and factorial validity. Criterion validity, where ADVS scores are used to represent the standard, was estimated by correlation using Spearman correlation coefficients, and by agreement or interchangeability, using Bland–Altman analysis. The VDA correlates well with the ADVS for overall scales -0.83

and distance scales -0.84 but less well for near scales -0.53 . These coefficients are negative because increasing disability yields an increased score on VDA, but a decreased ADVS value. Bland–Altman assessment of agreement between the ADVS and the VDA finds limits of agreement of -0.94 to 0.70 difference on the 1–4 VDA scale or -32 to 24 on the 1–100 ADVS scale. Assuming the VDA and the ADVS have identical anchoring and that there was perfect agreement, a score of 2.5 on the VDA would be 50 on the ADVS, but with these limits of agreement it varies from 18–74 (2 SD). These limits are almost equivalent to plus or minus one response category which is clearly too broad for the two instruments to be clinically interchangeable.

Unrotated factor analysis identifies one factor which explains 50% of the variance and shows excellent loadings (0.60–0.84) with all items except appreciating colours, with which it has a reasonable loading (0.30) (Table 3). The factor analysis findings give a Kaiser–Caffrey reliability coefficient of 0.94. Oblimin rotation identifies three other factors which correlate well with only some items. These factors can be categorised as representing the themes of these items. These three subvariables are best interpreted as a mobility factor, a distance/lighting/reading factor, and a near and related tasks factor (Table 4).

Discussion

The VDA takes less time to administer and score than the ADVS. This is the result of structural and methodological differences in the two questionnaires. All 18 VDA questions use the same format and are answered on the same scale, thus speeding up completion and score calculation. The questions are set up on the ADVS to ask whether the respondent performs the task and then to ask how well they perform the task. This means that for each of the 20 items, two questions are asked instead of one in the case of the VDA. The responses vary on some ADVS questions which lengthens observer explanation and slows patient response. Scoring is also slowed by the additional step of converting scale scores to a 1–100 scale. The VDA scoring method reports on a scale from 1 to 4 which is easily conceptualised. The ADVS is also slower to score because it is set out over seven pages and quite a lot of page turning is required to calculate the subscales. Previous reports suggest the VF-14 to be as quick to implement as the VDA.⁶⁴ The VDA is a quick and effective method of measuring visual disability. The items cover similar areas

as a normal history taking for cataract, although perhaps in more detail.

CONTENT VALIDITY

The reduced list of items seems to be representative of the universe of questions that could have been asked because accepted methodology was followed for arriving at the short list of items: consideration of patients' preferences on which disabilities they wanted cataract surgery to reverse, consideration of other authors' disability items, consideration of items in the literature on cataract and disability, consideration of the suggestions of experienced cataract surgeons, and item reduction by pilot study and principal components analysis.^{24 39 53 65} The other test for content validity is whether the instrument is clinically sensible,^{24 65} which it seems to be. This procedure for content validity defines the relevance of the tool to the population it was developed for. Since the cataract surgery outcome study that the VDA was developed for uses the same population, this process establishes relevance of the VDA to its target population.

RELIABILITY

The VDA exhibits very high interobserver reliability. The intraclass correlation coefficients (ρ) were 0.94 for total visual disability, 0.93 for mobility related visual disability score, and 0.92 for both distance/lighting/reading and near and related tasks visual disability. These are excellent reliability scores which suggest that the VDA is stable across different observers.⁵³ This would make the VDA suitable for outcome studies where different individuals may collect preoperative and postoperative data.

The VDA has excellent test-retest reliability. The intraclass test-retest correlation coefficients (ρ) were 0.98 for total and distance/lighting/reading visual disability, 0.97 for mobility related visual disability, and 0.96 for near and related tasks visual disability. These are exceptionally high scores.^{56 66} The error variance (random fluctuations) between the two performances is minimal.

High test-retest correlation is probably due to the use of a short four point scale. The gaps between responses are quite large in a short scale, so respondents are unlikely to give different responses. This approach gives excellent test-retest reliability but may sacrifice sensitivity to small changes in status over time.^{40 42} Finer scales, such as Rosser lines, may be more likely to give poorer test-retest correlation, but may be more sensitive to small changes in status.⁴⁰ This would be more suitable for questionnaires principally designed to scale individuals on a continuum of disability. However, this questionnaire was intended for looking at the impact of cataract surgery on visual disability, which should involve large changes in disability status, so good test-retest reliability is more important than sensitivity to subtle changes in disability status.⁶⁷ The creation of subscales by combining several four point scale items has the effect of improving the sensitivity

of the instrument to subtle clinical change by effectively increasing the number of categories on the scale.⁴² This has the advantage of creating an instrument that is both sensitive to major events without sacrificing the ability of the tool to discriminate between patients along a continuum of disability.

The excellent test-retest reliability implies that the VDA is very stable across time. High test-retest reliability is necessary for studies where the same individual is being retested at different times.⁵³ A reliability of greater than 0.90 allows comparison of differences between individual cases, whereas only comparisons between groups are appropriate for lower reliabilities.⁶⁸

The VDA has excellent internal consistency reliability. This is important since multiple items were combined into indices and such an approach is only valid if the multiple items measure the same thing, in this case, visual disability. Cronbach's α is 0.93 for total visual disability, 0.92 for mobility visual disability, 0.89 for distance/lighting/reading visual disability, and 0.80 near and related tasks visual disability. This suggests that the items in the overall index, as well as all three subscales, accurately detect the presence of their theme, which can be assumed to be domains of visual disability. It is generally held that Cronbach's α should be at least 0.80 to detect accurately the presence of the theme or to detect changes following intervention.⁵⁷ An α of 0.90 is necessary if results are to facilitate clinical decision making, but an α of 0.70 is acceptable under some circumstances—for example, exploratory research.⁶⁹

VALIDITY

Since there are no standards for visual disability, a surrogate standard is used to provide evidence for construct validity. That standard is the ADVS.¹⁷ The correlation coefficients which compare the VDA with the ADVS are negative because increasing disability yields an increasing score on the VDA but a decreasing score on the ADVS. The magnitude of the correlations are adequate at -0.83 for the overall scores, -0.84 for the distance scores, and -0.53 for the near scores. The ADVS does not have a mobility subscale. This does not prove the validity of the VDA, but simply shows that the VDA and the ADVS measure a similar concept, which is probably visual disability. However, it also does not prove that the VDA and the ADVS measure visual disability so similarly that they are interchangeable.

Bland-Altman assessment of agreement between the two measures demonstrates that the limits of agreement are too broad for the two instruments to be clinically interchangeable. This is not to say the VDA and the ADVS do not measure the same concept, but they measure and scale it in different ways. Both questionnaires measure visual disability, and both could be used for research on visual disability, but their scores cannot be compared within a single study because their limits of agreement are too broad. These differences reflect the different structures of the two

questionnaires. The VDA includes domains of visual disability which we were interested in quantifying to assess their relation with various measures of visual function, such as contrast sensitivity, glare loss, and colour vision. Therefore, the VDA has more questions on mobility related activities, whereas the ADVS probes vision for driving in more depth. The ADVS has specific questions on near tasks, whereas the VDA has general questions about near tasks which individuals can apply to their own situation. The ADVS has no specific questions about appreciating colour and the two instruments address glare disability in different ways. It should not be surprising that the two instruments do not have sufficient agreement to be clinically interchangeable, even though they both measure visual disability.

The ADVS has some practical differences from the VDA in addition to the content differences. The ADVS takes longer to administer and to calculate scores for the scales, whereas the VDA is easier to score because all items are scored in the same way. This uniformity aids interpretation.³⁹ The ADVS varies a little across questions; moreover, the final scales are converted to 100 point range whereas the individual questions are on a 1 to 5 scale. Although conversion is straightforward, it perhaps requires more thought for interpreting than the VDA. The ADVS is set out over seven pages, whereas the VDA requires only one page. The saving of both paper and time was important in the context of a large study into the outcome of cataract surgery where lots of other time and paper consuming data were collected.

An alternative instrument for use as a surrogate criterion was the VF-14.³⁷ However, the VF-14 paper was not published when development of the VDA began in May 1993. Although developed without the same rigour for item selection and reduction, the VF-14 contains many items similar to the VDA and a similar scoring system. The VF-14 also has good reliability and validity.^{37 70}

The construct validity of the VDA could also be explored by comparing VDA scores with practical tests of functional ability to perform the tasks listed. This was not attempted because the practical difficulties associated with physical measurement of a person's ability to perform such tasks as crossing a road or watching TV, including scoring and scaling problems, would cause significant variation in the relation between test scores and VDA scores. This approach would add nothing to the construct validity of the VDA.

Alternatively, the VDA could be compared with tests of vision if tests of vision were true and robust indicators of functional ability. However, there have been numerous reports that tests of vision do not capture visual disability very well.^{17 20 21 71} Furthermore, it was planned to use the VDA to explore the relation between measures of vision and visual disability, the use of measures of vision as part of construct validity created a circular logic which would defeat the aims. Again, clinical history

involves patients' subjective appraisal of their own abilities so this remains the most appropriate standard for validity testing rather than practical tests or vision tests.

The final aspect of construct validity is factorial validity. This provides empirical support for the instrument's scales by demonstrating how well the items in the VDA measure common themes.^{53 57} A large proportion of the variance is explained by the first factor which shows excellent communality with almost all items. This suggests that the instrument is valid for the measurement of one content area,⁵⁸ namely, visual disability. However, rotation of the factor analysis reveals that three factors can be identified. These can be classified as a mobility factor, a distance/lighting/reading factor, and a near and related tasks factor. The most strongly communal items were grouped into the subscales which have the common content areas listed above.⁵⁸ For the mobility index and the distance/lighting/reading subscale communality was greater than 0.70 for each item (Table 4). For the near and related tasks, communality was greater than 0.40 for each item (Table 4). The use of five to eight items in each subscale assists discrimination along a continuum of visual disability. In addition to factorial validity, each subscale has excellent internal consistency (Cronbach's $\alpha > 0.80$) which facilitates research using these subscales, such as their relation to measures of visual function. Interestingly, the VF-14 does not yield clinically coherent subscales when applied to a large number of cataract patients and examined with factor analysis.³⁷

Conclusions

The VDA is designed to quantify visual disability in the cataract patient. The 18 items included were carefully selected by robust methodology to ensure content validity. Patients are asked to assess the extent to which their vision interferes with their ability to carry out these activities. Answers are limited to four possible levels of visual interference: 1 not at all, 2 a little, 3 quite a bit, 4 a lot. The scaling system is designed to detect large changes in status that may occur with surgery rather than subtle changes that may occur with a slight increase in cataract severity. Scores for all 18 VDA items are combined to create an overall index of visual disability. Scores for subsets of items can also be combined to create subindices of visual disability. The VDA overall index and the three subscales have excellent reliability and validity. The questionnaire is quick to administer, easy to score, and has meaningful interpretation. The VDA is a suitable instrument for cataract surgery outcome studies where a measure of visual disability is required.

We wish to thank Professor John Keeves of the School of Education, Flinders University of South Australia, for helpful advice on questionnaire development, Bronwyn Krieg for interobserver data collection, and Wendy Laffer of the Department of Ophthalmology, Flinders Medical Centre, for editorial assistance.

1 Foster A, Johnson GJ. Blindness in the developing world. *Br J Ophthalmol* 1993;77:398-9.

- 2 Cataract Management Guidelines Panel. Management of functional impairment due to cataract in adults. *Ophthalmology* 1993;100:119S-49S.
- 3 Ninn-Pedersen K. Cataract patients in a defined Swedish population 1986-1990. *Acta Ophthalmol* 1997;75:1-23.
- 4 Stark WJ, Sommer A, Smith RE. Changing trends in intraocular implantation. *Arch Ophthalmol* 1989;107:1441-4.
- 5 Javitt JC, Kendix M, Tielsch JM, et al. Geographic variations in utilisation of cataract surgery. *Med Care* 1995;30:90-105.
- 6 Drummond MF, Yates JM. Clearing the cataract backlog in a (not so) developing country. *Eye* 1991;5:481-6.
- 7 Harbin TS. Visual impairment by sunlight in posterior subcapsular cataracts. *Ophthalmic Surg* 1973;4:34-6.
- 8 Hess RF, Woo G. Vision through cataracts. *Invest Ophthalmol Vis Sci* 1978;17:428-35.
- 9 Hess RF, Carney LG. Vision through abnormal cornea: a pilot study of the relationship between vision loss from corneal distortion, corneal edema, keratoconus and some allied corneal pathology. *Invest Ophthalmol Vis Sci* 1979;18:476-83.
- 10 Skalka HW. Arden grating test in evaluating "early" posterior subcapsular cataracts. *Southern Med J* 1981;74:1368-70.
- 11 Kupersmith M, Siegal I, Carr R. Subtle disturbances of vision with compressive lesions of the anterior visual pathway measured by contrast sensitivity. *Ophthalmology* 1982;82:68-72.
- 12 Hart WMJ, Burde RM. Three-dimensional topography of the central visual field: sparing of foveal sensitivity in macular disease. *Ophthalmology* 1983;90:1028.
- 13 Marron JA, Bailey IL. Visual factors in orientation-mobility performance. *Am J Optom Physiol Opt* 1982;59:413-26.
- 14 Ross JE, Bron AJ, Clarke DD. Contrast sensitivity and visual disability in chronic simple glaucoma. *Br J Ophthalmol* 1984;68:821-7.
- 15 Koch DD. Glare and contrast sensitivity testing in cataract patients. *J Cataract Refract Surg* 1989;15:158-64.
- 16 Lennerstrand G, Ahlstrom CO. Contrast sensitivity in macular degeneration and the relation to subjective visual impairment. *Acta Ophthalmol* 1989;67:225-33.
- 17 Mangione CM, Phillips RS, Seddon JM, et al. Development of the "activities of daily vision scale". A measure of visual functional status. *Med Care* 1992;30:1111-26.
- 18 Pesudovs K, Coster DJ. Visual performance and cataract morphology: the nexus between subjective disability and objective measures of visual function. *Asia Pacific J Ophthalmol* 1994;6:12-18.
- 19 Rubin GS, Bandeen Roche K, Prasada-Rao P, et al. Visual impairment and disability in older adults. *Optom Vis Sci* 1994;71:750-60.
- 20 Owsley C, Sloane ME. Contrast sensitivity, acuity, and the perception of "real world" targets. *Br J Ophthalmol* 1987;71:791-6.
- 21 Elliott DB, Hurst MA, Weatherill J. Comparing clinical tests of visual function in cataract with the patient's perceived disability. *Eye* 1990;4:712-7.
- 22 Newman DA, Houser BP. Visual disability inventory: documenting functional impairment caused by cataract. *J Cataract Refract Surg* 1991;17:244-5.
- 23 World Health Organisation. International classification of impairments, disabilities, and handicaps. Geneva: WHO, 1980.
- 24 Feinstein AR, Josephy BR, Wells CK. Scientific and clinical problems in indexes of functional disability. *Ann Intern Med* 1986;105:413-20.
- 25 Epstein AM. The outcomes movement—will it get us where we want to go? *N Engl J Med* 1990;323:266-70.
- 26 Gill TM, Feinstein AR. A critical appraisal of the quality of quality-of-life measurements. *JAMA* 1994;272:619-26.
- 27 Fletcher AE, Ellwein LB, Selvaraj S, et al. Measurement of visual function and quality of life in patients with cataracts in southern India. *Arch Ophthalmol* 1997;115:767-74.
- 28 Murphy SB, Donderi DC. Predicting the success of cataract surgery. *J Behav Med* 1980;3:1-14.
- 29 Bernth-Petersen P. Visual functioning in cataract patients—methods of measuring and results. *Acta Ophthalmol* 1981;59:198-205.
- 30 Donderi DC, Murphy SB. Predicting activity and satisfaction following cataract surgery. *J Behav Med* 1983;6:313-28.
- 31 OCTET. Subjective assessment of the effect of cataract surgery and a review of long term aims. *Eye* 1987;1:247-53.
- 32 Applegate WB, Miller ST, Elam JT, et al. Impact of cataract surgery with lens implantation on vision and physical function in elderly patients. *JAMA* 1987;257:1064-6.
- 33 Kosnik W, Winslow L, Kline D. Visual changes in daily life throughout adulthood. *J Gerontol* 1988;43:63-70.
- 34 Moorman C, Sommer A, Stark WJ, et al. Changing indications for cataract surgery. *Ophthalmic Surg* 1990;21:761-6.
- 35 Batterbury M, Khaw PT, Hands R, et al. The cataract explosion: the changing pattern of diagnoses of patients attending an ophthalmic outpatient department. *Eye* 1991;5:369-72.
- 36 Stenevi U, Lundstroem M, Thorburn W. A national cataract register. I. Description and epidemiology. *Acta Ophthalmol* 1995;73:41-4.
- 37 Steinberg EP, Tielsch JM, Schein OD, et al. The VF-14: an index of functional impairment in patients with cataract. *Arch Ophthalmol* 1994;112:630-8.
- 38 Mangione CM, Phillips RS, Lawrence MG, et al. Improved visual function and attenuation of declines in health related quality of life after cataract extraction. *Arch Ophthalmol* 1994;112:1419-25.
- 39 Jaeschke R, Guyatt GH. How to develop and validate a new quality of life instrument. In: Spilker B, ed. *Quality of life assessment in clinical trials*. New York: Raven Press, 1990:47-57.
- 40 Guyatt GH, Walter S, Norman G. Measuring change over time: assessing the usefulness of evaluative instruments. *J Chron Dis* 1987;40:171-8.
- 41 Deyo RA, Inui TS. Toward clinical applications of health status measures: sensitivity of scales to clinically important changes. *Health Serv Res* 1984;10:275-84.
- 42 Bailey IL, Bullimore MA, Raasch TW, et al. Clinical grading and the effects of scaling. *Invest Ophthalmol Vis Sci* 1991;32:422-32.
- 43 Elam JT, Graney MJ, Applegate WB, et al. Functional outcome one year following cataract surgery in elderly persons. *J Gerontol* 1988;43:M122-6.
- 44 Allan HF. Cataracts: functional disability is important in determining the clinical significance. *Geriatrics* 1975;30:47-50.
- 45 Hilbourne JFH. Social and other aspects of adjustment to single eye cataract extraction in elderly patients. *TOSUK* 1975;95:254-9.
- 46 Mackensen G, Bohn P. How do patients assess their visual acuity after intracapsular extraction and correction with eyeglasses. *Klin Monatsbl Augenheilkd* 1984;184:84-8.
- 47 Immonen I, Tuominen R, Raivo I. Visual results and social rehabilitation after cataract surgery. A retrospective comparison of the effect of non-IOL and IOL surgery. *Acta Ophthalmol* 1988;66:572-6.
- 48 Bruce DW, Gray CS. Beyond the cataract: visual and functional disability in elderly people. *Age Ageing* 1991;20:389-91.
- 49 Chan CW, Billson FA. Visual disability and major causes of blindness in NSW: a study of people aged 50 and over attending the Royal Blind Society 1984 to 1989. *Aust NZ J Ophthalmol* 1991;19:256-65.
- 50 Legro MW. Quality of life and cataracts: a review of patient-centred studies of cataract surgery outcomes. *Ophthalmic Surg* 1991;22:431-43.
- 51 Bergman B, Sjostrand J. Vision and visual disability in the daily life of a representative population sample aged 82 years. *Acta Ophthalmol* 1992;70:33-43.
- 52 Lowe KJ, Gregory DA, Jeffery RI, et al. Suitability for day case cataract surgery. *Eye* 1992;6:506-9.
- 53 McDowell I, Newell C. *Measuring health: a guide to rating scales and questionnaires*. New York: Oxford University Press, 1987:208-210.
- 54 Bravo G, Potvin L. Estimating the reliability of continuous measures with Cronbach's alpha or intraclass correlation coefficient: toward the integration of two traditions. *J Clin Epidemiol* 1991;44:381-90.
- 55 Cronbach LJ. Coefficient alpha in the internal structure of tests. *Psychometrika* 1951;16:297-334.
- 56 Spector WD. Functional disability scales. In: Spilker B, ed. *Quality of life assessment in clinical trials*. New York: Raven Press, 1990:47-57.
- 57 Turner RR. Rehabilitation. In: Spilker B, ed. *Quality of life assessment in clinical trials*. New York: Raven Press, 1990:247-67.
- 58 Comrey AL. Common methodologic problems in factor analytic studies. *J Consult Clin Psychol* 1978;46:648-59.
- 59 Granger CV. A conceptual model for functional assessment. In: Granger CV, Gresham GE, eds. *Functional assessment in rehabilitation medicine*. Baltimore: Williams and Wilkins, 1984:223-61.
- 60 Seiler LH. The 22-item scale used in field studies of mental illness: a question of method, a question of substance, and a question of theory. *J Health Soc Behav* 1973;14:252-64.
- 61 Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;iii:307-10.
- 62 Spearritt D. Factor Analysis. In: Keeves JP, ed. *Educational research, methodology, and measurement; an international handbook*. 2nd ed. Oxford: Pergamon, 1997:528-39.
- 63 Kaiser HF, Caffrey J. Alpha factor analysis. *Psychometrika* 1965;30:1-14.
- 64 Musch DC, Farjo AA, Meyer RF, et al. Assessment of health-related quality of life after corneal transplantation. *Am J Ophthalmol* 1997;124:1-8.
- 65 Feinstein AR, Wells CK, Joyce CM, et al. The evaluation of sensibility and the role of patient collaboration in clinimetric indexes. *Trans Assoc Am Phys* 1985;98:146-9.
- 66 Helmstadter GC. *Principles of psychological measurement*. London: Methuen, 1966.
- 67 Kirshner B, Guyatt GH. A methodological framework for assessing health indices. *J Chron Dis* 1985;38:27-36.
- 68 Ware JE. Methodological considerations in selection of health status assessment procedures. In: Wenger NK, Mattson ME, Furberg CD, eds. *Assessment of quality of life in clinical trials of cardiovascular therapies*. New York: Lejacq, 1984.
- 69 Nunnally JC. *Psychometric theory*. New York: McGraw-Hill, 1978:245.
- 70 Cassard SD, Patrick DL, Damiano AM, et al. Reproducibility and responsiveness of the VF-14. *Arch Ophthalmol* 1995;113:1508-13.
- 71 Ball K, Owsley C. The useful field of view test: a new technique for evaluating age-related declines in visual function. *J Am Optom Assoc* 1993;64:71-9.