Activities of Daily Vision Scale: What Do the Subscales Measure?

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

Purpose. Previous Rasch analysis of the Activities of Daily Vision Scale (ADVS) did not address psychometric properties of its subscales or provide detailed assessment of dimensionality (whether the ADVS measures one or multiple constructs). This study was designed to examine these properties.

Methods. Two hundred thirty-two participants (mean age, 74.2 years) awaiting cataract surgery self-administered the ADVS. Rasch analysis was used to assess the ADVS and its five subscales for unidimensionality (by principal components analysis, [PCA]), precision by person separation (discrimination between strata of participant ability), and targeting (matching of item difficulty to participant ability). Adequate person separation (minimum acceptable value, 2.0) is the fundamental requirement for measurement.

Results. Only the near vision subscale had adequate measurement properties (person separation, 2.30). The entire ADVS showed a misfit to the Rasch model and lacked unidimensionality. PCA confirmed the presence of two additional traits—driving and glare disability—but neither possessed adequate person separation when assessed individually. Deleting these traits restored unidimensionality, but additional items misfit, necessitating item reduction. Finally, an eight-item ADVS-Near Vision Scale showed good fit and unidimensionality. Its contents were identical with the original near vision subscale. Targeting was suboptimal (2.30 logits).

Conclusions. Only one subscale, near vision, met the criteria for measurement. The revised eight-item ADVS-Near Vision subscale is a unidimensional measure of visual disability in cataract patients with mild visual disability. However, it is limited by measurement of near visual ability only. For more comprehensive measurement of visual disability, other questionnaires such as Catquest-9SF are preferable for cataract surgery outcomes assessment. (*Invest Ophthalmol Vis Sci.* 2010; 51:694–700) DOI:10.1167/iovs.09-3448

The Activities of Daily Vision Scale (ADVS) was developed as a self-report instrument in patients with cataract to assess the need for surgery, as well as the outcomes after surgery.[1,2] In the earlier investigation of the psychometric properties of the ADVS using Rasch analysis on a relatively small sample of 43 patients with cataract, Pesudovs et al.[3] proposed a reduced 15-item version, albeit with misfitting items and poor targeting. However, such suboptimal properties do not meet the standards of a good measure. Furthermore, certain components of the ADVS have not been addressed so far in the literature. These include measurement properties of the subscales, assessment of dimensionality (whether the questionnaire taps a single latent construct or a number of constructs), and differential item functioning (DIF) that tests whether items perform consistently across population subgroups.

Subscales were included in the ADVS to provide a comprehensive assessment of activities that affect a typical cataract patient. The psychometric properties of subscales need to be assessed individually because performance of the overall questionnaire cannot infer adequate functioning of its subscales. Measurement precision (as determined by person separation with a minimum acceptable value of 2.0) is an important property of subscales, because subscales usually contain a small number of items that will limit the potential person separation. Unidimensionality (that the questionnaire taps a single latent construct) is a prerequisite for generating a summary score.[4–8] The presence of multiple subscales in the ADVS intended to capture different aspects of visual disability (e.g., glare, driving) may cause multidimensionality and therefore should be investigated. Previous Rasch analysis of the ADVS used fit statistics to examine the dimensionality. However, recent studies suggest that fit statistics alone are inadequate for determining unidimensionality and recommend performing principal components analysis (PCA) of the residuals for more detailed evidence of dimensionality (see the Methods section for further explanation of PCA).[9–13] Another key attribute required in a measure is the absence of differential item functioning (DIF). DIF occurs when different subgroups of participants (e.g., sex, age), despite equal levels of underlying trait (visual disability in the case of ADVS), respond differently to a given item.[14,15] For example, in the ADVS, if men and women scored differently on the item *playing cards*, this item would be said to be displaying sex-related DIF.

Furthermore, simplified conversion from raw to Rasch scaled scores were not provided in the earlier Rasch analysis of the ADVS. Such conversions, if made available would circumvent the need to perform Rasch analysis, enabling researchers to use the scoring benefits of the analysis. Similar conversion algorithms have been developed for other questionnaires revalidated by Rasch analysis.[16–18]

Given the above limitations of the earlier Rasch analysis of the ADVS, we set out to provide a more comprehensive anal-

From the [1]NH&MRC (National Health and Medical Research Council) Centre for Clinical Eye Research, Department of Optometry and Vision Science, Flinders Medical Centre and Flinders University of South Australia, Bedford Park, South Australia, Australia; the [2]Meera and L. B. Deshpande Centre for Sight Enhancement, Vision Rehabilitation Centres, L. V. Prasad Eye Institute, Hyderabad, India; the [3]Centre for Eye Research Australia, Department of Ophthalmology, University of Melbourne, Melbourne, Victoria, Australia; the [4]Vision CRC (Cooperative Research Center)-Sydney, Sydney, New South Wales, Australia; and the [5]Singapore Eye Research Institute, Singapore National Eye Centre, Singapore.

Supported by NHMRC Public Health Fellowship 359277 (ELL); Australian Research Council Linkage Grant LP0560779 (ELL); NHMRC Centre for Clinical Research Excellence in Ophthalmology Outcomes Research, 2005–2009 (KP); NHMRC Career Development Award, 2007–2011 (KP); Eye Foundation Australian and New Zealand Surveillance Ophthalmic Unit, 2008–2012 (KP); and FMC (Flinders Medical Center) Foundation (KP).

Submitted for publication January 23, 2009; revised May 26, and July 23 and 25, 2009; accepted August 1, 2009.

Disclosure: **V.K. Gothwal**, None; **T.A. Wright**, None; **E.L. Lamoureux**, None; **K. Pesudovs**, None

Corresponding author: Konrad Pesudovs, NH&MRC Centre for Clinical Eye Research, Department of Optometry and Vision Science, Flinders Medical Centre, Bedford Park, South Australia, 5042, Australia; konrad.pesudovs@flinders.edu.au.

Table 1. Sociodemographics of the Study Population for ADVS

Characteristic		*n* (%) or Mean ± SD
Age, y		74.2 ± 9.6
Sex		
Male		105 (45.3)
Female		127 (54.7)
Distance visual acuity, habitual		
Surgical eye	LogMAR (Snellen)	0.48 ± 0.30 (20/63^{+1}*)
	Range	−0.10 to 1.60 (20/16 to hand movements)
Fellow Eye	LogMAR (Snellen)	0.30 ± 0.31 (20/40)
	Range	−0.26 to 2.00 (20/10^{-2}† to light perception)
Binocular	LogMAR (Snellen)	0.22 ± 0.19 (20/32^{-1}‡)
	Range	−0.26 to 0.90 (20/10^{-2}§ to 20/160)
Awaiting second-eye surgery		93 (42.3)
Ocular comorbidity‖		
Present		106 (46.5)
Absent		122 (53.5)
Systemic comorbidity¶		
Present		171 (85.9)
Absent		28 (14.1)

$n = 232$.
* Participant read the 20/63 line completely and one extra letter from the next, smaller line.
† Participant missed two letters in the 20/10 line.
‡ Participant missed one letter in the 20/32 line.
§ Participant missed two letters in the 20/10 line.
‖ Includes glaucoma, diabetic retinopathy, and age-related macular degeneration; data were missing in four cases.
¶ Includes diabetes, hypertension, and angina, and data were missing in 33 cases.

ysis via the following three steps: First, apply the analysis to investigate the measurement properties of the native subscales of the ADVS in an Australian cataract population; second, assess the dimensionality of the entire ADVS, specifically using PCA of residuals and determine whether more appropriate subscales could be formed. If we found that the ADVS was not unidimensional, then we considered re-engineering to create a unidimensional scale; third, provide ready-to-use spread sheets that convert raw scores to Rasch-scaled scores for the ADVS as a whole and for its valid subscales.

Methods

Activities of Daily Vision Scale

The 22 items (20 activities) of the ADVS are categorized into five subscales: night driving, daytime driving, distance vision activities that do not include driving (far vision), near vision activities, and glare disability.[3] Three items contribute to two different subscales and therefore are included in both the subscales; thus the total number of items across subscales is 25. The questionnaire responses were organized and assigned ordinal values, as recommended by the developers.[2]

Study Population

Participants were patients with cataract attending the Flinders Medical Centre, Adelaide, South Australia. They were mailed the ADVS questionnaire for self-administration while on the waiting list (average waiting period, 3–4 months) for cataract extraction. The questionnaires were returned via a self-addressed, prepaid envelope.

Included patients were 18 years of age or older, were English-speaking, and had no severe cognitive impairment. Approval of the ethics of the protocol was obtained, and all patients who agreed to participate signed a consent form. The study was conducted in accordance with the tenets of the Declaration of Helsinki. We included patients with coexisting ocular and systemic comorbidities, as exclusion may provide an inaccurate picture of the elderly cataract population in Australia.[19] Characteristics of participants who completed the ADVS are shown in Table 1.

Clinical Assessment

Routine clinical assessments were performed by an ophthalmic team, and cataract was established as the principal cause for visual disability in each patient. All assessments were performed before cataract extraction. Habitual visual acuity was measured by using computerized testing based on logMAR principles with screen illumination of 150 cd/m^2. All assessments were performed monocularly and binocularly.

Rasch Analysis

The data were analyzed with Winsteps software (ver. 3.68)[20] using the Andrich rating scale model for polytomous data.[21]

In the first step, we assessed the response categories and the thresholds.[8,21] The threshold represents the intersection between any two adjacent categories (i.e., between 1 and 2, 2 and 3, and so on) where the probability of either category being chosen is equal. In the ADVS, there are four thresholds for five categories of each item. We used category probability curves (CPCs) to examine the ordering of thresholds graphically. Thresholds should demonstrate an order from most to least difficult category, but disordering can occur. Disordered thresholds suggest that the response categories are not efficient in discriminating between two ability levels; that is, participants with more ability could respond with the same category as another participant with lower ability. Disordering occurs because participants have difficulty discriminating between response categories. We reorganized the categories that showed disordered thresholds by combining certain categories. Once the response categories were found to perform as intended, we performed further Rasch analyses.

Measurement precision was assessed in terms of person separation, which gives an estimate of the spread or separation of persons by strata or groups along the measurement construct.[8,22] The minimum acceptable separation is 2.0, and this enables the distinction of three strata (for example, mild, moderate, and severe visual disability).

Rasch fit statistics in combination with PCA of residuals were used to test the dimensionality of the ADVS and each subscale.[23] As the Rasch model is probabilistic, some amount of deviation in scores is expected. This deviation in expected versus observed scores is captured by fit statistics (i.e., infit mean square, or MnSq).[22] The ideal value of Infit MnSq is 1.0 (indicates no deviation). In accordance with the literature, an infit MnSq between 0.7 and 1.3 was an indicator of acceptable fit. Items outside this range were considered misfits.[24] In essence, this range permitted observations to contain up to 30% less or more variation than predicted by the model. Misfitting items were removed iteratively (i.e., one at a time) starting with the most misfitting, until all remaining items fit the model.[25] Furthermore, when items fit the model's expectations, the residuals[26] (observed minus expected scores) should be randomly distributed, with all meaningful variance in the data accounted for by the Rasch dimension of item difficulty–person ability. In practice, however, some interitem correlations typically remain; PCA describes the additional factors that may be extracted from the data.[9–11] If 60% or more of the variance is accounted for by the principal component, then there is a low likelihood of additional components being present.[27] The first contrast in the residuals reveals whether there are any patterns within the variance unexplained by the principal component to suggest that a second trait is being measured. We used the criterion of an eigenvalue of >2.0 for the first contrast, which indicates that the contrast has the strength of at least two items to be sufficient evidence of a second construct, as this is greater than the magnitude seen with random data.[27]

In the present study, we performed PCA, an assessment for misfitting items. Thus, the iterative method to remove items that did not fit the model is different from the earlier Rasch analysis of the ADVS. We used this approach because items can misfit for several reasons, including poorly constructed wording. Fit statistics identify only items that misfit, not misfitting items that group to form additional constructs, so fit statistics alone are not as informative of multidimensionality as PCA. When PCA is performed first, it helps to more clearly identify additional construct(s) if they are present in the overall scale.

An ideal scale should function in the same way regardless of which group is assessed. DIF occurs when given the same level of the latent trait, the difficulty levels of items vary systematically based on sample characteristics, such as age and sex. The variables for DIF analysis, selected a priori, included age (<76 years vs. ≥76 years; median age, 76), sex, cataract status (first eye versus second eye surgery), systemic comorbidity and ocular comorbidity (present versus absent). Testing for DIF can occur based on either significance or magnitude. Because significance testing is highly sample-size dependent, we prefer testing for DIF magnitude.[28] Therefore, in the present study, we defined DIF based on magnitude: insignificant DIF as <0.50 logit, mild (but probably inconsequential) as between 0.50 and 1.00 logit, and notable as >1.00 logit.[29]

For a well-targeted instrument (i.e., item difficulty matched with participant ability), there would be no ceiling or floor effects in the person-item map.[30,31] Consequently, mistargeting implies lower person separation, leading to inability to differentiate between participants along the latent trait.[30] The person-item map illustrates targeting and further helps to identify gaps and redundancies in the item distribution. Appropriate items can then be added to fill the gaps, and redundant items can perhaps be deleted.

Adequate person separation constituted the minimum acceptable measurement properties of the Rasch models for the subscales and the entire ADVS to be termed a measure. If the subscales could not be repaired, full analysis of dimensionality using PCA was not performed.

Rasch analysis was conducted in two phases: assessment of performance of the subscales in phase I, and investigation of dimensionality of the entire ADVS to determine whether more appropriate subscales could be developed in phase II. Descriptive statistics were analyzed with commercial software (SPSS software ver. 15.0; SPSS, Chicago, IL).

Results

Of 478 questionnaires mailed, 232 were returned, providing an overall response rate of 48.5%.

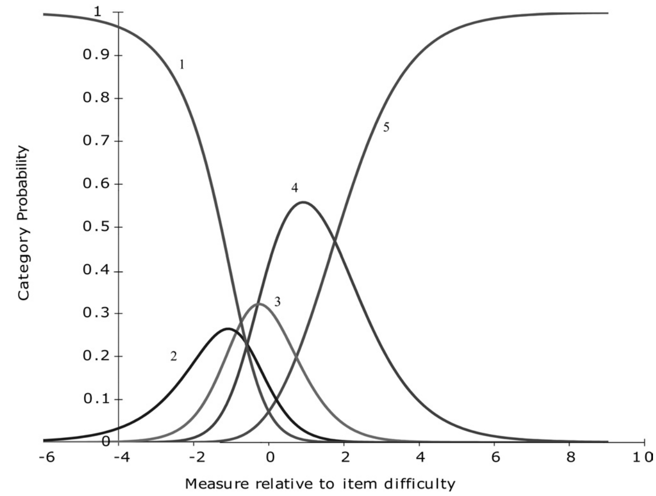

Figure 1. Rasch model category probability curves for all items together in the ADVS showing the likelihood that a participant with a particular visual disability will select a category. The scale (*x*-axis) from −6 to +10 symbolizes the latent trait of visual disability, with severity of level of difficulty increasing toward the *right*. The *y*-axis represents the probability of category being selected. Response categories: 1, unable; 2, extremely difficult; 3, moderately difficult; 4, a little difficult; 5, not at all difficult. For any given point along this scale, the category most likely to be chosen by a participant is shown by the category curve with the highest probability. At no point was category 2 the most likely to be chosen, resulting in disordered thresholds.

Phase I: Assessment of the Native Subscales

Assessment of Response Categories of the ADVS. Participants did not use the response categories as intended. The response categories were intended to cover a range of visual disability, whereby each category should be the most likely to be chosen for part of this range representing stepwise increase in severity. However, this was not the case. Figure 1 shows the category probability curves, which illustrate the range of visual disability for which each of the five response categories were most likely to be chosen. Category 2, extremely difficult is not the most likely category to be endorsed at any level of visual disability. This is described as disordered thresholds, because the threshold between categories 1 and 2 lay to the right (instead of the left) of the threshold between categories 2 and 3. Therefore, we combined category 2 with category 3 to form a new category (a lot of difficulty) and thus reduced the number of categories from five to four and ordered all thresholds.

Analyses of Subscales. Only one (near vision) of the five subscales possessed acceptable measurement properties (Table 2). The main problem was lack of person separation, which could not be remediated without the addition of items. However, addition of items is beyond the scope of the present study. The results of further analyses of the near vision subscale are presented in the next section.

Near Vision. Person separation was satisfactory (Table 2). One item misfit and was deleted, after which the remaining eight items fit the model; person separation continued to remain satisfactory, but targeting worsened by 0.26 logits. Mistargeting was evidenced by a mean participant ability of 2.30 logits. This result indicates that the participants had visual abilities that extended well beyond what could be captured by the items (Fig. 2). However, on visual inspection of the person-item map (Fig. 2), it is not readily apparent that some of the more able participants would be targeted by item thresholds. Each item has sublevels of difficulty and can be performed without difficulty, with a little difficulty, a lot of (moderate or extreme) difficulty, or cannot be performed at all, and each

TABLE 2. Performance of the Subscales of the ADVS

	Subscales				
Parameter	**Far Vision**	**Near Vision**	**Glare Disability**	**Night Driving**	**Day Driving**
Items, *n*	6	9	3	4	3
Misfitting items, *n*	0	1	0	0	0
Person separation	1.62	2.30	0	1.89	1.39
Mean item location	0	0	0	0	0
Mean person location	0.71	1.40	0.28	1.37	1.92
Principal components analysis, eigenvalue	—	2.0	—	—	—

level feeds into measurement. Thus, an eight-item near vision subscale can represent 4×8 separate levels of difficulty. Nevertheless, in this case, targeting was suboptimal. PCA of the residuals showed that the variance explained by the measures was 68.0%, and the unexplained variance explained by the first contrast was 2.0 eigenvalue units. There were no significant additional contrasts. Only one item showed DIF by sex. Males rated the item *read ingredients on cans of food* 0.60 logits easier relative to other tasks than did their female counterparts (Table 3).

Phase II: Assessment of the Dimensionality of the Entire ADVS

The person separation was good (>2.0), the targeting was reasonable, and two items misfit (Table 4). PCA of the residuals showed that the variance explained by the measures (57.2%) was less than ideal, and the unexplained variance explained by the first and second contrasts was 2.5 and 2.1 eigenvalue units respectively (Table 4). Four items loaded (correlation, >0.4) positively onto the first contrast and belonged to night and daytime driving (two items each). Four items loaded (correlation, >0.4) positively onto the second contrast and belonged to glare disability (2 items), far vision, and daytime driving (1 item each). Seven (31.8%) items showed DIF by sex and age (Table 5). Taken together, these findings indicate that the ADVS was not unidimensional. The items measuring different traits had to be removed. Therefore, items from these contrasts were deleted to restore unidimensionality. After deletion, 14 items, predominantly related to near vision, remained as the core of the ADVS, which had adequate person separation and was unidimensional by PCA.

However, item misfit existed in the 14-item core scale and a few iterations were necessary to optimize its performance. Six misfitting items were removed, one at a time, starting with the most misfitting item. After this, the remaining eight items fit the Rasch model. These items were identical with that of the eight-item near vision subscale and therefore shared the same psychometric properties. Henceforth, this reduced version is referred to as the ADVS-Near Vision Scale. There was no role for any of the originally proposed subscales in this reduced version.

As noted in the introduction, the previous Rasch analysis of the ADVS proposed a 15-item version (although not ideal due to retention of misfitting items). The authors had used a different iterative method (using fit statistics only) to delete items that did not fit the model. In the present study, we conducted PCA first, followed by removal of misfitting items. However, to be consistent with the earlier study, we also tried eliminating misfitting items in the first step followed by PCA. The results, however, were the same. Therefore, we determined that there was only one solution to measurement with the ADVS—an eight-item Near Vision Scale.

Instead of discarding the items from the two contrasts found in the PCA, we investigated whether these items could be used to form separate subscales with valid measurement properties. Person separation was inadequate for both the scales (1.49 and 1.35). Thus, the decision to delete these items was appropriate.

Criterion Validity

The correlation between mean participant ability and visual acuity in the worse eye (i.e., eye to be operated on) was not significant ($r = -0.03$, $P = 0.71$). However, low, but statisti-

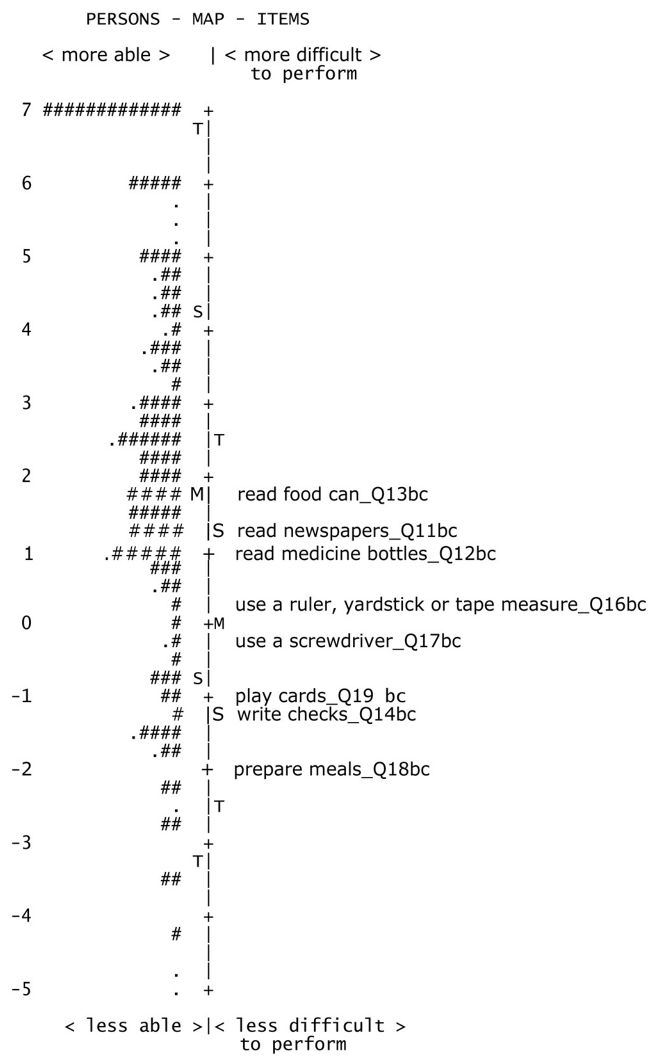

FIGURE 2. Person-item map for the eight-item ADVS-Near Vision scale ($n = 232$) in cataract assessment. *Vertical line*: the measure of the visual disability variable, in logit units. Participants appear in ascending order of ability (on the *left* side of the map), whereas the items appear in ascending order of difficulty (on the *right* side of the map). Alongside each item, its number is indicated, as in the 22-item original ADVS. Item names have been abbreviated to fit the space; the correct description of items can be found in Mangione et al.[2] Each #, two participants; each dot, one to three participants; M, mean; S, 1 SD from the mean; T, 2 SD from the mean. By convention, the mean item difficulty is set at 0 logits (indicated with M). Accordingly, mean visual ability of participants is indicated with M.

TABLE 3. Items Showing Differential Item Functioning in All the Five Subscales of the ADVS

	Demographic Variable				
Item	Sex	Age	Cataract Status	Systemic Comorbidity	Ocular Comorbidity
Walk down steps without handrails or help in dim light	Males* (0.90)	Younger† (0.57)	Awaiting surgery in first eye† (0.65)	—	—
Use public transportation	—	Older† (0.97)	—	—	—
Walk down steps without handrails or help during day light	—	—	—	Without† (0.97)	—
Watch television	Females* (0.87)	—	—	With† (0.89)	—
Read ingredients on cans of food	Males‡ (0.60)	—	—	—	—
See peoples' faces from across the street	—	Older§ (0.78)	—	—	—
Drive in unfamiliar areas	—	—	—	—	With‖ (1.10)

All values are in logits (i.e., log of odds ratio or the log-odds of the level of difficulty of an item relative to the difficulty of the total set of items analyzed) and the listed subgroup rated these items as easier relative to other tasks by the amount of logits indicated in parentheses.
* Night driving.
† Far vision subscale.
‡ Near vision subscale.
§ Glare disability.
‖ Day driving

cally significant, correlation was obtained between mean participant ability and visual acuity in the better eye ($r = -0.20$, $P = 0.01$).

Conversion of Raw Scores to Rasch Measure

Ideally, users of the revised versions of the ADVS should perform Rasch analysis on their own data, as populations may vary. However, for those who wish to use the scoring benefits of Rasch analysis (but may not be familiar with the process), we have developed ready-to-use spread sheets for conversion of raw scores to Rasch-scaled scores for the ADVS-Near Vision Scale. These sheets can be obtained by contacting the corresponding author or can be downloaded as Supplementary Material at [INSERT URL]. However, we caution that these conversions can be applied only if the sample is similar to that of the present study.

DISCUSSION

The first goal was to determine whether the five proposed subscales of the ADVS possess the properties of a measure. Only one subscale (near vision) fulfilled the criteria. Among the desirable features of an optimally functioning instrument include an ability to discriminate as many strata or groups of participant ability as possible, simulating the gradations on a ruler; the finer the gradations, the better the measurement properties.[19,31] The four dysfunctional subscales (far vision, glare, and daytime and night driving) lacked adequate discriminative ability, in that they could distinguish only between two strata (i.e., able versus unable) of participant ability. Given that person separation is sample dependent, the finding of dysfunctional subscales is therefore only applicable in similar populations. Therefore, subscales should be tested in other populations. Assuming that this sample is typical of a cataract population in the developed world, the likelihood of finding adequately performing subscales would be low. Other instruments, such as the Impact of Vision Impairment (IVI) questionnaire, which were revalidated by Rasch analysis in a similar cataract population, have optimal functioning subscales due to a sufficient number of well-targeted items.[16]

The second goal was to assess the dimensionality of the ADVS, specifically using the PCA of residuals. This analysis revealed that the ADVS was not unidimensional, thereby, invalidating the use of an overall or summary score.[5–7] The PCA of residuals indicated the presence of two additional dimensions: driving and glare disability. This finding suggested that the ADVS was measuring more than one trait, violating one of the essential requirements for measurement: unidimensionality. Other researchers have also expressed concerns about items related to driving (and mobility) that do not fit the core set of items.[3,32] Item misfit also confirmed the lack of unidimensionality of the ADVS. However, the finding of item misfit is not novel. Earlier Rasch analysis of the ADVS by Pesudovs et al.[19] also showed misfitting items, along with the presence of

TABLE 4. Overall Performance of the Original and Revised Versions of the ADVS

Versions	Phase I Original ADVS	Phase II Revised Version (ADVS: Near Vision)
Items, *n*	22	8
Misfitting items, *n*	2	0
Person separation	3.00	2.32
Reliability	0.90	0.84
Mean item location	0	0
Mean person location	0.90	1.66
Principal components analysis, eigenvalue	2.5 (first contrast 2.1 (second contrast)	2.0
Number of valid subscales	1	0

Table 5. Items Showing Differential Item Functioning in Entire (Native Version) ADVS

Item	Demographic Variable				
	Sex	Age	Cataract Status	Systemic Comorbidity	Ocular Comorbidity
Drive at night	Men (0.59)	—	—	—	—
Walk down steps without handrails or help in daylight	Men (1.24)	—	—	—	—
Walk down steps without handrails or help in dim light	Men (1.01)	Younger (0.61)	—	—	—
Watch television	Women (0.51)	—	—	—	—
Thread a needle without using threading device	Women (0.57)	—	—	—	—
See peoples' faces from across the street	—	Older (1.02)	—	—	—

All values are in logits (i.e., log of odds ratio or the log-odds of the level of difficulty of an item relative to the difficulty of the total set of items analyzed) and the listed subgroup rated these items as easier relative to other tasks by the amount of logits indicated in the parentheses. There was no differential item functioning by cataract status or by systemic or ocular comorbidity.

two chief dimensions (driving and reading). This analysis, however, lacked comprehensive assessment of dimensionality using PCA.[3] The presence of misfitting items in this study indicates that the items in the ADVS were not clearly understood, or were measuring some other trait, and therefore added noise (inaccuracy) to the measurement scale.[3,33–35] Finally, the existence of notable DIF further established that the ADVS was not unidimensional. The DIF was notable by age and sex for three items. Consequently, comparison of scores for these different participant groups may not be appropriate. Although the reason for DIF of these items is not entirely clear, one explanation for DIF by age, may be that older participants were influenced by physical limitations due to age or comorbidity, so that it was more difficult to perform activities such as walking down steps without handrails or help in dim light, compared with their younger counterparts.

The lack of unidimensionality calls into question the validity of the native ADVS. Reestablishing unidimensionality is vital to optimizing the properties of the ADVS. Unidimensionality was restored by deleting items from the contrasts found in the PCA and items that misfit. This process resulted in the formation of an eight-item ADVS-Near Vision Scale which was identical with that of the reduced version of the near vision subscale in the ADVS. Unlike the earlier Rasch analysis of the ADVS which used only 43 patients, the larger sample size in the present study helped to estimate item difficulty more precisely, enabling the determination of a satisfactory solution. Thus, the total score generated from the ADVS—Near Vision scale is valid with interval level properties. Therefore, its scores can be managed by using parametric statistics where required (i.e., comparison of preoperative and postoperative data).

The ADVS-Near Vision scale is not without limitations. First, it can effectively measure near visual ability, but not other areas of visual ability. Its limited scope is a shortcoming, as it measures only a subset of the visual disability issues facing the cataract patient. Second, targeting was poor in this preoperative cohort, and one could speculate that targeting would only get worse after surgery as a result of the improved visual functioning expected after cataract surgery. The poor targeting would cause a ceiling effect to the measurement which would cause underestimation of the real change that occurs with surgery.[19,30] Mistargeting indicates that the items of the ADVS were too easy for the visual abilities of the participants. This finding means that the population either did not have visual disability or that the questions that formulate the ADVS represent tasks that are relatively easy, and our population had visual disability on the more difficult tasks. We argue that it is the latter, because at our center, visual disability is the indication for cataract surgery; no patients are listed for cataract surgery unless they report visual disability. However, patients may have visual difficulties with personally relevant tasks (i.e., threading a needle) that are not included in the ADVS. Thus, the activities in the ADVS that were included when it was developed almost two decades ago appear unsuitable for the current patient with cataract who is undergoing surgery in Australia, where there has been a considerable lowering of the threshold for cataract surgery in the past few years.[36,37]

Except for the Catquest-9SF,[29] nonlinearity (i.e., ceiling and floor) effects are common in the visual function questionnaires examined with Rasch analysis, which may result in underestimation of the clinical improvements for the participants with mild visual disability.[16,17,38] Results of the present study indicated that ADVS could be refined by enriching the upper extreme of the scale with more difficult items such as those requiring fine resolution. Although new items can be generated and added to legacy instruments such as the ADVS, this approach requires revalidation in a new population. Revalidation is possible, but a comparatively superior strategy would be formation of item banks and computer adaptive testing (CAT).[39–41] Item banks contain Rasch-calibrated items pooled from different questionnaires that can be administered to participants by a computerized algorithm that targets the ability of the participant according to his or her response (CAT). Such a strategy would help eliminate the limitation of poor targeting. Furthermore, a relatively smaller number of items would be needed to specifically target a given participant, with the resultant effect of reduced participant burden. Item banking and CAT have been created and used in other areas of health assessment.[42,43] Results such as those seen in the present study suggest that the ophthalmic community should be engaged in the development of such an item bank.

In conclusion, the new, reduced version of the ADVS is unidimensional, essentially measuring near visual ability. With superior measurement properties, the ADVS-Near Vision Scale can be used in place of the original ADVS in patients with cataract as they present today in the developed world. Its brevity may make clinical application easier, but it is limited by measuring only one aspect of visual disability. The ADVS is still in use,[44] and this is appropriate if only measuring near visual disability is required. However, if measurements of other aspects of visual disability are desired, for instance, scoring of overall visual disability, then questionnaires such as the Catquest-9SF[29] are more appropriate. Similarly, if a measure of emotional well-being is required, then the IVI questionnaire is a better choice.[16] Therefore, these results highlight the need for researchers to determine the content under measurement before selecting questionnaires.

References

1. Mangione CM, Orav EJ, Lawrence MG, Phillips RS, Seddon JM, Goldman L. Prediction of visual function after cataract surgery: a prospectively validated model. *Arch Ophthalmol.* 1995;113:1305–1311.
2. Mangione CM, Phillips RS, Seddon JM, et al. Development of the 'Activities of Daily Vision Scale': a measure of visual functional status. *Med Care.* 1992;30:1111–1126.
3. Pesudovs K, Garamendi E, Keeves JP, Elliott DB. The Activities of Daily Vision Scale for cataract surgery outcomes: re-evaluating validity with Rasch analysis. *Invest Ophthalmol Vis Sci.* 2003;44: 2892–2899.
4. Rasch G. *Probabilistic Models for Some Intelligence and Attainment Tests*. Copenhagen, Denmark: Institute of Educational Research; 1960.
5. Tesio L. Measuring behaviours and perceptions: Rasch analysis as a tool for rehabilitation research. *J Rehabil Med.* 2003;35:105–115.
6. Conrad KJ, Smith EV Jr. International conference on objective measurement: applications of Rasch analysis in health care. *Med Care.* 2004;42:11–16.
7. Tennant A, McKenna SP, Hagell P. Application of Rasch analysis in the development and application of quality of life instruments. *Value Health.* 2004;7(suppl 1):S22–S26.
8. Wright BD, Masters GN. *Rating Scale Analysis.* Chicago: MESA Press; 1982.
9. Linacre JM. Structure in Rasch residuals: why principal components analysis? *Rasch Meas Trans.* 1998;12:636.
10. Linacre JM. Detecting multidimensionality: which residual datatype works best? *J Outcome Meas.* 1998;2:266–283.
11. Smith EV Jr. Detecting and evaluating the impact of multidimensionality using item fit statistics and principal component analysis of residuals. *J Appl Meas.* 2002;3:205–231.
12. Smith EV Jr, Johnson BD. Attention deficit hyperactivity disorder: scaling and standard setting using Rasch measurement. *J Appl Meas.* 2000;1:3–24.
13. Smith RM, Miao CY. Assessing unidimensionality for Rasch measurement. In: Wilson M, ed. *Objective Measurement*. Norwood, NJ: Ablex Publishing; 1994:316–327.
14. Smith RM, Suh KK. Rasch fit statistics as a test of the invariance of item parameter estimates. *J Appl Meas.* 2003;4:153–163.
15. Scheuneman JD. A method for assessing bias in test items. *J Educ Meas.* 1979;16:143–152.
16. Pesudovs K, Caudle LE, Rees G, Lamoureux EL. Validity of a visual impairment questionnaire in measuring cataract surgery outcomes. *J Cataract Refract Surg.* 2008;34:925–933.
17. Lamoureux EL, Pesudovs K, Pallant JF, et al. 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;15:224–233.
18. 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;47:4732–4741.
19. Pesudovs K, Burr JM, Harley C, Elliott DB. The development, assessment and selection of questionnaires. *Optom Vis Sci.* 2007; 84:663–674.
20. Linacre JM. *WINSTEPS. Rasch Measurement Computer Program*. Chicago: Winsteps.com; 2009.
21. Andrich DA. A rating scale formulation for ordered response categories. *Psychometrika.* 1978;43:561–573.
22. Bond TG, Fox CM. *Applying the Rasch model: Fundamental Measurement in the Human Sciences*. Mahwah, NJ: Lawrence Erlbaum Associates; 2001.
23. Linacre JM. *Learning from Principal Components Analysis of Residuals.* Chicago, IL: COMET; 1999.
24. Wright BD, Linacre JM. Reasonable mean-square fit values. *Rasch Meas Trans.* 1994;8:370.
25. Hagquist C, Bruce M, Gustavsson JP. Using the Rasch model in nursing research: an introduction and illustrative example. *Int J Nurs Stud.* 2009;46:380–393.
26. Moore D, McCabe G. *Introduction to the Practice of Statistics.* New York: WH Freeman & Co; 1993.
27. Linacre JM. *A User's Guide to WINSTEPS.* Chicago, IL: Winsteps.com; 2009.
28. Wang WC, Yao G, Tsai YJ, Wang JD, Hsieh CL. Validating, improving reliability, and estimating correlation of the four subscales in the WHOQOL-BREF using multidimensional Rasch analysis. *Qual Life Res.* 2006;15:607–620.
29. 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;35:504–513.
30. Mallinson T, Stelmack J, Velozo C. A comparison of the separation ratio and coefficient alpha in the creation of minimum item sets. *Med Care.* 2004;42:I17–24.
31. Mallinson T. Why measurement matters for measuring patient vision outcomes. *Optom Vis Sci.* 2007;84:675–682.
32. Globe D, Varma R, Azen SP, Paz S, Yu E, Preston-Martin S. Psychometric performance of the NEI VFQ-25 in visually normal Latinos: the Los Angeles Latino Eye Study. *Invest Ophthalmol Vis Sci.* 2003;44:1470–1478.
33. White LJ, Velozo CA. The use of Rasch measurement to improve the Oswestry classification scheme. *Arch Phys Med Rehabil.* 2002; 83:822–831.
34. Smith RM. Person fit in the Rasch model. *Edu Psychol Meas.* 1986;46:359–372.
35. Massof RW, Fletcher DC. Evaluation of the NEI visual functioning questionnaire as an interval measure of visual ability in low vision. *Vision Res.* 2001;41:397–413.
36. Taylor HR, Vu HT, Keeffe JE. Visual acuity thresholds for cataract surgery and the changing Australian population. *Arch Ophthalmol.* 2006;124:1750–1753.
37. McCarty CA, Keeffe JE, Taylor HR. The need for cataract surgery: projections based on lens opacity, visual acuity, and personal concern. *Br J Ophthalmol.* 1999;83:62–65.
38. Velozo CA, Lai JS, Mallinson T, Hauselman E. Maintaining instrument quality while reducing items: application of Rasch analysis to a self-report of visual function. *J Outcome Meas.* 2000;4:667–680.
39. Fayers PM. Applying item response theory and computer adaptive testing: the challenges for health outcomes assessment. *Qual Life Res.* 2007;16(suppl 1):187–194.
40. Hays RD, Lipscomb J. Next steps for use of item response theory in the assessment of health outcomes. *Qual Life Res.* 2007; 16(suppl 1):195–199.
41. Cook KF, O'Malley KJ, Roddey TS. Dynamic assessment of health outcomes: time to let the CAT out of the bag? *Health Serv Res.* 2005;40:1694–1711.
42. Haley SM, Ni P, Hambleton RK, Slavin MD, Jette AM. Computer adaptive testing improved accuracy and precision of scores over random item selection in a physical functioning item bank. *J Clin Epidemiol.* 2006;59:1174–1182.
43. Haley SM, Coster WJ, Andres PL, Kosinski M, Ni P. Score comparability of short forms and computerized adaptive testing: simulation study with the activity measure for post-acute care. *Arch Phys Med Rehabil.* 2004;85:661–666.
44. Denoyer A, Denoyer L, Halfon J, Majzoub S, Pisella PJ. Comparative study of aspheric intraocular lenses with negative spherical aberration or no aberration. *J Cataract Refract Surg.* 2009;35:496–503.