An Evaluation of the Amblyopia and Strabismus Questionnaire Using Rasch Analysis

Marta Vianya-Estopa, David B. Elliott, and Brendan T. Barrett

Purpose. To evaluate whether the Amblyopia and Strabismus Questionnaire (A&SQ) is a suitable instrument for the assessment of vision-related quality-of life (VR-QoL) in individuals with strabismus and/or amblyopia.

Methods. The A&SQ was completed by 102 individuals, all of whom had amblyopia, strabismus, or both. Rasch analysis was used to evaluate the usefulness of individual questionnaire items (i.e., questions); the response-scale performance; how well the items targeted VR-QoL; whether individual items showed response bias, depending on factors such as whether strabismus was present; and dimensionality.

Results. Items relating to concerns about the appearance of the eyes were applicable only to those with strabismus, and many items showed large ceiling effects. The response scale showed disordered responses and underused response options, which improved after the number of response options was reduced from five to three. This change improved the discriminative ability of the questionnaire (person separation index increased from 1.98 to 2.11). Significant bias was found between strabismic and nonstrabismic respondents. Separate Rasch analyses conducted for subjects with and without strabismus indicated that all A&SQ items seemed appropriate for individuals with strabismus (Rasch infit values between 0.60 and 1.40), but several items fitted the model poorly in amblyopes without strabismus. The AS&Q was not found to be unidimensional.

Conclusions. The findings highlight the limitations of the A&SQ instrument in the assessment of VR-QoL in subjects with strabismus and especially in those with amblyopia alone. The results suggest that separate instruments are needed to quantify VR-QoL in amblyopes with and without strabismus. (*Invest Ophthalmol Vis Sci.* 2010;51:2496–2503) DOI:10.1167/iovs.09-4381

Three broad approaches have been used to understand the impact of amblyopia and strabismus on the individuals who live with these conditions. The ability of individuals with amblyopia/strabismus to perform everyday tasks such as reading,[1] reaching and grasping,[2,3] hand-eye co-ordination,[4] gait,[5] and driving[6,7] have been examined and compared with that of persons with normal vision. Another approach involves population-based studies[8,9] aimed at understanding the impact of the conditions on, for example, education and employment.[10–14] A third approach is to use questionnaires[15–18] or interviews,[19] thus allowing an assessment of any psychosocial impact that the conditions may exert.[19–30] To date, there is very little evidence concerning the impact that amblyopia alone may exert on vision-related quality of life (VR-QoL).[31,32] On the other hand, there have been several previous investigations of the self-reported psychosocial impact of strabismus or strabismus surgery[18,20–24,26–30] and the views and judgments of employers[10,12,13] or dating agency[14] personnel about the employment and dating prospects of those with strabismus. Some studies have also claimed a psychosocial effect of amblyopia,[28,30,31] although these have typically included a large majority of subjects who also had strabismus.

The purpose of this study was to examine whether the first VR-QoL questionnaire to be developed for amblyopia and strabismus, the Amblyopia and Strabismus Questionnaire (A&SQ),[33] is a suitable instrument for the assessment of VR-QoL in individuals with amblyopia and/or strabismus. The A&SQ has been tested and validated by its developers,[16,33–35] but it is not clear to what extent it can identify the effects on VR-QoL of amblyopia separately from those associated with strabismus. Although approximately two thirds of amblyopes have strabismus,[36,37] amblyopia is commonly present without strabismus, and strabismus can be present without amblyopia.[38] Although the functional consequences of amblyopia and strabismus may be similar—for example, in relation to the consequences of having reduced or absent stereopsis[38]—the psychosocial aspects relating to the two conditions could differ markedly; in strabismus, the eyes have an abnormal appearance, whereas the same does not apply to eyes with amblyopia alone. In this study, A&SQ data were gathered in individuals with strabismus only, in individuals with amblyopia only, and in those with both strabismus and amblyopia.

Methods

Basic Approach

The A&SQ was evaluated by using traditional psychometric approaches combined with Rasch analysis, an item-response theory–based model.[39] This methodology is similar to that used to assess VR-QoL questionnaires in patients with cataract[40,41] and in individuals with different forms of refractive correction.[42–44] Rasch analysis[45–47] uses the responses to all questions (often called items) from a group of subjects and models the data to provide an item-person map, on which respondents are placed on a logit scale of low to high VR-QoL (in this case), and items are placed on the same scale as determined by their degree of difficulty (see Fig. 2 for examples). A logit (derived from log-odds unit) is the log of the odds ratio and is calculated as log (P/1 − P), where P is the probability that a subject will choose a particular response category for an item. Rasch analysis can provide an indication of the appropriateness of the response categories. The A&SQ has a five-point response scale for each item, and Rasch analysis can determine whether subjects respond as though categories 0, 1, 2, 3, and 4 are in ascending order or whether some are disordered and whether some categories could be usefully merged if responses suggest that neighboring categories appear indistinguishable to respondents. The

From the School of Optometry and Vision Science, University of Bradford, Bradford, United Kingdom.

Submitted for publication July 27, 2009; revised November 23, 2009; accepted December 15, 2009.

Disclosure: **M. Vianya-Estopa**, None; **D.B. Elliott**, None; **B.T. Barrett**, None

Corresponding author: Brendan T. Barrett, School of Optometry and Vision Science, University of Bradford, Bradford BD7 1DP, West Yorkshire, UK; b.t.barrett@bradford.ac.uk.

Investigative Ophthalmology & Visual Science, May 2010, Vol. 51, No. 5

person-item map can indicate whether the items are at an appropriate level of difficulty for the subjects: Is there a close match between items and subjects, or does the map suggest that items are typically too difficult or too easy? The difficulty can be quantified using the difference between the mean logit score for the items and subjects. Rasch analysis can also indicate how well each individual item fits the model. A poorly fitting item may not be measuring the same thing as other items used, or it may be poorly worded and therefore poorly understood by respondents. Some items may provide redundant information, perhaps because other items in the questionnaire provide very similar information. Rasch analysis can indicate whether the questionnaire can discriminate well between respondents and thus allows the precision of the instrument to be assessed as changes to the questionnaire, such as removing items or changing the response categories, are made. Finally, Rasch analysis can also indicate whether the data are providing information that is unidimensional or whether the responses suggest that more than one dimension is being assessed. For example, the data may suggest that all items of a questionnaire measure VR-QoL, or they may indicate that some items measure functional vision and others assess psychosocial difficulties and that the two groups of items are entirely separate and should not be combined into one score.

Participants

One-hundred eighty individuals from the University of Bradford Eye Clinic, from the orthoptic clinics at Leeds Teaching Hospitals NHS (National Health Service) Trust, and from general optometric practice were invited to complete the A&SQ questionnaire. They received copies of the A&SQ questionnaire, either by mail or handed to them by their clinicians. They were also asked to provide demographic details about themselves relating to their age, sex, and ethnicity. The inclusion criteria for the study were 16 years of age or older and the presence of amblyopia, strabismus, or both. Amblyopia was defined as a difference in best corrected visual acuity of at least one line between the right and left eyes, the presence of amblyogenic factors, and/or a history of treatment for amblyopia. Exclusion criteria included any current or past ocular disease or abnormality, apart from strabismus and amblyopia, that could alter visual function. In cases in which the clinician issued the questionnaire, he or she provided a report including a note of the clinical details (presence/absence of amblyopia) and of the best corrected visual acuity of the right and left eyes. In cases in which the questionnaire was distributed by mail, clinical data were obtained from the clinical notes once a written informed consent form for participation in the study had been completed. The study adhered to the tenets of the Declaration of Helsinki and had the approval of the Leeds Teaching Hospitals and the University of Bradford's research ethics committees.

Amblyopia and Strabismus Questionnaire

The A&SQ[33] was downloaded from http://www.retinafoundation.org/questionniare.html/.

The individual items of the self-administered A&SQ are shown in abbreviated form in Table 2. The A&SQ has three items (1, 17, and 24) that act as filters only and thus were not included in any analyses. All but three items (4, 5, and 26) had the following five-category Likert response scale: *none of the time* (score 4), *a little of the time* (score 3), *some of the time* (score 2), *most of the time* (score 1), and *all of the time* (score 0). The responses for item 26 were: *definitely false* (score 4), *mostly false* (score 3), *not sure* (score 2), *mostly true* (score 1), and *definitely true* (score 0). Items 4 and 5 had three response options: *yes* (score 4), *moderate* (score 2) and *no* (score 0). The suggested scoring rule of the A&SQ is to score all missing data and not-relevant responses with the highest possible score of 4.[33] However, the validity of this approach has been called into question,[17–19] and it is not used in most quality-of-life questionnaire scoring systems.[39] Rasch analysis can overcome the problem of missing data, as it can accurately predict responses for individual items from individual patients, based on their responses to all other items, because it is based on a modeling system.[45,46] For this reason, in the present study, missing data and not-relevant responses were not assigned any score and were thus left blank for all analyses.

Statistical Analysis

The A&SQ questionnaire data were analyzed by using descriptive statistics, including the percentage of missing responses, skew and kurtosis, and ceiling and floor effects (SPSS for Windows, ver. 15.0; SPSS, Chicago, IL).

Rasch Analysis

All Rasch analyses were performed with a computer program (Winsteps, ver. 3.66) and by applying the Andrich rating scale model with an item-grouping level in which all items that share the same response scale (all except 4, 5, and 26) were analyzed together as a single group.[47] First, Rasch analysis was used to investigate how the response categories were used and whether any were disordered or whether neighboring categories were indistinguishable to respondents and could be merged. Second, we established whether the A&SQ behaves consistently among predefined subgroups within the sample. We were particularly interested in whether individuals with and without strabismus would respond differently to any item within the A&SQ. These differences were tested by using differential item functioning (DIF) and establishing whether item response bias was found within different subgroups, despite having equal levels of the underlying trait (i.e., VR-QoL).[45,47] The following predefined subgroups were evaluated: strabismus (presence or absence), sex (male or female), age (presbyopic, >45 years; prepresbyopic, 45 years or younger), and the presence or absence of spectacles for distance viewing. The χ^2 test comparing the mean level of item difficulty between each of these subgroups was used to identify significant levels of DIF.[47] Third, the usefulness of each item from the questionnaire (i.e., how well it fit the model) was determined; we examined how well each item matched the VR-QoL of the participants and whether removal of some items and/or merging of some response scales could lead to an improved instrument. Finally, Rasch analysis was used to determine whether the data were unidimensional.[39,45,46,41–43] Principal components analysis (PCA) of the residuals (i.e., observed responses minus their expected responses, as indicated by the Rasch model) was used to assess the dimensionality of the A&SQ.[48] PCA decomposes the item correlation matrix to first indicate what proportion of variance of the residuals is explained by the principal component. If it explains a large amount of the variance of the data, say >60%, then it is likely that the dataset is unidimensional. Patterns within the variance that are unexplained by the principal component suggest that a second construct is being measured. An eigen-value greater than 2.0 suggests a second dimension, as this is greater than the magnitude seen with random data. We finally examined any items that could form a second dimension and determined whether their number and inherent qualities suggest that an important second dimension is being measured (Linacre JM, personal communication, October 2008).

Results

Participant's Characteristics

One hundred two participants completed the questionnaire, for a response rate of 57% (Table 1). Clinical characteristics of the sample including proportions of participants with amblyopia only, strabismus only, and both amblyopia and strabismus, as well as median visual acuities for the better and worse eyes are presented in Table 1. We did not collect demographic data from the nonrespondents.

Descriptive Statistics

As expected, none of the 33 nonstrabismic amblyopes reported problems related to misalignment of their eyes (items

TABLE 1. Characteristics of the 102 Participants Who Completed the A&SQ[33]

Characteristic	Result
Mean age ± SD, y	48.3 ± 17.5
Sex, *n* (%)	
Male	30 (29)
Female	72 (71)
Ethnicity, *n* (%)	
White	99 (97)
Other	3 (3)
All participants: median VA better eye, logMAR; range	0.00; −0.20 to 0.48
All participants: median VA worse eye, logMAR; range	0.48; −0.10 to 2.30
Anisometropic amblyopia, *n* (%); median VA of worse eye, logMAR	33 (32); 0.38
Strabismic amblyopia, *n* (%); median VA of worse eye, logMAR	39 (38); 0.60
Mixed amblyopia, *n* (%); median VA of worse eye, logMAR	19 (19); 0.70
Strabismus without amblyopia, *n* (%); median VA of worse eye, logMAR	11 (11); 0.00

25 and 26); they gave no score for these items. Descriptive data are therefore presented separately in Table 2 for participants with strabismus (with or without amblyopia, $n = 69$) compared with those with amblyopia only ($n = 33$). Many of the strabismic subgroup items were not normally distributed, with statistically significant skewness and kurtosis (items 6, 7, 9, 10, 11, 14, 15, and 16) principally due to very high ceiling effects; these ranged from 72.5% to 91.3% with most subjects having

TABLE 2. Descriptive Statistics of the A&SQ

	Strabismus with/without Amblyopia ($n = 69$)						Amblyopia Only ($n = 33$)					
A&SQ Items	Skew	Kurtosis	Missing Data (%)	Ceiling (%)	Floor (%)	Infit	Skew	Kurtosis	Missing Data (%)	Ceiling (%)	Floor (%)	Infit
1. I can see equally well with both eyes						(Filter)						
2. Worry about losing better eye	−0.5	−0.4	8.7	20.3	7.2	1.35	−0.7	0.1	3.0	27.3	6.1	0.89
3. Worry something might get into better eye	−0.9	0.1	8.7	30.4	4.3	1.20	−1.1	0.7	3.0	39.4	6.1	0.72
4. Distance estimation	−0.4	−1.2	0	42.0	21.7	0.93	−0.6	−0.6	3.0	45.5	9.1	1.49
5. Depth perception	−0.4	−1.2	4.3	40.6	20.3	0.95	−0.4	−0.8	3.0	48.5	3.0	1.06
6. Feel unsure when putting something on a table	−2.4	4.5	0	84.1	0	0.87	−3.1	10.1	0	84.8	0	0.60
7. Miss the other person's hand when shaking hands	−3.0	7.2	0	91.3	0	0.68	−3.0	7.3	0	90.9	0	0.46
8. Difficulty parking car	−0.9	−0.7	30.4	40.6	0	1.27	−1.6	2.4	18.2	51.5	0	1.16
9. Difficulty putting the cap on a pen	−2.9	8.1	0	87.0	0	0.64	−3.0	7.3	0	90.9	0	0.82
10. Difficulty putting a power plug into a socket	−2.7	5.4	0	89.9	0	0.69	−2.0	2.3	0	84.8	0	0.91
11. Difficulties pouring drinks	−2.6	8.1	0	73.9	1.4	0.75	−3.7	14.0	0	90.9	0	0.77
12. Difficulties walking down stairs	−2.0	3.4	0	65.2	4.3	0.95	−1.7	1.4	0	75.8	0	1.06
13. Difficulties playing ball games	−1.1	0	37.7	31.9	7.2	0.83	−1.4	0.9	42.4	39.4	0	1.59
14. Difficulties in a shopping mall	−2.1	3.2	4.3	73.9	1.4	0.86	−1.5	1.1	6.1	66.7	0	1.26
15. Difficulties in a supermarket	−2.2	3.5	0	81.2	0	0.99	−1.8	2.5	0	75.8	0	1.21
16. Difficulties in a train station	−2.0	3.1	−4.3	72.5	0	0.82	−1.9	4.3	3.0	63.6	0	0.84
17. I see double						(Filter)						
18. Double vision disturbs me in my daily activities	−0.9	0.1	66.7	14.5	0	0.66	−0.1	−1.1	60.6	12.1	0	0.69
19. When tired, careful not to miss what I reach for	−1.7	2.1	0	68.1	0	0.70	−1.9	2.8	0	72.7	0	1.08
20. When tired, I have to do things more slowly	−1.5	1.5	0	62.3	1.4	0.86	−1.9	3.5	0	66.7	0	0.79
21. Squint or shut one eye in bright sunlight	−0.2	−0.8	2.9	18.8	11.6	1.4	−0.4	−1.1	0	36.4	15.2	1.20
22. Difficulties making eye contact in a one-to-one conversation	−1.2	0.3	1.4	53.6	4.3	1.28	−2.5	7.1	0	75.8	3.0	1.13
23. Difficulty making eye contact in a group conversation	−1.3	0.8	1.4	56.5	4.3	1.12	−3.5	13.9	0	81.8	3.0	0.94
24. My eyes are misaligned						(Filter)						
25. I feel insecure because of the eye misalignment	−0.7	−0.6	40.6	20.3	5.8	1.27	—	—	100	—	—	—
26. Decreased self-confidence because of the eye misalignment	0.4	−1.2	40.6	8.7	20.3	1.20	—	—	100	—	—	—

The infit values were determined after response-scale reduction but prior to item reduction.

TABLE 3. Summary of Rasch Outputs Used during Response Scale Reduction

	Strabismus and/or Amblyopia (n = 102)			
Response Scale	**Logit Patient**	**Logit Item**	**Person Separation Index**	**Person Separation Reliability**
01234	1.49	0.00	1.98	0.80
01234	1.25	0.00	2.02	0.80
01234	0.94	0.00	2.11	0.98

01234, original 5-point Likert score; 01234, 4-point Likert score combining the two least-able response options; 01234, 3-point Likert score combining the three least-able response options.

no problem with these task-based items. Significant skewness and kurtosis were present in the amblyopia-only subgroup dataset for items 6, 7, 9, 10, 11, 12, 15, 19, 20, 22, and 23 and due to large ceiling effects that ranged from 72.7% to 90.9%.

Of the 69 individuals with strabismus, only 4 (5.8%) reported double vision all the time in response to item 17. Seven (10.1%) and 13 (18.8%) individuals reported diplopia some of the time and a little of the time, respectively. Forty-five (65.2%) individuals with strabismus had never experienced diplopia. In contrast, only 2 (6.1%) of the 33 amblyopes without strabismus reported having diplopia some of the time. Eleven (33.3%) of the 33 experienced double vision a little of the time, and the remainder (60.6%) had diplopia none of the time.

Rasch Analysis: Response Scale Assessment

Examination of the 20 items with a five-response option for all 102 participants showed that category 0 (typically relating to difficulties with various tasks all the time) was chosen by <10% of the participants for 83% of the items (Table 2). This finding indicates that very few participants had substantial difficulty with a large majority of A&SQ items. In addition, several items showed disordered responses for categories 0, 1, and 2, indicating that participants had difficulty discriminating among these three options, which have similar labeling (*all the time*, *most of the time*, and *some of the time*). Thus, according to suggested guidelines for response scale reduction,[39] we combined the two least-able response-scale categories (scores 0 and 1). This reduction of scoring categories improved the discriminatory ability of the A&SQ, as indicated by an increase in the person separation index (Table 3), which is used to indicate the reliability of the scale to discriminate between people of different abilities; the minimum recommended value is 2.00.[41] A further reduction of the rating scale category (Table 3) reduced the mismatch between person and item scores (see the next section) with a further improvement of the person separation index. Therefore, the five-response option was reduced to a three-response option representing *some of the time* (or more), *a little of the time*, and *none of the time* for all items except item 26 where it represents *definitely true*, *not sure*, and *definitely false*. Finally, examination of the two items that already had a three-point scale (items 4 and 5) showed that the response options were used appropriately. Thus, no further scale-reductions were made in the case of these items.

Differential Item Functioning

DIF analysis (i.e., item response bias) was performed on the whole dataset (i.e., from all 102 participants). There was no significant DIF between items for sex or the wearing of spectacles. However, three items exhibited DIF for age (Fig. 1). Items 2 (*worry about losing good eye*) and 12 (*difficulties walking down stairs*) were more difficult in those aged 45 years and older (χ^2 with $P = 0.009$ and 0.001, respectively). Item 22 (*difficulties making eye contact in a one-to-one conversation*) was more difficult in the younger participants (χ^2, $P = 0.005$).

In relation to the presence or absence of strabismus, items 25 and 26 (*I feel insecure because of the eye misalignment* and *decreased self-confidence because of the eye misalignment*) could not be analyzed with DIF, because all nonstrabismic individuals reported that these were not applicable. However, DIF analysis revealed that participants with strabismus found item 23 (*difficulties making eye contact in a group conversation*) significantly more difficult than was the case for nonstrabismic amblyopes (χ^2, $P = 0.029$).

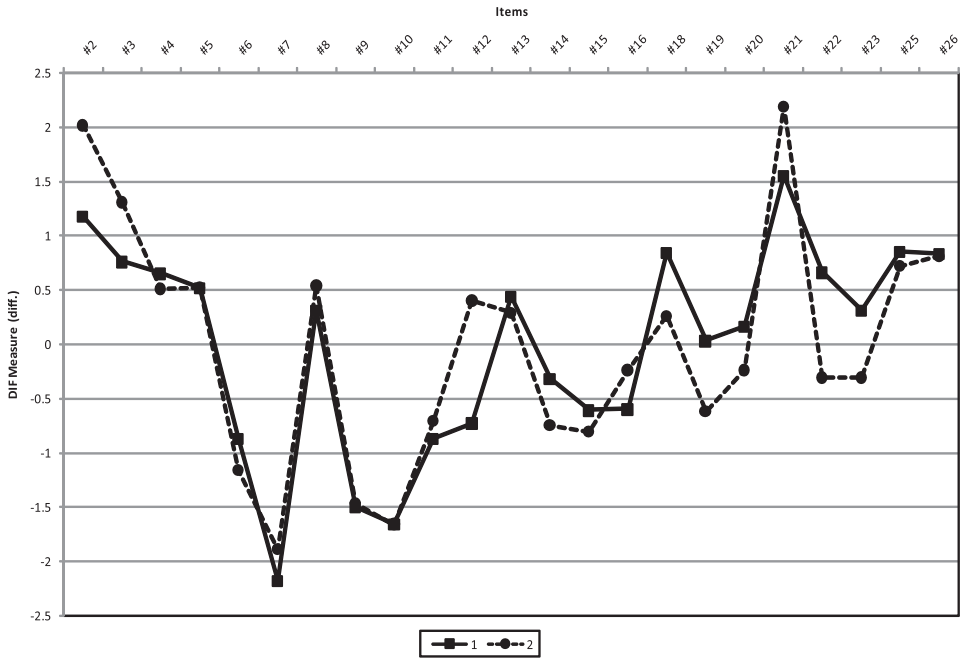

FIGURE 1. DIF plot showing measures for subjects aged <45 years (*solid line*) versus those aged 45 years or older (*dashed line*). Items 2, 12, and 22 exhibited a statistically significant DIF for age.

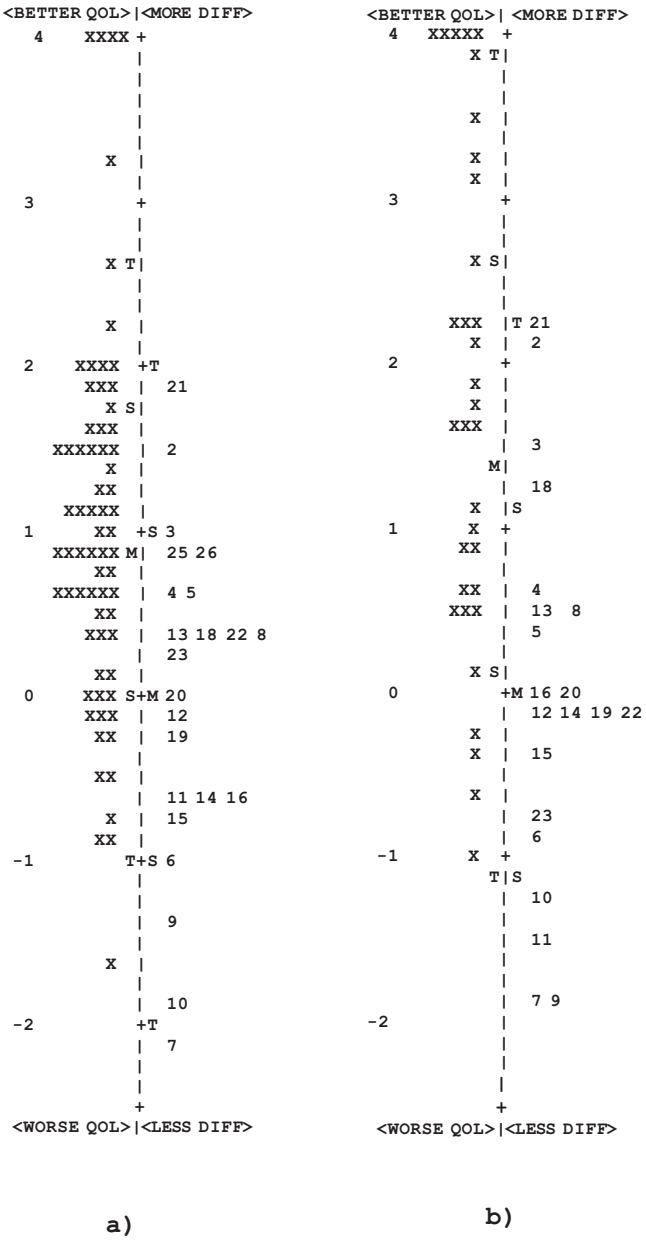

FIGURE 2. Person-item maps for the 23-item A&SQ presented separately for subjects with (**a**) strabismus (with and without amblyopia) and (**b**) amblyopia-only. On the *left* of the *dashed line* are the subjects, represented by X. On the *right* are the questionnaire items represented by their numbers. Participants with better vision-related quality of life and more difficult items are near the top of the map. The scale in is logits. M, mean; S, 1 SD from the mean; T, 2 SD from the mean.

Fit of the A&SQ Items to the Rasch Model

Because of the significant differences between the strabismic and nonstrabismic amblyope subgroups shown by the descriptive statistics and DIF analysis, we examined the usefulness of each item (i.e., how well it fit the Rasch model) for each of the subgroups. Figure 2 shows the person-item map (after response-scale reduction) where item difficulty and person ability are plotted along the same linear scale in logit units. Participants are shown as X's on the left of the vertical line, and items appear on the right. More difficult items and participants with better VR-QoL appear in descending order from the top of the map.

Item Reduction

Rasch fit statistics (outfit and infit) were used to identify how well each item contributes to VR-QoL and whether the *removal* of items that contribute little to the A&SQ would lead to better performance. We focused on the evaluation of infit values, since they are weighted to take less notice of extreme responses. The infit mean square values for each item after response-scale reduction are reported in Table 2. During item elimination, items were removed one at a time, and the effect of their removal was assessed in terms of any change in the person-separation index and in relation to the difference between the person- and item-means. Items were removed if they improved the person-separation index and/or decreased the difference between the person- and item-means. The process is iterative, and the analysis was rerun to identify the next candidate item for removal until all remaining items appeared useful or the removal of an item resulted in a person-separation index below 2.00.[39,45] We used a lenient definition of good fit of infit between 0.60 and 1.40,[49] rather than the stricter but more commonly used 0.70 to 1.30,[39] due to the relatively small number of subjects in the amblyopia-only subgroup ($n = 33$), a factor that limits the reliability of the Rasch analysis results in this subgroup. Infit values <0.60 indicate that the responses are too predictable, and values greater than 1.40 indicate too much variation in the responses.[39,41] Hence, items with infit mean squares outside the range of 0.60 to 1.40 were considered for removal by applying the following additional criteria[39]: item mean farthest from the subject mean, missing data for greater than 50% of cases, a ceiling effect greater than 80%, and both skewness and kurtosis greater than 2.0.

All items in the strabismic group had infit mean square values in the range of 0.60 to 1.40 (Table 2), and thus no items were removed from the A&SQ. Items 25 and 26 were not available for inclusion in the analysis of the nonstrabismic group, as no participant provided a scoring response. In addition, several items in the amblyopia-only subgroup showed infit values outside the 0.60 and 1.40 range, and items 13, 7, and 4 were removed (in that order), after which all the remaining items fitted the model. After removal of these three items, the difference between person and item means was 0.33 and the person separation was 2.20.

Dimensionality Assessment

The analysis showed that only 46.9% of the variance was explained by the model. The unexplained variance explained by the first contrast had an eigen-value of 3.2, which is significantly higher than the chance value of 2.0 (Ref. 48 and Linacre JM, personal communication, October 2008). Similarly, the fact that the raw variance explained by the items (21%) is only three times the variance explained by the first contrast (7.3%) suggests a noticeable secondary dimension in the items. The presence of several items (26, 2, 25, 22, and 23 or A, B, C, D, and E as shown on the PCA factor plot) vertically separated from the rest in Figure 3 indicates that respondents treated these items differently from the remaining items on the A&SQ. This relatively large number of items (5/23) and inherent qualities (mainly psychosocial) suggest that an important second dimension is being measured.

DISCUSSION

To our knowledge, this represents the first independent evaluation of the A&SQ instrument. Our results highlight several limitations of the A&SQ that could guide the development of this and future instruments. First, the questionnaire does not appear to be unidimensional, but instead provides two measures. One is a visual function measure and the other is a

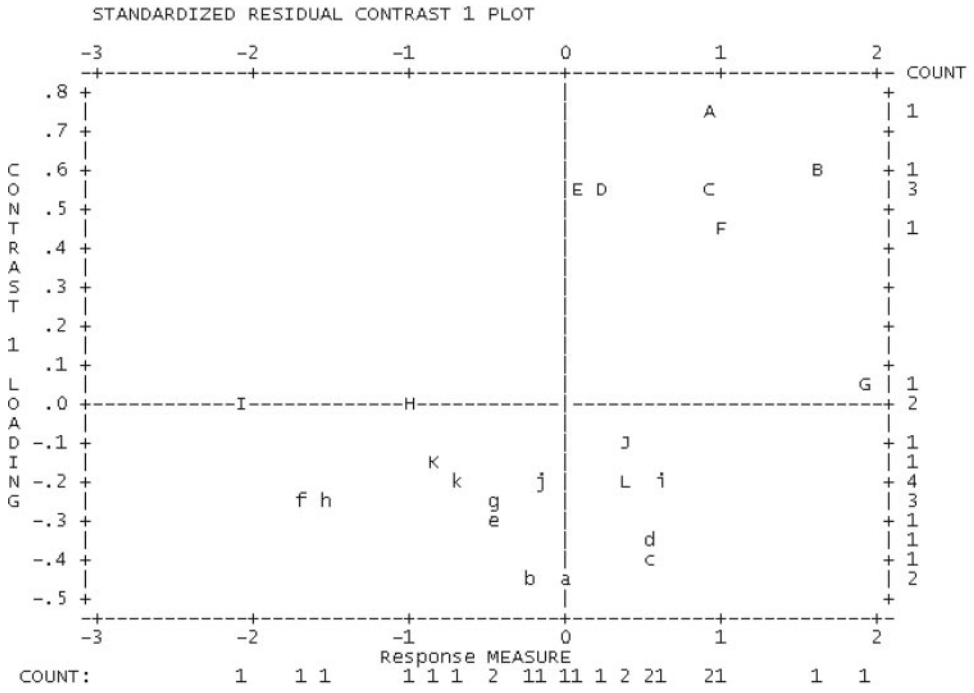

FIGURE 3. Plot of PCA of residuals. Each letter describes an item: A, 26; B, 2; C, 25; D, 22; E, 23; F, 3; G, 21; H, 6; I, 7; J, 13; K, 11; L, 8; a, 20; b, 19; c, 4; d, 5; e, 14; f, 10; g, 16; h, 9; i, 18, j, 12, and k, 15.

psychosocial measure. Scores from the items that make up these two measures should not be added together to provide a single summary score of VR-QoL, as they are measuring two very different things. Thus, we suggest that it would be more appropriate to use the A&SQ to provide two scores.

The A&SQ is also limited in its assessment of VR-QoL in nonstrabismic amblyopes. The finding that none of the nonstrabismic amblyopes reported problems relating to ocular misalignment (items 25 and 26) was not unexpected, since the cosmetic appearance of an eye with amblyopia alone does not differ from that of a normal, healthy eye. The inclusion of items within the A&SQ that are not relevant to those with amblyopia alone limits the clinical usefulness of this questionnaire in such individuals.[42] The developer's strategy to overcome this problem was to assign the highest score to items with missing or not-relevant responses.[33] This approach has been criticized because the Likert scoring method involves adding up scores from individual items to obtain an overall score and because higher overall scores suggest a better VR-QoL.[17–19] We support others' concerns about this issue. The A&SQ developers suggest that individuals who, for example, do not drive or play ball games be given the highest score of 4 for items 8 *difficulty parking the car* and 13 *difficulty playing ball games*, regardless of their level of VR-QoL, because these subjects would give a not-applicable response to these questions.

The significant DIF in the comparison of strabismic and nonstrabismic participants further highlights the inability of the A&SQ to be used as a tool appropriate for the assessment of VR-QoL in individuals with strabismus and/or amblyopia. As a result, we conducted separate analyses for participants with strabismus ($n = 69$) and for those with amblyopia only ($n = 33$). Rasch fit statistics indicated that all A&SQ items fit the Rasch model in subjects with strabismus, but several items misfitted the Rasch model in the amblyopia-only group, and three items (4, 7, and 13), plus the two items (25 and 26) that did not apply to the amblyopia-only participants, could be removed with minimal effect. Similarly, the difference in locations between the mean-person and -item logit values (represented as M on the maps) was larger in nonstrabismic amblyopes than in the strabismic population, indicating that the A&SQ was less able to capture the VR-QoL of respondents with amblyopia only. It is interesting to note that a recently developed VR-QoL instrument was designed for adults with strabismus only.[17] Our results suggest that a specific instrument is necessary to evaluate VR-QoL in nonstrabismic amblyopia.

A limitation of our study is the relatively small sample size for Rasch analysis of the amblyopia-only subgroup ($n = 33$). For this reason, we did not draw any specific conclusions (such as which items could usefully be removed from the questionnaire) based solely on the Rasch analysis of these data. The relatively strong sex bias in our sample combined with the 60% response rate may also mean that the VR-QoL of our sample differs from the VR-QoL of the amblyopia/strabismus population at large. However, DIF analysis showed no effect of the participants' sex in our results.

Future instruments to assess VR-QoL in strabismus and/or amblyopia would ideally include items that better target the visual impact and concerns of individuals with these conditions. The Rasch person-item maps (Fig. 2) show an uneven spread of items across the full range of participant's scores, suggesting poor targeting of the A&SQ items for both the strabismic and amblyopia-only populations. A ceiling effect can be seen in Figure 2, which indicates that many subjects had an excellent VR-QoL (showed by the many X's at the top of the person-item maps), but that few items assessed VR-QoL at this level (shown by the lack of items at the top of the maps). This result indicates that many participants did not have problems, even with the most difficult questionnaire items (i.e., those highest on the person-item maps in Fig. 2). This large ceiling effect cannot be overcome by an alteration to the instrument and requires the addition of items that better target the impact of amblyopia and/or strabismus on those who live with the conditions. Hatt et al.[19] have recently conducted in-depth interviews to identify specific problems affecting VR-QoL in adults with strabismus. A wide range of quality-of-life concerns were found, including adjustments made to cope with strabismus (e.g., *saying people's names so they know I'm talking to them*) and efforts to reduce symptoms (e.g., *trying to straighten eye*), which are not included in the A&SQ. Similarly, a patient-centered approach using interviews or focus groups in adults with amblyopia alone may help establish difficulties

that are not addressed by the A&SQ and that could lead to the development of an instrument in the same way that Hatt et al. developed their questionnaire for evaluating VR-QoL in adults with strabismus.[17–19] In contrast, the A&SQ appears to have been developed with limited patient input, and the fact that the pilot questionnaire and final A&SQ contained the same number of items suggests that item reduction was not part of the development of the questionnaire.[33]

In summary, our findings highlight the limitations of the A&SQ instrument in the assessment of VR-QoL in subjects with strabismus and especially in those with amblyopia alone. Many of the tasks asked about in the A&SQ are not perceived as being in any way difficult for most of the subjects. We suggest that separate instruments are needed to quantify VR-QoL in individuals with strabismus and nonstrabismic amblyopia. A newly designed questionnaire for adults with strabismus appears to have addressed many of the problems of that group,[17–19] but there remains a need to identify the concerns and perceptions of those with amblyopia alone, to enable the impact of amblyopia and strabismus to be separated. The development of appropriate condition-specific instruments will allow a comprehensive understanding of the functional and psychosocial consequences of living with these common conditions.

Acknowledgments

The authors thank Mike Linacre and Estibaliz Garamendi for help with various aspects of Rasch analysis and Alison Bruce, Ian Simmons, Janice Hoole, Eve Panesar, and Niall O'Kane for help during data collection.

References

1. Stifter E, Burggasser G, Hirmann E, Thaler A, Radner W. Monocular and binocular reading performance in children with microstrabismic amblyopia. *Br J Ophthalmol.* 2005;89:1324–1329.
2. Grant S, Melmoth DR, Morgan MJ, Finlay AL. Prehension deficits in amblyopia. *Invest Ophthalmol Vis Sci.* 2007;48:1139–1148.
3. Melmoth D, Finlay AL, Morgan M, Grant S. Grasping deficits and adaptations in adults with stereo vision losses. *Invest Ophthalmol Vis Sci.* 2009;50:3711–3720.
4. Webber AL, Wood JM, Gole GA, Brown B. The effect of amblyopia on fine motor skills in children. *Invest Ophthalmol Vis Sci.* 2008; 49:594–603.
5. Buckley JG, Panesar GK, MacLellan MJ, Pacey IE, Barrett BT. Changes to control of adaptive gait in individuals with long-standing reduced stereoacuity. *Invest Ophthalmol Vis Sci.* 2010;51: 2487–2495.
6. Bauer A, Dietz K, Kolling G, Hart W, Schiefer U. The relevance of stereopsis for motorists: a pilot study. *Graefes Arch Clin Exp Ophthalmol.* 2001;239:400–406.
7. Tijtgat P, Mazyn L, De Laet C, Lenoir M. The contribution of stereo vision to the control of braking. *Accident Anal Prevent.* 2008;40: 719–724.
8. Rahi JS. Does amblyopia affect educational, health, and social outcomes? Findings from a 1958 British birth cohort. *BMJ.* 2006; 332:820–825.
9. Chua B, Mitchell P. Consequences of amblyopia on education, occupation, and long term vision loss. *Br J Ophthalmol.* 2004;88: 1119–1121.
10. Coats DK, Paysse EA, Towler AJ, Dipboye RL. Impact of large angle horizontal strabismus on ability to obtain employment. *Ophthalmology.* 2000;107:402–405.
11. Goff MJ, Suhr AW, Ward JA, Croley JK, O'Hara MA. Effect of adult strabismus on ratings of official U.S. Army photographs. *J AAPOS.* 2006;10:400–403.
12. Mojon-Azzi SM, Mojon DS. Opinion of headhunters about the ability of strabismic subjects to obtain employment. *Ophthalmologica.* 2007;221:430–433.
13. Mojon-Azzi SM, Mojon DS. Strabismus and employment: the opinion of headhunters. *Acta Ophthalmol.* 2009;87:784–788.
14. Mojon-Azzi SM, Potnik W, Mojon DS. Opinions of dating agents about strabismic subjects' ability to find a partner. *Br J Ophthalmol.* 2008;92:765–769.
15. Sabri K, Knapp CM, Thompson JR, Gottlob I. The VF-14 and psychological impact of amblyopia and strabismus. *Invest Ophthalmol Vis Sci.* 2006;47:4386–4392.
16. van de Graaf ES, van der Sterre GW, van Kempen-du Saar H, Simonsz B, Looman CW, Simonsz HJ. Amblyopia and strabismus questionnaire (A&SQ): clinical validation in a historic cohort. *Graefes Arch Clin Exp Ophthalmol.* 2007;245:1589–1595.
17. Hatt SR, Leske DA, Bradley EA, Cole SR, Holmes JM. Development of a quality-of-life questionnaire for adults with strabismus. *Ophthalmology.* 2009;116:139–144.
18. Hatt SR, Leske DA, Bradley EA, Cole SR, Holmes JM. Comparison of quality-of-life instruments in adults with strabismus. *Am J Ophthalmol.* 2009;148:558–562.
19. Hatt SR, Leske DA, Kirgis PA, Bradley EA, Holmes JM. The effects of strabismus on quality of life in adults. *Am J Ophthalmol.* 2007; 144:643–647.
20. Hrisos S, Clarke MP, Wright CM. The emotional impact of amblyopia treatment in preschool children: randomized controlled trial. *Ophthalmology.* 2004;111:1550–1556.
21. Jackson S, Harrad R, Morris M, Rumsey N. The psychosocial benefits of corrective surgery for adults with strabismus. *Br J Ophthalmol.* 2006;90:883–888.
22. Menon V, Saha J, Tandon R, Mehta M, Khokhar S. Study of the psychosocial aspects of strabismus. *J Pediatr Ophthalmol Strabismus.* 2002;39:203–208.
23. Nelson BA, Gunton KB, Lasker JN, Nelson LB, Drohan LA. The psychosocial aspects of strabismus in teenagers and adults and the impact of surgical correction. *J AAPOS.* 2008;12:72–76.
24. Olitsky SE, Sudesh S, Graziano A, Hamblen J, Brooks SE, Shaha SH. The negative psychosocial impact of strabismus in adults. *J AAPOS.* 1999;4:209–211.
25. Satterfield D, Keltner JL, Morrison TL. Psychosocial aspects of strabismus study. *Arch Ophthalmol.* 1993;111:1100–1105.
26. Tolchin JG, Lederman ME. Congenital (infantile) esotropia: psychiatric aspects. *J Pediatr Ophthalmol Strabismus.* 1978;15:160–163.
27. Uretmen O, Egrilmez S, Kose S, Pamukçu K, Akkin C, Palamar M. Negative social bias against children with strabismus. *Acta Ophthalmol Scand.* 2003;81:138–142.
28. Webber AL, Wood JM, Gole GA, Brown B. Effect of amblyopia on self-esteem in children. *Optom Vis Sci.* 2008;85:1074–1081.
29. Burke JP, Leach CM, Davis H. Psychosocial implications of strabismus surgery in adults. *J Pediatr Ophthalmol Strabismus.* 1997; 34:159–164.
30. Akay AP, Cakoloz B, Berk AT, Pasa E. Psychosocial aspects of mothers of children with strabismus. *J AAPOS.* 2005;9:268–273.
31. Packwood EA, Cruz OA, Rychwalski PJ, Keech RV. The psychosocial effects of amblyopia study. *J AAPOS.* 1999;3:15–17.
32. Koklanis K, Abel LA, Aroni R. Psychosocial impact of amblyopia and its treatment: a multidisciplinary study. *Clin Exp Ophthalmol.* 2006;34:743–750.
33. van de Graaf ES, van der Sterre GW, Polling JR, van Kempen H, Simonsz B, Simonsz HJ. Amblyopia & Strabismus Questionnaire: design and initial validation. *Strabismus.* 2004;12:181–193.
34. Felius J, Beauchamp GR, Stager DR Sr, Van De Graaf ES, Simonsz HJ. The amblyopia and strabismus questionnaire: English translation, validation, and subscales. *Am J Ophthalmol.* 2007;143:305–310.
35. van de Graaf ES, Felius J, van Kempen-du Saar H, et al. Construct validation of the amblyopia and strabismus questionnaire (A&SQ) by factor analysis. *Graefes Arch Clin Exp Ophthalmol.* 2009;247: 1263–1268.
36. Flom MC, Bedell HE. Identifying amblyopia using associated conditions, acuity, and nonacuity features. *Am J Optom Physiol Opt.* 1985;62:153–160.

37. Attebo K, Mitchell P, Cumming R, Smith W, Jolly N, Sparkes R. Prevalence and causes of amblyopia in an adult population. *Ophthalmology.* 1998;105:154–159.
38. Von Noorden GK. *Binocular Vision and Ocular Motility*: *Theory and Management of Strabismus.* London: Mosby; 2002.
39. Pesudovs K, Burr JM, Harley C, Elliott DB. The development, assessment and selection of questionnaires. *Optom Vis Sci.* 2007; 84:664–675.
40. 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.
41. 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.
42. Pesudovs K, Garamendi E, Elliott DB. The Quality of Life Impact of Refractive Correction (QIRC) Questionnaire: development and validation. *Optom Vis Sci*. 2004;81:769–777.
43. Pesudovs K, Garamendi E, Elliott DB. A quality of life comparison of people wearing spectacles or contact lenses or having undergone refractive surgery *J Refract Surg*. 2006;22:19–27.
44. Garamendi E, Pesudovs K, Stevens MJ, Elliott DB. The Refractive Status and Vision Profile: evaluation of psychometric properties and comparison of Rasch and summated Likert-scaling. *Vision Res.* 2006;46:1375–1383.
45. Bond TG, Fox CM. *Applying the Rasch Model*: *Fundamental Measurement in the Human Sciences*. London: Lawrence Erlbaum Associates; 2001.
46. Massof RW. The measurement of vision disability. *Optom Vis Sci.* 2002;79:516–552.
47. Linacre JM. WINSTEPS Rasch measurement (computer program). Chicago: Winstepscom; 2008.
48. Smith EV. Detecting and evaluating the impact of multidimensionality using item fit statistics and principal component analysis of residuals. *J Appl Meas.* 2002;3:205–231.
49. Linacre JM, Wright, BD. Reasonable mean-square fit values. *Rasch Meas Trans*. 1994;8 370.