tvst

Article

Validation of the Symptom Questionnaire for Visual Dysfunctions (SQVD): A Questionnaire to Evaluate Symptoms of any Type of Visual Dysfunctions

Pilar Cacho-Martínez[1], Mario Cantó-Cerdán[1], Francisco Lara-Lacárcel[2], and Ángel García-Muñoz[1]

[1] Departamento de Óptica, Farmacología y Anatomía, Universidad de Alicante, Alicante, Spain
[2] Departamento de Oftalmología, Optometría, Otorrinolaringología y Anatomía Patológica, Universidad de Murcia, Murcia, Spain

Correspondence: Pilar Cacho-Martínez, Departamento de Óptica, Farmacología y Anatomía, Apartado 99, Universidad de Alicante, 03080 Alicante, Spain.
e-mail: cacho@ua.es

Received: July 26, 2021
Accepted: January 5, 2022
Published: February 3, 2022

Keywords: visual symptomatology; questionnaires; visual dysfunctions; Rasch analysis; refractive, accommodative, and binocular dysfunctions

Citation: Cacho-Martínez P, Cantó-Cerdán M, Lara-Lacárcel F, García-Muñoz Á. Validation of the symptom questionnaire for visual dysfunctions (SQVD): A questionnaire to evaluate symptoms of any type of visual dysfunctions. Transl Vis Sci Technol. 2022;11(2):7, https://doi.org/10.1167/tvst.11.2.7

Purpose: To assess psychometric properties of the Symptom Questionnaire for Visual Dysfunctions (SQVD) questionnaire, including accuracy, validity, and reliability, in a clinical sample of patients having any type of visual dysfunction.

Methods: A clinical sample of 306 patients self-administered the SQVD. Rasch analysis was performed to analyze the functionality of the response categories, fit statistics, differential item functioning (DIF), person and item reliability, targeting, local dependency, unidimensionality, and transformation table. Accuracy was assessed by means of receiver operating characteristic (ROC) curves, using symptoms reported in each patient's clinical record as the gold standard for classifying patients with and without symptoms. The concurrent validity, known group validity, and test–retest reliability (repeatability, using the intraclass correlation coefficient [ICC]) were also examined.

Results: SQVD showed orderly category responses. The 14 items fit the Rasch model without significant DIF for gender, presbyopia, and dysfunctions. Person and item reliabilities were 0.81 and 0.85, respectively. Targeting was −1.49 logits. Yen's Q3 statistic showed no local dependency. SQVD was unidimensional (first contrast of the residual = 1.852 eigenvalue with a variance explained by measures of 52.23%). The area under the ROC curve was 0.836 (95% confidence interval [CI], 0.792–0.879) with a cutoff of $\geq$6 showing good accuracy (sensitivity = 0.759; specificity = 0.783). SQVD showed good concurrent and known group validity and high repeatability (ICC, 0.857; 95% CI, 0.710–0.933) when administered twice 1 week apart.

Conclusions: SQVD has shown good psychometric properties. It can be considered an accurate, valid, and reliable questionnaire to detect visual symptoms related to any type of refractive, accommodative, and binocular dysfunction.

Translational Relevance: SQVD may be used for diagnostic purposes, as it can accurately detect symptoms related to any sort of visual dysfunction. It may also be useful to monitor the treatment outcomes of these conditions.

Introduction

Visual dysfunctions (refractive, accommodative, and binocular) are common in clinical practice among patients undergoing visual examinations. These anomalies may cause symptoms that can impact the patient's comfort level when performing visual tasks.[1] The scientific literature has shown that these visual dysfunctions are related to various symptoms that may be shared by different visual anomalies.[2,3] However, it has also been shown that there are differences in the way information is collected and in the methodology that is followed to classify anomaly-related symptoms.[2,3] The most common way to record the symptoms reported by a patient undergoing a visual examination is the patient's clinical record, although several questionnaires related to visual symptoms have

been developed.[4–15] However, scientific evidence has shown that only three of them have been psychometrically validated.[6–9] The Conlon survey was developed to detect visual discomfort, but it is not related to any specific visual dysfunction,[7] whereas the Convergence Insufficiency Symptom Survey (CISS) and its version for parents are both specific for convergence insufficiency.[6,8,9]

Additional questionnaires have also been developed focusing on other visual conditions. Thus, in the literature, we can find scales for patients with low vision,[16] cataracts,[17,18] keratoconus,[19] or vertical deviations[20]; for computer users[21,22]; and for contact lens wearers,[23] as well as instruments assessing quality of vision[24] and quality of life related to various aspects of vision.[25–28] However, to date, there have been no questionnaires available to capture visual symptoms related to any type of visual dysfunction (refractive, accommodative, or binocular). For this reason, we developed a questionnaire on visual symptomatology referred to as the Symptom Questionnaire for Visual Dysfunctions (SQVD). This process included an initial step in which we compiled from the scientific evidence[3] a list of symptoms that could be triggered by these anomalies. Following that systematic review,[3] which resulted in a comprehensive list of symptoms related to any type of visual anomalies, we considered including those symptoms in a questionnaire using a Delphi methodology.[29] The results of this Delphi study led to the design of the initial scale, which was tested in a small patient sample (including comprehensive patient consultations[30]) and was used to develop the pilot version of SQVD, which consisted of 33 items. It was tested in a clinical sample of 125 patients and assessed by means of Rasch analysis.[31] The outcomes of the Rasch analysis showed a reduction in the number of items, which resulted in the final version of SQVD having 14 items to detect the presence and frequency of visual symptoms related to visual anomalies.

As it was necessary to analyze and confirm the appropriateness of this questionnaire in a larger sample, the purpose of the present study was to validate the SQVD in a clinical sample of patients with any type of visual anomalies. The aim was to assess psychometric properties of the SQVD, including accuracy, validity, and test–retest reliability (repeatability).

Methods

Subjects and Procedure

Patients 14 years of age and older consecutively attending visual examinations in a private optometric clinic were selected as potential participants in this study; they were later recruited if they met all of the inclusion and exclusion criteria. Each patient underwent a complete eye examination. Ocular health and refractive status were evaluated, and accommodative and binocular tests were performed in order to diagnose any type of visual anomaly (be it refractive, accommodative, or binocular).[32] The presence of symptoms was confirmed by the optometrist assessment carried out during the case history. The diagnosis of each visual dysfunction relied on the criteria of García-Muñoz et al.[32] When the visual examination was completed, those subjects with any type of ocular disease, a history of refractive surgery, or dry eye or who were taking medication that could alter visual function were excluded from the study.

As a result of the abovementioned process, 306 patients between the ages of 14 and 87 years (mean age, 38.38 ± 17.29 years) were included in the study. All of the participants were Spanish, and 184 of them were women (60.1%). Among the participants, 204 patients had some type of visual dysfunction (66.7%): 132 cases of refractive, 14 accommodative (one accommodative insufficiency and 13 accommodative excess), 50 binocular (17 convergence insufficiency, 10 convergence excess, four insufficiency divergence, five basic esophoria, four basic exophoria, one hyperphoria, three amblyopia, and six strabismus), and eight with both accommodative and binocular anomalies (four accommodative excess plus convergence insufficiency; one accommodative excess plus convergence excess; one accommodative excess plus divergence insufficiency; two accommodative insufficiency plus convergence insufficiency). Consequently, 102 subjects in the sample did not show any anomaly (33.3%).

The study was approved by the University of Alicante's Ethics Committee and followed the tenets of the Declaration of Helsinki. Informed consent was obtained from all subjects after giving then an explanation of the nature of the study. For underaged participants (i.e., under 18 years of age), it was their parents or legal guardians who accepted the study's participation principles and signed the informed consent; the participants also gave their assent to participate.

The SQVD was completed by all 306 subjects included in the sample. It was self-administered. Before they undertook their visual examination, patients were given a printed copy of the questionnaire (containing the list of questions and answer options) and were accompanied by a researcher in case they had any questions while answering the questionnaire. This researcher was a different person than the examiner who performed the visual examination, to avoid bias in the study.

The SQVD is shown as Supplementary Material. The tool is presented in its original validated language (Spanish). We have also included its translation into English for non-Spanish-speaking readers, but it must be kept in mind that this is not a cross-cultural adaptation for the English language: the tool in English would also require further cross-validation. The questionnaire has 14 items addressing blurred vision, binocular problems, ocular irritation, headache, concentration difficulties, reading problems, and postural aspects. Each item has three response options (Likert scale) to indicate the frequency of the symptom (propensity of its presence):

- No: the symptom never occurs (0 point).
- Occasionally or often: the symptom occurs sporadically (at least once every 15 days) or once or twice a week (1 point).
- Almost always: the symptom occurs almost every day (2 points).

This explanation is also included in the printed questionnaire. As shown above, the answer for each item is assigned a score between 0 and 2 points, and the total SQVD score is then obtained by adding the 14 individual item scores; thus, scores can range from 0 to 28.

Data Analysis

Rasch analysis was performed and included all parameters that the scientific evidence has shown to be essential when developing and validating a questionnaire.[33] Rasch analysis is a probabilistic model that estimates the difficulty of items (item difficulty) and the relative abilities of the respondents (person ability) and aligns these two parameters in an invariant interval-level scale. It transforms simple ordinal categorical data into interval-level data on a linear logit scale.[25,34,35] The Andrich Rating Scale Model for Rasch analysis was employed[36] using Winsteps 4.8.1. The parameters analyzed included response category functioning, fit statistics, differential item functioning, person reliability, item reliability, targeting, local item dependency, unidimensionality, and transformation table. In addition to the Rasch analysis, a receiver operating characteristic (ROC) analysis was carried out to assess the accuracy of the questionnaire and to obtain the cutoff value for the instrument. The concurrent validity and known group validity were also evaluated; moreover, the repeatability was calculated based on the intraclass correlation coefficient (ICC).

Response Category Functioning

Response options were analyzed by means of category probability curves[37] to test whether the response categories were ordered. These curves show the likelihood for a subject to choose a given category. Because the threshold is the midpoint between adjacent response categories, it represents the point where the likelihood of choosing either response category is the same.[38] To assess whether the categories are adequately ordered, the scientific literature recommends calculating Andrich thresholds, which should be spaced at least 1.4 logits apart.[37] If this situation does not occur, thus showing a disordered threshold, it may be necessary to collapse adjacent categories.[33,39–41]

Fit Statistics

Rasch fit statistics (infit and outfit mean square [MNSQ]) were obtained to explore whether the data fit the Rasch model expectations.[33,35,42] Infit and outfit MNSQ values closer to 1 indicate a good fit to the model; that is, that more difficult items are less likely to be affirmed successfully (and vice versa). Values below 0.70 suggest that there may be a redundancy of items, whereas values above 1.30 could mean that items are measuring something different from the overall scale. For this reason, fit statistics values must range between 0.70 and 1.30 logits.[39]

Differential Item Functioning

Differential item functioning (DIF) analysis is used to determine whether different subgroups respond differently to particular items.[19] It is important to undertake this analysis, as the presence of DIF may impact the fit of the data to the model and may corrupt measures.[33] In this study, DIF analysis was assessed for gender, presbyopia (presbyopes vs. non-presbyopes, considering presbyopia as the need for near addition), and dysfunction (having vs. not having a visual dysfunction). It had been established that mean differences in person measures between groups should be less than 1.0 logit.[25] For values above 1.0 logit, a notable DIF must be considered.[18] Following these criteria, in this study, DIF was considered to be present when the findings showed a statistically significant ($P < 0.05$) DIF contrast and a difference above 1.0 logit.[19] The Rasch–Welch t-test method was then used to establish the significance of the DIF contrast.

Person and Item Reliability

To test the overall performance of the instrument, we assessed the SQVD person and item reliability (separation index).[33,39] Person and item reliability determines the replicability of the person and item locations along the trait continuum.[19] The values can

range between 0 and 1, with higher values indicating better reliability. A person reliability value > 0.80 is considered to be acceptable (person separation index > 2 logits) and implies that the measure can stratify the population into at least three groups.[34] Accordingly, because the person separation index is used to classify people, a low person separation value suggests that the instrument may not be sensitive enough to discriminate between high and low performance, thus indicating that more items are needed. For item reliability, a value above 0.90 (item separation index > 3 logits) is considered appropriate.[39,43] In this case, the item separation index is used to evaluate the item hierarchy, such that a low item separation implies that the sample is not large enough to prove the item difficulty hierarchy of the scale.

Targeting

Targeting refers to the difference between the person ability mean and the item difficulty mean. It can be assessed with Rasch analysis exploring the person–item map,[44] which depicts the spread and hierarchy of subjects and items; the closer the person ability mean is to the item difficulty mean, the better the targeting. A difference of zero between both values is considered perfect targeting of the scale, whereas a difference greater than 1 logit is considered mistargeting.[38]

Local Dependency

Local dependency determines whether the response to any item has a direct influence on the response to any other item,[45] and for that reason local independence of items is a requirement of the model. Yen's Q3 statistic was used to detect local dependence by means of the residual correlation matrix. Christensen et al.[45] showed that no singular critical value can be appropriate to indicate dependency; however, simulations have proved that the Q3 critical value appears to be reasonably stable around a value of 0.2 above the average correlation. Consequently, this is the recommended reference to use, and any residual correlation > 0.2 above the average correlation may be considered to indicate local dependency.

Unidimensionality

Unidimensionality refers to the assumption that the questionnaire measures a single construct; that is, that the items summed together form a unidimensional scale.[33] Principal component analysis (PCA) of the residuals was used to explore this property. The residuals were considered to be the differences between observed data and model-derived estimates. It has been suggested in the literature that the first contrast of the residual, should not be above 2 eigenvalue, and the variance explained by the measures must be >50%.[33]

Transformation Table

Using Rasch analysis, it is possible to transform the ordinal scores of a questionnaire into an interval scale without modifying the original responses of the instrument.[33] In the present study, this conversion table was obtained by considering the raw scores of the ordinal measures, and then the corresponding interval-level scores in logits and ordinal scale were obtained. This transformation table helps other authors use the questionnaire and obtain Rasch scores, although the patient populations should be similar to that included in this study.

ROC Analysis

The accuracy of SQVD was assessed by means of ROC curves, sensitivity (S) and specificity (Sp).[46] ROC curves depict the true positive rate (S) versus the false positive rate (1 − Sp) over a range of cutoff values. The overall accuracy of a test can be quantified by means of the area under the ROC curve; thus, the larger the area, the better the test.[46] In order to analyze the diagnostic validity of the SQVD, a ROC analysis was performed using the symptoms reported by the subjects in their case history as the gold standard and subsequent patient classification as subjects with symptoms and asymptomatic subjects. The ROC analysis relied on the original raw score of the SQVD for both groups of patients (i.e., symptomatic and asymptomatic) and was used to choose the area under the ROC curve and the coordinates of the curve. The cutoff point was chosen by means of a balance between S and Sp.[46] This cutoff is necessary to take into account when a patient passes or fails the SQVD.

Validity

The concurrent validity and known group validity were evaluated. The concurrent validity represents the correlation level between the questionnaire score and the score of clinical measures. In the present study, the SQVD score was correlated with visual acuity based on the Spearman's rho coefficient value. A correlation between 0.3 and 0.9 is considered to be adequate.[47] As for the known group validity, which is the extent to which the instrument can discriminate between clinically different groups, it was assessed by analyzing the statistically significant differences found between the two groups of patients: those with visual anomalies and those that did not show any visual dysfunction. The Mann–Whitney *U* test was used with a significance level of 0.05.

Test–Retest Reliability

The test–retest reliability refers to how repeatable the results are when the instrument is administered by the same observer.[39] It was quantified by means of the ICC[48] and corresponding 95% confidence intervals. The literature recommends an ICC value $\geq$ 0.80 to demonstrate temporal stability when administered in two different periods.[39,47,48] Considering an expected ICC of 0.85, two measurements, a lowest acceptable ICC of 0.7, a significance level of 0.05, and a power of 80%, the required sample size turned out to be 42 patients.[49] Thus, in order to assess the SQVD repeatability, 50 randomly chosen participants were administered the questionnaire a second time, 1 week after the first.[50] ROC curves, validity, and test–retest reliability were performed using SPSS Statistics 20.0 for Windows (IBM, Armonk, NY).

Results

The category probability curves depicted in Figure 1 show good category discrimination and orderly category responses. The Andrich threshold revealed separations greater than 1.4 logits between adjacent categories (see Table 1), which indicates ordered thresholds and implies that each category response had equal probability to be endorsed by the patients.

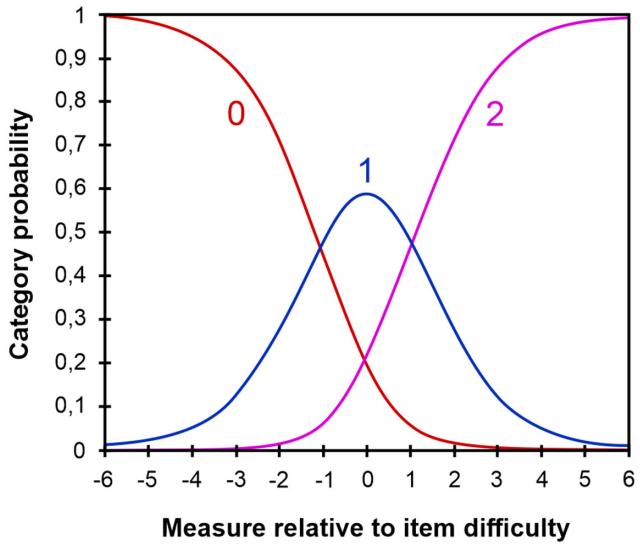

Figure 1. Category probability curves (CPCs) for the SQVD. Each curve in the CPC graph represents one response category (no = 0; occasionally/often = 1; almost always = 2). The figure shows the performance of the three response categories of the SQVD, which asked about the frequency of each of the symptoms under evaluation. The point where two adjacent curves overlap is the threshold, and the two corresponding categories have equal likelihood of being chosen.

Table 1. SQVD Categories and Andrich Threshold Values (Logits)

SQVD Category	Andrich Threshold
0	None
1	−1.34
2	1.34

Table 2 summarizes the infit and outfit MNSQ values for each of the 14 items included in the SQVD. Fit statistics revealed that all 14 items fitted the Rasch model, as they fell within the range values suggested. DIF contrast results are also shown in Table 2. Values greater than 1 logit reveal the presence of DIF. As can be seen in the table, there were no statistically significant DIF differences for gender, dysfunction, or presbyopia (<1 logit) except for items 1 and 2 (related to headache and blurred near vision) which showed DIF for presbyopia. Table 3 shows the overall infit and outfit MNSQ statistics for the SQVD, which indicate that mean infit and outfit (for both subjects and items) fit the Rasch model. Reliability values (person and idem reliability) are also shown in Table 3; the person reliability value of 0.81 showed a good person separation index.

With regard to targeting, Figure 2 shows the person–item map for SQVD. Person ability (in logits) is illustrated in the left-hand column; patients with higher ability are shown at the top of the figure. For this patient sample, mean person ability was –1.49 logits (shown in the figure as the left M). The right-hand column displays item difficulty, for which the mean is always 0 (shown as the right M). This poor targeting (–1.49 logits) means that an important number of patients are located at the bottom of the figure, indicating a floor effect. The person–item map reveals that items were targeting the more symptomatic patients; therefore, patients having fewer visual symptoms (especially those at the floor) could not be properly differentiated by the SQVD items.

Yen's Q3 statistic confirmed that none of the 14 items included in the SQVD showed local dependency, as the residual correlation values were always <0.2 above the average correlation. Furthermore, PCA of the residuals confirmed the unidimensionality of the SQVD, as the magnitude of the first contrast of the residual was 1.852 eigenvalue and the variance explained by the measures was 52.23%.

The transformation table results are provided in Table 4, which shows conversion of the ordinal scale (0 to 28) into the corresponding interval-level scores

Table 2. Rasch Analysis Results for Each SQVD Item

Item	Infit MNSQ	Outfit MNSQ	Gender DIF Contrast	Presbyopia DIF Contrast	Visual Dysfunction DIF Contrast
1	0.93	0.94	0.25	1.15[a]	0.20
2	1.07	1.06	0.02	1.10[a]	0.26
3	0.93	0.95	0.00	0.09	0.26
4	1.25	1.25	0.10	0.49	0.00
5	0.89	0.98	0.10	0.14	0.35
6	1.00	0.93	0.15	0.48	0.13
7	0.90	0.91	0.00	0.24	0.45
8	0.90	0.87	0.47	0.21	0.36
9	1.14	1.10	0.58	0.92	0.13
10	1.04	1.05	0.35	0.08	0.51
11	1.05	1.14	0.10	0.00	0.10
12	1.15	1.12	0.52	0.55	0.00
13	0.73	0.70	0.24	0.60	0.47
14	0.96	1.03	0.14	0.80	0.31

[a]Statistically significant ($P < 0.05$).

Table 3. Summary of Global Fit Statistics for SQVD Persons and Items, and Reliability Parameters

Persons		Items		Reliability (Separation Index)	
Infit MNSQ	Outfit MNSQ	Infit MNSQ	Outfit MNSQ	Person	Item
1.01	1.00	1.00	1.00	0.81 (2.11)	0.85 (2.41)

(logits) and their subsequent rescaling into the ordinal scale range (0–28) of the tool.

Figure 3 depicts the ROC curves, with an area under the curve of 0.836 ($P < 0.001$; 95% CI, 0.792–0.879). A score cutoff of ≥6 had the best balance for S and Sp (0.759 and 0.783, respectively), which means that patient scores ≥ 6 indicate the presence of visual symptoms.

As for concurrent validity, there was a significant correlation (correlation coefficient = 0.246; $P < 0.001$) between logMAR visual acuity and the score obtained in the questionnaire. With regard to known group validity, the mean SQVD score for subjects with some type of visual dysfunction was 8.41 ± 4.25 versus 3.82 ± 2.19 for subjects with no visual dysfunctions, indicating statistically significant differences ($P < 0.001$) between the two groups. If we break down the data by dysfunction type, the mean score for subjects with refractive dysfunctions was 8.61 ± 4.36, whereas for subjects with accommodative and binocular anomalies it was 8.06 ± 4.07. There were statistically significant differences between each of these two subgroups and the group of patients with no visual anomalies ($P < 0.001$).

Furthermore, as for test–retest reliability, the ICC was 0.857 ($P < 0.001$; 95% CI, 0.710–0.933).

Discussion

The outcomes of this study confirm that the SQVD has good psychometric properties and is an accurate, valid, and repeatable questionnaire to detect symptoms in patients with any type of visual dysfunction.

The tool has shown that data adequately fit the Rasch model, as all fit statistics fell within the appropriate intervals established by the scientific literature.[26,35,37,39,51] There was no local dependency for items, and the tool demonstrated unidimensionality.[33]

Furthermore, the tool performed similarly for all of the sample subgroups, because, in general, there was no DIF for gender, dysfunction, or presbyopia. The fact that item 1 (headache) and item 2 (blurred near vision) showed DIF for presbyopia indicates that these two symptoms are presbyopia related, thus suggesting a particular symptomatology for this population. Consequently, these two items should not be removed from the questionnaire as they provide information about symptoms of an important visual condition.

The scale is also reliable. The good person reliability value of the SQVD implies that the instrument is able to stratify subjects into at least three groups based on the

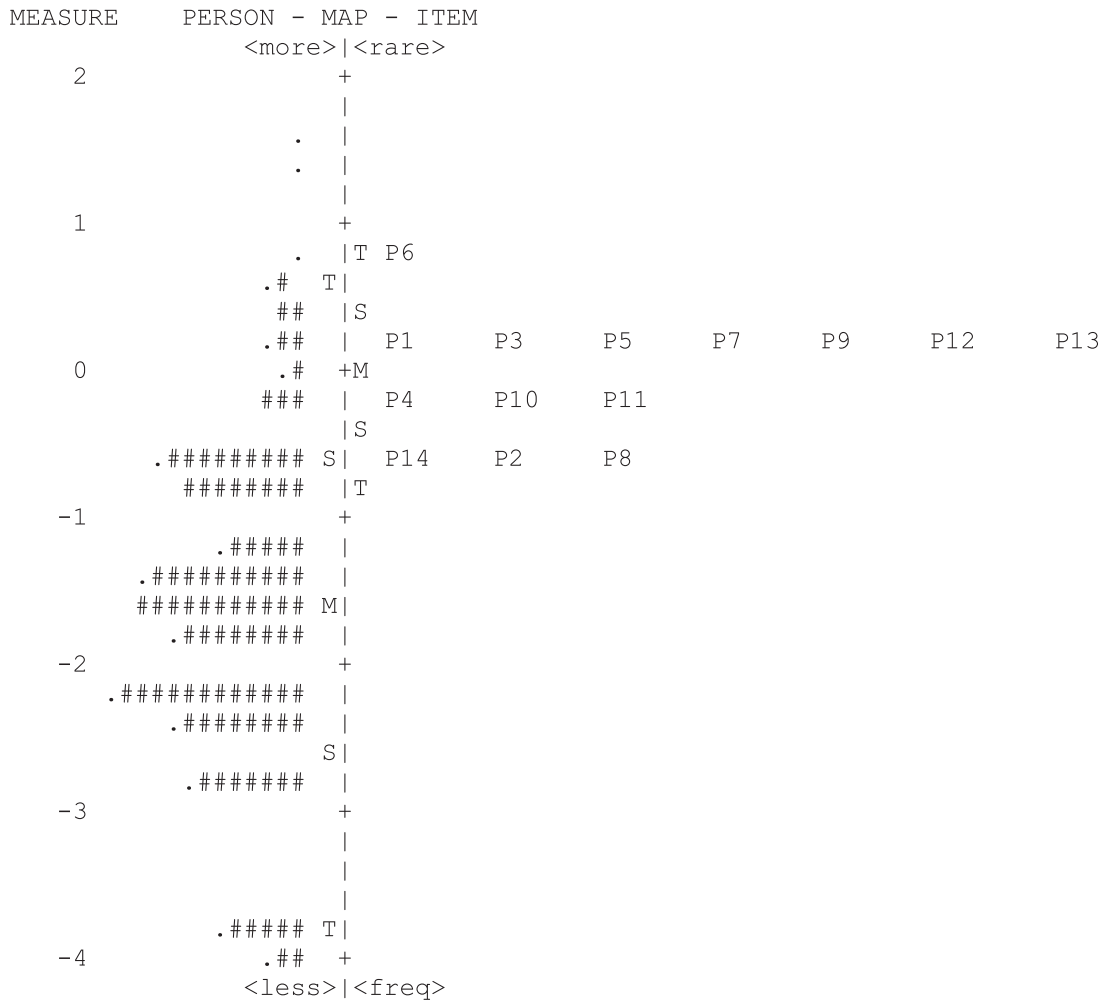

Figure 2. Person–item map for the SQVD. Patients (person ability) are represented on the left of the *dashed line*. The # symbol represents three subjects, and the • symbol indicates one or two subjects. The items of SQVD with their number (Pnumber of item) are shown on the *right* of the *dashed line* of the figure. "M" indicates the mean measure (*left*, person ability; *right*, item difficulty). "S" shows 1 SD from the mean, and "T" indicates 2 SDs, all expressed in logits. The person–item map orders the symptoms of the patients and the item difficulty. Higher ability for persons (in this study, higher frequency of symptoms) and more difficult items are at the top of the figure.

latent trait being assessed. However, because the item reliability was 0.85, a larger sample would be required to prove the item difficulty hierarchy of the scale.

The poor targeting by the SQVD, which reveals a floor effect, can be explained by the high number of patients in the sample that reported few symptoms. Similar results have been reported in other studies using different symptom instruments, as many patients say they have no symptoms at all[22,24] or underreport their discomfort.[21] Because the same scenario may occur with the SQVD, the targeting may be considered to be reasonable. Future studies could analyze larger samples made up of patients showing higher symptom levels so as to test this targeting.

The SQVD was also shown to be a valid instrument. Concurrent validity outcomes proved that there is a significant statistical correlation between visual acuity and SQVD score, which suggests that the instrument is able to detect symptoms in these patients. However, the low correlation value implies that these visual anomalies are not always related to a problem affecting visual acuity. For example, although accommodative and binocular anomalies may cause symptoms, they do not necessarily lead to a drop in visual acuity. With regard to known group validity, the instrument has proven to be able to differentiate symptoms of subjects with and without dysfunctions. These outcomes also imply that the SQVD has good validity.

Rasch analysis has also allowed development of a conversion table that transforms ordinal data into interval-level data.[33] With this transformation table, users may improve the precision of the SQVD. For example, a patient with an original ordinal score of 15 will have a corresponding interval score of 14.49,

Table 4. Conversion of Raw SQVD Scores (0–28) to Interval Scale (Logit Units) Using the Original Scale Metrics

	Interval Measure	
Ordinal Measure, Raw Score	Logit	Scale
0	−4.91	0.00
1	−3.66	3.55
2	−2.91	5.70
3	−2.44	7.04
4	−2.09	8.05
5	−1.80	8.88
6	−1.54	9.60
7	−1.31	10.25
8	−1.10	10.85
9	−0.90	11.42
10	−0.71	11.96
11	−0.53	12.48
12	−0.35	12.99
13	−0.18	13.50
14	0.00	13.99
15	0.17	14.49
16	0.35	15.00
17	0.53	15.51
18	0.71	16.03
19	0.90	16.57
20	1.10	17.14
21	1.31	17.74
22	1.54	18.39
23	1.80	19.11
24	2.09	19.95
25	2.44	20.96
26	2.91	22.30
27	3.67	24.45
28	4.91	28.00

using the same scale range. By applying this conversion, clinicians can report changes in the variable under evaluation as it better allows use of parametric statistics. The only requirement for applying this conversion is that the patient must complete the entire questionnaire, providing an answer for each and every item.

The SQVD also showed good diagnostic accuracy to detect symptoms related to any type of visual dysfunction. The ROC analysis cutoff value indicates that a score ≥ 6 suggests that the subject's visual symptoms are related to some type of visual dysfunction. Finally, this instrument has also shown high repeatability, which is an important property, for example, when evaluating changes in symptoms when considering a specific treatment.

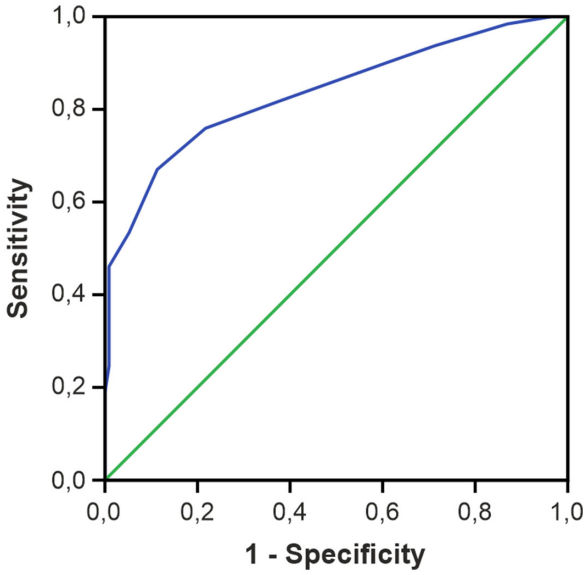

Figure 3. ROC curve for the SQVD questionnaire. The ROC curve summarizes the results of sensitivity and specificity analyses for various cutoff points. The greater the area under the curve, the better the accuracy of the SQVD to detect symptoms.

All of these findings highlight the contributions of this study, as this new scale provides a clinical tool to be used in clinical practice for the exploration of visual anomalies. To the best of our knowledge, no other similar questionnaire can be compared with the SQVD. There is a novel scale for schoolchildren with nonstrabismic binocular vision anomalies,[27] but it is designed to quantify quality of life. Even though it has several items that ask about visual symptoms, the goal of the tool is to measure quality of life in a pediatric population, thus comparison with the SQVD would be difficult. The Conlon questionnaire[7] was developed to quantify visual discomfort, but its description of visual discomfort is neither related nor comparable to symptoms associated with visual dysfunction. Items in the Conlon survey include several symptoms within the same question, which makes it difficult to compare this instrument with other scales. In any case, it does have several questions that are similar to some SQVD items, such as those about headache, blurred vision, ocular irritation, and reading problems. Again, when comparing the SQVD with the CISS survey,[9] symptoms related to headache, blurred vision, red eyes, the need to reread the text, and sleepy feeling are common to both questionnaires. This could suggest that the CISS could be used to evaluate symptoms due to any type of visual anomaly; however, it was developed for convergence insufficiency only. Although a cross-cultural adaptation for the Spanish language has been made using the Rasch method,[52] the original questionnaire[9] was not validated using this method. Several

authors have used the CISS for other visual dysfunctions,[53–56] but this is not an appropriate approach because a questionnaire should only be used for the specific condition for which it was developed and validated.[57,58] In fact, other studies in the literature have reported difficulties when applying the CISS to other visual conditions, different from convergence insufficiency.[59–61]

Such difficulties with existing questionnaires emphasize the need for a survey specifically developed and validated for particular types of visual condition (be it refractive, accommodative, or binocular), such as the SQVD.

The present study does have several limitations. The targeting and item reliability outcomes imply that larger samples of patients with higher symptom levels would be desirable. The pilot study on which the present study was based[31] analyzed a sample made up 125 subjects; this study, having a larger clinical sample of 306 subjects, has shown improved targeting and item reliability while preserving the reliability and validity of the SQVD. Another study limitation has to do with convergent validity. This property has not been tested, as there is no other questionnaire similar to the SQVD that captures the symptomatology of all visual dysfunctions. However, in future research projects, it would be interesting to test the discriminant validity and compare the SQVD with other instruments that focus on other visual conditions. Despite these limitations, the strength of this study lies in the fact that the recruited patients came from a real-world clinical practice population. For clinical purposes, an instrument that is devised to detect the presence and frequency of symptoms caused by any type of visual dysfunction should be tested in a sample having characteristics similar to those of the population in which it will be applied, which is what we did in the present study. Consequently, this tool will be useful in detecting symptoms in clinical practice, which is the main contribution of this study. The SQVD would be useful not only for diagnostic purposes, as it can accurately detect the presence of symptoms, but also for monitoring symptom severity and frequency in patients who are undergoing treatment for their visual anomalies.

In conclusion, the results of this study show that the SQVD has good psychometric properties and good diagnostic accuracy and is a valid and repeatable questionnaire to detect the presence and frequency of visual symptoms related to any type of visual dysfunction, be it refractive, accommodative, or binocular. Hence, this tool can be used in clinical practice to identify patients with symptoms that are due to visual anomalies and is also valid for research studies.

Acknowledgments

Disclosure: **P. Cacho-Martínez**, None; **M. Cantó-Cerdán**, None; **F. Lara-Lacárcel**, None; **Á. García-Muñoz**, None

References

1. Cacho-Martínez P, García-Muñoz A, Ruiz-Cantero MT. Is there any evidence for the validity of diagnostic criteria used for accommodative and nonstrabismic binocular dysfunctions? *J Optom*. 2014;7:2–21.
2. Cacho-Martínez P, Cantó-Cerdán M, Carbonell-Bonete S, García-Muñoz A. Characterization of visual symptomatology associated with refractive, accommodative, and binocular anomalies. *J Ophthalmol*. 2015;2015:895803.
3. García-Muñoz A, Carbonell-Bonete S, Cacho-Martínez P. Symptomatology associated with accommodative and binocular vision anomalies. *J Optom*. 2014;7:178–192.
4. Russell GE, Wick B. A prospective study of treatment of accommodative insufficiency. *Optom Vis Sci*. 1993;70:131–135.
5. Birnbaum MH, Soden R, Cohen AH. Efficacy of vision therapy for convergence insufficiency in an adult male population. *J Am Optom Assoc*. 1999;70:225–232.
6. Borsting E, Rouse M, De Land P. Prospective comparison of convergence insufficiency and normal binocular children on CIRS symptom surveys. Convergence Insufficiency and Reading Study (CIRS) group. *Optom Vis Sci*. 1999;76:221–228.
7. Conlon EG, Lovegrove WJ, Chekaluk E, Pattison PE. Measuring visual discomfort. *Vis Cogn*. 1999;6:637–663.
8. Borsting E, Rouse M, Mitchell G, et al. Validity and reliability of the revised Convergence Insufficiency Symptom Survey in children aged 9 to 18 years. *Optom Vis Sci*. 2003;80:832–838.
9. Rouse M, Borsting E, Mitchell G, et al. Validity and reliability of the revised Convergence Insufficiency Symptom Survey in adults. *Ophthalmic Physiol Opt*. 2004;24:384–390.
10. Adler P. Efficacy of treatment for convergence insufficiency using vision therapy. *Ophthalmic Physiol Opt*. 2002;22:565–571.
11. Gallaway M, Scheiman M, Malhotra K. The effectiveness of pencil pushups treatment for

convergence insufficiency: a pilot study. *Optom Vis Sci*. 2002;79:265–267.
12. Borsting E, Rouse M, Deland P, et al. Association of symptoms and convergence and accommodative insufficiency in school-age children. *Optometry*. 2003;74:25–34.
13. Sterner B, Gellerstedt M, Sjöström A. Accommodation and the relationship to subjective symptoms with near work for young school children. *Ophthalmic Physiol Opt*. 2006;26:148–155.
14. Vaughn W, Maples WC, Hoenes R. The association between vision quality of life and academics as measured by the College of Optometrists in Vision Development Quality of Life questionnaire. *Optometry*. 2006;77:116–123.
15. Rouse M, Borsting E, Mitchell GL, et al. Academic behaviors in children with convergence insufficiency with and without parent-reported ADHD. *Optom Vis Sci*. 2009;86:1169–1177.
16. Marella M, Pesudovs K, Keeffe JE, O'Connor PM, Rees G, Lamoureux EL. The psychometric validity of the NEI VFQ-25 for use in a low-vision population. *Invest Ophthalmol Vis Sci*. 2010;51:2878–2884.
17. Steinberg EP, Tielsch JM, Schein OD, et al. The VF-14. An index of functional impairment in patients with cataract. *Arch Ophthalmol*. 1994;112:630–638.
18. Lundstrom 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.
19. Kandel H, Pesudovs K, Ferdi A, et al. Psychometric properties of the Keratoconus Outcomes Research Questionnaire: a Save Sight Keratoconus Registry Study. *Cornea*. 2020;39:303–310.
20. Feinberg DL, Rosner MS, Rosner AJ. Validation of the Binocular Vision Dysfunction Questionnaire (BVDQ). *Otol Neurotol*. 2021;42:e66–e74.
21. González-Pérez M, Susi R, Antona B, Barrio A, González E. The Computer-Vision Symptom Scale (CVSS17): development and initial validation. *Invest Ophthalmol Vis Sci*. 2014;55:4504–4511.
22. Seguí MM, Cabrero-García J, Crespo A, Verdú J, Ronda-Pérez E. A reliable and valid questionnaire was developed to measure computer vision syndrome at the workplace. *J Clin Epidemiol*. 2015;68:662–673.
23. Arroyo-Del Arroyo C, Fernández I, López-de la Rosa A, Pinto-Fraga J, González-García MJ, López-Miguel A. Design of a questionnaire for detecting contact lens discomfort: the Contact Lens Discomfort Index [published online ahead of print March 18, 2021]. *Clin Exp Optom*, doi:10.1080/08164622.2021.1896945.
24. McAlinden C, Pesudovs K, Moore JE. The development of an instrument to measure quality of vision: the Quality of Vision (QoV) questionnaire. *Invest Ophthalmol Vis Sci*. 2010;51:5537–5545.
25. McAlinden C, Lipson M. Orthokeratology and Contact Lens Quality of Life Questionnaire (OCL-QoL). *Eye Contact Lens*. 2018;44:279–285.
26. Pesudovs K, Garamendi E, Elliott DB. The Quality of Life Impact of Refractive Correction (QIRC) Questionnaire: development and validation. *Optom Vis Sci*. 2004;81:769–777.
27. Hu J, Wang G, Zhou Z, et al. Evaluation of a novel quality of life scale for schoolchildren with nonstrabismic binocular vision anomalies. *Biomed Res Int*. 2020;2020:4723402.
28. 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.
29. Cantó-Cerdán M, Cacho-Martínez P, García-Muñoz A. Delphi methodology for symptomatology associated with visual dysfunctions. *Sci Rep*. 2020;10:19403.
30. Perneger TV, Courvoisier DS, Hudelson PM, Gayet-Ageron A. Sample size for pre-tests of questionnaires. *Qual Life Res*. 2015;24:147–151.
31. Cantó-Cerdán M, Cacho-Martínez P, Lara-Lacárcel F, García-Muñoz A. Rasch analysis for development and reduction of Symptom Questionnaire for Visual Dysfunctions (SQVD). *Sci Rep*. 2021;11:14855.
32. García-Muñoz A, Carbonell-Bonete S, Cantó-Cerdán M, Cacho-Martínez P. Accommodative and binocular dysfunctions: prevalence in a randomised sample of university students. *Clin Exp Optom*. 2016;99:313–321.
33. Leung YY, Png ME, Conaghan P, Tennant A. A systematic literature review on the application of Rasch analysis in musculoskeletal disease–a special interest group report of OMERACT 11. *J Rheumatol*. 2014;41:159–164.
34. Kandel H, Khadka J, Watson SL, Fenwick EK, Pesudovs K. Item banks for measurement of refractive error-specific quality of life. *Ophthalmic Physiol Opt*. 2021;41:591–602.
35. 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.

36. Masters GN. A Rasch model for partial credit scoring. *Psychometrika*. 1982;47:149–174.
37. Linacre JM. Optimizing rating scale category effectiveness. *J Appl Meas*. 2002;3:85–106.
38. McAlinden C, Khadka J, de Freitas Santos Paranhos J, Schor P, Pesudovs K. Psychometric properties of the NEI-RQL-42 questionnaire in keratoconus. *Invest Ophthalmol Vis Sci*. 2012;53:7370–7374.
39. Pesudovs K, Burr JM, Harley C, Elliott DB. The development, assessment, and selection of questionnaires. *Optom Vis Sci*. 2007;84:663–674.
40. Andrich D. An expanded derivation of the threshold structure of the polytomous Rasch model that dispels any "threshold disorder controversy". *Educ Psychol Meas*. 2013;73:78.
41. Andrich D. The legacies of R. A. Fisher and K. Pearson in the application of the polytomous Rasch model for assessing the empirical ordering of categories. *Educ Psychol Meas*. 2013;73:553.
42. Linacre J. What do Infit and Outfit, mean-square and standardized mean? *Rasch Meas Trans*. 2002;16:2:878.
43. Bond T, Fox CM. *Applying the Rasch model: fundamental measurement in the human sciences*. 3rd ed. New York: Routledge; 2015.
44. Stelmack J, Szlyk JP, Stelmack T, et al. Use of Rasch person-item map in exploratory data analysis: a clinical perspective. *J Rehabil Res Dev*. 2004;41:233–241.
45. Christensen KB, Makransky G, Horton M. Critical values for Yen's Q3: identification of local dependence in the Rasch model using residual correlations. *Appl Psychol Meas*. 2017;41:178–194.
46. Cacho-Martínez P, García-Muñoz A, Ruiz-Cantero MT. Diagnostic validity of clinical signs associated with a large exophoria at near. *J Ophthalmol*. 2013;2013:549435.
47. Khadka J, McAlinden C, Pesudovs K. Quality assessment of ophthalmic questionnaires: review and recommendations. *Optom Vis Sci*. 2013;90:720–744.
48. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull*. 1979;86:420–428.
49. Walter SD, Eliasziw M, Donner A. Sample size and optimal designs for reliability studies. *Stat Med*. 1998;17:101–110.
50. Marx RG, Menezes A, Horovitz L, Jones EC, Warren RF. A comparison of two time intervals for test–retest reliability of health status instruments. *J Clin Epidemiol*. 2003;56:730–735.
51. Wright BD, Linacre JM. Observations are always ordinal; measurements, however, must be interval. *Arch Phys Med Rehabil*. 1989;70:857–860.
52. González-Pérez M, Pérez-Garmendia C, Barrio AR, García-Montero M, Antona B. Spanish cross-cultural adaptation and Rasch analysis of the Convergence Insufficiency Symptom Survey (CISS). *Transl Vis Sci Technol*. 2020;9:23.
53. Davis AL, Harvey EM, Twelker JD, Miller JM, Leonard-Green T, Campus I. Convergence insufficiency, accommodative insufficiency, visual symptoms, and astigmatism in Tohono O'odham students. *J Ophthalmol*. 2016;2016:6963976.
54. Wajuihian SO. Correlations between clinical measures and symptoms: report 2: accommodative and vergence measures with symptoms. *J Optom*. 2021;14:142–155.
55. Scheiman M, Grady MF, Jenewein E, et al. Frequency of oculomotor disorders in adolescents 11 to 17 years of age with concussion, 4 to 12 weeks post injury. *Vision Res*. 2021;183:73–80.
56. Krarup TG, Nisted I, Christensen U, Kiilgaard JF, la Cour M. The tolerance of anisometropia. *Acta Ophthalmol*. 2020;98:418–426.
57. Juniper EF. Validated questionnaires should not be modified. *Eur Respir J*. 2009;34:1015–1017.
58. Tsang S, Royse CF, Terkawi AS. Guidelines for developing, translating, and validating a questionnaire in perioperative and pain medicine. *Saudi J Anaesth*. 2017;11:S80–S89.
59. Horan LA, Ticho BH, Khammar AJ, Allen MS, Shah BA. Is the Convergence Insufficiency Symptom Survey specific for convergence insufficiency? A prospective, randomized study. *Am Orthopt J*. 2015;65:99–103.
60. Horwood AM, Toor S, Riddell PM. Screening for convergence insufficiency using the CISS is not indicated in young adults. *Br J Ophthalmol*. 2014;98:679–683.
61. Clark TY, Clark RA. Convergence Insufficiency Symptom Survey scores for reading versus other near visual activities in school-age children. *Am J Ophthalmol*. 2015;160:905–912.e2.