Published in final edited form as:

Optom Vis Sci. 2014 June ; 91(6): 624–633. doi:10.1097/OPX.0000000000000270.

Rasch Analysis of the Student Refractive Error and Eyeglass Questionnaire

Mabel Crescioni, DrPH, Dawn H. Messer, OD, MH, Terri L. Warholak, RPh, PhD, Joseph M. Miller, MD, MPH, J. Daniel Twelker, OD, PhD, FAAO, and Erin M. Harvey, PhD

Department of Ophthalmology and Vision Science (MC, DHM, JMM, JDT, EH), the College of Pharmacy (TLW), the College of Public Health (JMM, EH), and the College of Optical Sciences (JMM), The University of Arizona, Tucson, Arizona

Abstract

Purpose—To evaluate and refine a newly developed instrument, the Student Refractive Error and Eyeglasses Questionnaire (SREEQ), designed to measure the impact of uncorrected and corrected refractive error on vision-related quality of life (VRQoL) in school-aged children. **Methods.** A 38 statement instrument consisting of two parts was developed: Part A relates to perceptions regarding uncorrected vision and Part B relates to perceptions regarding corrected vision and includes other statements regarding VRQoL with spectacle correction. The SREEQ was administered to 200 Native American 6th through 12th grade students known to have previously worn and who currently require eyeglasses. Rasch analysis was conducted to evaluate the functioning of the SREEQ. Statements on Part A and Part B were analyzed to examine the dimensionality and constructs of the questionnaire, how well the items functioned, and the appropriateness of the response scale used.

Results—Rasch analysis suggested two items be eliminated and the measurement scale for matching items be reduced from a 4-point response scale to a 3-point response scale. With these modifications, categorical data were converted to interval level data, to conduct an item and person analysis. A shortened version of the SREEQ was constructed with these modifications, the SREEQ-R, which included the statements that were able to capture changes in VRQoL associated with spectacle wear for those with significant refractive error in our study population.

Conclusions—While the SREEQ Part B appears to have less than optimal reliability to assess the impact of spectacle correction on VRQoL in our student population, it is also able to detect statistically significant differences from pretest to posttest on both the group and individual levels to show that the instrument can assess the impact that glasses have on VRQoL. Further modifications to the questionnaire, such as those included in the SREEQ-R, could enhance its functionality.

Keywords

questionnaire development; vision-related quality of life; refractive error; psychometrics; Native American

Multiple vision-related quality of life (VRQoL) instruments exist, but questionnaires designed specifically for children that are appropriate for assessing the impact of refractive error are rare. For example, the Children's Visual Function Questionnaire¹ and the Impact of Vision Impairment on Children Instrument² have been validated to measure VRQoL in children with significant non-correctable visual impairment^{3, 4} but the questions are unsuitable for measuring the impact of uncorrected refractive error.¹ Additionally, the generic Pediatric Quality of Life Inventory was not found to be specific enough to show the impact of refractive error on quality of life in adolescents.⁵ Five validated questionnaires have been developed specifically for the evaluation of VRQoL related to refractive error: The National Eye Institute Refractive Error Quality of Life Instrument (NEI-RQL-42),⁶ the Refractive Status and Vision Profile Survey (RSVP)⁷, the Quality of Vision questionnaire⁸, the Vision Function and Quality of Life (VFQoL) questionnaire⁹ and the Quality of Life Impact of Refractive Correction (QIRC) questionnaire.¹⁰ Khadka et.al.¹¹ reported that the QIRC demonstrated the highest quality psychometric properties, but the QIRC, as well as with the other 4 instruments, was designed for adults and includes questions inappropriate for assessment of children (questions regarding driving, paying bills, etc.). The only questionnaire we were able to identify designed to measure VRQoL related to refractive error correction in children was the Pediatric Refractive Error Profile (PREP).¹² The PREP, while not psychometrically validated, has been shown to be sensitive enough to detect differences in visual quality of life between spectacle correction and contact lens correction in children. Thus, we used the PREP as the framework for developing a new questionnaire to address the issue of VRQoL with and without spectacles appropriate for use in a school-aged population. The resulting questionnaire is the Student Refractive Error and Eyeglasses Questionnaire (SREEQ)(Table 1).

The purpose of this study was to use Rasch analysis to evaluate SREEQ validity and to determine if this instrument is suitable for assessing the impact of spectacle utilization on the VRQoL in a sample of middle and high school Native American students from a tribe with a high prevalence of astigmatism. This study is unique in that it captures and allows for comparison of the quality of life of children with and without spectacles. Rasch analysis provides insight into the dimensionality and constructs of the questionnaire, and provides information about how well items discriminate between respondents, how well item difficulty targets person perceptions, and how appropriate the response scale is. Rasch analysis also allows transformation of raw ordinal scores into interval-level measurement (expressed in log of the odds units, or logits) when the data fit the model.

METHODS

Questionnaire Development

The PREP questionnaire was used as a framework for the development of the SREEQ.¹² The PREP consists of 10 subscales and 26 statements for each of two scenarios: "When I wear my eyeglasses..." and "When I wear my contact lenses..." Fifteen of the statements were adopted without modification (with exception of replacing "contact lenses" with "glasses"), 11 statements were adopted with modification, and 12 new statements were created to address areas felt to be of particular relevance to our highly astigmatic study

population. The SREEQ was developed to measure the benefit in VRQoL that school-aged children experience with spectacle wear. To measure pre-post change matching Part A and B questions were included along with additional items relevant to VRQoL, but not measurable using a pre/post format. Our multi-disciplinary (optometry, ophthalmology, psychology and public health) research team, who have spent years serving the visions care needs of this highly astigmatic population, used this experience to formulate new questions, modify PREP questions and review the questionnaire for face and content validity prior to first administration of the SREEQ.

The result was a 38-item, Likert-scaled instrument consisting of two parts (see Appendix, available at [LWW insert link]). Part A consists of 15 statements and relates to student perceptions regarding their uncorrected vision and Part B consists of 23 statements and relates to student perceptions regarding their corrected vision and includes other statements regarding VRQoL with spectacle correction. Both Parts A and B used two different 4-point scales (“All of the Time”, “Most of the Time”, “Some of the Time”, “None of the Time” and “Very Much”, “Somewhat”, “A little Bit”, “Not at All”) (Table 1). Twelve of the statements were paired (nearly identical for Part A and B). An additional three statements were added to Part A and an additional eleven statements to Part B to address variables that may be relevant to VRQoL with spectacle correction (e.g., satisfaction with frames and eyeglass fit) and to those with astigmatism correction. These statements were unable to be phrased appropriately for both questionnaires (e.g., the statements regarding frames could not be worded to correspond to the Part A items regarding “When I don’t wear my glasses...”).

Study Population

Students in 6th through 12th grade who were attending school during the 2010–2011 academic school year either at a school on the Tohono O’odham Reservation or at an Arizona State Board of Education charter school that focuses on Tohono O’odham language and culture, were recruited for participation in a longitudinal study of visual development. Subjects were students who presented for an eye examination, reported they were currently wearing glasses or had previously worn eyeglasses (i.e., glasses were currently lost or broken or student chose not to wear them), and who met study requirements for spectacle prescription (Astigmatism: 1.00 D in either eye; Myopia: 1.00 D on any meridian, in either eye; Hyperopia: 2.50 D on any meridian, in either eye; Anisometropia: 1.50 D) upon eye examination. Subjects who presented to the examination wearing contact lenses were excluded from analysis.

This research complied with the Declaration of Helsinki and was approved by the Institutional Review Board of the University of Arizona. Parents provided written informed consent for all students under the age of 18 and all students provided written assent before any testing was conducted. Subjects over the age of 18 provided written informed consent.

Eye Examination

Uncorrected visual acuity was measured with the 62 × 65 cm ETDRS Series 2000 letter charts (Charts 1, 2 and 3, Precision Vision, Inc. LaSalle, IL) at 4 meters. Charts were displayed in a cabinet that provides LED back-illumination and automatically calibrated to a

photopic light level of 85 cd/m² (ESV3000, Good-Lite Co., Elgin, IL). Acuity was recorded as the logMAR size of the smallest line on which the student correctly identified at least 3/5 letters. Cycloplegic autorefraction was performed with the Retinomax K+2 (Nikon Corporation, Tokyo, Japan) at least 30 minutes after subjects received one drop of 0.5% proparacaine followed by a drop of 1% tropicamide and one drop of 1% cyclopentolate. Autorefraction was then refined with retinoscopy and subjective refraction and a dilated fundus examination was performed. Spectacles were selected by and ordered for students who met study criteria for spectacle correction, and were dispensed at a second visit that occurred several weeks after the exam.

Eyeglass Dispensing and SREEQ Administration

The SREEQ was administered at the second visit (several weeks after the exam) by the study coordinator, who is a member of the Tohono O'odham Nation. The study coordinator gave each student the survey and provided simple verbal instructions. She then read each statement to the subject while the subject followed along on their survey. After each statement the study coordinator prompted the subject to record a response before she continued on to the next statement. Survey administration took approximately 15 minutes. Best corrected acuity with their new spectacles was also tested at the second visit using the same method that was used for testing of uncorrected acuity (see above).

Data Analysis

Rasch analysis was performed to evaluate scale and item functioning. Negatively worded statements (2 and 7 in Part A and B and 16 and 20 in Part B) were reverse coded so that all scoring was oriented in the same direction. Twelve of the 15 statements in Part A are identical to statements in Part B with the exception of the introductory clause. We considered these twelve statements to be a retrospective pre- and post-assessment of changes in VRQoL related to correction of refractive error with spectacles. Retrospective pretest-posttest is defined as a "self-report during the course or at the end of treatment that measured subjects' recall of how they were functioning before program outset".¹³ A retrospective pretest-posttest analysis was chosen as a method of limiting response-shift bias, a phenomenon that may occur when the internal construct changes as a result of the experience.¹⁴ Retrospective pretest-posttest methods have demonstrated their usefulness in supporting validity when used to obtain attitudinal responses.^{15–18}

Rasch analysis was used to calculate item difficulty (item measure) in relation to person ability (person measure) by placing both on the same linear continuum. A high item measure (in logits) indicates a high level of the assessed latent trait (e.g., high vision functioning/fewer symptoms) while a low item measure (in logits) indicates a low level of the assessed latent trait (e.g., low vision functioning/more symptoms). A high person measure (in logits) indicates that a person possesses a high level of the assessed latent trait (e.g., high vision functioning/fewer symptoms) while a low person measure (in logits) indicates that a person possesses a low level of the assessed latent trait (e.g., low vision functioning/more symptoms). When the data fit the model, Rasch analysis generates a summative score for each person, which tells us each person's responses for Part A and B. This score is a log odds (logit-score) ratio calculated for each individual that is used to determine which

statements were the hardest to agree with (i.e., high person measure, more difficult to respond in a manner that reflects higher VQOL and/or fewer symptoms). The Wolfe and Chiu procedure was used to compare item and person values on Parts A and B for those questions that were matched. The Wolfe and Chiu procedure is an anchoring technique for measuring pretest-posttest change with a Rasch rating scale and has been useful in interpreting data using a retrospective pretest-posttest data collection technique.¹⁹ Stacking of Part A and Part B data were done as part of the Wolfe and Chui procedure to obtain common step calibrations. This allowed us to obtain corrected item and person calibrations for Part A and Part B data. A t-test was used to test whether there was a statistically significant difference in the person summative scores from the Part A versus Part B questionnaires.

In addition to the 12 matching items, Part A included 3 other items to evaluate distance vision (item 13) and academic performance (item 14 and 15). Part B included 11 other items to measure quality of life with regard to distance vision (item 14), comfort with spectacles (items 11, 15 and 19), academic performance (items 16 and 21), appearance (items 18 and 20), and peer and family support (items 17, 22 and 23). We grouped the statements not part of the pre-post assessment according to their response scale and examined the dimensionality and constructs of these items, how well the items functioned, and the appropriateness of the response scale used.

In summary, we divided the questionnaire into four groups: items matching from Part A and B (Part A 1–12 and Part B 1–10, 12–13), items 14–15 from Part A, items 11 and 14–17 from Part B and items 18–23 from Part B. Item 13 on Part A was not analyzed because it did not match any other items in terms of wording or rating scale. Unidimensionality, error, and local independence were assessed by looking at “item” and “person” fit statistics for each group as well as conducting principal components analysis of the residuals. Fit statistics provide evidence necessary to show that we have one construct measured for each group of statements that were analyzed together. Content validity was evaluated according to the following: (1) item calibration value distribution range, (2) gaps in the measurement continuum (the distance between each item’s calibration value), (3) instrument targeting relative to sample population (variance among sample, i.e., no ceiling or floor effects), and (4) reliability of item and person placement within the hierarchy (item separation, i.e., items should have some range and no major gaps).

Differential Item Functioning (DIF), which examines the levels of response ability of different subgroups of the same study population,¹¹ was assessed by pre and post responses. We used the following criteria for DIF assessment: small or absent if the difference in logits was less than 0.50 logits, 0.50 to 1.0 logits as minimal (but probably inconsequential) DIF and greater than 1.0 logit as notable DIF.²⁰

Descriptive statistics were calculated with SPSS Version 19 (IBM Corp., Armonk, New York) and Rasch analysis was conducted with WINSTEPS version 3.74.0 software (www.winsteps.com).

RESULTS

Subjects

The SREEQ was administered to 200 students. Six students were excluded because they had strabismus or an ocular condition other than refractive error that affected vision. Five were excluded because it was later determined that they had never previously worn eyeglasses and eight subjects because they presented to the eye exam wearing contact lenses. Thus, the final sample for our analysis included 181 subjects.

A summary of subject characteristics is shown in Table 2. The average uncorrected binocular visual acuity was 0.25 logMAR and average corrected binocular visual acuity was -0.05 logMAR (20/35.5 and 20/18 Snellen equivalent, respectively). The majority of our subjects ($n = 100$, 55%) were myopic or myopic astigmats. Sixty-three subjects (35%) were mixed astigmats and 18 (10%) were hyperopic astigmats. No subjects were prescribed eyeglasses for simple hyperopia. Thirty-five percent presented wearing glasses, while 65% reported having previously worn eyeglasses (i.e., glasses were currently lost or broken or student chose not to wear them).

Rasch Analysis

Rasch analysis for the matching Part A and B questions found the requirements for proper rating scale function were met as follows: (1) the number of observations in each category was greater than 10; (2) the average category measures increased with the rating scale categories; (3) INFIT and OUTFIT MNSQ statistics for the measured steps were 0.92 and 1.09, respectively; (4) category thresholds increased with the rating scale categories; (5) category thresholds were at least 1.4 logits apart; and (6) the shape of each rating scale distribution was peaked after changing the rating scales from 4- to 3- point scale.²¹ The rating scales were combined by pulling together the response categories “some of the time” and “most of the time”, into “some/most of the time” thus reducing the rating scale to 3-points. Figure 1 depicts the changes in the category probability curve for the 10 matching items in Part A and B before and after combining the rating scales from 4- to 3- point scales. The remaining three sets of questions met all the requirements listed above for proper rating scale function using a 4-point scale, thus their rating scales were not combined.

Item fit is an index of whether items function logically and fit in the same construct as other items, and do not collect too much error.²² After examining item fit statistics, it was necessary to eliminate items 2 and 7 from the matching questions in Parts A and B because the Infit and Outfit statistics did not fall within the recommended item mean square range for Infit and Outfit statistics (0.6 to 1.4).²³ Item “misfit” may be due to the item being too complex, confusing, or it may indicate the item measures a different construct than the other items. After eliminating these items, the INFIT and OUTFIT mean square (MNSQ) and standardized mean square (Zstd) values were within the recommended threshold, 0.6–1.4 and -2.0 – 2.0 , respectively. These findings indicate that these data exhibited good fit and supported the unidimensionality and local independence requirements of the model, as recommended by Wright & Panchapakesan, demonstrating that the 10 items measured the same constructs (i.e., unidimensionality) and produced additivity of measures (i.e., true

interval level data).²⁴ Principal Components Analysis of the residuals gave the unexplained variance accounted for by the first contrast as 1.8 eigenvalue units for Part A and 1.6 eigenvalue units for Part B. We used the criterion that the first contrast should have an eigenvalue of >2.0 to cause us to reject the assumption of unidimensionality. The SREEQ showed DIF from Part A to Part B for three items. The only item that showed notable DIF was item 12 from Part A “When I don’t wear my glasses, I feel dizzy” (1.15 logits). Categorical data were converted to interval level data i.e. to logits, to test the differences between responses to the 10 matching items remaining in the analysis from Part A and Part B.

The item separation index (item measure, i.e., the extent that items are sufficiently spread out to define distinct levels of perceived VRQoL) was 7.60 and 2.40, which translate to an item reliability of 0.98 and 0.85, respectively, for the 10 matching items of Part A and Part B. The person separation index was 2.43 and 0.57, which translate to a person reliability of 0.86 and 0.25, respectively (analogous to Cronbach alpha), for the 10 items in Parts A and B.

Item and Person Measures Pretest to Posttest—The item measures showed that for three of the ten matching items the responses for Part A were significantly different than for Part B ($p < 0.05$). The specific items that were significantly different, after performing a t-test and Bonferroni correction, from Part A to B were items stating: “My vision is blurry” (Item 3 in Part A and B, $p < 0.001$); “I get headaches or my head or eyes hurt when I read” (Item 10 in Part A and B, $p < 0.001$); and “I feel dizzy” (Item 12 on Part A and Item 13 on Part B, $p < 0.001$).

The group means for the 10 matching items were 0.61 logits (± 1.8 logits) and 1.85 logits (± 1.6 logits) for the Part A and Part B, respectively, reflecting a statistically higher VRQoL with glasses than without glasses for the group as a whole (Wilcoxon signed rank test, $z = -10.859$, $p < 0.001$). Statistical significance was measured using non-parametric statistics because our data were not normally distributed.

Rasch analysis also allowed evaluation of individual person measures and each item’s contribution to the overall instrument. The use of this model in evaluating responses from pretest to posttest presents an advantage over other statistical methods, allowing quantification of perceived VRQoL for each student. Individually, 104 (58.1%) of 181 students completing the questionnaire demonstrated a self-perceived improvement in their VRQoL with spectacles as measured by the attitudinal questionnaire ($p < 0.05$).

Item Evaluation Pretest to Posttest (10 Matching Items in Part A and B)—Figure 2 depicts the expected score maps and student normative distribution and what can be expected from each person/item interaction (expected score map) with items at the bottom of the hierarchy being the easiest to endorse (i.e., mean responses closer to the side of scale which represents higher VRQoL and/or fewer symptoms) and items at the top being the most difficult for students to endorse (i.e. responses closer to the side of scale which represents poorer VRQoL and/or more symptoms). Figure 2 also depicts the frequency of the responses (“all of the time”, “some of the time”, and “none of the time”) to each of the

ten matching items in Part A and B and reflects the differences in responses from Part A and B which demonstrate the change in VRQoL without glasses (pre) to with glasses (post).

Item Evaluation for Unmatched Part A Items—Two non-matching items on Part A (Items 14 and 15) used the response scale: “very much”, “somewhat”, “a little bit”, and “not at all”. Both items relate to school and as Figure 3 shows there is a significant gap between these two items which points to the need for additional items related to academic performance (i.e., a gap in content). Table 3 summarizes the fit statistics and reliability index for all unmatched questions.

Item Evaluation for Unmatched Part B items—There were 11 items on Part B not included in the pre/post assessment. Items 11 and 14–17 used the same response scale as the pre/post items. Items 18 to 23 used the “very much”, “somewhat”, “a little bit”, and “not at all” scale. Item evaluation was done separately for these two groups of items (see Figures 4 & 5).

For items 11 and 14–17, item 11 (“When I wear my glasses, my nose or ears hurt”) was easiest to endorse positively (i.e. most likely to respond toward “none of the time” side of scale), while item 16 (“When I wear my glasses, I do better on tests”) was the most difficult to endorse positively (i.e., most likely to respond toward “all of the time” side of scale).

For items 18–23, a “very much” response corresponded with better VRQoL for all items except Item 20, which was therefore reverse coded. According to the item hierarchy, item 22, “When I wear my glasses, my friends like the way I look”, was the most difficult of the items to endorse positively (i.e., respond toward the “very much” side of scale) while item 23, “When I wear my glasses, my family members like the way I look” was the least difficult to endorse positively when subjects are wearing glasses, making a clear distinction between perception of peers and family.

DISCUSSION

While there are several validated instruments available to measure refractive error-specific VRQoL in adults,^{6, 25, 26} there are currently no existing validated instruments specifically designed to assess the impact of corrected and uncorrected refractive error on school-aged children’s VRQoL. The SREEQ was developed to serve as such an instrument and this study was conducted to assess the functioning of this new questionnaire in our subject population of Native American middle and high school students, primarily from the Tohono O’odham Nation.

This study demonstrated that overall the SREEQ Part B has less than optimal reliability to assess the impact of spectacle correction of refractive error on this population. Rasch analysis found that a better fitting model could be obtained by changing the response options of the 10 matching items in Part A and B from a 4-point to a 3-point scale by combining the “some of the time” and “most of the time” categories. We also identified two items to be removed from the analysis. Items 2 and 7 (Part A and B) were phrased in a different

direction (i.e., positively) than the rest of the survey and we believe this may have caused confusion for the students.

For the 10 matching items in Parts A and B, the person separation index was very good for Part A (0.86), but low for Part B (0.25), indicating that Part B did not successfully distinguish between subjects. This may be because the subjects' VRQoL without glasses was dependent on their uncorrected acuity which varied widely, whereas with glasses, the subjects had uniformly good visual acuity, resulting in more similar responses across subjects in Part B, and a low separation index. In future versions of SREEQ, adding more detailed questions regarding critical visual functioning could increase the person separation index of Part B.

Rasch analysis also found the hierarchical ordering of the items in the keyform maps in Figure 2–5 show that while the items of the SREEQ are fairly well distributed in terms of difficulty, some gaps are present, suggesting that future versions the questionnaire could be enhanced by the addition of statements on certain topics. For instance, a gap is seen in Figure 3, indicating additional statements related to academics and effort needed to complete an academic task that use the same scale could be added. Gaps observed in Figure 5, which show statements mostly related to self-perception of appearance and perception of appearance by peers and family, suggests that it might be helpful to ask more specific statements about the aesthetics and comfort of the frames.

Our analysis of the 10 matching Part A statements shows that when subjects are asked about whether they have to squint to see things clearly when they are not wearing their glasses (Item 4), the majority of subjects responded “all of the time”. Our analysis suggests that this was a “difficult” statement to answer “all of the time”. The difficulty may be because many of our subjects have to squint to see clearly without their glasses. Whereas, Item 12 which asks about whether they feel dizzy when they are not wearing their glasses was “easy” or higher VQoL because not many of them actually experience dizziness when they are not wearing their glasses. Although there was notable DIF for Item 12 from Part A, “When I don’t wear my glasses, I feel dizzy”, it was not eliminated because doing so could lead to lowering the reliability and validity²⁰ of Part A.

We observed a statistically significant change from Part A to Part B for three of the ten matching statements: “My vision is blurry” (Item 2); “I get headaches or my head or eyes hurt when I read” (Item 8); and “I feel dizzy” (Item 9). The items that are statistically significant appear to be those related to specific problems that are more identifiable by subjects. These items are also those that fall right in the middle of the keyform map, i.e., these were neither the easiest nor the hardest items to respond positively to (report higher VQOL). The finding that the SREEQ was also able to detect statistically significant differences from pretest to posttest on both the group and individual levels show that the instrument can detect the impact that glasses have on VRQoL and indicate that the instrument can have broad applicability.

Based on our analyses, we have constructed a new version of the survey, using a 3-point rather than the 4-point scale, and including only the 10 statements that are repeated in Part A

and B, deleting Items 2 and 7. This shortened SREEQ (SREEQ-R) focuses on measuring the difference in VRQoL with versus without glasses and is available online. As this survey has not been validated in its present form, it should be used cautiously until it can be validated in a diverse population. However, we have included it here because we believe this simplified version should prove useful for determination of change in VRQoL with versus without glasses, as these 10 matching (pre-post) statements and the assessment of change using these 10 statements were validated in the present paper, although not in the SREEQ-R format.

Our study has several limitations. The potential for recall bias in our study is of significant concern. Health-related quality of life issues have been found to be more susceptible to recall bias than more objective events²⁷ and the longer the recall period the greater the potential for recall error.²⁸ The SREEQ asks subjects to report both on visual symptoms when they don't wear their glasses (Part A) as well as when they do wear their glasses (Part B). Subjects who faithfully wear their eyeglasses during all waking hours, may not accurately recall what vision symptoms they experienced when not wearing their eyeglasses. Likewise, subjects who had lost or broken their eyeglasses may not recall accurately how well they functioned or what symptoms they experiences when wearing their eyeglasses. While errors in recall could be non-differential in nature, the possibility of recall bias cannot be ruled out.

Second, we did not have a wide variety of refractive errors represented in our sample. The majority of our subjects were myopic and mixed astigmats. This prevented us from being able to determine if the SREEQ distinguished between different refractive error groups as other VRQoL studies have done.^{7, 25} Finally, the generalizability of our analysis on the reliability and validity of this questionnaire is limited, as we did not have a diverse subject population, with almost all of our subjects being Native American students with astigmatic refractive error. Future studies including a wider range of refractive errors and a more diverse population would allow determination of the generalizability of the instrument.

In conclusion, our study indicated that modifications to the SREEQ were needed to improve its functionality. A shortened version of the SREEQ was constructed incorporating these modifications: using a 3-point rather than the 4-point scale and including only the 10 statements that are repeated in Part A and B. The SREEQ-R should be a feasible and reliable instrument to assess the impact of spectacle correction on VRQoL in our student population. However, future versions of the SREEQ could be further enhanced with the addition of statements aimed at reducing the gaps identified in the key form maps.

Acknowledgments

Supported by grant U10 EY13153 (EMH) from the National Eye Institute of the National Institutes of Health, Department of Health and Human Services, by unrestricted funds to the Department of Ophthalmology and Vision Science from Research to Prevent Blindness (JMM), Walt and Lilly Disney Award for Amblyopia Research from Research to Prevent Blindness (JMM), and by a Career Development Award from Research to Prevent Blindness (EMH). The authors thank the Tohono O'odham Nation, Indian Oasis/Baboquivari School District, the Bureau of Indian Affairs Office of Indian Education Programs the San Xavier Mission School, Ha:san Preparatory and Leadership School, and the students who participated in the study. This study is overseen by a NIH/NEI Data Monitoring and Oversight Committee [Robert Hardy, PhD (chair), Tina Aguilar, Morgan Ashley, Donald Everrett, MA, Jeanette Francisco, Jonathan Holmes, MD, Ronald Johnson, MD, Rosemary Lopez, and Karla Zadnik, OD, PhD].

References

1. Feliuss J, Stager DR Sr, Berry PM, Fawcett SL, Stager DR Jr, Salomao SR, Berezovsky A, Birch EE. Development of an instrument to assess vision-related quality of life in young children. *Am J Ophthalmol*. 2004; 138:362–72. [PubMed: 15364217]
2. Cochrane G, Lamoureux E, Keeffe J. Defining the content for a new quality of life questionnaire for students with low vision (the Impact of Vision Impairment on Children: IVI_C). *Ophthalmic Epidemiol*. 2008; 15:114–20. [PubMed: 18432495]
3. Cochrane GM, Marella M, Keeffe JE, Lamoureux EL. The Impact of Vision Impairment for Children (IVI_C): validation of a vision-specific pediatric quality-of-life questionnaire using Rasch analysis. *Invest Ophthalmol Vis Sci*. 2011; 52:1632–40. [PubMed: 21178144]
4. Gothwal VK, Sumalini R, Irfan SM, Giridhar A. Rasch analysis of Impact of Vision Impairment for Children questionnaire. *Optom Vis Sci*. 2013; 90:820–7. [PubMed: 23748842]
5. Wong HB, Machin D, Tan SB, Wong TY, Saw SM. Visual impairment and its impact on health-related quality of life in adolescents. *Am J Ophthalmol*. 2009; 147:505–11. [PubMed: 19056077]
6. Hays RD, Mangione CM, Ellwein L, Lindblad AS, Spritzer KL, McDonnell PJ. Psychometric properties of the National Eye Institute-Refractive Error Quality of Life instrument. *Ophthalmology*. 2003; 110:2292–301. [PubMed: 14644710]
7. Vitale S, Schein OD, Meinert CL, Steinberg EP. The refractive status and vision profile: a questionnaire to measure vision-related quality of life in persons with refractive error. *Ophthalmology*. 2000; 107:1529–39. [PubMed: 10919904]
8. 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; 51:5537–45. [PubMed: 20505205]
9. Brady CJ, Keay L, Villanti A, Ali FS, Gandhi M, Massof RW, Friedman DS. Validation of a visual function and quality of life instrument in an urban Indian population with uncorrected refractive error using Rasch analysis. *Ophthalmic Epidemiol*. 2010; 17:282–91. [PubMed: 20868254]
10. Pesudovs K, Garamendi E, Elliott DB. The Quality of Life Impact of Refractive Correction (QIRC) Questionnaire: development and validation. *Optom Vis Sci*. 2004; 81:769–77. [PubMed: 15557851]
11. Khadka J, McAlinden C, Pesudovs K. Quality assessment of ophthalmic questionnaires: review and recommendations. *Optom Vis Sci*. 2013; 90:720–44. [PubMed: 23873034]
12. Walline JJ, Jones LA, Chitkara M, Coffey B, Jackson JM, Manny RE, Rah MJ, Prinstein MJ, Zadnik K. The Adolescent and Child Health Initiative to Encourage Vision Empowerment (ACHIEVE) study design and baseline data. *Optom Vis Sci*. 2006; 83:37–45. [PubMed: 16432471]
13. Davis NM. A medication error prevention educational retreat. *Hosp Pharm*. 2000; 35:466–7.
14. Nicholson T, Belcastro PA, Gold RS. Retrospective pretest- posttest analysis versus traditional pretest-posttest analysis. *Psychol Rep*. 1985; 57:525–6.
15. Skeff KM, Stratos GA, Bergen MR. Evaluation of a medical-faculty development program - a comparison of traditional pre/post and retrospective pre/post self-assessment ratings. *Eval Health Prof*. 1992; 15:350–66.
16. Howard GS. Response-shift bias - a problem in evaluating interventions with pre-post self-reports. *Evaluation Rev*. 1980; 4:93–106.
17. Aiken LS, West SG. Invalidity of true experiments - self-report pretest biases. *Evaluation Rev*. 1990; 14:374–90.
18. Sprangers M, Hoogstraten J. Pre-testing effects in retrospective pretest-posttest designs. *J App Psychol*. 1989; 74:265–72.
19. Wolfe EW, Chiu CW. Measuring pretest-posttest change with a Rasch Rating Scale Model. *J Outcome Meas*. 1999; 3:134–61. [PubMed: 10204324]
20. Gothwal VK, Wright TA, Lamoureux EL, Pesudovs K. Rasch analysis of the quality of life and vision function questionnaire. *Optom Vis Sci*. 2009; 86:836–44.

21. Linacre JM. Investigating rating scale category utility. *J Outcome Meas.* 1999; 3:103–22. [PubMed: 10204322]
22. Green, KE.; Frantom, CG. Survey development and validation with the Rasch model. Paper presented at the International Conference on Questionnaire Development, Evaluation, and Testing; Charleston, SC. November 14–17, 2002.;
23. Bond, TG.; Fox, CM. *Applying the Rasch Model: Fundamental Measurement in the Human Sciences.* 2. Mahwah, NJ: L. Erlbaum; 2007.
24. Wright B, Panchapa N. A procedure for sample-free item analysis. *Educ Psychol Meas.* 1969; 29:23.
25. Nichols JJ, Mitchell GL, Saracino M, Zadnik K. Reliability and validity of refractive error-specific quality-of-life instruments. *Arch Ophthalmol.* 2003; 121:1289–96. [PubMed: 12963612]
26. Gothwal VK, Wright TA, Elliott DB, Pesudovs K. The refractive status and vision profile: Rasch analysis of subscale validity. *J Refract Surg.* 2010; 26:912–5. [PubMed: 20506960]
27. Schmier JK, Halpern MT. Patient recall and recall bias of health state and health status. *Expert Rev Pharmacoecon Outcomes Res.* 2004; 4:159–63. [PubMed: 19807519]
28. Stull DE, Leidy NK, Parasuraman B, Chassany O. Optimal recall periods for patient-reported outcomes: challenges and potential solutions. *Curr Med Res Opin.* 2009; 25:929–42. [PubMed: 19257798]

APPENDIX

The appendix, a 38-item, Likert-scaled instrument consisting of two parts, is available at **[LWW insert link]**.

Measuring Children's Vision Related Quality of Life

A newly developed instrument designed to measure the impact of uncorrected and corrected refractive error on vision-related quality of life (VRQoL) in school-aged children is evaluated using Rasch analysis. While one section of the instrument has less than optimal reliability to assess the impact of spectacle correction on VRQoL in our student population, it is also able to detect statistically significant differences from pretest to posttest on both the group and individual levels to show that the instrument can assess the impact that glasses have on VRQoL. Further modifications to the questionnaire, such as those included in the revised instrument, could enhance its functionality.

Figure 1.

Category Probability Curve for Part A (Without Glasses) and Part B (With Glasses) using 4-point scale (top) and Category Probability Curve for Part A (Without Glasses) and Part B (With Glasses) using 3-point scale (bottom).

Figure 2.
Subject responses to statements with the All, Some and None response scale WITH and WITHOUT glasses statements for the 10 matching statements from Part A and B. (N = 181).

Figure 3.
 Subject responses to statements with the Very Much, Somewhat and Not at all response scale WITHOUT glasses (part A) statements 14 and 15 (n = 181).

Figure 4.
Patient responses to statements with the All, Some and None response scale WITH glasses
(Part B) statements 11– 17 (N = 181).

KEY
 V = Very much
 S = Somewhat/A little bit
 N = Not at all
 * = Reverse keyed

Figure 5.
 Patient responses to statements with the Very Much, Somewhat and Not at all response scale
 WITH glasses statements (Part B) 18–23 (N=181).

Table 1

Part A and B Matching Items and Rating Scale.

Part A Questions	Rating Scale*	Pairs with Item # in Part B
1. When I don't wear my glasses, I have problems seeing clearly.	1	1
2. When I don't wear my glasses, my vision is very clear	1	2
3. When I don't wear my glasses, my vision is blurry.	1	3
4. When I don't wear my glasses, I have to squint to see things clearly.	1	4
5. When I don't wear my glasses, I have problems seeing the computer or video games clearly.	1	5
6. When I don't wear my glasses, I have problems reading.	1	6
7. When I don't wear my glasses, I am able to see clearly far away	1	7
8. When I don't wear my glasses, I have problems seeing the board, at the movies, or other things far away.	1	8
9. When I don't wear my glasses, I have problems recognizing people across the street or down the hall.	1	9
10. When I don't wear my glasses, I get headaches or my head or eyes hurt when I read	1	10
11. When I don't wear my glasses, things look distorted, slanted or double.	1	12
12. When I don't wear my glasses, I feel dizzy.	1	13
13. When I don't wear my glasses, I have problems seeing things when I play outdoors.		-
14. When I don't wear my glasses, my schoolwork is harder to do	2	-
15. When I don't wear my glasses, it is harder to do well on tests	2	-

Part B Questions	Rating Scale*	Pairs with Item # in Part A
1. When I wear my glasses, I have problems seeing clearly.	1	1
2. When I wear my glasses, my vision is very clear	1	2
3. When I wear my glasses, my vision is blurry.	1	3
4. When I wear my glasses, I have to squint to see things clearly.	1	4
5. When I wear my glasses, I have problems seeing the computer or video games clearly.	1	5
6. When I wear my glasses, I have problems reading.	1	6
7. When I wear my glasses, I am able to see clearly far away	1	7
8. When I wear my glasses, I have problems seeing the board, at the movies, or other things far away.	1	8
9. When I wear my glasses, I have problems recognizing people across the street or down the hall.	1	9
10. When I wear my glasses, I get headaches or my head or eyes hurt when I read	1	10
11. When I wear my glasses, my nose or ears hurt	1	-
12. When I wear my glasses, things look distorted, slanted or double.	1	11
13. When I wear my glasses, I feel dizzy.	1	12
14. I have a problem wearing my glasses when I play outdoors	1	-
15. I am bothered by my glasses when I play sports, dance or do other activities.	1	-
16. When I wear my glasses, I do better on tests.	1	-
17. When I wear my glasses, my classmates make fun of me	1	-
18. When I wear my glasses, I like how I look	2	-
19. I like my frames		-
20. If I didn't wear glasses, I would look better	2	-
21. When I wear my glasses, my schoolwork is easier	2	-

Part B Questions	Rating Scale*	Pairs with Item # in Part A
22. When I wear my glasses, my friends like the way I look	2	-
23. When I wear my glasses, my family members like the way I look	2	-

<u>*Scale 1</u>	<u>*Scale 2</u>
All of the Time	Very Much
Most of the Time	Somewhat
Some of the Time	A little Bit
None of the Time	Not at All

Table 2

Demographic and Clinical Characteristics of Subjects.

Validation Analyses	
Demographics	
Number of Patients	181
Gender, N(%)	
Female	99 (54.7)
Male	82 (45.3)
Age (Years)	
Mean (SD)	14.7 (2.13)
Range	10.93 to 20.44
Grade, N(%)	
6th	43 (23.8)
7th	33 (18.2)
8th	26 (14.4)
9th	29 (16.0)
10th	15 (8.3)
11th	24 (13.3)
12th	11 (6.1)
Clinical Characteristics	
Spherical Equivalent (Diopters)	
Mean (SD)	−1.30 (2.05)
Range	−8.13 to 3.69
Astigmatism (Diopters)	
Mean (SD)	2.38 (1.6)
Range	0 to 6.50
Uncorrected BVA* (logMAR) N(%)	
Better than 0.3	101 (55.8)
0.3 or worse	80 (44.2)
Corrected BVA* (logMAR), N(%)	
Better than 0.1	160 (88.4)
0.1 or worse	21 (11.6)

* Binocular Visual Acuity

Table 3

Rasch Item Summary Scores for Unpaired Questions

	Items 14–15 (Part A)	Items 11, 14–17 (Part B)	Items 18–23 (Part B)
Person Separation Index	2.05	0.87	1.36
Subject Reliability	0.81	0.43	0.65
Item Separation Index	2.77	6.45	3.32
Item Reliability	0.88	0.98	0.92
INFIT MNSQ (Mean)	0.97	1.01	1.02
INFIT ZSTD (Mean)	−0.1	−0.1	−0.3
OUTFIT MNSQ (Mean)	0.93	0.95	1.00
OUTFIT ZSTD (Mean)	−0.4	−0.4	−0.4
INFIT MNSQ (SD)	0.00	0.26	0.42
INFIT ZSTD (SD)	0.00	2.5	3.9
OUTFIT MNSQ (SD)	0.01	0.26	0.46
OUTFIT ZSTD (SD)	0.1	2.2	3.7