Psychometric Performance of the NEI VFQ-25 in Visually Normal Latinos: The Los Angeles Latino Eye Study

Denise Globe,[1] *Rohit Varma,*[2,3,4] *Stanley P. Azen,*[2,3,4] *Sylvia Paz,*[2] *Elaine Yu,*[3,4] *Susan Preston-Martin,*[3] *and the Los Angeles Latino Eye Study Group*[5]

Purpose. To characterize the psychometric performance of the National Eye Institute 25-item Visual Function Questionnaire (NEI VFQ-25) in visually normal Latinos (Mexican Americans).

Methods. The Los Angeles Latino Eye Study (LALES) is a population-based study to assess the prevalence of eye disease and self-reported visual functioning in Latinos aged 40 or more years. Self-reported visual functioning was assessed by using English and Spanish versions of the NEI VFQ-25. Psychometric properties of the NEI VFQ-25, including internal consistency of the subscales and the individual items, were assessed through the Multi-trait Analysis Program-Revised (MAP-R) analysis. Adjusted mean and median subscale scores were compared between English and Spanish speakers to identify any systematic differences.

Results. Of the 1917 participants from two census tracts, 1171 participants with no visual impairment were included in this analysis. The mean age of the participants was 52.3 years, 57% of the participants were female, and 67.5% of the participants were Spanish speaking. Median scores for Spanish-speaking participants were significantly lower than those of the English-speaking participants on four subscales: Ocular Pain, General Vision, Vision-Specific Mental Health, and General Health ($P < 0.05$). Internal consistency for three of eight measurable subscales for the study group was poor (Cronbach $\alpha < 0.6$).

Conclusions. This study reveals psychometric inconsistencies in the NEI VFQ-25 when administered to visually normal Latinos. The difference in mean subscale scores between Spanish and English speakers must be integrated into the development of population norms of visual function. Further detailed psychometric evaluation is needed to determine the validity of this instrument in Latino populations. (*Invest Ophthalmol Vis Sci.* 2003;44:1470–1478) DOI:10.1167/iovs.02-0292

The Los Angeles Latino Eye Study (LALES) is a population-based prevalence survey of ocular disease in Latinos, aged 40 or more years, in the city of La Puente in Los Angeles County, California. In addition, the LALES is designed to determine the level of self-reported visual functioning in this population-based sample, as measured by the National Eye Institute 25-item Visual Function Questionnaire (NEI VFQ-25).[1] The NEI VFQ-25 is a vision targeted measure of health related quality of life that produces a single overall visual function score that rates the patient's perceived visual functioning. In addition, 12 subscale scores are generated as independent, function-specific measures of visual functioning.[2,3] The NEI VFQ-25 has been used to assess visual functioning in individuals with and without ocular disease.[4,5]

To characterize visual functioning accurately in a population-based study and to assess the association between ocular disease and the patient's perception of visual functioning, it is essential to have an instrument that is culturally sensitive and appropriate for Latinos.[6–8] The purpose of assessing the cultural appropriateness of an instrument is to ensure that different participants experience the questionnaire equally so that responses are consistent between individuals with different characteristics. Language and acculturation are intimately connected.[9] One major construct in the measurement of acculturation is the language in which an individual speaks, writes, reads, and thinks.[10,11] To date, few studies have assessed measures of visual functioning in a visually normal Latino population.

The NEI VFQ-25 field test excluded non–English-speaking participants.[3] This field test included both individuals with ocular disease and a reference group of 122 participants who were examined at one of seven ophthalmology practices and who had no clinical evidence of ocular disease. Although a Spanish translation is currently available, only one study has validated this translation in a large, population-based Latino cohort.[4] That study provides an initial assessment of the psychometric properties of the NEI VFQ-25 in a Latino population. Although differences in responses to items are reported, comparing Spanish- and English-speaking participants and those with visual impairment more than or equal to 20/40 with those with visual impairment worse than 20/40, Cronbach α and convergent and divergent validity (psychometric performance) are not assessed between English and Spanish speakers.

The NEI VFQ-25 was originally developed to assess visual functioning in individuals with ocular disease. However, it is important to establish benchmarks in visually normal individuals as an important comparator for visual functioning for those with ocular disease. These benchmarks can be acquired in two ways. First, they can be obtained from a clinic-based sample. This method was used in the NEI VFQ-51 field test, in which a reference group without ocular disease or visual impairment was selected from patients attending an eye clinic.[1] A second method is to obtain the benchmark from a population-based sample. This approach has been used for the NEI VFQ-25 in a population-based sample in Arizona.[4] In addition, it has been used by the International Quality of Life Assessment Project (IQOLA), which has established methods to translate a general quality-of-life instrument, the short form (SF)-36. In the current study, we used methodologies set forth by the IQOLA

From the [1]Department of Pharmaceutical Economics and Policy, School of Pharmacy, the [2]Doheny Eye Institute, and the Departments of [3]Ophthalmology and [4]Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, California.

[5]Los Angeles Latino Eye Study Group members are listed in the Appendix.

Supported by National Eye Institute Grants U10-EY11753 and EY03040 and an unrestricted grant for the Research to Prevent Blindness.

Submitted for publication May 17, 2002; revised August 9 and December 9, 2002; accepted December 30, 2002.

Commercial relationships policy: N.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Corresponding author: Rohit Varma, Doheny Eye Institute, 1450 San Pablo St., DEI 4803, Los Angeles, CA 90033; rvarma@usc.edu.

Investigative Ophthalmology & Visual Science, April 2003, Vol. 44, No. 4

project to assess the psychometric properties in a general population. These norms can be used as a benchmark for comparison of visual functioning in persons with visual impairment or ocular disease across the continuum of visual acuity in both epidemiologic and clinic-based studies.

For this article, we assessed only visually normal participants, comparing the psychometric performance of the English and Spanish versions of the NEI VFQ in an entirely Latino population. It expands on the work by Broman et al.[4] on a population-based sample of Hispanics from Arizona. Our results, and the results of Broman et al., can be used as a benchmark for comparison of visual functioning in persons with visual impairment or ocular disease.[12,13]

METHODS

Study Design

LALES is a population-based prevalence survey of ocular disease in noninstitutionalized, self-identified Latinos aged 40 or more years in six adjacent census tracts in and around the city of La Puente, California. The study cohort consisted of 1917 self-identified Latinos, aged 40 or more years, living in La Puente. The study protocol was approved by the Institutional Review Board at the University of Southern California and adhered to the tenets of the Declaration of Helsinki. Demographic information, ocular disease, risk factors, history of ocular and medical disease, access to care, acculturation, and insurance status were collected through interviews in the participants' homes. The participants also completed health-related quality-of-life (Medical Outcomes Study 12-Item Short Form Health Survey SF-12)[14] and visual functioning questionnaires (NEI VFQ-25),[1] which were administered in a standardized manner by two, trained, female interviewers at the LALES Local Eye Examination Center. The interview was followed by a medical examination and a complete ocular examination. Interviews were administered with a computer-assisted program designed specifically for the LALES, which facilitated consistency in administration of the interview.

Study Group

Similar to methods used by the IQOLA project,[12,15] a visually normal subgroup of LALES participants without evidence of visual impairment or ocular disease (e.g., visually impairing cataract, diabetic retinopathy, macular degeneration, or glaucoma) were identified for this analysis. Participants in the visually normal group had binocular visual acuity better than 20/40 and normal visual fields at initial examination. Visual acuity was measured binocularly using the Early Treatment Diabetic Retinopathy Study (ETDRS) distance charts (Lighthouse International, New York, NY) with the participant's normal refractive correction (see Azen et al.[16] for details). Visual field analysis was performed to assess peripheral vision (24-2 full-threshold or the 24-2 Swedish interactive test algorithm [SITA] standard test; Humphrey Field Analyzer II; Zeiss-Humphrey Systems, Dublin, CA). Participants with reliable visual fields and with no evidence of visual field loss were included in this study. The analysis was then performed again in all participants, both with and without visual impairment.

Self-Reported Visual Function

The NEI VFQ-25 is composed of 12 vision-specific subscales. Each subscale contains a minimum of one and a maximum of four items. The subscales include: General Health (one item), General Vision (one item), Near Vision Activities (three items), Distance Vision Activities (three items), Ocular Pain (two items), Vision-Specific Social Function (two items), Vision-Specific Role Difficulties (two items), Vision-Specific Mental Health (4 items), Vision-Specific Dependency (3 items), Driving Difficulties (two items), Color Vision (one item), and Peripheral Vision (one item). The NEI VFQ-25 is scored using standard algorithms.[1] Each item was scored on a scale from 0 (lowest visual functioning) to 100 (best visual functioning). Items were then reverse coded, when appropriate, so that the directionality of the items and subscales were comparable. Item scores within a subscale were averaged to yield the subscale score (range, 0–100). Interviewers administered the questionnaire in either English or Spanish, according to participant's preference. Spanish speakers were defined as those who completed the questionnaire in Spanish; English speakers were defined as those who completed the questionnaire in English.

Covariates

The covariates included age, gender, income, education, employment status, and number of comorbidities. The number of comorbid medical conditions was computed as a summation of a list of 13 self-reported,[10] non–ocular-related medical conditions, including diabetes, arthritis, stroke or brain hemorrhage, high blood pressure, angina, heart attack, heart failure, asthma, skin cancer, other cancer, back problems, hearing problems, and other major health problems.

Statistical Analysis

First, descriptive statistics were generated to determine the distribution of both demographic and clinical characteristics. Demographic and clinical characteristics were then compared between Spanish and English speakers. The χ^2 test was used for discrete variables and Student's *t*-test was used for continuous variables.

Psychometric Analysis

Psychometric performance was assessed through the Multi-trait Analysis Program-Revised (MAP-R), a program designed to evaluate questionnaires that have ordered-response choices.[17] First, item-specific missing rates (number of times an item was not answered) for the NEI VFQ-25 subscales were compared between English- and Spanish-speaking participants. Next, the distribution of scores obtained for each scale were examined by calculating the percentage of the sample achieving the lowest possible score of zero (floor) and highest possible score of 100 (ceiling).[18] In a heterogeneous sample, subscale scores should have less than 20% of observations at the theoretical floor (floor effect) or ceiling (ceiling effect).[19] Even fewer floor effects were expected in this study, because this sample was specifically selected to have no evidence of visual impairment or ocular disease.

Cronbach α (a measure of the extent to which items within a single subscale correlate with the subscale score) was then calculated for each subscale (for the entire sample and for each language subgroup) as a measure of reliability of the subscale's internal consistency.[20] The acceptable minimum Cronbach α is 0.70.[21]

Item internal consistency (the degree to which each individual item measures the underlying construct) was measured through Pearson correlation of each item with the subscale to which it was assigned. If the hypothesis that the item measures the underlying construct represented by the subscale to which it is assigned was correct, then the correlation between that item and the subscale would be greater than 0.40. Item discriminant validity (the degree to which different constructs are correlated) was also evaluated with Pearson correlation coefficients for the entire sample and for Spanish- and English-speaking Latinos. Item discriminant validity evaluates the association of an individual item with the subscale that it is part of, compared with other subscales. An item should correlate most strongly with its own subscale when compared with every other subscale in the questionnaire. The acceptable cutoff for item discriminant validity is less than 0.40.[17]

Comparison of English and Spanish Subscale Scores

Analysis of covariance was used to compare the mean self-reported visual functioning subscale scores between Spanish- and English-speaking Latinos to determine whether there was a systematic difference in responses. The mean NEI VFQ-25 subscale scores were adjusted for the standard covariates (age, gender, income, education, employment sta-

TABLE 1. Demographic Characteristics of the Study Group Stratified by Language Spoken

	Complete Group (n = 1171)	English (n = 380)	Spanish (n = 791)	P*	Adjusted P†
Age (y)	52.3 (9.6)	53.9 (10.9)	51.5 (8.8)	<0.001	—
Female	667 (57)	218 (33)	449 (67)	0.84	—
Mean Number of comorbidities	1.4	1.6	1.3	<0.0001	<0.005
Income ≥ $20,000	570 (49)	212 (56)	358 (45)	<0.01	0.05
Education					
Less than high school	710 (61)	100 (26)	610 (77)		
High School graduate	461 (39)	280 (74)	181 (23)	<0.0001	0.26
Employment					
Working	677 (58)	208 (55)	469 (59)		
Not working	498 (42)	169 (45)	320 (41)	0.17	0.12
Currently driving	909 (78)	336 (88)	573 (72)	<0.0001	0.57

Age data are the mean ± SD. The remaining data are the Number of subjects, with the percentage of the total group in parentheses.

* Probabilities are from χ^2 (discrete variables) and Student's t-tests (continuous variables) contrasting Spanish-speaking versus English-speaking Latinos.

† Age-adjusted probabilities are from Mantel-Haenszel χ^2 test for categorical variables contrasting Spanish-speaking versus English-speaking Latinos.

tus, and number of comorbidities). We then assessed the normality of the residuals in the parametric ANCOVA and determined that after adjusting for covariates, the distribution of the subscales remained nonnormal.

Because the self-reported visual functioning subscales were skewed, nonparametric statistics were used to compare the NEI VFQ-25 subscale scores between Spanish- and English-speaking Latinos to determine any systematic difference. For this analysis, subscale scores were first rank ordered. This rank was used as the dependent variable in a traditional ANCOVA, adjusting for the standard covariates (age, gender, income, education, employment status, and number of comorbidities), to determine whether differences in median scores were statistically significant.[22]

RESULTS

The Questionnaire

Of the initial 1917 participants who completed both a home questionnaire and a clinical examination, 1171 fulfilled the visually normal criteria (no visual impairment, normal visual field, and no ocular disease; Table 1). The participation rate for the sample in this analysis was 82% (nonparticipants were more likely to be younger, uninsured, and satisfied with their vision). Most of the participants were Mexican American (11% were born in countries other than the United States or Mexico). The majority of participants (67%) chose to complete the questionnaire in Spanish. None of the participants changed the language spoken during the administration of the questionnaire. Spanish-speaking Latinos were slightly younger ($P < 0.001$) and had a lower income than English-speaking Latinos ($P < 0.01$). Spanish speakers were less likely to have finished college and less likely to be employed, and they had significantly more comorbid conditions compared with English speakers (all $P < 0.0001$). After adjusting for age, differences in the number of comorbidities and household income remained statistically significant ($P \leq 0.05$).

Psychometric Analysis

Tables 2 through 5 summarize the findings from the psychometric analysis. We had complete data for more than 99% (1160) of the 1171 participants. Because 23% of participants in the reference group were not currently driving, only 77% of the participants had a Driving subscale score. Psychometric results were similar when visually impaired participants were included in the analysis.

Ceiling and Floor Effects. A ceiling effect (>20% of scores at the highest possible score) was evident for 9 of the 12 subscales (Table 2). English-speaking Latinos had a significantly higher ceiling effect on three subscales: General Vision, Ocular Pain, and Vision-Specific Mental Health. Conversely, Spanish speakers had a significantly higher ceiling effect on the Distance Vision and Vision-Specific Role Function subscales, compared with English speakers. In contrast, no subscales had a floor effect (>20% at the lowest possible score), as expected.

Subscale Internal Consistency Reliability. The Cronbach α[21] for assessing reliability of the subscale's internal consistency ranged from a low of 0.24 (Driving subscale, English speaker), to a high of 0.82 (Vision-Specific Dependency, Spanish speakers; Table 3). The acceptable minimum Cronbach α of 0.70 was achieved for only two of the eight subscales for which this could be computed. The Vision-Specific Dependency subscale met the minimum criteria for internal consistency in both the English and the Spanish versions, as well as the complete cohort. Vision-Specific Mental Health met the minimum criteria for the English speakers, whereas Vision-Specific Role Function achieved the minimum criteria for the Spanish speakers and the complete cohort. The subscale's internal consistency was marginal (Cronbach $\alpha \geq 0.60$ and < 0.70) for two subscales in both the English- and Spanish-speaking groups (Ocular Pain and Near Vision). Subscale internal consistency was also marginal for the Vision-Specific Mental Health subscale in Spanish-speaking Latinos and for Vision-Specific Role Function in English-speaking Latinos. In addition, the subscale's internal consistency was poor (Cronbach $\alpha < 0.60$) in three subscales (Driving Difficulties, Vision-Specific Social Function, and Distance Vision) in both the English- and Spanish-speaking groups.

Item Internal Consistency and Discriminant Validity. The correlation coefficients for item internal consistency and discriminant validity ranged from 0.18 to 0.78 (Tables 4, 5). Seven of 21 (33%) items did not have item–subscale correlations that correlated more strongly with their own subscales than with every other subscale in the questionnaire. When item validity was assessed, Spanish-speaking Latinos were less likely to achieve minimum internal consistency (62% achieved the minimum of 0.4) compared with English-speaking Latinos (81% achieved the minimum). Similarly, items completed by Spanish speakers were less likely to achieve minimum item discriminant validity (62%) compared with items completed by English-speaking Latinos (76%). Correlations between all items and scales are presented in Table 5. Two Distance Vision items

TABLE 2. Distribution of NEI VFQ-25 Subscale Ceiling Effects in the Study Group Stratified by Language Spoken

Subscale	Complete Group (n = 1171)	English (n = 380)	Spanish (n = 791)	P
Ceiling effect (%)*				
General Health	**7.6**	**8.7**	**7.1**	0.33
General Vision	**9.3**	**13.7**	**7.2**	<0.001
Ocular Pain	26.4	34.0	22.8	<0.0001
Near Vision	32.6	30.0	33.9	0.18
Activities				
Distance Vision Activities	50.0	40.3	54.7	<0.0001
Vision-Specific Social Function	80.1	80.3	80.0	0.92
Vision-Specific Mental Health	**13.5**	**18.2**	**11.3**	<0.01
Vision-Specific Role Function	69.8	59.2	74.8	<0.0001
Vision-Specific Dependency	72.0	70.5	72.7	0.44
Driving Difficulties‡	39.2	39.2	39.2	0.99
Color Vision	88.9	86.8	89.6	0.16
Peripheral Vision	74.4	72.6	75.2	0.34
Floor Effect†				
No floor effects were found (range = 0.0 and 5.0%)				

*Ceiling effect: percent of participants with scores at the maximum scale score (100), indicating no difficulty in this area. For instance, a score of 100 on the Ocular Pain scale indicates that the respondent has no ocular pain. A ceiling effect is evident when 20% or more of the participants score at the highest possible score. Subscales without ceiling effect are set in bold.

† Floor effect: 20% percent or more of participants with scores at the minimum scale score (0), indicating extreme difficulty in this area.

‡ Scores could be generated for only 336 participants in the English-speaking group and 574 patients in the Spanish-speaking group for the Driving Difficulties subscale. Only those participants who are currently driving complete this subscale.

and both Driving items did not achieve either minimum reliable internal consistency or discriminant validity for both Spanish- and English-speakers. The Vision-Specific Mental Health scale did not achieve minimum criteria for the Spanish-speakers. One item ("How much of the time do you worry about your eyesight?") did not have minimum reliable internal consistency, whereas the other three Mental Health items ("I feel frustrated a lot of the time because of my eyesight"; "I have much less control over what I do because of my eyesight"; "I worry about doing things that will embarrass others or myself because of my eyesight") correlated more strongly with the Vision-Specific Dependency subscale than with the Mental Health subscale. One Vision-Specific Mental Health item ("I worry about doing things that will embarrass others or myself because of my eyesight") was more strongly correlated with Vision-Specific Dependency in the English speakers as well.

Comparison of Subscale Scores in English- and Spanish-Speaking Latinos. Adjusted NEI VFQ-25 median scores were significantly different between visually normal Spanish- and English-speaking Latinos (Table 6). Spanish speakers had significantly lower median General Health, General Vision, Vision-Specific Mental Health, and Ocular Pain scores after adjusting for differences in age, gender, income, employment status, education, and the number of comorbidities between the two groups. The magnitude of these differences ranged from 4 to 14.1 percentage points. No differences between median subscale scores were evident between interviewers.

DISCUSSION

Quality-of-life visual-functioning instruments must be culturally specific to assess reliably the association between ocular dis-

TABLE 3. Distribution of Cronbach α for NEI VFQ-25 Subscales for the Study Group Stratified by Language Spoken

Subscale	Complete Group (n = 1171)	English (n = 380)	Spanish (n = 791)
Ocular Pain	0.64*	0.66*	0.62*
Near Vision	0.64*	0.67*	0.63*
Distance Vision	0.51†	0.56*	0.51†
Vision-Specific Social Function	0.50†	0.59†	0.47†
Vision-Specific Mental Health	0.68*	0.79	0.63*
Vision-Specific Role Function	0.77	0.65*	0.81
Vision-Specific Dependency	0.81	0.78	0.82
Driving Difficulties	0.30†	0.24†	0.34†

Cronbach's α is a measure of internal consistency of the subscale. It is calculated for subscales with two or more items. Acceptable minimum Cronbach α = 0.70.

* Marginal subscale internal consistency, Cronbach $\alpha \geq 0.60$ and < 0.70

† Poor subscale internal consistency, Cronbach $\alpha < 0.60$.

Cronbach α cannot be calculated for General Health, General Vision, Color Vision and Peripheral Vision, because each of these subscales contains only one item.

TABLE 4. Item Internal Consistency and Item Discriminant Validity for Each Question of the NEI VFQ-25

Subscale	Items	Item Internal Consistency			Item Discriminant Validity		
		Complete Group	English	Spanish	Complete Group	English	Spanish
Driving	1. How much difficulty do you have driving during the daytime in familiar places?	0.28*	0.24*	0.31*	0.36†	0.36†	0.41†
	2. How much difficulty do you have driving at night?	0.28*	0.24*	0.31*	0.43†	0.49†	0.42†
Ocular Pain	1. How much pain or discomfort have you had in and around your eyes?	0.47	0.50	0.45			
	2. How much does pain or discomfort in or around your eyes keep you from doing what you'd like to be doing?	0.47	0.50	0.45			
Near Vision Activities	1. How much difficulty do you have reading ordinary print in newspapers?	0.40	0.48	0.38*			
	2. How much difficulty do you have doing work or hobbies that require you to see well up close?	0.54	0.59	0.53			
	3. How much difficulty do you have finding something on a crowded shelf?	0.46	0.42	0.48			
Distance Vision Activities	1. How much difficulty do you have reading street signs or the names of stores?	0.37*	0.42	0.38*			
	2. How much difficulty do you have going down steps, stairs, or curbs in dim light or at night?	0.40	0.38*	0.43	0.58†	0.56†	0.58†
	3. How much of the time does your vision limit you in recognizing people or objects across the street?	0.33*	0.32*	0.35*			
Social Functioning	1. How much difficulty do you have seeing how people react to things you say?	0.35*	0.45	0.32*	0.50†		0.53†
	2. How much difficulty do you have visiting with people in their homes, at parties, or in restaurants?	0.35*	0.45	0.32*	0.42†		0.42†
Mental Health	1. How much of the time do you worry about your eyesight?	0.25*	0.45	0.18*			
	2. I feel frustrated a lot of the time because of my eyesight.	0.60	0.74	0.55			0.66†
	3. I have much less control over what I do because of my eyesight.	0.59	0.68	0.55			0.66†
	4. I worry about doing things that will embarrass others or myself because of my eyesight.	0.53	0.56	0.51	0.67†	0.71†	0.66†
Role Function	1. Do you accomplish less than you would like because of your vision?	0.63	0.49	0.70		0.60†	
	2. Are you limited in how long you can work or do other activities because of your vision?	0.63	0.49	0.70			
Dependency	1. I stay home most of the time because of my eyesight.	0.51	0.50	0.52			
	2. Because of my eyesight, I have to rely too much on what other people tell me.	0.75	0.68	0.78			
	3. I need a lot of help from others because of my eyesight.	0.75	0.72	0.76			

Item internal consistency and item discriminant validity represent the fit of each question (item) with the subscale to which it has been assigned. Subscales must contain more than one item for this item level analysis to be completed. Only subscales meeting this criteria (more than one item) are displayed in the table. For less than desirable item internal consistency, item-scale correlation is less than 0.4 (*). Item discriminant validity is less than desirable if item correlation with another subscale is significantly higher than item correlation with its own subscale (†). Every item should correlate with its own subscale for the entire instrument to achieve discriminant validity. Item internal consistency and item discriminant validity could not be calculated for Peripheral Vision, Color Vision, General Vision, and General Health because these subscales have only one item. When no data are shown, item discriminant validity met minimum criteria—that is, item correlation with any other subscale is not significantly higher than item correlation with its own scale. *n* as in Tables 1–3.

ease and self-reported health status. These measures can also be used to evaluate the effectiveness of treatment and assess the impact of disease on individual functioning. These evaluations can be performed in various ethnic groups and may be used as an alternative outcome measure in clinical intervention studies. They may also help determine reimbursement levels by health insurance plans and guide public health policies.[23] The first step in determining whether a particular self-reported health status instrument can be used, especially in diverse populations and communities, is to assess the reliability and validity of the measure. Although a psychometric analysis has been conducted in English speakers [2]and an analysis has been performed in Spanish-speaking Latinos,[4] this study expands on the previous work by presenting a detailed psychometric analysis between Spanish- and English-speaking, visually normal Latinos. Limited- or non–English speakers have been identified

TABLE 5. Pearson Item-Scale Correlations Corrected for Overlap

	Driving		Distance Vision Activities		Social Function		Ocular Pain		Near Vision Activities		Mental Health		Role Function		Dependency		General Health		General Vision		Color Vision		Peripheral Vision	
Item	**Eng**	**Span**	**Eng**	**Span**	**Eng**	**Span**	**Eng**	**Span**	**Eng**	**Span**	**Eng**	**Span**	**Eng**	**Span**	**Eng**	**Span**	**Eng**	**Span**	**Eng**	**Span**	**Eng**	**Span**	**Eng**	**Span**
Driving																								
DR1	**0.24***	**0.31***	0.36†	0.36	0.26	0.33	0.21	0.30	0.17	0.27	0.24	0.31	0.25	0.41†	0.21	0.34	0.03	0.15	0.17	0.19	0.11	0.29	0.29	0.39
DR2	**0.24***	**0.31***	0.49†	0.41†	0.25	0.36	0.20	0.36	0.26	0.30	0.31	0.29	0.31	0.42†	0.30	0.36	0.16	0.18	0.17	0.30	0.18	0.19	0.40†	0.39
Distance Vision Activities																								
DV1	0.35	0.37	**0.42**	**0.38***	0.29	0.34	0.24	0.26	0.42	0.51†	0.31	0.29	0.33	0.26	0.28	0.29	0.07	0.24	0.34	0.28	0.21	0.31	0.40	0.43
DV2	0.51†	0.45†	**0.38***	**0.43**	0.35	0.40	0.30	0.31	0.37	0.44	0.32	0.33	0.39	0.36	0.40	0.30	0.23	0.29	0.17	0.27	0.20	0.35	0.56†	0.58†
DV3	0.32	0.28	**0.32***	**0.35***	0.45†	0.38	0.25	0.23	0.35	0.28	0.32	0.25	0.34	0.25	0.36	0.22	0.19	0.18	0.19	0.21	0.22	0.20	0.29	0.27
Social Functioning																								
SF1	0.26	0.34†	0.47	0.37	**0.45**	**0.32***	0.28	0.28	0.39	0.41†	0.35	0.32	0.35	0.38	0.37	0.34	0.08	0.19	0.17	0.20	0.46	0.53†	0.45	0.53†
SF2	0.23	0.31	0.40	0.42†	**0.45**	**0.32***	0.34	0.26	0.33	0.29	0.36	0.22	0.43	0.33	0.49	0.25	0.19	0.17	0.11	0.14	0.42	0.26	0.35	0.29
Ocular Pain																								
OP1	0.22	0.26	0.31	0.23	0.23	0.25	**0.50**	**0.45**	0.32	0.28	0.34	0.31	0.33	0.26	0.35	0.22	0.12	0.15	0.28	0.18	0.26	0.18	0.19	0.25
OP2	0.18	0.42	0.32	0.35	0.43	0.31	**0.50**	**0.45**	0.37	0.35	0.41	0.48	0.45	0.51	0.46	0.40	0.20	0.21	0.19	0.30	0.36	0.27	0.30	0.36
Near Vision Activities																								
NV1	0.11	0.20	0.31	0.33	0.22	0.20	0.28	0.22	**0.48**	**0.38***	0.31	0.19	0.30	0.20	0.29	0.22	0.11	0.18	0.41	0.33	0.26	0.17	0.17	0.20
NV2	0.27	0.30	0.42	0.39	0.38	0.36	0.32	0.33	**0.59**	**0.53**	0.33	0.36	0.42	0.39	0.31	0.29	0.15	0.26	0.33	0.34	0.35	0.28	0.37	0.44
NV3	0.30	0.33	0.51	0.45	0.40	0.47	0.34	0.34	**0.42**	**0.48**	0.38	0.41	0.44	0.41	0.42	0.39	0.13	0.26	0.23	0.31	0.39	0.37	0.46	0.50
Mental Health																								
MH1	0.22	0.06	0.28	0.12	0.21	0.09	0.36	0.17	0.28	0.09	**0.45**	**0.18***	0.35	0.13	0.33	0.14	0.13	0.06	0.39	0.12	0.16	0.06	0.24	0.13
MH2	0.24	0.29	0.35	0.31	0.29	0.25	0.34	0.41	0.39	0.35	**0.74**	**0.55**	0.51	0.48	0.61	0.66†	0.15	0.20	0.33	0.26	0.22	0.17	0.37	0.25
MH3	0.33	0.41	0.35	0.35	0.45	0.32	0.32	0.43	0.35	0.40	**0.68**	**0.55**	0.59	0.53	0.61	0.66†	0.13	0.19	0.29	0.24	0.28	0.22	0.44	0.33
MH4	0.26	0.33	0.40	0.32	0.40	0.31	0.35	0.33	0.33	0.33	**0.56**	**0.51**	0.46	0.48	0.71†	0.65†	0.16	0.18	0.24	0.23	0.33	0.23	0.37	0.30
Role Function																								
RF1	0.33	0.46	0.42	0.34	0.35	0.39	0.39	0.41	0.45	0.38	0.60†	0.49	**0.49**	**0.70**	0.49	0.49	0.11	0.22	0.29	0.32	0.30	0.29	0.45	0.36
RF2	0.25	0.45	0.42	0.34	0.46	0.41	0.39	0.40	0.38	0.36	0.43	0.51	**0.49**	**0.70**	0.46	0.53	0.15	0.22	0.24	0.31	0.33	0.32	0.41	0.34
Dependency																								
DP1	0.28	0.36	0.31	0.32	0.44	0.35	0.32	0.38	0.24	0.28	0.50	0.57	0.32	0.44	**0.50**	**0.52**	0.24	0.18	0.18	0.19	0.28	0.26	0.35	0.26
DP2	0.28	0.36	0.45	0.28	0.44	0.28	0.46	0.28	0.39	0.33	0.62	0.63	0.55	0.50	**0.68**	**0.78**	0.16	0.17	0.23	0.24	0.33	0.23	0.50	0.27
DP3	0.23	0.36	0.42	0.27	0.42	0.30	0.36	0.28	0.42	0.35	0.62	0.60	0.50	0.48	**0.72**	**0.76**	0.13	0.16	0.19	0.23	0.37	0.27	0.40	0.32

Bold values are item-scale correlation corrected for overlap (relevant item removed from its scale for correlation). The correlations are also hypothesized to be highest in the same row. For less than desirable item internal consistency, item-scale correlation is less than 0.4 (*). For less than desirable item discriminant validity, item correlation with competing scale is significantly higher than its correlation with its own scale (†). For item labels, please refer to Table 4. Eng, English; Span, Spanish.

TABLE 6. Adjusted Median NEI VFQ-25 Subscale Scores for the Study Group Stratified by Language Spoken

Subscale	English (n = 380)	Spanish (n = 791)	P‡
Color Vision	95.7	97.0	0.31
Vision-Specific Dependency	94.0	93.1	0.50
Driving Difficulties*	88.9	90.9	0.29
Distance Vision	**88.2**	**90.4**	0.02
General Health	**56.7**	**48.7**	<0.0001
General Vision	**74.2**	**68.2**	<0.0001
Vision-Specific Mental Health	**82.2**	**78.9**	0.006
Near Vision	83.0	82.4	0.97
Ocular Pain	**83.0**	**78.8**	0.007
Peripheral Vision	91.2	91.7	0.98
Vision-Specific Role Function	**89.6**	**92.0**	0.0038
Vision-Specific Social Function	96.0	95.4	0.30
Composite†	86.3	85.1	0.12

*Scores could be generated for only 336 participants in the English-speaking group and 574 patients in the Spanish-speaking group for the Driving Difficulties subscale. Only those participants who are currently driving completed this subscale.

† Composite score is an unweighted mean of the 25 individual item scores.

‡ Adjusted for age, gender, comorbidities, education, employment status, and income.

Data are expressed as the median.

as groups that have not been extensively studied when assessing the reliability and validity of health-related quality-of-life instruments.[24] In this study, we assessed the reliability and comparative performance of the NEI VFQ-25 in English- and Spanish-speaking, visually normal Latinos.

Overall, the psychometric performance of the NEI VFQ-25 in terms of reliable internal consistency in both Spanish- and English-speaking Latinos was found to be acceptable. The Cronbach α was acceptable for two subscales in both English (Vision-Specific Dependency and Vision-Specific Mental Health) and Spanish (Vision-Specific Dependency and Vision-Specific Role Function) at the greater than 0.70 level. The Cronbach α was low for a few scales. In particular, the Driving scale had a very low Cronbach α. Similarly, Broman et al.[4] found Cronbach $\alpha < 0.70$ in the Driving and Ocular Pain Subscales in the entire cohort of the population-based Proyecto VER (Video en Red) study. However, Broman et al. did not provide separate data on internal consistency for the 80% Spanish speakers, compared with the 20% English speakers. Recently, in an effort to improve internal consistency of the Driving subscale, Mangione et al.[1] have added an additional item to the Driving subscale in both the English and Spanish versions.

The marginal and poor reliability of internal consistency may suggest that the subscale constructs are not homogenous or that the items are not appropriate measures of these constructs for Latinos.[25] It is likely that supplementing other subscales of the NEI VFQ-25 with additional items from the NEI VFQ-51 and eliminating poor items may improve the reliability of the internal consistency of this questionnaire when used in both Spanish- and English-speaking Latinos.[26]

In our study, psychometric analysis at the item level revealed that 7 of 21 items (number of items in subscales with at least two items) failed to display reliable item internal consistency or item discriminant validity. This trend was more notable for Spanish-speaking than for English-speaking Latinos. Broman et al.[4] report that there were differences between Spanish and English speakers in their responses to the NEI VFQ-25. Although they also found poor reliability of item internal consistency, they were able to identify only one item that had poor item discriminant validity (worry about eyesight). A direct comparison of the two studies is difficult, because Broman et al. did not report reliable internal consistency or item discriminant validity at the item level, by language spoken. Some of the possible explanations for our findings include cultural inappropriateness of the instrument, poor translation of the instrument, too many or too few item response choices, and a visually normal sample. As suggested earlier, one suggestion for improving item consistency is to increase the number of items in each subscale. Increasing the number of items in a subscale may improve the item consistency by providing a more accurate representation of the underlying construct that the subscale is measuring. Too few items may not reflect the full breadth of the construct and thus appear to be disparate. Of note, when Broman et al. used an inflation factor to adjust the reliability of item internal consistency for subscales with very few items, the one item that had poor reliability did not achieve acceptable reliability (0.30). The shorter version of the NEI VFQ, the NEI VFQ-25, was developed to decrease the time burden of administering the NEI VFQ without sacrificing any of the essential subscales. Our analysis suggests that this reduced number of test items has poor reliable internal consistency among Latinos.[14,17,26,] Other explanations for these findings include problems with translation, respondents' inability to understand question stems or response choices, and use of items that are not culturally appropriate. Our findings that these problems are present in both English- and Spanish-speaking Latinos suggest that the issue is less language dependent and more a function of culture or the sample. The finding that the comprehensibility of the Spanish translation and item stems was found to be appropriate by the LALES focus groups[27] supports this conclusion. However, both Spanish- and English-speaking respondents in these focus groups reported that the response choices for items in the NEI VFQ-25 were too numerous and confusing. Another possible explanation may be the visually normal sample used for this analysis. When participants with visual impairment and/or abnormal visual fields were included in an analysis the same subscales continued to be below the threshold for minimal reliability.

The differences between our findings and those of Broman et al.[4] may be explained by differences in the heritage of the Latinos in the two studies. The heritage of the LALES population is predominantly Mexican, whereas the Proyecto VER population from Arizona is likely to have a higher proportion of Native American heritage.[32] Further evaluation of the impact of

these issues on the psychometric properties of the NEI VFQ-25 should be conducted.

That there was no ceiling effect in some of the NEI-VFQ subscales in this visually normal sample is surprising. Because the participants in this analysis had no evidence of visual impairment, we would expect more than 20% of the participants to score at the highest possible range on all 12 subscales. Indeed, in the initial NEI VFQ-51 field test, 23% to 95% of the subscale scores of individuals with good visual acuity and visual fields were at the ceiling (all subscales except General Health [12% at ceiling] and General Vision [9% at ceiling]). Broman et al.[4] did not quantify ceiling effects by subscale in the Proyecto VER population, although they reported that items with response choices measuring degree of difficulty and time spent on an activity were more likely to exhibit a ceiling effect. In the LALES, scores for visually normal participants did not exhibit a ceiling effect for three of the 12 subscales. In addition, in five subscales, a higher proportion of English-speaking Latinos' scores were at the ceiling that of Spanish-speaking Latinos. English speakers were less likely to report any Ocular Pain or Vision-Specific Mental Health problems and were more likely to report problems with Distance Vision and Vision-Specific Role Function. These differences may reflect a difference in the cultural appropriateness and cultural relevance of different items to Spanish- and English-speaking participants. English-speaking Latinos may find certain items more or less meaningful in the context of their visual functioning compared with Spanish speakers. This would be reflected in differences, by language spoken, in which items are more or less frequently rated to reflect impairment or perfect visual functioning.[24] The interpretation of this result would be enhanced by comparison of NEI VFQ-25 data in an English-speaking, non-Latino cohort of similar demographic and clinical characteristics. This comparison would clarify whether Latinos with no visual impairment have lower scores, regardless of language spoken, or whether this trend is limited to Spanish-speaking Latinos.

One explanation of the lower health status reported by Latinos is that this perceived poor health status is the result of racial/ethnic discrimination, culture, and other factors.[19] Indeed, in our study, Spanish-speaking participants had statistically significantly lower median scores on 4 of the 12 NEI VFQ-25 subscales compared with English-speaking Latinos. This is consistent with the report of lower self-reported general quality of life, as measured by the SF-36, in other Latino cohorts.[26] The association between language and attitudes and beliefs toward health and health care has been well documented in the literature.[29,30] Few quality-of-life studies have assessed the relationship with ethnicity, although an individual's perception and response to illness is strongly determined by cultural factors.[31] The value that different cultures allocate to the constructs measured by quality of life may differ, and the findings from our study may reflect these cultural differences. Because spoken language is a component of most measures of acculturation,[29,32] the trend for Spanish speakers to score lower suggests that sociocultural issues play a significant role in determining self-perceived visual function.

One suggested explanation for the less than optimal scale of reliability for internal consistency, item internal consistency, and discriminant validity for some of the NEI-VFQ subscales in this sample may be a large proportion of cognitive impairment that is associated with increasing age. The Cognitive Abilities Screening Instrument (CASI)[33] was administered as a cognitive status screen to a subsample of participants in the LALES. The frequency of cognitive impairment was 2.5% ($n = 58$). This low prevalence of cognitive impairment is unlikely to impact the results of this analysis.

Our findings demonstrate less than optimal performance of the NEI VFQ-25 in this visually normal Latino cohort. Several subscales achieved a less than optimal reliable subscale internal consistency, reliable item internal consistency, and discriminant validity. From the current analysis, it is difficult to determine whether these poor psychometric properties are due to the number and phrasing of the items and item response choices or to the cultural appropriateness of the items in general. Potential solutions to help improve the performance of the NEI VFQ-25 include revision or reduction of items, reduction of response choices, and controlling for other characteristics such as clinical health or mental status. In addition, item response analysis provides more information concerning item difficulty and the performance of individual item response choices.[34] These studies are needed to conclude whether the NEI VFQ-25 is a valid measure of self-reported visual functioning for Latinos in population-based studies. The NEI VFQ-25 was designed to capture the impact of visual impairment on visual functioning along the entire spectrum of visual acuity.[2] It is as important to understand the performance of this instrument as a measure of visual functioning along this continuum, including individuals without visual impairment.

References

1. Mangione C, Lee P, Guiterrez P, Spitzer K, Berry S, Hays R. Development of the 25-Item National Eye Institute Visual Function Questionnaire. *Arch Ophthalmol.* 2001;119:1030–1058.
2. Mangione C, Lee P, Pitts J, Gutierrez P, Berry S, Hays R. Psychometric properties of the National Eye Institute Visual Function Questionnaire (NEI-VFQ). *Arch Ophthalmol.* 1998;116:1496–1504.
3. Mangione C, Berry S, Spritzer K, et al. Identifying the content area for the 51-item National Eye Institute Visual Function Questionnaire: results from focus groups with visually impaired persons. *Arch Ophthalmol.* 1998;116:227–233.
4. Broman AT, Munoz B, West SK, et al. Psychometric properties of the 25-item NEI-VFQ in a Hispanic population: Proyecto VER. *Invest Ophthalmol Vis Sci.* 2001;42:606–613.
5. National Center for Health Statistics. National Health and Nutrition Examination Survey: Survey questionnaires, examination components and laboratory components. Hyattsville, MD: National Center for Health Statistics; 1999.
6. Revicki DA, Osoba D, Fairclough D. Recommendations on health-related quality of life research to support labeling and promotional claims in the United States. *Qual Life Res.* 2000;9:887–900.
7. Cox DR, Fitzpatrick R, Fletcher AE, Gore SM, Speifelhalter DJ, Jones DR. Quality of life assessment: can we keep it simple? *J R Stat Soc.* 1992;155:353–393.
8. Scott-Lennox JA, Wu AW, Boyer JG, Ware JE. Reliability and validity of French, German, Italian, Dutch, and UK English translations of the Medical Outcomes Study HIV Health Survey. *Med Care.* 1999;37:908–925.
9. Berry JW. On cross-cultural comparability. *Int J Psychol.* 1969;4:207–229.
10. Solis J, Marks G, Garcia M, Shelton D. Acculturation, access to care and use of preventive services by Hispanics: findings from HHANES 1982–1984. *Am J Public Health.* 1990;80:11–19.
11. Marin G, Marin BO. *Applied Social Research Methods Series.* Newbury Park, CA: Sage; 1991:66–79.
12. Gandek B, Ware JD, Group IP. Methods for validating and norming translations of health status questionnaires: the IQOLA project approach. *J Clin Epidemiol.* 1998;51:953–959.
13. Varni JW, Seid M, Kurtin PS. PedsQL 4.0: reliability and validity of the Pediatric Quality of Life Inventory Version 4.0 Generic Core Scales in healthy and patient populations. *Med Care.* 2001;39:800–812.
14. Ware J, Kosinsk M, Keller S. A 12-item short-form health survey: construction of scales and preliminary test of reliability and validity. *Med Care.* 1996;34:220–226.
15. Leplege A, Ecosse E, Verdier A, Perneger T. The French SF-36 health survey: translation, cultural adaptation and preliminary psychometric evaluation. *J Clin Epidemiol.* 1998;51:1013–1023.

16. Azen S, Preston-Martin S, Globe D, Ying-Lai M, Hahn S, Varma R. Binocular visual acuity summation and inhibition in an ocular epidemiological study: the Los Angeles Latino Eye Study. *Invest Ophthalmol Vis Sci.* 2002;43:1742–1748.
17. Hays R, Hayashi T. Beyond internal consistency reliability: rationale and user's guide for Multitrait Analysis Program on the microcomputer. *Behav Res Methods Instruct.* 1990;22:167–175.
18. Cunningham WE, Bozzette SA, Hays RD, Knouse DE, Shapiro MF. Comparison of health-related quality of life in clinical trials and non clinical trial human immunodeficiency virus-infected cohorts. *Med Care.* 1995;33:AS15–AS25.
19. Cunningham WE, Hays RD, Burton TM, Kington RS. Health status measurement performance and health status differences by age, ethnicity, and gender: assessment in the medical outcomes study. *J Health Care Poor U.* 2000;11:58–76.
20. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika.* 1951;16:297–334.
21. Streiner DL, Norman GR. *Health Measurement Scales: A Practical Guide to Their Development and Use*. 2nd ed. New York: Oxford University Press; 1998.
22. Conover WJ, Iman RL. Rank transformations as a bridge between parametric and nonparametric statistics. *Am Stat.* 1981;35:124–129.
23. Patrick DL, Chiang Y. Measurement of health outcomes in treatment effectiveness evaluations; conceptual and methodological. challenges. *Med Care.* 2000;38:II14–II25.
24. Stewart A, Napoles-Springer A. Health-related quality-of-life assessments in diverse population groups in the United States. *Med Care.* 2000;38:II102–II124.
25. Kressin NR, Clark JA, Whittle J, et al. Racial differences in health-related beliefs, attitudes, and experiences of VA cardiac patients: scale development and application. *Med Care.* 2002;40:I72–I85.
26. Hays RD, Anderson RT, Revicki D. Assessing reliability and validity of measurement in clinical trials. In: Staquet M, Hays R, Fayers P, eds. *Quality of Life Assessment in Clinical Trials: Methods and Practice.* Oxford: Oxford University; 1998:169–175.
27. Globe D, Shoua-Glusberg A, Paz S, et al. Using focus groups to develop a culturally sensitive methodology for epidemiological surveys in a Latino population: findings from the Los Angeles Latino Eye Study Focus Groups. *J Ethnicity Dis.* 2002;12:259–266.
28. West SK, Munoz B, Klein R, et al. Risk factors for type II diabetes and diabetic retinopathy in a Mexican-American population: Proyecto VER. *Am J Ophthalmol.* 2002;134:390–398.
29. Deyo R, Diehl AK, Hazuda H, Stern MP. A simple language-based acculturation scale for Mexican Americans: validation and application in health care research. *Am J Public Health.* 1985;75:51–55.
30. Meyerowitz BE, Richardson J, Hudson S, Leedham B. Ethnicity and cancer outcomes: behavioral and psychosocial considerations. *Psychol Bull.* 1998;123:47–70.
31. Hutchinson JF. Quality of life in ethnic groups. In: Spilker B, ed. *Quality of Life and Pharmacoeconomics in Clinical Trials.* 2nd ed. Philadelphia: Lippincott-Raven; 1996:587–593.
32. Marin G, Sabogal F, Marin BV, Otero-Sabogal R, Perez-Stable EJ. Development of a short acculturation scale for Hispanics. *Hisp J Behav Sci.* 1987;9:183–205.
33. Teng EL, Hassegawa K, Homma A, et al. The Cognitive Abilities Screening Instrument (CASI): a practical test for cross-cultural epidemiological studies of dementia. *Int Psychogeriatr.* 1994;6:45–58.
34. Zhu W, Updyke WF, Lewandowski C. Post-hoc Rasch analysis of optimal categorization of an ordered-response scale. *J Outcome Meas.* 1997;1:286–304.

APPENDIX I

The Los Angeles Latino Eye Study Group

LaVina Abbott, Elisa Arevalo, Stanley P. Azen, Lupe Cisneros, Elizabeth Corona, Carolina Cuestas, Denise R. Globe, Sora Hahn, Ronald Klein, Mei Lai, George Martinez, Marta Mora, Sylvia Paz, Susan Preston-Martin, Ronald E. Smith, Mina Torres, Natalia Uribe, Rohit Varma, and Myrna Zuniga.