Otolaryngology–Head and Neck Surgery (2006) 135, 603-607

ORIGINAL RESEARCH

Evaluating the validity of the Voice Handicap Index-10 (VHI-10) among Hebrew speakers

Ofer Amir, PhD, Yael Tavor, BA, Tali Leibovitzh, BA, Odelia Ashkenazi, BA, Orit Michael, BA, Adi Primov-Fever, MD, and **Michael Wolf, MD,** Tel-Aviv and Tel-Hashomer, Israel

OBJECTIVE: To evaluate the validity of a translated version of the Voice Handicap Index-10 (VHI-10).

STUDY DESIGN AND SETTING: In a parallel group design, 221 patients with different laryngeal pathologies and 172 people with no laryngeal pathology completed a Hebrew version of the VHI-10. Validity and reliability were assessed as well as group differences.

RESULTS: Statistical analyses demonstrated high reliability values (Cronbach's Alpha r = 0.949). Responses were not affected by age ($P = 0.373$) or gender ($P = 0.360$). The control group received significantly lower scores than all pathological groups ($P < 0.05$). Within the pathological groups, the "neurogenic" and "mucosa irregularity" groups were rated higher than all other pathological groups ($P < 0.05$).

CONCLUSION: The VHI-10 questionnaire maintains its validity and reliability across translation to Hebrew. Moreover, although the VHI-10 is essentially a unidimensional tool, it provides partial information on the 3 subjective dimensions of the full VHI.

Voice disorders can be evaluated with the use of instrumental as well as perceptual approaches. The clinical merit of these approaches for voice evaluation has been established previously.[1-4] However, measures such as stroboscopy, electromyography, acoustic analyses, and even listeners' evaluations were shown to be insufficient for assessing the level of disability experienced by the speaker as a function of a voice disorder.[5] Therefore, because voice problems are viewed as a multidimensional disorder, the quantification of the patient's self-perception of the problem is considered an important factor in the assessment of the voice disorders, along with the physiological and perceptual assessment.

Although various instruments have been developed for self-evaluation of voice problems, the Voice Handicap Index[6] (VHI) has been widely recognized and accepted for research as well as for clinical application.[7] Since then, the VHI has been translated and adapted to many languages[8-13] and was demonstrated to reliably quantify the subjective perception of handicap experienced by voice patients.

Recently, a shortened form of the VHI has been published, the VHI-10.[14] This 10-item questionnaire was shown to provide a valid representation of the patient's self-evaluation of his or her handicap that was comparable with the results obtained in the full version of the VHI.[14] In addition, the VHI-10 was assessed in comparison with the Vocal Performance Questionnaire (VPQ)[15] and was shown to be a consistent and valid unidimensional tool. However, although the VHI provides an overall handicap score and 3 subscores (functional, physical, and emotional), the VHI-10 provides only an overall score. Furthermore, although the VHI was translated and adapted to many different languages, the validity of the VHI-10 was never evaluated in a language other than English. Therefore, the purpose of the present study was 2-fold. First, it was intended to evaluate whether the VHI-10 would maintain its validity across trans-

From the Department of Communication Disorders (Drs Amir and Ms Tavor, Leibovitzh, Ashkenazi, and Mr Michael); Sackler School of Medicine, Tel-Aviv University, and the Department of Otorhinolaryngology–Head and Neck Surgery (Drs Primov-Fever and Wolf); The Sheba Medical Center, Tel-Hashomer, Israel.

This study was partially supported by the Shreiber Internal Fund, Tel-Aviv University; PI: Ofer Amir.

Reprint requests: Ofer Amir, PhD, Department of Communication Disorders, Tel-Aviv University, The Sheba Medical Center, Tel-Hashomer, 52621 Israel.

E-mail address: oferamir@post.tau.ac.il.

0194-5998/$32.00
doi:10.1016/j.otohns.2006.06.1252

lation. Second, we were interested to learn if the VHI-10 could provide information that would be comparable with the scores of the 3 subscales of the original VHI.

METHODS

Translation and Adaptation

The translation procedure of the VHI-10 to Hebrew was adopted from the procedure used previously for the VHI.[12] In essence, the original version of the VHI-10[14] was translated, independently, by 4 native-speakers of Hebrew, who are also highly proficient in spoken and written English. The resulting 4 Hebrew versions were then translated back to English by 4 independent native-speakers of English, who are also highly proficient in spoken and written Hebrew. The 4 retranslated English versions of the questionnaire were compared with the original version. Of the 4 translations, the items that were translated most accurately from English to Hebrew and then back to English were selected for inclusion in the final version of the questionnaire. Finally, the assembled Hebrew version was presented to 4 English-Hebrew bilingual judges, along with the original English questionnaire, to confirm that the 2 versions are comparable.

Subjects

A total of 393 participants (164 men and 229 women), with a minimum age of 16 years, were included in this study, after obtaining the approval from our institutional review board. Of these participants, 221 were included in the pathological group and 172 were included in the control group. The control group consisted of nondysphonic individuals who were recruited in the Tel-Aviv area. These individuals' responses were included in the study only if the subjects denied having any history of voice problems, complaints of voice function, loss of working days due to voice problems, or history of speech or voice therapy as well as singing lessons. In contrast, all patients included in the pathological group had a reported voice complaint and were examined by laryngologists in different medical centers in the Tel-Aviv area. All patients completed the questionnaire before they received any medical or voice treatment.

The participants in the pathological group were assigned to 6 pathological sub-groups on the basis of the diagnoses reported by the laryngologists after the laryngeal examinations. The "mass lesions" group consisted of patients diagnosed with nodules, polyps, cysts, or granuloma. The "inflammation" group consisted of patients with Reinke's edema, chronic laryngitis, swelling of the laryngeal mucosa, and gastro-esophageal reflux (GERD). The "mucosa irregularities" group consisted of patients who were diagnosed with disturbances or asymmetric mucosal wave patterns, sulcus vocalis, and superficial vocal fold scar. The "neurogenic" group consisted of patients with uni- or bilateral vocal fold paralysis or paresis and spasmodic dysphonia. The "functional" group consisted of patients who were diagnosed with normal structural larynx, but hyper- or hypofunctional laryngeal mobility patterns. Finally, the "other" group consisted of patients with different laryngeal pathologies or abnormalities that could not be assigned to any of the other groups or where the number of patients in each category was relatively small. Patients in this group were diagnosed with presbiphonia, glottic gap, laryngectomy, and patients who were postoperational with no current specific laryngeal findings. Table 1 presents data on gender, age, height, and weight distribution within each of the experimental groups.

Table 1
Gender, age (years), height (cm), and weight (kg) distribution within the different study groups

		Age		Height		Weight	
Group	Gender	Mean	SD	Mean	SD	Mean	SD
Mass lesions	Male (n = 20)	41.10	(12.33)	168.11	(24.37)	82.84	(25.63)
	Female (n = 58)	36.07	(10.62)	163.73	(8.53)	65.69	(18.67)
Inflammation	Male (n = 15)	42.60	(16.34)	171.21	(5.55)	81.93	(16.13)
	Female (n = 17)	46.35	(16.03)	163.46	(7.30)	61.00	(9.50)
Mucosa irregularity	Male (n = 4)	58.75	(0.96)	165.50	(7.59)	96.75	(55.34)
	Female (n = 10)	39.00	(13.63)	161.20	(4.39)	64.90	(10.60)
Neurogenic	Male (n = 20)	52.45	(13.81)	175.56	(7.82)	84.78	(25.24)
	Female (n = 7)	50.29	(12.02)	161.67	(9.31)	74.67	(24.99)
Functional	Male (n = 15)	34.67	(17.52)	172.81	(8.73)	73.00	(14.84)
	Female (n = 22)	36.23	(16.67)	161.73	(5.98)	61.81	(10.03)
Other	Male (n = 20)	60.45	(13.51)	172.65	(5.74)	77.56	(9.69)
	Female (n = 13)	50.23	(14.62)	160.60	(5.19)	61.82	(8.98)
Control	Male (n = 70)	39.50	(14.60)	176.43	(8.33)	79.64	(12.62)
	Female (n = 102)	37.25	(13.55)	162.73	(6.26)	62.46	(14.83)
Overall	Male (n = 164)	44.14	(16.31)	173.80	(11.39)	80.35	(18.54)
	Female (n = 229)	38.74	(13.96)	162.73	(6.90)	63.64	(15.27)

Table 2
Group mean scores, standard deviations, and range of scores on the VHI-10 and VHI questionnaires

	VHI-10			VHI		
Group	Mean	SD	Range	Mean	SD	Range
Control	2.43	2.99	0-16	7.51	7.55	0-37
Inflammation	14.54	10.71	0-40	37.50	29.10	2-114
Other	17.32	11.04	0-40	42.06	26.34	0-104
Functional	20.10	10.89	1-40	53.16	29.94	7-119
Mass lesions	19.69	10.12	0-40	55.08	26.30	0-115
Mucosa irregularity	20.71	9.83	5-40	59.54	24.43	16-116
Neurogenic	26.96	8.07	4-37	76.57	24.26	4-113

Administration of the Questionnaire

Each participant completed the Hebrew version of the VHI-10 independently. Only participants who were Hebrew speakers and readers were included in the study. Thus, a small number of candidates who were either illiterate or immigrants who could not read Hebrew proficiently enough by themselves to complete the questionnaire were excluded from the study. In addition to completing the VHI-10, all participants completed the following: a) the complete version of the Voice Handicap Index,[12] b) a medical and voice history information form, and c) 2 general questions with respect to their satisfaction with their voice. The first general question, "How much are you troubled by your voice?" was presented before the administration of the VHI-10. Participants were asked to respond to this question on a 7-point scale, where 1 was labeled "not at all," and 7 was labeled "very much." The second general question, "How satisfied are you with your voice?" was presented to the participants at the end, after completing the VHI-10, the medical and voice information form, and the VHI. Participants were asked to respond to this question on a 10-point scale, where 1 was labeled "completely dissatisfied" and 10 was labeled "highly satisfied."

RESULTS

Overall Scores and Group Differences

Mean scores and standard deviations of the VHI-10 obtained by the participants in the different experimental groups as well as their overall VHI scores are presented in Table 2. Several preliminary observations can be made. The VHI and the VHI-10 presented group differences similarly. With both questionnaires, the control group received markedly lower values than all pathological groups. The scores obtained by the control group ranged between 0 to 16, on the VHI-10; and between 0 to 37, on the VHI. Conversely, the scores obtained by the 6 pathological groups ranged between 0 to 40 on the VHI-10 and 0 to 119 on the VHI. Furthermore, the control group appeared more homogeneous than the pathological groups. This is reflected by the fact that the overall standard deviation within the control group is approximately 3 times smaller than those obtained by the pathological groups. Finally, handicap severity ratings of the 6 pathological groups maintained the same hierarchy with both questionnaires.

Statistical analysis, using 2 separate univariate analyses-of-variance, for the VHI-10 and VHI scores, revealed a significant main effect for group [($F_{6,\ 341} = 83.16$, $P < 0.001$) and ($F_{6,\ 379} = 84.91$, $P < 0.001$), respectively]. Post-hoc analyses were performed using Tamhane's test since, based on the data presented in Table 2, intergroup equality of variances could not be assumed. Results of the post-hoc analyses revealed a significant group difference between the control group and all 6 pathological groups with the use of both questionnaires ($P < 0.05$). Furthermore, in both analyses, the neurogenic and the mucosa irregularity groups were rated significantly higher than all other groups ($P < 0.05$).

Validity and Reliability

Validity of the VHI-10 was assessed first by computing Pearson correlation coefficients between the scores of the VHI-10 and VHI. In addition, Spearman correlation coefficients were calculated among the scores of the 2 questionnaires and the responses to the 2 general voice-satisfaction questions. Table 3 presents the results matrix of these anal-

Table 3
Correlation coefficient matrix among the VHI-10, VHI questionnaires and the 2 general voice satisfaction questions with Pearson correlation for comparing the 2 questionnaires, and Spearman correlation for comparing the 2 general questions with the 2 questionnaires

	VHI-10	VHI	Question 1	Question 2
VHI-10	–			
VHI	0.970*	–		
Question 1	0.809*	0.831*	–	
Question 2	−0.784*	−0.780*	−0.760*	–

*Correlation is significant at the $P = 0.01$ level.

yses. Data show a strong relationship between the VHI-10 and the VHI and among the 2 general questions and the questionnaires scores. It is noted that the highest correlation coefficient value was obtained between the 2 questionnaires, whereas the lowest value was obtained between the 2 general questions.

The reliability of the translated VHI-10 version was examined, first, by evaluating internal consistency (Cronbach's alpha coefficient). The resulting overall alpha coefficient for the VHI-10 was $r = 0.949$. Then to evaluate parallel form reliability, the 10 items were divided into 2 equal parts, similar in content and number of items. Spearman correlation coefficient was calculated between the 2 similar halves, resulting in $r = 0.918$ ($P < 0.001$).

Finally, the effects of gender and age on the participants' responses to the VHI-10 were evaluated. Based on an independent-sample t test, results revealed no statistically significant gender difference ($P = 0.36$). In addition, no significant correlation was found between respondents' scores on the VHI-10 and their ages ($r = 0.048$, $P = 0.373$).

Evaluation of Subscores

The relationship between the scores of the subscales of the original VHI and the scores of the corresponding questions in the VHI-10 was assessed with Pearson correlation coefficients. For this purpose, the scores of the functional, physical, and emotional subscales of the VHI were examined in association with the scores of the corresponding questions from the VHI-10. Specifically, questions 1 through 5 in the VHI-10 were compared with the functional subscale, questions 6, 7, and 10 were compared with the physical subscale, and questions 8 and 9 were compared with the emotional subscale. The resulting correlation coefficient matrix is presented in Table 4.

Results of the analysis indicate a strong correlation between the scores of the 3 subscales obtained from the original VHI and the scores of the corresponding questions in the VHI-10. Pair-wise comparisons, using Hotelling adjusted t test,[16] revealed that the correlation between the matching categories of the functional and physical subscales were significantly higher than those obtained between all other pairs of categories ($P < 0.05$). In contrast, the correlation coefficients of the emotional subscale were not significantly higher than the other pairs of categories.

Table 4
Correlation coefficient matrix among the 3 subscales of the VHI (functional, physical, and emotional) and the corresponding questions in the VHI-10

	VHI-10		
	Functional (items 1-5)	Physical (items 6,7,10)	Emotional (items 8,9)
VHI			
Functional	0.947*	0.768*	0.780*
Physical	0.792*	0.937*	0.844*
Emotional	0.873*	0.830*	0.886*

*Correlation is significant at the $P = 0.01$ level.

DISCUSSION

The present study provides a large data set that for the first time compares the results of the VHI and VHI-10 in a language other than English. The primary incentive for this study stemmed from the fact that the VHI questionnaire was previously shown to maintain high reliability and validity across translation to various languages. In contrast, the validity of the VHI-10, which is a shorter and less time-consuming questionnaire, has never been assessed in other languages. Our results show that, in a wide range of voice pathologies, the VHI-10 maintained its validity across translation. Based on the group distribution, as well as the correlation between the 2 questionnaires, it is concluded that the Hebrew version of the VHI-10 provides a clinically valid measure for the patient's handicap self-perception.

The VHI-10 is intended to provide a unifactorial self-assessment of voice handicap.[14] Nonetheless, our secondary incentive for this study was to learn whether the VHI-10 could provide some information on the 3 dimensions addressed by the full-length VHI. Results suggest that in addition to the overall score of the VHI-10, preliminary information could be obtained, which is highly correlated with the scores of the full length VHI's functional and physical subscales. This result can be interpreted to suggest that although the VHI-10 is primarily intended to provide a unidimensional evaluation of vocal handicap, it might provide the clinician with additional insight into specific aspects of the subjective experience of the voice patient.

CONCLUSIONS

The Hebrew version of the VHI-10 was shown to be a valid and reliable measure of voice handicap, similar to the original English version. Moreover, although the VHI-10 is basically intended to provide a unidimensional evaluation of voice handicap, results suggest that it could provide additional information on specific dimensions of the subjective perception of voice handicap.

We thank the speech-language pathologists and ENT physicians at the Sheba Medical Center and other medical centers across Israel for their assistance in recruiting candidates for participation in this study. Esther Shabtai, MSc, is greatly acknowledged for her assistance and guidance in the statistical analyses. The senior author thanks Dr Michael S. Benninger and Dr. Thomas Murry for

their professional and methodological advice in the preliminary stages of this study.

REFERENCES

1. Woo P. Quantification of videostrobolaryngoscopic findings: measurements of the normal glottal cycle. Laryngoscope 1996;106:1–16.
2. Heman-Ackah YD. Diagnostic tools in laryngology. Curr Opin Otolaryngol Head Neck Surg 2004;12:549–52.
3. Baken RJ. Clinical measurement of speech and voice. Needham Heights, MA: Allyn and Bacon; 1987.
4. Yamaguchi H, Shrivastav R, Andrews ML, et al. A comparison of voice quality ratings made by Japanese and American listeners using the GRBAS scale. Folia Phoniatr Logop 2003;55:147–57.
5. Benninger MS, Atiuja AS, Gardner G, et al. Assessing outcomes for dysphonic patients. J Voice 1998;12:540–50.
6. Jacobson BH, Johnson A, Grywalski C, et al. The voice handicap index (VHI): development and validation. Am J Speech Lang Pathol 1997;6:66–70.
7. Criteria for determining disability in speech-language disorders. Agency for Healthcare Research and Quality. Evidence Report/Technology Assessment. January 2002, Number 52.
8. Nawka T, Wiesmann U, Gonnermann U. Validierung des Voice Handicap Index (VHI) in der deutschen fassung. HNO 2003;51:921–9.
9. Hsiung MW, Lu P, Kang BH, et al. Measurement and validation of the Voice Handicap Index in voice disordered patients in Taiwan. J Laryngol Otol 2003;117:478–81.
10. Guimarães I, Abberton E. An investigation of the Voice Handicap Index with speakers of Portuguese: preliminary data. J Voice 2004; 18:71–82.
11. Pruszewicz A, Obrebowski A, Wiskirska-Woznica B, et al. Complex voice assessment: Polish version of the Voice Handicap Index (VHI). Otolaryngol Pol 2004;58:547–9.
12. Amir O, Ashkenazi O, Leibovitzh T, et al. Applying the Voice Handicap Index (VHI) to dysphonic and nondysphonic Hebrew speakers. J Voice 2006;20:318–24.
13. Verdonck-de Leeuw IM, Debodt M, Woisard V, et al. (2005). Validation of the Voice Handicap Index by assessing equivalence of 7 European translations. Presented in the 6th Pan European Voice Conference (PEVOC6), London, England.
14. Rosen CA, Lee AS, Osborne J, et al. Development and validation of the Voice Handicap Index-10. Laryngoscope 2004;114:1549–56.
15. Deary IJ, Webb A, Mackenzie K, et al. Short, self-report voice symptom scales: psychometric characteristics of the Voice Handicap Index-10 and the Vocal Performance Questionnaire. Otolaryngol Head Neck Surg 2004;131:232–5.
16. Glass GV, Hopkins KD. Statistical methods in education and psychology, 2nd ed. Needham Heights, Massachusetts: Allyn and Bacon; 1984. pp 310–1.