Cross-Cultural Adaptation and Validation of the Voice Handicap Index Into Estonian☆

*Aune Ello, †,‡Linda Sõber, *,§Merje Viigand, and *,§Lagle Lehes, *†‡§Tartu, Estonia

Summary: Objectives. This study aimed to validate the Voice Handicap Index (VHI-30) for the Estonian language and determine its psychometric properties.

Study design. Cross-sectional and comparative study, combining subjective and objective methods.

Methods. The objective data included voice recordings analyzed using the PRAAT software, which calculates the Acoustic Voice Quality Index (AVQI) to assess voice quality. Subjective data were gathered using the VHI-30 (Estonian version) (VHI-30-EST) to evaluate voice-related quality of life, and videolaryngostroboscopy was performed to examine vocal fold function in patients with dysphonia. Ten percent of the participants completed the VHI again after 2-4 weeks to assess test-retest reliability. The questionnaire's internal consistency, test-retest reliability, and clinical validity were evaluated.

Results. The VHI-30-EST demonstrated excellent internal consistency, with a Cronbach's α of 0.977 in the test group and 0.893 in the control group. The test-retest reliability was strong, with an intraclass correlation coefficient of 0.968 for the total score. Validity analysis showed significant differences between the test and control groups for both VHI-30-EST and AVQI scores. Receiver operating characteristic curve analysis identified effective cut-off scores of 15.5 for VHI-30-EST and 2.8 for AVQI, with high sensitivity for both. Moderate correlations were found between VHI-30-EST scores and AVQI in the test group.

Conclusions. The VHI-30-EST demonstrated strong internal consistency, excellent test-retest reliability, and solid clinical validity. Significant differences were observed between the test and control groups for both VHI-30-EST and AVQI scores, with moderate correlations between them. Effective cut-off scores were identified for both VHI-30-EST and AVQI, showing high sensitivity in distinguishing voice disorders. These results confirm the reliability and validity of VHI-30-EST in assessing voice-related quality of life in Estonian-speaking individuals.

Key Words: Voice Handicap Index—Quality of life—Validity—Reliability-Estonian language.

INTRODUCTION

The assessment of voice-related quality of life has gained significant importance in the field of speech and language pathology.[1,2] A widely used instrument for evaluating the impact of voice disorders on individuals' daily lives is the Voice Handicap Index (VHI). Developed by Johnson et al in 1997,[3] the VHI was among the first tools designed to characterize and measure the impact of voice impairment on a patient's quality of life. Since then, it has become an invaluable resource for clinicians and researchers.

Validating the VHI in different languages and cultural contexts is crucial to ensure its efficacy and reliability across diverse populations. For example, the psychosocial impact of voice disorders can vary significantly across cultures.[4] Validation ensures that the VHI and its subscales address relevant concerns regarding voice handicaps in a manner that is meaningful to the population being assessed. Without cultural adaptation, questionnaires may fail to capture the true impact of voice disorders on an individual's life, leading to inaccurate assessments. Moreover, psychometric features, such as reliability and validity, are essential for a clinical assessment tool.[5] Validating the VHI in multiple languages ensures that it maintains internal consistency, test-retest reliability, and construct validity. Different cultural contexts may influence how individuals interpret items; therefore, validation helps to confirm that the VHI functions consistently regardless of language or cultural background.

Moreover, validating the VHI in multiple languages enables researchers to compare data systematically, thereby contributing to a more comprehensive understanding of voice disorders globally. In 2002, the Agency for Healthcare Research and Quality recognized the VHI as a reliable and valid diagnostic tool.[6] Since then, the VHI has been translated and validated into numerous languages, including Latvian,[7] Slovak,[8] Croatian,[9] Italian,[10] Czech,[11] Chinese,[12] Polish,[13] Hebrew,[14] Brazilian Portuguese,[15] Greek,[16] and Arabic,[17] among others. These studies, along with others, have shown that the VHI effectively measures how individuals subjectively perceive limitations related to voice issues.[14]

Estonian is a unique Finno-Ugric language, distinct from the Indo-European language family. Therefore, linguistic nuances, idiomatic expressions, and culturally specific

Accepted for publication January 10, 2025.

☆ No funding sources were involved in the conduct of this research or the preparation of this article.

From the *Department of Special Education and Speech Therapy, University of Tartu, Tartu, Estonia; †Department of Oto-Rhino-Laryngology, University of Tartu, Tartu, Estonia; ‡Ear Clinic, Tartu University Hospital, Tartu, Estonia; and the §Sports Medicine and Rehabilitation Clinic, Tartu University Hospital, Tartu, Estonia.

Address correspondence and reprint requests to: Aune Ello, Department of Special Education and Speech Therapy, University of Tartu, Jakobi 5, Tartu, Estonia 51005. E-mail: aune.ello@gmail.com

Journal of Voice, Vol xx, No xx, pp. xxx–xxx

0892-1997

https://doi.org/10.1016/j.jvoice.2025.01.017

health- and medicine-related terminology may not be translated directly from English into Estonian. This study aimed to validate the VHI in Estonian and to determine its psychometric properties. Upon validation, the test could be implemented by speech-language pathologists and other professionals who treat voice disorders in clinical practice.

METHODS

Ethical considerations

This study was approved by the Research Ethics Committee of the University of Tartu (License no. 388/T-16). All participants were briefed about the objectives of the study and assured of the confidentiality and anonymity of their data. Participation was voluntary, and written consent was obtained from all participants before their involvement.

Participants

Based on the aims of the study, 64 asymptomatic healthy adults and 34 patients with dysphonia, all of whom self-reported their sex assigned at birth, were recruited as control and test groups, respectively. Participants had not received any prior voice therapy.

The dysphonic group consisted of seven male and 27 female patients with an age range of 23-79 years (with mean age, 55.8 years; median age, 59.5 years). Following laryngeal videostroboscopy, patients with dysphonia were diagnosed with either functional voice disorders (n = 10) or organic voice disorders (n = 24). The underlying conditions of the organic voice disorders included vocal fold paresis (n = 11), Reinke's edema (n = 2), vocal nodules (n = 1), spasmodic dysphonia (n = 2), dysarthria (n = 5), cancerous processes (n = 4), gastroesophageal reflux disease (n = 2). The selection of these subjects was based on the following criteria: participants were 18 years or older, had a diagnosis of dysphonia validated by an otolaryngologist or a speech-language pathologist, were fluent in Estonian, and had the capability to complete the questionnaire.

The age range of the control group was 27-88 years (mean age, 59.7 years; median age, 60 years). The control group included individuals over 18 years of age, without any identified voice disorders or voice complaints, who were required to have Estonian language fluency and the ability to complete the questionnaire. The subjects came from diverse professional backgrounds and were matched with the dysphonic group for key demographic characteristics, including age and sex.

The exclusion criteria for both the test group and the control group were as follows: the presence of any factors (eg, mental or sensory conditions) that could interfere with data collection or the inability to complete the voice quality questionnaire independently.

Instrumentation

This was a cross-sectional and comparative study. Data were collected using a combination of objective (PRAAT) and subjective (VHI, laryngeal videostroboscopy) methods, performed by our multidisciplinary voice team.

We used the PRAAT program for acoustic voice analysis. The PRAAT is a computer program that provides information on various acoustic parameters.[18] The Acoustic Voice Quality Index (AVQI) is a measure of voice quality calculated using a script within the PRAAT software. Developed by Maryn et al,[19] the AVQI is a six-factor model that allows the objective evaluation of dysphonia severity in both sustained vowels and continuous speech. This index is derived from a weighted combination of six acoustic measures, including shimmer local, shimmer local dB, harmonic-to-noise ratio, general slope of the spectrum, tilt of the regression line through the spectrum, and smoothed cepstral peak prominence, all integrated into a linear regression formula.

Flexible fiberoptic nasopharyngolaryngoscopy (Olympus model CS-170, ENF-V4 video rhino-laryngoscope) was used to perform laryngeal videostroboscopy (VLS) in the test group.

Procedure

First, a linguistic validation process was performed. The original English version of the VHI[3] was translated into Estonian by two qualified professional translators. Discrepancies in the translation of certain items were resolved through discussion with speech and language therapists until a unanimous agreement was reached. The survey was then back-translated into English by two professional Estonian-English language practitioners who were unaware of the original questionnaire and had not been involved in the initial translation. The Estonian version of the VHI was pilot-tested on 11 male and 20 female participants, all with normal voice function. Modifications were made based on the feedback from patients in the pilot study.

After the validation, a conducted voice-related quality of life was assessed using the VHI questionnaire. The VHI is a 30-item patient-based self-administered tool designed to measure the effect of voice disorders on the daily activities of individuals experiencing dysphonia.[3] This self-assessment survey examines the effects of voice issues across three primary areas: functional (F), physical (P), and emotional (E).[3] Each statement is rated on a five-point Likert scale (0 = never; 1 = almost never; 2 = sometimes; 3 = almost always; 4 = always). All participants were asked to fill out the VHI-30-EST questionnaire. Ten percent of the participants in the test group completed the VHI-30-EST questionnaire again after 2-4 weeks to assess the test-retest reliability of the questionnaire. Test-retest reliability was evaluated for both the overall score and each of the three subsections of VHI-30-EST. During this timeframe, the participants did not undergo any therapeutic interventions; therefore, significant changes in the patients' health statuses were not anticipated.

Next, the AVQI was calculated for each participant in both groups using PRAAT on the same day. Audio recordings were captured using a Shure SM48 microphone positioned 10 cm from the patient's mouth at a 45° angle. The microphone was connected to an HP laptop running Windows 11 and all voice samples were evaluated using the AVQI script version *03.0131* in PRAAT *(5.3.55)*. During the evaluation, all the participants phonated a sustained /a/ vowel three times at a comfortable pitch and loudness, and the best attempt was selected for analysis. Additionally, patients read three phonetically balanced sentences, developed by Mihkla and Piits of the Estonian Language Institute, at a comfortable pitch and loudness (*Ühe kapteni avar mantel oli õlgadele visatud. Ära keskendu me kassi abitusele. Ma ju kiigutan poissi ja looma*).

The procedure, which was conducted in collaboration with an otolaryngologist, was incorporated into the patient's routine clinical visit to confirm and characterize the nature of the voice disorder. The VLS was used to visualize vocal-fold vibrations and closures during phonation and speech. The participants in the test group were instructed to sustain Estonian vowels /i/ and /e/, repeat a sentence, "*Juulikuus suur tuul,"* and count numbers ranging from 1 to 10. Laryngeal videostroboscopy was only performed in the test group.

Statistical analysis

Statistical analysis was conducted using JASP 0.19 (JASP, University of Amsterdam) and SPSS 20 (SPSS, Inc., Chicago, IL) software. Given the non-normal distribution of the data, as indicated by the Shapiro–Wilk test ($P < 0.05$), nonparametric tests were applied for further analysis.

The internal consistency of the VHI was evaluated using Cronbach's alpha coefficient.

Cronbach's alpha values below 0.70 indicated minimal reliability, 0.70-0.79 indicated moderate, 0.80-0.90 indicated high, and above 0.90 indicated very high reliability.[20]

Test-retest reliability was evaluated using the intraclass correlation coefficient (ICC) by comparing the results from the initial and subsequent questionnaire administrations. The ICC values were categorized according to the 95% confidence interval (CI): below 0.5 indicated poor reliability, between 0.5 and 0.75 indicated moderate reliability, between 0.75 and 0.9 indicated good reliability, and above 0.9 indicated excellent reliability.[21]

Construct validity was evaluated using Pearson's correlation coefficient to determine the strength and direction of the relationship between VHI scores and relevant constructs.

The correlation between the VHI-30 and AVQI scores in both groups was examined using Pearson's correlation coefficient.

Receiver operating characteristic (ROC) curve analysis was employed to identify the cut-off points for the VHI-30-EST and AVQI. The cut-off point was established by locating the point closest to the upper left corner of the ROC curve for positive cases and the upper right corner for negative cases. The effectiveness of the cut-off point was validated by ensuring that the area under the ROC curve (AUC) was at least 0.7, indicating an acceptable level of sensitivity (true positive rate) and specificity (true negative rate).

The Mann-Whitney U test was used to compare VHI and AVQI results between the control group and test groups.

This comprehensive statistical approach provides robust measures of the reliability, validity, and ability of the VHI to differentiate between groups.

RESULTS

Internal consistency analysis of the VHI

Cronbach's α coefficient was calculated to evaluate the homogeneity of the questionnaire items, showing the extent to which they measure a common underlying construct. The analysis revealed very high internal reliability of the questionnaire in the test group (α = 0.977) and high internal reliability in the control group (α = 0.893) (Table 1).

Test-retest reliability

The stability of the VHI-30-EST over time was confirmed through a test-retest method, where participants completed the VHI-30-EST twice with an average interval of 4 weeks. The reliability of the VHI-30-EST was validated by calculating the (ICC), which demonstrated consistent and dependable results across the testing period. The ICC was strong for the total score (r = 0.968), with a 95% CI ranging from 0.732 to 0.997. The ICC estimates for the three subscales were: 0.907 (95% CI: 0.361-0.990) for the functional subscale, 0.942 (95% CI: 0.552-0.994) for the physical subscale, and 0.950 (95% CI: 0.602-0.995) for the emotional subscale.

Clinical validity of the VHI and AVQI

Information on voice-related quality of life was gathered from both groups using the VHI-30-EST questionnaire, and voice acoustic measures were collected using an AVQI script within the PRAAT software. The mean VHI-30-EST total score for the healthy group was 7.58 (standard deviation [SD] 7.49), whereas that for the test group was 42.32 (SD 32.23). The highest mean scores were calculated for the physical subscale in both groups. The mean AVQI total score for the control group was 1.72 (SD 0.67), and that for the test group was 4.17 (SD 1.98). The total AVQI and VHI scores and its subscales are shown in Table 2. Significant differences were found in all subscales ($P < 0.05$).

TABLE 1.
Internal Consistency of the VHI-EST Subscales

		Cronbach Alpha (α)	
VHI scale	Number of items	Test Group	Control Group
VHI T	30	0.977	0.893
VHI F	10	0.940	0.701
VHI P	10	0.935	0.834
VHI E	10	0.965	0.751

TABLE 2.
AVQI and VHI-30-EST Subscales and Total Score for the Test (*N* = 34) and the Control (*N* = 64) Groups

	Test group (*N* = 34)		Control group (*N* = 64)		M-W Test	
	Mean	SD	Mean	SD	*U*	*P* Value
VHI T	42.324	32.233	7.578	7.489	208.500	< 0.001
VHI F	13.412	11.062	2.969	2.834	393.000	< 0.001
VHI P	17.941	11.292	3.313	3.550	203.500	< 0.001
VHI E	11.500	11.965	1.453	2.582	266.000	< 0.001
AVQI	4.165	1.980	1.724	0.670	127.500	< 0.001

Mean, standard deviation. Mann-Whitney *U* test.

Correlation between VHI-30-EST scores and AVQI

Correlations between VHI-30-EST scores and AVQI measures indicated moderate but significant correlations in the test group and poor but significant correlations in the control group except for the emotional subscale, which showed no correlation with AVQI. The summarized correlation results are presented in Table 3.

Construct validity

Pearson's correlation analysis revealed a very strong positive correlation ($r > 0.9$) between the total VHI-30-EST score and its subscale scores, indicating excellent construct validity (Table 4).

Cut-off score for VHI-30-EST and AVQI

The ROC curve analysis indicated that both VHI-30-EST and AVQI could effectively distinguish between the test and control groups. For VHI-30-EST, the AUC was 0.904 ($P < 0.001$). A cut-off score of 15.5 was established, with a sensitivity of 0.824 and 1-specificity of 0.141. For AVQI, the AUC was 0.941 ($P < 0.001$), with a cut-off value of 2.8, sensitivity of 0.824, and 1-specificity of 0.047 (Figure 1).

DISCUSSION

This study used the VHI questionnaire to assess voice-related quality of life in both healthy controls and patients diagnosed with dysphonia. The results provided a comprehensive comparison between the two groups, revealing substantial differences in their perceived vocal handicap, as measured by the total VHI scores and subscales, along with strong reliability and validity indicators for the VHI-30-EST tool.

TABLE 3.
Correlation Between VHI-30-EST and AVQI for the Test and the Control Groups

		Test group	Control group
		AVQI	AVQI
VHI T	Pearson's *r*	0.447	0.300
	P value	0.008	0.016
VHI F	Pearson's *r*	0.507	0.298
	P value	0.002	0.017
VHI P	Pearson's *r*	0.399	0.337
	P value	0.019	0.007
VHI E	Pearson's *r*	0.349	0.049
	P value	0.043	0.699

Pearson's *r*, *P* value.

TABLE 4.
Pearson Moment Correlation Coefficients Between VHI-EST Total Scale and the Subscales in Both the Test and the Control Groups

Scale	Total	Functional	Physical	Emotional
Total	*	*	*	*
Functional	0.960	*	0.767	0.855
Physical	0.959	*	*	0.800
Emotional	0.916	*	*	*

In the VHI-30-EST study comparing individuals with and without voice disorders, the mean total score for the test group (42.32) was significantly higher than that for the control group (7.58), reflecting a greater perception of voice-related impairment among those with vocal issues. The physical subscale, which measures sensations and discomfort related to voice use, had the highest average score in both groups, suggesting that both healthy individuals and those with voice disorders were the most aware of the physical aspects of their voice. Similar findings have been reported in several other transcultural studies,[7,8,10,14,16,22–25] suggesting that patients primarily associate voice problems with physical symptoms rather than emotional or functional difficulties. The significant difference between the two groups ($P < 0.05$) highlights how voice disorders impact the quality of life and demonstrates that the VHI questionnaire effectively distinguishes healthy individuals and those with vocal problems.

In addition to VHI validation, we also established cut-off scores for AVQI in Estonian. The AVQI provides an objective assessment of the acoustic parameters of voice. This index provides additional value to the relatively subjective VHI score. The AVQI measurements corroborated the findings from the VHI-30-EST, with the control group averaging a score of 1.72 compared with the test group's score of 4.17. These findings reinforce the idea that acoustic measures can effectively reflect subjective experiences reported in the VHI-30-EST. The significant differences

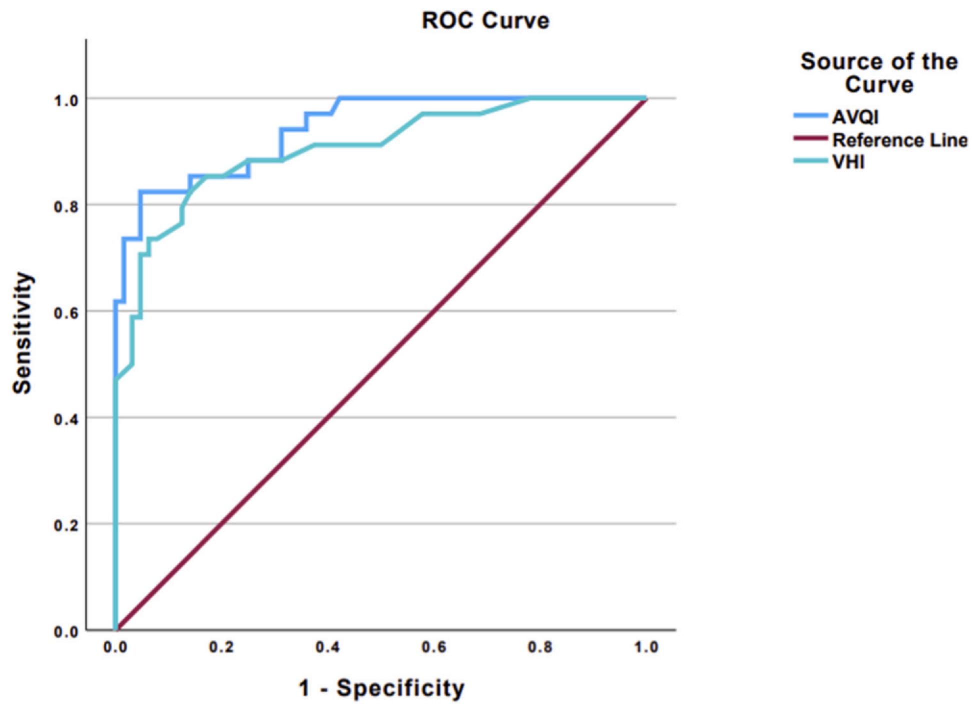

FIGURE 1. Receiver operating characteristic (ROC) curve for VHI-30-EST and AVQI. VHI-30-EST, Voice Handicap Index-30 (Estonian version); AVQI, Acoustic Voice Quality Index.

across all groups ($P < 0.05$) indicate that voice quality issues are not only subjective but also measurable through acoustic analysis. These results underscore the importance of integrating subjective and objective assessments when evaluating voice disorders. The combination of VHI-30-EST and AVQI data offers a comprehensive understanding of how voice issues affect an individual's quality of life.

Consistent with earlier findings,[10,12,16,26,27] the VHI-30-EST questionnaire showed outstanding internal consistency, indicating that each subscale effectively measures a specific aspect of voice impairment. Cronbach's α showed high reliability in both groups, particularly in the test group ($\alpha = 0.977$), and moderately high reliability in the control group ($\alpha = 0.893$). The study findings show that each scale is consistent and accurately evaluates a particular aspect of the condition.

In our study, the test-retest group consisted of individuals with voice complaints, which aligns with the approach used in most studies.[3,10,12,14–17] The high intraclass ICC values, particularly for the total score ($r = 0.968$), indicated that the questionnaire produced consistent results across repeated administrations. This discovery adds to the existing evidence,[8–10,16,17,23,26,28,29] suggesting that the VHI-30-EST is a stable tool for measuring voice-related quality of life over time. Strong ICC estimates across the three subscales (functional, physical, and emotional) also supported the consistency of the questionnaire. This strong temporal stability is essential in clinical settings, as it allows for reliable tracking of voice-related quality of life over time, which is important for assessing treatment outcomes and monitoring voice conditions.

As far as we know, the VHI-30 has not been validated for any Finno-Ugric languages. Therefore, we are not able to compare the results across related languages. However, we can still compare the results within regions and cultural context. In Estonian, the cut–off point for differentiating between healthy and impaired individuals was determined to be 15.5, with high sensitivity (0.824) and specificity (1-specificity of 0.141). This cut–off value served as a threshold for distinguishing between individuals with and without voice disorders. Johnson et al[3] reported an approximate cut-off score of 30 points for the original VHI, though not determined by ROC curves. As determined by ROC curves, the chosen cut-off of 15.5, which falls within the range of findings from other studies,[8,11,24] highlights both the variability in results and the complexity of defining a universal standard for diagnosing voice disorders. Our research show similarities with the results of studies conducted in geographically proximate districts, eg, Czech Republic (13)[11] and Poland (17).[13] In contrast, the VHI-EST-30 cut-off score revealed similarities also with languages from cultures further away, such as Brazilian Portuguese (19)[30] and Iran (14.5).[31] Consequently, given the data, we cannot definitively conclude that cultural and linguistic similarities affect VHI cut-off scores. For Estonian, we can draw conclusions within our own language and cultural context. Moreover, we must agree with the Slovakian study,[8] which points out that it is questionable whether the VHI can unequivocally confirm the presence of a voice disorder. Therefore, while the VHI can be an important clinical assessment tool, specialists must approach patients holistically. The value of VHI lies in assessing therapy dynamics, as it helps to objectively measure changes.

The AUC of 0.904 indicated that the VHI-30-EST is a good diagnostic tool for distinguishing between individuals with and without voice impairment. The high AUC value ($P < 0.001$) demonstrates that the questionnaire accurately identified those with significant voice-related quality of life issues.

Pearson's correlation analysis revealed a strong positive linear correlation ($r > 0.9$) between the total VHI score and its subscales, indicating strong construct validity for the VHI-30-EST. This finding suggests that the questionnaire's subscales (physical, functional, and emotional) are highly representative of the overall voice-related handicaps experienced by individuals. These results align with several previous studies.[12,16,26]

Correlations between VHI-30-EST scores and AVQI measures demonstrated moderate but significant relationships in the test group. These findings are consistent with those of previous studies.[27,28] In contrast, the correlations in the control group were weaker but still significant, except for the emotional domain, where no significant correlation with the AVQI was found. These findings suggest that while the AVQI provides objective voice quality data, it does not fully capture the emotional aspects of voice disorders, which are more subjective and may require self-reported measures, such as the VHI.

Similar to studies in Croatia,[9] Latvia,[7] and Sweden,[24] the primary limitation of this study was its small sample size. The responsiveness of the VHI-30-EST to changes following intervention was not assessed and could be explored in future studies.

CONCLUSION

This study demonstrated that the VHI-30-EST is a highly reliable and valid tool for assessing the impact of voice disorders on quality of life. The tool demonstrated strong internal consistency and test-retest reliability, while Pearson's correlation analysis further validated the questionnaire's construct validity. Overall, the VHI-30-EST is effective in distinguishing between healthy and impaired individuals, and its stability across different populations makes it an essential tool for the diagnosing and evaluating of voice disorders.

Declaration of Competing Interest

The authors have no conflicts of interest.

Acknowledgments

The authors would like to thank the patients who participated in the study.

References

1. Wheeler KM, Collins SP, Sapienza CM. The relationship between VHI scores and specific acoustic measures of mildly disordered voice production. *J Voice*. 2006;20:308–317. https://doi.org/10.1016/j.jvoice.2005.03.006.
2. Guimarães I, Abberton E. An investigation of the Voice Handicap Index with speakers of Portuguese: preliminary data. *J Voice*. 2004;18:71–82. https://doi.org/10.1016/j.jvoice.2003.07.002.
3. Johnson A, Jacobson B, Grywalski C, Silbergleit A, Jacobson G, Benninger M. The Voice Handicap Index (VHI): development and validation. *Am J Speech Lang Pathol*. 1997;6:66–70.
4. Yiu EML, Ho EM, Ma EPM, et al. Possible cross-cultural differences in the perception of impact of voice disorders. *J Voice*. 2011;25:348–353. https://doi.org/10.1016/j.jvoice.2009.10.005.
5. D'Este C. New views of validity in language testing. *Educ Linguist Lang Educ*. 2012;1. https://doi.org/10.14277/2280-6792/5p.
6. Biddle AK, Watson LR, Hooper CR, Lohr KN, Sutton SF. Criteria for Determining Disability in Speech-Language Disorders: Summary. In: AHRQ Evidence Report Summaries. Agency for Healthcare Research and Quality (US); 2002. Available at: https://www.ncbi.nlm.nih.gov/books/NBK11866/. Accessed February 19, 2024.
7. Trinite B, Sokolovs J. Adaptation and validation of the Voice Handicap Index in Latvian. *J Voice*. 2014;28:452–457. https://doi.org/10.1016/j.jvoice.2014.01.008.
8. Frajkova Z, Krizekova A, Missikova V, Tedla M. Translation, cross-cultural validation of the Voice Handicap Index (VHI-30) in Slovak language. *J Voice Off J Voice Found*. 2022;36:145.e1–145.e6. https://doi.org/10.1016/j.jvoice.2020.04.003.
9. Bonetti A, Bonetti L. Cross-cultural adaptation and validation of the Voice Handicap Index into Croatian. *J Voice Off J Voice Found*. 2013;27:130.e7–130.e14. https://doi.org/10.1016/j.jvoice.2012.07.006.
10. Schindler A, Ottaviani F, Mozzanica F, et al. Cross-cultural adaptation and validation of the Voice Handicap Index into Italian. *J Voice Off J Voice Found*. 2010;24:708–714. https://doi.org/10.1016/j.jvoice.2009.05.006.
11. Krtickova J, Svec JG, Haviger J, et al. Validation of the Czech Version of the Voice Handicap Index. *J Voice Off J Voice Found*. 2023. https://doi.org/10.1016/j.jvoice.2023.04.020. Published online June 19, S0892-1997(23)00145-5.
12. Lam PKY, Chan KM, Ho WK, Kwong E, Yiu EM, Wei WI. Cross-cultural adaptation and validation of the Chinese Voice Handicap Index-10. *Laryngoscope*. 2006;116:1192–1198. https://doi.org/10.1097/01.mlg.0000224539.41003.93.
13. Miaśkiewicz B, Gos E, Dębińska M, et al. Polish translation and validation of the Voice Handicap Index (VHI-30). *Int J Environ Res Public Health*. 2022;19:10738. https://doi.org/10.3390/ijerph191710738.
14. Amir O, Ashkenazi O, Leibovitzh T, Michael O, Tavor Y, Wolf M. Applying the Voice Handicap Index (VHI) to dysphonic and nondysphonic Hebrew speakers. *J Voice Off J Voice Found*. 2006;20:318–324. https://doi.org/10.1016/j.jvoice.2005.08.006.
15. Behlau M, Alves Dos Santos L de M, Oliveira G. Cross-cultural adaptation and validation of the voice handicap index into Brazilian Portuguese. *J Voice Off J Voice Found*. 2011;25:354–359. https://doi.org/10.1016/j.jvoice.2009.09.007.
16. Helidoni ME, Murry T, Moschandreas J, Lionis C, Printza A, Velegrakis GA. Cross-cultural adaptation and validation of the voice handicap index into Greek. *J Voice Off J Voice Found*. 2010;24:221–227. https://doi.org/10.1016/j.jvoice.2008.06.005.
17. Malki KH, Mesallam TA, Farahat M, Bukhari M, Murry T. Validation and cultural modification of Arabic Voice Handicap Index. *Eur Arch Oto-Rhino-Laryngol Off J Eur Fed Oto-Rhino-Laryngol Soc EUFOS Affil Ger Soc Oto-Rhino-Laryngol-Head Neck Surg*. 2010;267:1743–1751. https://doi.org/10.1007/s00405-010-1296-x.
18. Boersma P, Van Heuven V. Speak and unSpeak with PRAAT. *Glot Int*. 2001;5:341–347.
19. Maryn Y, De Bodt M, Roy N. The acoustic Voice Quality Index: toward improved treatment outcomes assessment in voice disorders. *J Commun Disord*. 2010;43:161–174. https://doi.org/10.1016/j.jcomdis.2009.12.004.
20. Cohen L, Manion L, Morrison K. *Research Methods in Education*. 6th ed., Oxfordshire, U.K.: Routledge; 2007. https://doi.org/10.4324/9780203029053.

21. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med.* 2016;15:155–163. https://doi.org/10.1016/j.jcm.2016.02.012.
22. Sorensen JR, Printz T, Mehlum CS, Heidemann CH, Groentved AM, Godballe C. Cross-cultural Adaption and Validation of the Danish Voice Handicap Index. *J Voice Off J Voice Found.* 2019;33:441–444. https://doi.org/10.1016/j.jvoice.2018.01.010.
23. Lam PKY, Chan KM, Ho WK, Kwong E, Yiu EM, Wei WI. Cross-cultural adaptation and validation of the Chinese Voice Handicap Index-10. *Laryngoscope.* 2006;116:1192–1198. https://doi.org/10.1097/01.mlg.0000224539.41003.93.
24. Ohlsson AC, Dotevall H. Voice handicap index in Swedish. *Logoped Phoniatr Vocol.* 2009;34:60–66. https://doi.org/10.1080/14015430902839185.
25. Núñez Batalla F, González Márquez R, Peláez González MB, González Laborda I, Fernández Fernández M, Morato Galán M. Acoustic voice analysis using the Praat programme: comparative study with the Dr. Speech Programme. *Acta Otorrinolaringol Engl Ed.* 2014;65:170–176. https://doi.org/10.1016/j.otoeng.2014.05.007.
26. Núñez-Batalla F, Corte-Santos P, Señaris-González B, Llorente-Pendás JL, Górriz-Gil C, Suárez-Nieto C. Adaptation and validation to the Spanish of the Voice Handicap Index (VHI-30) and its shortened version (VHI-10). *Acta Otorrinolaringol Engl Ed.* 2007;58:386–392. https://doi.org/10.1016/S2173-5735(07)70376-9.
27. Nawka T, Wiesmann U, Gonnermann U. Validierung des Voice Handicap Index (VHI) in der deutschen Fassung. *HNO.* 2003;51:921–930. https://doi.org/10.1007/s00106-003-0909-8.
28. Sotirović J, Grgurević A, Mumović G, et al. Adaptation and validation of the Voice Handicap Index (VHI)-30 into Serbian. *J Voice.* 2016;30:758.e1–758.e6. https://doi.org/10.1016/j.jvoice.2015.09.002.
29. Behlau M, Alves Dos Santos L de M, Oliveira G. Cross-cultural adaptation and validation of the voice handicap index into Brazilian Portuguese. *J Voice Off J Voice Found.* 2011;25:354–359. https://doi.org/10.1016/j.jvoice.2009.09.007.
30. Behlau M, Madazio G, Moreti F, et al. Efficiency and cutoff values of self-assessment instruments on the impact of a voice problem. *J Voice Off J Voice Found.* 2016;30:506.e9–506.e18. https://doi.org/10.1016/j.jvoice.2015.05.022.
31. Moradi N, Pourshahbaz A, Soltani M, Javadipour S. Cutoff point at voice handicap index used to screen voice disorders among persian speakers. *J Voice Off J Voice Found.* 2013;27:130.e1–130.e5. https://doi.org/10.1016/j.jvoice.2012.08.007.