Vocal Fatigue Index in Finnish-Speaking Population

*,†,1Eliina Kankare, †Leena Rantala, and †Anne-Maria Laukkanen, *†*Tampere, Finland*

SUMMARY: Background and objective. Vocal fatigue is an important complaint that may indicate a voice disorder or a risk thereof. There is a need for a reliable tool to detect and quantify vocal fatigue and distinguish dysphonic and vocally healthy speakers. The Vocal Fatigue Index (VFI) questionnaire has been found valid and reliable among speakers of different languages. This study aims to validate it for speakers of Finnish.

Study design. Experimental comparative study.

Methods. The VFI questionnaire was translated from English to Finnish according to the WHO recommendations. Next, it was subjected to the validation procedure. In total, 160 Finnish speakers volunteered to participate in the study. Hundred-and-eight were voice patients (83 males, 25 females) and 52 were vocally healthy controls (37 females, 15 males). As a comparison, the Voice Handicap Index (VHI) questionnaire was completed and voice samples were recorded to enable Acoustic Voice Quality Index (AVQI03.01$_{FIN}$) analysis.

Results. Results from the first and second completions of the VFI(F) questionnaire correlated strongly (Spearman's rho 0.901, $P = 0.01$). Answers to the individual questions the VFI(F) also correlated strongly, showing high internal consistency. Factor 1 (*Tiredness of voice and avoidance of voice use*) of the VFI correlated strongly with the VHI, and the two other factors (*Physical discomfort associated with voicing* and *Improvement of symptoms*) correlated moderately with the VHI. Factor one of the VFI(F) correlated moderately with AVQI03.01$_{FIN}$ and its sub-parameters, CPPS, HNR, and shimmer. The VFI(F) showed good construct validity, differentiating voice patients and controls at cut-off 13.5, with sensitivity of 0.963 and specificity of 0.885. Discriminatory power was strong for all factors: F1 $A_{ROC} = 0.985$, F2 $A_{ROC} = 0.864$, and F3 $A_{ROC} = 0.821$.

Conclusion. The VFI(F) correlates with the VHI and with AVQI01.01$_{FIN}$ and it is a valid and reliable tool for detecting vocal fatigue in Finnish speakers.

Key words: Subjective evaluation−VFI−VHI−AVQI.

INTRODUCTION

Voice problems are common, particularly among professional voice users, such as teachers, kindergarten teachers, telemarketers, salespersons, priests, actors, and singers.[1-10] The main complaint is often vocal fatigue, which in turn manifests via many subjective, auditory-perceptual, acoustic, and physiological characteristics. Subjective sensations include discomfort or pain in the neck, irritation of the larynx, the feeling of a lump in the throat, a need to cough or clear the throat, and increased effort required for voicing.[11] Perceived signs of deterioration of vocal function include hoarseness, breathiness, instability of voice, pitch breaks, loss of high pitches, and increased effort.[12] Acoustic features related to vocal fatigue include increased perturbation (jitter, shimmer) and decreased harmonic-to-noise ratio (HNR), which may indicate hoarseness.[13] Signs of increased effort (rise in f0, SPL, decreased spectral tilt, or decreased

open time of the glottis) are often reported after a vocal loading task (a laboratory test with prolonged, loud reading or a teacher's long working day).[14-17] Laryngeal findings related to prolonged voice use include edema of the vocal folds, the appearance of glottal chink, and changes in vocal fold vibration and closure patterns.[12]

However, the findings of different studies on the changes in various parameters often have been either opposite or insignificant and uncorrelated, and as such, vocal fatigue has been difficult to define comprehensively. Suggested explanations include that the loading tasks used have not been sufficient to induce vocal fatigue or that there can be many sources of vocal fatigue, causing different manifestations.[12,18] Voice users are also capable of compensating for the potential effects of vocal loading by changing their type of voicing (eg, breathiness due to impaired glottal closure can be avoided by increasing vocal effort). Difficulties in finding a comprehensive subjective-objective definition of vocal fatigue have resulted in consensus that vocal fatigue essentially means the subjective sensation of local tiredness and weakness of voice after a period of voice use.[19] This, in turn, has led to the need to develop a reliable tool to sufficiently describe and quantify vocal fatigue and to distinguish normophonic and dysphonic voice users in this respect.

Nanjundeswaran et al[20] presented and validated the Vocal Fatigue Index (VFI), which is a questionnaire that consists of 19 questions concerning three factors: (F1) *Tiredness of voice and avoidance of voice use*, (F2) *Physical discomfort related to voice use*, and (F3) *Improvement of symptoms with rest*. Since differences across languages and

cultures may cause differences in the results obtained via questionnaires, the VFI must be validated in a particular language in order to form a clinically usable tool. So far, the VFI has been validated in German (in 2019 [21]; a modified version was validated in 2021),[22] Brazilian Portuguese,[23,24] Hong-Kong Chinese,[25] Turkish,[26] Spanish,[27] Persian,[28] Croatian,[29] Malayalam,[30] and Polish.[31]

The present study investigates the reliability and validity of the Finnish translation of the VFI questionnaire in Finnish-speaking subjects. In addition to collecting and comparing VFI results from normophonic and dysphonic speakers, we also compare VFI results with those obtained via another questionnaire, the Voice Handicap Index (VHI), which has been used to disclose characteristics and communicative impairment due to vocal dysfunction.[32,33] Additionally, we compare VFI results with acoustic results obtained using the Acoustic Voice Quality Index (AVQI; see Maryn et al 2010),[34] which consists of six acoustic parameters (shimmer in dB and percentage, CPPS, HNR, and spectral slope and tilt) measured for a sample consisting of several syllables of standard text reading and three seconds of sustained vowel /a:/. AVQI has been found to correlate well with the degree of hoarseness (G in GRBAS scale)[35-37] and to differentiate between normophonic and dysphonic voices.[38,39] It has also shown moderate correlations with the VHI.[40,41] AVQI has been validated in many languages, including Finnish.[42,43] By combining VFI and AVQI results, we hope to gain a better understanding of the background of vocal fatigue in the participants.

METHODS

Participants

The Ethics Committee of Tampere University Hospital provided approval for the present study (R19069). The participants were 108 voice patients from a voice clinic (females n = 83, males n = 25, mean age 53 years, range 21−82 years) and 52 vocally healthy controls (females n = 37, males n = 15, mean age 51 years, range 20−85 years). Voice patients were recruited for the study from the Tampere University Hospital department of phoniatrics. The patients were diagnosed by the phoniatrician or otolaryngologist. Diagnoses for the voice patients were vocal fold paresis (n = 37), spasmodic dysphonia (n = 37), and functional dysphonia (n = 32). Vocally healthy controls were recruited to join the study from Tampere University and from the staff of the phoniatrics department of Tampere University Hospital. Inclusion criteria for the control group were less than 30 points in the total score of the VHI questionnaire[32] and normal sounding voice as stated by speech therapists who were collecting the data. In the Table 1 summarizes the demographics of the two participant groups.

Procedure

The VFI questionnaire was translated from English to Finnish according to the WHO (2014) recommendations for the translation for clinical research tools. A bilingual translator who was familiar with the VFI questionnaire was consulted in the translating process. Discussion on the questionnaire was conducted with a monolingual group, and afterwards, corrections were made with a bilingual translator. A professional independent translator conducted back-translation, and then small corrections were made.[44] This tool is hereafter designated the VFI(F).

From the beginning of data collection for the research, every patient who came to the speech therapist's office and had a diagnosis of vocal paresis, spasmodic dysphonia, or functional dysphonia was asked to participate in the study. The data collection started in the beginning of January 2020 and ended in the beginning of May 2022. Only one patient declined to participate in the study.

In the present study, participants completed the VFI and VHI questionnaires two times. The first time was when they were asked to participate in the research, and the second time was after two weeks. After they had completed the questionnaire a second time, they were asked to send it to the researcher by mail in a return envelope. If the second questionnaire had not arrived after about four weeks, the research nurse placed a reminder call to the participant. In some cases, the time between the first and second completion of the questionnaires rose to nine weeks. Five patients were excluded from the study because they did not complete the questionnaires a second time. In the first appointment, all participants also filled in a preliminary information form as used in the phoniatric clinic (name, age, gender, occupation, "how is your voice usually," "is your voice getting tired when you talk," "do you frequently experience reflux, headaches, neck pain," "do you have allergies," and other questions about their voice and voice use). Simultaneously with the first completion of the questionnaire, the participants also gave three voice samples: 1) reading aloud a standard text, 2) speaking spontaneously for one minute, and 3) sustaining /a:/ three times, all three samples in habitual speaking voice loudness and pitch. The spontaneous speaking samples were not used in this study.

Recordings and acoustic analysis

Recordings of the voice patients and of some of the healthy controls were made in a quiet surgery room or office, while some of the healthy controls were recorded in a sound-treated studio at Tampere University. All recordings were made using a similar device and procedure: an AKG C544L head-mounted condenser microphone (4 cm from the lip corner), digitization at 44100 samples per second, and 16 bits amplitude quantization using the Focusrite iTrack Solo soundcard. All the recordings were consistent with the recommended norm of SNR > 30 dB, for acceptable conditions for acoustic recordings and analysis (the mean SNR of the recordings was 38.5 dB with SD of 6 dB).

The voice samples were analyzed with Praat (6.1.16) software[45] and AVQI03.01$_{FIN}$ script.[43,46] In the AVQI analyses, the first 31 syllables of the loud reading samples and three

TABLE 1.
Demographic Characteristics of Participants in Three Dysphonic and Vocally Healthy Groups

Participants Dysphonic/Healthy Controls Group Female/Male	Number of Participants (Mean age in Years ±SD)	VHI Score Points ± SD (Range) First Filling	Second Filling	Sig. Level
Dysphonic group				
Total	108 (52.8 ± 13.1)	58 ± 21 (15-106)	54 ± 23 (1-107)	$P = 0.009$
Female	83 (51.8 ± 13.1)	59 ± 22 (17-106)	55 ± 24 (1-107)	$P = 0.014$
Male	25 (56.5 ± 12.6)	55 ± 18 (15-88)	53 ± 23 (15-106)	NS
Vocal fold paresis total (dg. J38.0)	37 (54.8 ± 13.2)	59 ± 23 (15-106)	55 ± 25 (11-107)	$P = 0.036$
Female	24 (53.2 ± 13.7)	60 ± 24 (17-106)	58 ± 26 (11-107)	NS
Male	13 (57.6 ± 12.2)	57 ± 20 (15-88)	50 ± 23 (15-86)	$P = 0.028$
Spasmodic dysphonia total (dg. R49.02)	38 (56.0 ± 13.8)	62 ± 17 (28-97)	61 ± 19 (20-106)	NS
Female	29 (55.3 ± 14.5)	64 ± 18 (28-97)	62 ± 18 (20-103)	NS
Male	9 (58.9 ± 11.4)	56 ± 16 (31-78)	60 ± 24 (36-106)	NS
Functional dysphonia total (dg. R49.01)	33 (46.9 ± 10.1)	52 ± 22 (23-99)	47 ± 24 (1-88)	NS
Female	30 (47.3 ± 9.9)	53 ± 22 (23-99)	47 ± 25 (1-88)	NS
Male	3 (43.3 ± 14.0)	42 ± 18 (25-61)	44 ± 22 (25-68)	NS
Healthy controls				
Total	52 (51.2 ± 14.5)	8 ± 6 (0-25)	6 ± 8 (0-29)	$P = 0.005$
Female	37 (49.4 ± 14.6)	8 ± 7 (0-25)	7 ± 8 (0-30)	$P = 0.005$
Male	15 (55.7 ± 3.7)	7 ± 6 (0-20)	6 ± 5 (0-15)	NS

Means, SDs and range in the VHI score in the first and second filling of the questionnaire. Difference between the first and the second score was tested with Wilcoxon Signed Ranks Test.

seconds from the middle of the sustained vowel phonation were used. The multiparametric tool AVQI gives an index value scaled between 0 and 10, as well as results for the sub-parameters smoothed cepstral peak prominence (CPPS), harmonics-to-noise ratio (HNR), shimmer local (SL), shimmer local dB (SLdB), general slope of the spectrum (Slope), and tilt of the regression line through the spectrum (Tilt).[46]

Statistical analyses
Statistical analyses were carried out with IBM SPSS Statistics for Windows (Version 26, IBM Corp.). Test-retest reliability was studied using Spearman's rank order correlation. The internal consistency of the questions in three parts of the VFI (F) questionnaire were calculated with Cronbach's alpha. The VHI questionnaire was used as a reference tool to study validity of the VFI questionnaire. Spearman's rank order correlation coefficient was used to find correlation between the results of the VFI(F) and VHI questionnaires and to compare the VFI(F) and the acoustic results. The interpretation used in this study for Spearman's rho correlation coefficient is as follows: r = 0.10−0.39 = weak correlation; r = 0.40−0.69 = moderate correlation; r = 0.70−0.89 = strong correlation; r = 0.90−1.00 = very strong correlation.[47] The Mann−Whitney *U* test was used to compare the VFI(F) and the results of acoustic analyses between patients and healthy controls. The Wilcoxon Signed Ranks Test was used to compare the results of the first and the second time questionnaire fillings. Nonparametric tests

were used because the distribution in some of the VFI(F) parameters was asymmetric. The receiver operating characteristic (ROC) statistic was used to evaluate the VFI(F)'s discrimination power between dysphonic and normophonic speakers and to define the threshold values for the VFI(F).

RESULTS
Correlation in all VFI(F) factors' test-retest (first and second VFI completion) results were strong (Spearman's rho). Factor 1 (*Tiredness of voice and avoidance of voice use*) correlated with rho 0.90 ($P < 0.01$), Factor 2 (*Physical discomfort associated with voicing*) showed rho 0.84 ($P < 0.01$), and Factor 3 (*Improvement of symptoms with rest*) had rho 0.76 ($P < 0.01$). The internal consistency of the VFI(F) was also high in all three factors of the questionnaire. Cronbach's alpha for Factor 1 was 0.97, for Factor 2 0.95, and for Factor 3 0.92.

The VFI(F) results were compared with the VHI results to study the content validity of the VFI(F). Factor 1 of the VFI(F) correlated strongly with the VHI questionnaire, tested with the Spearman's rank order correlation coefficient. It yielded rho = 0.93 ($P < 0.01$). Factors 2 and 3 of the VFI(F) correlated moderately with the VHI, (Factor 2 rho = 0.66, $P < 0.01$; Factor 3 rho = 0.53, $P < 0.01$).

The discriminant (construct) validity of the VFI(F) was tested between the voice patients and the vocally healthy controls using a nonparametric independent samples' *t* test (Mann−Whitney *U* test). The test showed a significant

TABLE 2.
Score Results of VFI(F) in Voice Patients and Healthy Controls in all Three Parts of VFI Questionnaire

Group	Patients (N = 108) Mean ± SD (Median)	Controls (N = 52) Mean ± SD (Median)	Sig. Level
Factor 1	24.8 ± 8.2 (25)	3.8 ± 3.9 (0)	$P = 0.000$
Tiredness of voice and avoidance of voice use			
Factor 2	7.0 ± 5.1 (7)	1.1 ± 2.2 (6)	$P = 0.000$
Physical discomfort associated with voicing			
Factor 3	6.6 ± 3.0 (7)	2.6 ± 3.1 (2)	$P = 0.000$
Improvement of symptoms with rest			

Difference between groups was tested with Mann−Whitney U test.

difference between the VFI(F) results of the patients and the healthy controls in all three factors of the VFI(F) ($P = 0.000$). The results indicate that the VFI(F) scores of the patients were clearly higher than the VFI(F) scores of the vocally healthy controls (Table 2). The results of all acoustic parameters also differed significantly between voice patients and healthy controls.

Participants filled in the VFI(F) questionnaire two times. Table 3 presents the results. The repeated results of VFI(F) total and of Factors II and III did not differ significantly either in the dysphonic patients or the healthy controls, while the results of Factor I differed significantly in both groups. To the best of our knowledge, repeated results of the VFI questionnaire have not been reported in any previous VFI validation studies.

Factor 1 of the VFI(F) correlated moderately with AVQI03.01$_{FIN}$ and the AVQI sub-parameters CPPS, HNR, Shim%, and ShdB, while no correlation was seen with SlopedB and TiltdB. Factor 2 of the VFI(F) correlated weakly with AVQI03.01$_{FIN}$, CPPS, and HNR. Factor 3 correlated weakly with AVQI03.01$_{FIN}$, CPPS, HNR, and Shim%. The correlation matrix of VFI(F) and AVQI03.01$_{FIN}$ is presented in Table 4. Figure 1 illustrates correlations between Factor 1 of the VFI(F) and AVQI03.01$_{FIN}$ (Figure 1 **A**) and CPPs (Figure 1 B).

Diagnostic accuracy in differentiating voice patients and healthy subjects was acceptable in all three factors of the VFI(F). Factor 1 of the VFI(F) had the highest discrimination power ($A_{ROC} = 0.985$), and for Factors 2 and 3 of the VFI(F), discrimination power was slightly lower but still strong ($A_{ROC} = 0.864$ and $A_{ROC} = 0.821$). To find the optimal threshold levels of VFI(F) and all three of its factors, the sensitivity and specificity were tested with ROC curves (Figure 2). The cut-off value for VFI(F) total score was 13.5, with sensitivity 0.963 and specificity 0.885. The threshold values of VFI(F) factors were as follows: Factor 1 (*Tiredness of voice and avoidance of voice use*) > 13 points; Factor 2 (*Physical discomfort associated with voicing*) > 2 points; and Factor 3 (*Improvement of symptoms with rest*) < 4 points. Diagnostic accuracy test results are summarized in Table 5.

DISCUSSION

This study investigated the reliability and validity of the Finnish translation of the VFI questionnaire. For this purpose, first, the test-retest reliability was measured by asking the participants to complete the VFI(F) questionnaire twice at an interval of two weeks. Next, the internal consistency of the questionnaire was studied by running correlation

TABLE 3.
Score Results of Filling the VFI(F) two Times in Voice Patients and Healthy Controls in all Three Parts of VFI Questionnaire

Group	VFI Score Points First Filling Mean ± SD (Range)	VFI Score Points Second Filling Mean ± SD (Range)	Sig. Level
Patients (N = 108)			
Factor 1	25 ± 8 (3-42)	23 ± 9 (0-43)	0.012
Factor 2	7 ± 5 (0-20)	7 ± 5 (0-18)	NS
Factor 3	7 ± 3 (0-12)	7 ± 3 (0-12)	NS
VFI(F) total	38 ± 13 (4-65)	37 ± 14 (7-64)	NS
Controls (N = 52)			
Factor 1	4 ± 4 (0-17)	3 ± 3 (0-16)	0.035
Factor 2	1 ± 2 (0-8)	1 ± 2 (0-8)	NS
Factor 3	3 ± 3 (0-12)	3 ± 3 (0-12)	NS
VFI(F) total	7 ± 8 (0-33)	7 ± 7 (0-30)	NS

Difference between the results of two filling was tested with the Wilcoxon Signed Ranks Test.

TABLE 4.
Correlation Between the Three VFI(F) Factors and AVQI03.01$_{FIN}$ and its sub-parameters Tested with Spearman's rho

	AVQI03.01FIN	CPPs$_{AVQI}$	HNR	Shim %	ShdB	Slope dB	Tilt dB
VFI(F) Part 1	0.60*	−0.59*	−0.43*	0.46*	0.46*	−0.11	0.36*
VFI(F) Part 2	0.23*	−0.22*	−0.16[†]	0.14	0.13	0.08	0.10
VFI(F) Part 3	0.21*	−0.22*	−0.17[†]	0.17[†]	0.14	0.04	0.06

* Correlation is significant at the level 0.01.
[†] Correlation is significant at the level 0.05.

analyses between the answers related to each of the three factors of the questionnaire. Content validity was tested by running correlation analyses between the results of the VFI (F) and the VHI(F) and with acoustic analysis using AVQI. Construct validity was tested with *t* tests conducted between VFI(F) results from vocally healthy and voice-disordered participants. Finally, the sensitivity and specificity of VFI (F) was studied with ROC analysis. This yielded threshold values to differentiate between vocally healthy and voice-disordered participants.

According to the results, in all three factors, the answers given during the second completion of the questionnaire correlated strongly with the answers given during the first (rho 0.76−0.90). Similar results have been obtained in other languages, such as English,[20] German,[21] and Turkish.[26] In a

Hong Kong study, Factor 3 yielded only "acceptable" test-retest correlation (rho 0.702), according to the interpretation of Kwong et al.[25] On the other hand, rho 0.70−0.89 can be interpreted as strong, as in the present study, following Schober et al.[47] However, in our study, Factor 3 was also the one with the lowest test-retest correlation (rho 0.76). Similar results were also found for German[21] and Turkish.[26] We also compared the results of repeated completions of the questionnaire in more detail than the previous validation studies. We observed that repeated results of the total VFI and those of Factors II and III did not differ significantly, but results of Factor I did. The same observations were made for the dysphonic group and the control group. Factor I, i.e., *Tiredness of voice and avoidance of voice use*, thus seems to show somewhat more variation from time to time than the factors *Physical discomfort associated with voicing* and *Improvement of symptoms.*

Internal consistency was good in the present study: Cronbach's alpha varied from 0.92 to 0.97 in all three factors. Similar results have been obtained in other studies.[20,21,25,26]

Correlation between the results of the VFI(F) and the VHI (F) showed good content validity. Correlation was strong for Factor 1 (*Tiredness of voice and avoidance of voice use*; rho
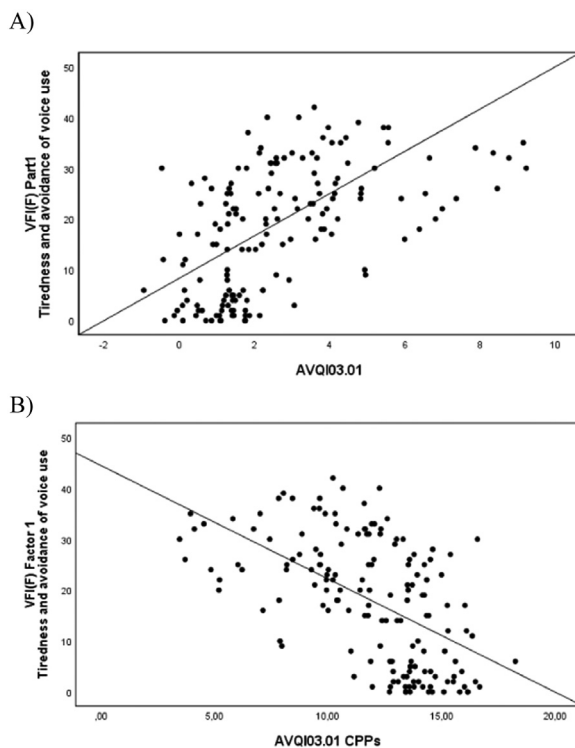

FIGURE 1. A) Relations between VFI(F) Factor 1 and AVQI03.01$_{FIN}$ score points (Spearman's rho 0.60, *P* = 0.01), and B) VFI(F) Factor 1 and AVQI03.01$_{FIN}$ CPPs depicted with scatterplots and linear regression lines (Spearman's rho -0.59, *P* = 0.01).
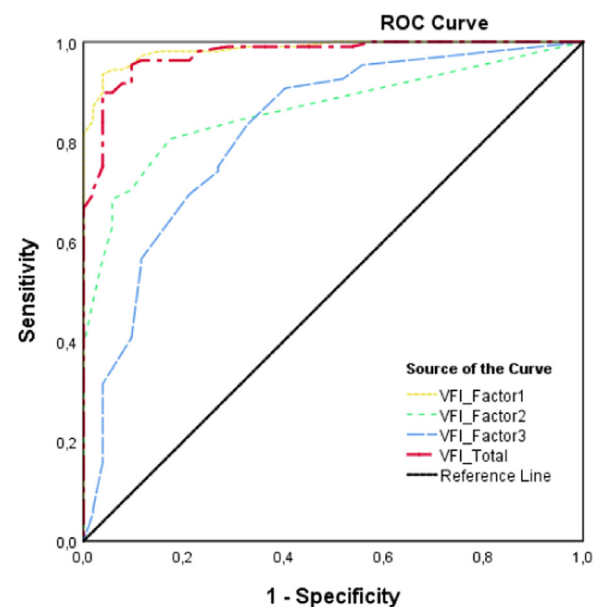

FIGURE 2. ROC curve analysis for the three parts of the VFI(F) to discriminate voice patients and vocally healthy participants.

TABLE 5.
Diagnostic Accuracy Results of the VFI(F)

VFI subscale	A_{ROC}	Threshold value	Sensitivity	Specificity
Factor 1	0.985	≥ 13	0.935	0.962
Factor 2	0.864	≥ 1.5	0.806	0.827
Factor 3	0.821	≤ 4.25	0.750	0.731
VFI(F) total	0.975	≥ 13.5	0.963	0.885

Summary of area under curve (AUC) values of the receiver operating characteristic (ROC) curves, cutoff values, sensitivity, and specificity data of Factors 1, 2, and 3 of the VFI(F).
Factor 1 questions 1−11, Factor 2 questions 12−16, Factor 3 questions 17−19.

0.93). Factors 2 (*Physical discomfort related to voice use*) and 3 (*Improvement of symptoms with rest*), in turn, correlated moderately with VHI (rho 0.66 and 0.53, respectively). Earlier studies have investigated the content validity of the VFI by comparing it with the V-RQOL (Turkish and Polish) and with the VHI (Spanish and Polish). The V-RQOL has shown high correlations with the VFI (Turkish and Polish). The correlations between the VFI (total and all three factors) and the VHI were good in Polish, while low correlation has been found between Factor 3 of the VFI and the VHI in Spanish.[27] The lower correlations between Factors 2 and 3 of the VFI and the VHI are understandable, since the VHI is more concerned with social handicap due to voice disorder than physical discomfort and recovery of voice symptoms.

Construct validity was high in the present study: VFI(F) values differed significantly between vocally healthy and voice-disordered participants in all three factors. This has been found in many other studies as well,[21,26,31] while in Hong Kong Chinese, Factor 3 did not differentiate between the two groups.[25] In the present study, the discriminating power was excellent for Factor 1 (Aroc 0.985) and strong for the other two factors (Aroc 0.864 and 0.821, respectively). Similar results have been reported in most other validation studies. The cut-off values vary in different validation studies; this is potentially related to cultural aspects and differences in the participants. Our cut-off values for the factors were >13, >1.5, and <4. The cut-off values are slightly lower than in the English,[20] German,[21] and Malayalam[30] studies but more similar to those in the Turkish study, where the values were >16, >4, and <4.[26] The Turkish study included a large number of patients (N = 285) with a wide spectrum of diagnoses (sulcus vocalis, nodule, polyp, cyst, edema, presbyphonia, premalignant lesion).[26] Similarly, the variation of diagnoses was large in the English,[20] Polish,[31] and German[21] studies. The Hong Kong study had a smaller number of participants with vocal fatigue (N = 87), and moreover, they were self-referred and thus not representative of patients with specified diagnoses.[25] The Malayalam study had a large number of participants (N = 528), but they were teachers without any diagnoses.[30] In the present study, the voice disorders included paresis, spasmodic dysphonia, and functional dysphonia. These groups represent large diagnostic groups in the clinics in Finland. The patients in this study also represented dysphonia severity from mild to severe. It is possible to speculate whether a wider range of diagnoses eg, would raise specificity and cut-off value for Factor 2. However, despite of the fact that only three diagnosis groups were included, the results did not differ substantially from those reported for patients with a very large variation of diagnoses, which indicates that the results of the present study give solid evidence of the usability of VFI in Finnish population.

In the present study, the VFI(F) values were compared with the results from acoustic analysis using AVQI.[46] According to the results, Factor 1 had moderate correlation with AVQI and its sub-parameters CPPS, HNR, and Shimmer (measured both in % and in dB). Other factors correlated only weakly with AVQI parameters. It seems plausible that tiredness of voice correlates with the acoustic characteristics of hoarseness, while spectral slope and tilt may vary more with type of phonation. For instance, it is known that slope value increases with SPL[48,49] and with firmness of phonation.[50,51]

Other studies have mainly investigated relations between total scores of AVQI and VHI and have not concerned themselves with sub-parameters of AVQI. However, Faham et al[52] observed that VHI correlated not only with AVQI but also with CPPS, although the correlation was low. However, the participants in the study by Faham et al. were ordinary university students with no known voice disorders.[52] The results of the present study suggest that VFI results are also supported by the results of acoustic analysis. Participants with a greater amount of hoarseness-related acoustic characteristics score higher in Factor 1 (*Tiredness of voice and avoidance of voice use*).

CONCLUSION

VFI(F) is a valid and reliable tool to distinguish voice-disordered and normal-voiced speakers. VFI(F) correlated with VHI. Tiredness of voice and avoidance of voice use (Factor 1 in the VFI) correlates also with acoustic characteristics related to hoarseness (AVQI and its sub-parameters CPPS, HNR, and Shimmer).

DECLARATION OF COMPETING INTEREST

None.

APPENDIX A
Vocal Fatigue Index (VFI)

ÄÄNENVÄSYMISINDEKSI

Nimi _____Henkilötunnus

_____

Pvä _____

Alla on joitakin oireita, joita tavallisesti liittyy ääniongelmiin. Merkitse rasti taulukon väittämiin aina sen vaihtoehdon kohdalle (0−4), joka kuvaa sitä, miten usein koet kyseistä oiretta.

Asteikko: 0 = ei koskaan, 1 = ei juuri koskaan, 2 = joskus, 3 = melkein aina, 4 = aina

1.	Minun ei tee mieli puhua, kun olen käyttänyt ääntäni jonkin aikaa	0	1	2	3	4
2.	Ääneni väsyy, kun puhun paljon.	0	1	2	3	4
3.	Minusta tuntuu, että joudun ponnistelemaan puhuessani.	0	1	2	3	4
4.	Ääneni käheytyy, kun käytän sitä.	0	1	2	3	4
5.	Äänenkäyttö tuntuu vaivalloiselta.	0	1	2	3	4
6.	Rajoitan yleensä puhumistani, kun olen käyttänyt ääntäni jonkin aikaa.	0	1	2	3	4
7.	Vältän sosiaalisia tilanteita, joissa tiedän joutuvani puhumaan enemmän.	0	1	2	3	4
8.	Minusta tuntuu, etten voi jutella perheeni kanssa työpäivän jälkeen.	0	1	2	3	4
9.	Jo lyhyen äänenkäytön jälkeen äänentuottoni on työlästä.	0	1	2	3	4
10.	Minun on vaikea saada ääntäni kuulumaan, kun olen käyttänyt sitä jonkin aikaa.	0	1	2	3	4
11.	Jo lyhyen käytön jälkeen ääneni tuntuu heikolta.	0	1	2	3	4
12.	Minulla on illalla kipua kaulalla, kun olen puhunut päivän aikana.	0	1	2	3	4
13.	Minulla on illalla kipua kurkussa, kun olen puhunut päivän aikana.	0	1	2	3	4
14.	Äänenkäyttö tuottaa kipua, kun puhun paljon.	0	1	2	3	4
15.	Kurkkuuni sattuu, kun käytän ääntäni.	0	1	2	3	4
16.	Kaulallani on epämiellyttävä tunne, kun käytän ääntäni.	0	1	2	3	4
17.	Ääneni tuntuu paremmalta levon jälkeen.	0	1	2	3	4
18.	Äänentuoton työläys vähenee lyhyen levon myötä.	0	1	2	3	4
19.	Käheys vähenee, kun annan ääneni levätä.	0	1	2	3	4

REFERENCES

1. Jones K, Sigmon J, Hock L, et al. Prevalence and risk factors for voice problems among telemarketers. *Arch Otolaryngol Head Neck Surg*. 2002;128:571–577.
2. Roy N, Merrill RM, Thibeault S, et al. Prevalence of voice disorders in teachers and the general population. *J Speech Lang Hear Res*. 2004;47:281–293.
3. Laukkanen AM, Kankare E. Vocal loading related changes in male teachers' voices investigated before and after a working day. *Folia phoniatr logop*. 2006;58:229–239.
4. Kankare E, Geneid A, Laukkanen AM, et al. Subjective evaluation of voice and working conditions and phoniatric examination in kindergarten teachers. *Folia Phoniatr Logop*. 2012;64:12–19.
5. Cantor Cutiva L. Association between occupational voice use and occurrence of voice disorders: a meta-analysis. *Areté*. 2018;18:1657–2513.
6. Cantor Cutiva L, Vogel I, Burdorf A. Voice disorders in teachers and their associations with work-related factors: a systematic review. *J com dis*. 2013;46:143–155.
7. Boltežar L, Bahar MŠ. Voice disorders in occupations with vocal load in Slovenia. *Slovenian J Public Health*. 2014;53:304–310.
8. Hagelberg AM, Simberg S. Prevalence of voice problems in priests and some risk factors contributing to them. *J Voice*. 2015;29:389.e11–389.e18.
9. Pestana PM, Vaz-Freitas S, Manso MC. Prevalence of voice disorders in singers: systematic review and meta-analysis. *J Voice*. 2017;31:722–727.
10. Prebil N, Boltežar IH, Bahar MŠ. Risk factors for voice problems in professional actors and singers. *Slovenian J Public Health*. 2020;59:92–98.
11. Solomon NP. Vocal fatigue and its relation to vocal hyperfunction. *Int J Speech Lang Pathol*. 2008;10:254–266.
12. Welham N, Maclagan M. Vocal fatigue: current knowledge and future directions. *J Voice*. 2003;17:21–30.
13. Scherer RC, Titze IR, Raphael BN, et al. Vocal Fatigue in a Trained and an Untrained Voice User. In: Baer T, Sasaki C, Harris K, eds. *Laryngeal Function in Phonation and Respiration*. San Diego, Calif: Singular Publishing Group; 1991:533–555.
14. Lauri ER, Alku P, Vilkman E, et al. Effects of prolonged oral reading on time-based glottal flow waveform parameters with special reference to gender differences. *Folia Phoniatr Logop*. 1997;49:234–246.
15. Rantala L, Vilkman E. Relationship between subjective voice complaints and acoustic parameters in female teachers' voices. *J Voice*. 1999;13:484–495.
16. Jónsdóttir V, Laukkanen AM, Siikki I. Changes in teachers' voice quality during a working day with and without electric sound amplification. *Folia phoniatr logop*. 2003;55:267–280.
17. Laukkanen AM, Ilomäki I, Leppänen K, et al. Acoustic measures and self-reports of vocal fatigue by female teachers. *J Voice*. 2008;22:283–289.
18. Titze IR. *Principles of Voice Production*. 2nd ed. Iowa City: Prentice Hall; 2000.
19. Verdolini K, Rosen CA, Branski RC. *Classification Manual of Voice Disorders-I*. 1st Ed. New York: Psychology Press; 2014.
20. Nanjundeswaran C, Jacobson BH, Gartner-Schmidt J, et al. Vocal Fatigue Index (VFI): development and validation. *J Voice*. 2015;29:433–440.
21. Barsties v. Latoszek B, Auner M, Graf S. Cross-cultural adaption and validation of the Vocal Fatigue Index in German. *J Voice*. 2019;35:161.e1–161.e13.
22. Barsties v. Latoszek B, Göllner M, et al. The German restructured vocal fatigue index and characteristics of dysphonic and vocally healthy populations. J Voice.2021. In press. https://doi.org/10.1016/j.jvoice.2021.01.024.

23. Zambon F, Moreti F, Nanjundeswaran C, et al. Cross-cultural adaptation of the Brazilian version of the Vocal Fatigue Index - VFI. *Codas.* 2017;29: e20150261.

24. Zambon F, Moreti F, Ribeiro VV, et al. Vocal Fatigue Index: validation and cut-off values of the Brazilian version. *J Voice.* 2022;36:434. e17–434.e24.

25. Kwong E, Tsang MLC. Cross-cultural adaptation and validation of the Hong Kong-Chinese version of Vocal Fatigue Index (VFI (HK)). *J Voice.* 2021 In press. https://doi.org/10.1016/j.jvoice.2021.03.019.

26. Şirin S, Öğüt MF, Bilgen C. Reliability and validity of the Turkish version of the Vocal Fatigue Index. *Turkish J Medical Sciences.* 2020;50:902–909.

27. Contreras-Regatero S, Vila-Rovira J, Verdejo C. Validity and reliability of Spanish version of two questionnaires of vocal fatigue in female teachers. *J Voice.* 2021;35:159.e1–159.e9.

28. Naderifar E, Moradi N, Farzadi F, et al. Cross-cultural adaptation and validation of the Vocal Fatigue Index into Persian. *J Voice.* 2019;33:947.e35–947.e41.

29. Bonetti A, Bonetti L, Cipcic O. Self-assessment of vocal fatigue in muscle tension dysphonia and vocal nodules: a preliminary analysis of the discriminatory potential of the Croatian adaptation of the Vocal Fatigue Index (VFI-C). *J Voice.* 2021;35:325.e1–325.e15.

30. Athira U, Devadas U. Adaptation and validation of Vocal Fatigue Index (VFI) to Malayalam language. *J Voice.* 2020;34:810.e19–810.e24.

31. Morawska J, Niebudek-Bogusz E, Pietruszewska, W. Linguistic and cultural adaptation and validation of the Vocal Fatigue Index into Polish. *J Voice.* In press. https://doi.org/10.1016/j.jvoice.2022.09.012.

32. Jacobson BH, Alex J, Cynthia G, et al. The Voice Handicap Index (VHI). *Am J Speech Lang Pathol.* 1997;3:66–70.

33. Alaluusua S, Johansson M. *Psycho-Social Handicap of Voice Disorder and its Rehabilitation: A Pilot Study of the Finnish Version of the Voice Handicap Index [In Finnish]. Master Thesis,* Helsinki: Institute of Behavioral Sciences, University of Helsinki; 2003.

34. Maryn Y, De Bodt M, Roy N. The Acoustic Voice Quality Index: toward improved treatment outcomes assessment in voice disorders. *J Com Dis.* 2010;43:161–174.

35. Hirano M. *Psycho-Acoustic Evaluation of Voice.* New York: Springer-Verlag; 1981.

36. Reynolds V, Buckland A, Bailey J, et al. Objective assessment of pediatric voice disorders with the Acoustic Voice Quality Index. *J Voice.* 2012;26. 672.e1-672.e7.

37. Maryn Y, Kim HT, Kim J. Auditory-perceptual and acoustic methods in measuring dysphonia severity of Korean speech. *J Voice.* 2016;30:587–594.

38. Jayakumar T, Benoy JJ. Acoustic Voice Quality Index (AVQI) in the measurement of voice quality: a systematic review and meta-analysis. *J Voice.* 2022. https://doi.org/10.1016/j.jvoice.2022.03.018. In press.

39. Batthyany C, Barsties v, Latoszek B, et al. Meta-analysis on the validity of the Acoustic Voice Quality Index. *J Voice.* 2022. https://doi.org/10.1016/j.jvoice.2022.04.022. In press.

40. Englert M, Barsties v, Latoszek B, et al. Exploring the validity of acoustic measurements and other voice assessments. *J Voice.* 2022. https://doi.org/10.1016/j.jvoice.2021.12.014. In press.

41. Pommée T, Maryn Y, Finck C, et al. The acoustic voice quality index, version 03.01, in French and the Voice Handicap Index. *J Voice.* 2020;34. 646-e1-646.e10.

42. Kankare E, Barsties v, Latoszek B, et al. The acoustic voice quality index in Finnish-speaking population. *Logoped Phoniatr Vocol.* 2019;30:1–8.

43. Kankare E, Rantala L, Ikävalko T, et al. Akustisen äänenlaatuindeksin (AVQI) version 03.01 validointi suomenkielisille puhujille. *Puhe ja kieli.* 2020;40:165–182.

44. Ustun TB, Kostanjesek N, Chatterji S, eds. *Measuring Health and Disability: Manual for WHO Disability Assessment Schedule (WHODAS 2.0).* Geneva: World Health Organization; 2010.

45. Boersma P, Weenink, D. Praat [computer program]: Doing Phonetics by Computer (Version 6.0.39), www.praat.org, 2023, March 26. https://www.fon.hum.uva.nl./praat/.

46. Maryn Y, Weenink D. Objective dysphonia measures in the program Praat: smoothed cepstral peak prominence and Acoustic Voice Quality Index. *J Voice.* 2015;29:35–43.

47. Schober P, Boer C, Schwarte LA. Correlation coefficients: appropriate use and interpretation. *Anesth Analg.* 2018;126:1763–1768.

48. Nordenberg M, Sundberg J. Effect on LTAS of vocal loudness variation. *Logoped Phoniatr.* 2004;29:183–191.

49. Pabon P, Ternström S. Feature maps of the acoustic spectrum of the voice. *J Voice.* 2020;34:161.e1–161.e26.

50. Kitzing P. LTAS criteria pertinent to the measurement of voice quality. *J Phon.* 1986;14:477–482.

51. Laukkanen AM, Rantala L. Does the Acoustic Voice Quality Index (AVQI) correlate with perceived creak and strain in normophonic young adult Finnish females? *Folia Folia phoniatr logop.* 2022;74:62–69.

52. Faham M, Laukkanen AM, Ikävalko T, et al. Acoustic Voice Quality Index as a potential tool for voice screening. *J Voice.* 2021;35:226–232.