

Cross-cultural Adaptation and Validation of the Italian Version of the Vocal Tract Discomfort Scale (I-VTD)

*Carlo Robotti, *Francesco Mozzanica, *Ilaria Pozzali, *Laura D'Amore, *Patrizia Maruzzi, †Daniela Ginocchio, †Stafania Barozzi, *Rosaria Lorusso, ‡Francesco Ottaviani, and *Antonio Schindler, *†‡Milan, Italy

Summary: Objective. To evaluate the reliability and validity of the Italian version of the VTD scale (I-VTD scale).

Study Design. Cross-sectional, nonrandomized, prospective study with controls.

Methods. For the item generation, a cross-cultural adaptation and translation process was performed following the back translation process (phase 1). For reproducibility analysis (phase 2), 102 patients with dysphonia were recruited (internal consistency analysis); 57 of them completed the I-VTD scale twice (test-retest reliability analysis). Seventy-three vocally healthy participants completed the I-VTD scale for normative data generation (phase 3). For validity analysis (phase 4), the scores obtained by patients with dysphonia and by vocally healthy participants were compared (construct validity analysis); in addition, 45 patients with dysphonia completed both the I-VTD scale and the Italian version of the Voice Handicap Index for criterion validity analysis. Finally, for responsiveness analysis (phase 5), a cohort of 30 patients with muscle tension dysphonia was recruited, and scores of the I-VTD scale before and after voice therapy were compared.

Results. Both the internal consistency and the test-retest reliability of the I-VTD scale were satisfactory. The scores obtained by patients with dysphonia and vocally healthy participants were significantly different. Moderate correlations between the Italian version of the Voice Handicap Index and the I-VTD scores were found. Finally, the scores of the I-VTD scale obtained in pretreatment conditions appeared to be significantly higher than those obtained after successful voice therapy.

Conclusion. The I-VTD scale appears a reliable and valid instrument for the assessment of vocal tract discomfort in Italian-speaking patients.

Key Words: Quality of life–Dysphonia–Vocal tract discomfort–Voice–Cross-cultural adaptation.

INTRODUCTION

Voice-related quality of life (VR-QOL) questionnaires are essential instruments in outcome research. These instruments assist clinicians in quantifying the level of handicap perceived by the patients because of voice problems, being useful in the evaluation of treatment success and representing an essential part of the multidimensional assessment of patients with dysphonia.^{1–3}

The importance of VR-QOL questionnaires lies in the fact that no objective measurements or perceptual ratings are able to quantify the handicap perceived by patients as a direct consequence of their vocal impairment.^{4,5} Several questionnaires are available, such as the Voice Handicap Index (VHI), the Voice Symptom Scale (VoiSS), and the VR-QOL.^{6–8} Most of them focus on specific alterations of commonly screened vocal parameters (pitch, pitch range, vocal note quality, loudness, flexibility, stamina) and their impact on QOL. However, patients can sometimes experience a feeling of discomfort other than dysphonia and specifically referred to the vocal tract. Vocal tract discomfort (VTD) is usually reported as a variable combination of symptoms of inflammation (sore, burning, tickling, irritability), musculoskeletal symptoms (aching, constriction, tightness), and

feeling of increased secretions,⁹ which can be described by patients as a low-level pain (on a scale of “no pain” to “unbearable pain”). VTD does not necessarily represent a symptom of tissue damage, but it is considered as the consequence of an impairment of the pain system as a whole, affecting both peripheral and central neural pathways.¹⁰

VTD is a common complaint in patients with hyperfunctional dysphonia,^{11,12} in teachers^{13,14} and, more generally, among professional voice users.¹⁵ In fact, frequency and severity of the VTD appear to be higher in settings characterized by a high vocal load.¹⁶ In addition, patients affected by benign vocal folds lesions and gastroesophageal reflux disease report VTD more frequently than do patients with neurologic vocal disturbances.¹⁷

Occasionally, VTD can be disregarded, improperly assuming that, as an intrinsic component of voice disorders, it will resolve autonomously as the intervention progresses.⁹ However, the presence of VTD must be taken into account as it is usually associated with a negative perception of the vocal experience and with particular signs and symptoms that might imply an emerging vocal disorder.¹³ For these reasons, VTD can be more distressing and cause more concern than the actual vocal changes, partly because of the unpleasant perception itself but also because this discomfort may give rise to fear of severe underlying pathologies, namely cancer of the larynx.⁹ To obtain information regarding the patients' VTD, Mathieson and colleagues^{9,10,12} developed the VTD scale (Appendix). The latter is a self-rating questionnaire comprising eight patient-derived qualitative items describing common symptoms or sensations related to the vocal tract (burning, tight, dry, aching, tickling, sore, irritable, lump in the throat). Each symptom or sensation is rated on a 7-point Likert scale (from 0 to 6), both for its frequency (from never

Accepted for publication September 7, 2017.

From the *Department of Biochemical and Clinical Sciences, Luigi Sacco Hospital, University of Milan, Milan, Italy; †Department of Clinical Sciences and Community, Fondazione IRCCS Ca' Granda, Ospedale Maggiore Policlinico, University of Milan, Milan, Italy; and the ‡Department of Clinical Sciences and Community, San Giuseppe Hospital, IRCCS Multimedica, University of Milan, Milan, Italy.

Address correspondence and reprint requests to Francesco Mozzanica, Department of Biochemical and Clinical Sciences, Luigi Sacco Hospital, University of Milan, Via G.B. Grassi 74, Milan 20157, Italy. E-mail: francesco.mozzanica@gmail.com

Journal of Voice, Vol. 33, No. 1, pp. 115–123

0892-1997

© 2017 The Voice Foundation. Published by Elsevier Inc. All rights reserved.

<https://doi.org/10.1016/j.jvoice.2017.09.001>

to always) and its severity (from none to extreme). The score for each subscale ranges from 0 to 48 points, whereas the total VTD score ranges from 0 to 96.

The VTD scale has already been applied to several cultural and linguistic contexts^{13,16,18–20} and tested within different clinical settings.^{10,13,15} To date, however, a validated Italian version of this instrument is not available. The aim of this study is to evaluate the reliability, validity, and the responsiveness to changes of the Italian version of the VTD scale (I-VTD scale). The underlying hypothesis is that the VTD scale can be adapted to Italian and that its reliability, validity, and responsiveness to changes are satisfactory. The relevance of this research lies in the fact that the I-VTD scale, allowing the evaluation of dysphonia-related sensory symptoms, could be particularly helpful in the management of patients with dysphonia, providing further important information for both the decision-making process and the outcome evaluation.¹⁶ In addition, the availability of I-VTD scale could assure a prompt dealing of patients' discomfort and this could increase their confidence in the clinician and facilitate their response to further intervention.¹⁰ Finally, an Italian validated version of the VTD could allow the accomplishment of cross-cultural and cross-national multicenter studies.

MATERIALS AND METHODS

This prospective cross-sectional study was structured through five different phases: five-step back translation (phase 1); reproducibility analysis (phase 2); normative data generation (phase 3); validity analysis (phase 4); and responsiveness analysis (phase 5). The Consensus-based Standards for the Selection of Health Measurement Instruments checklist²¹ was followed, to obtain proper conclusions regarding the measurement properties of the I-VTD scale. The present study was conducted in accordance with the Declaration of Helsinki and it was previously approved by the institutional review board of our hospital.

Participants

Different groups of subjects were recruited through the different phases of this study (Table 1). A written informed consent was obtained by all study participants. All data were collected prospectively. The inclusion criteria were age 18 years or older;

good understanding of Italian (written and spoken); absence of relevant cognitive impairment (Mini-Mental State Examination score >24 for participants older than 65 years); and preserved reading skills.

Phase 1: Five-step I-VTD item generation

A cross-cultural adaptation process was performed following standard techniques.^{22,23} The items of the original VTD scale were first translated into Italian by a bilingual phoniatrician and a professional translator (stage 1: forward translation). Semantic, conceptual, and idiomatic issues of the two translations were then discussed by the abovementioned translators and a speech and language pathologist (SLP) to produce a preliminary Italian version of the questionnaire (phase 2: synthesis). Subsequently, two native English-speaking professional translators with no medical background and not involved in the previous stage literally translated the VTD scale back to English (stage 3: back translation). This stage was performed to assess possible information bias and to make sure that the translated version of the VTD scale was clearly reflecting the same item content of the original instrument. Afterward, the translators and an expert committee reviewed all reports up to this point to create a prefinal version of the instrument (stage 4: expert committee review). Lastly, 30 subjects (15 males and 15 females; mean age 48.2 years; age range 18–75) reporting symptoms of dysphonia were recruited for a pilot study (stage 5: pretesting). The etiology of dysphonia was as follows: laryngopharyngeal reflux (seven patients); vocal fold polyp (four patients); vocal fold nodules (two patients); vocal fold cyst (one patient); functional dysphonia (six patients); Reinke edema (four patients); unilateral recurrent laryngeal nerve injury (one patient); vocal fold leukoplakia (three patients); and laryngeal cancer (two patients). Each patient autonomously completed the prefinal version of the I-VTD scale and then discussed wording and meaning of each item with a senior clinician. The wording of the items was then examined and adjusted to increase the questionnaire readability. Overall, three of the eight items of the VTD scale were adjusted. Subsequently, the committee members and all the professional translators analyzed all the items to confirm the cross-cultural equivalence (semantic, idiomatic, experimental, and conceptual)

TABLE 1.
Clinical and Demographic Characteristics of Patients and Vocally Healthy Participants

| Phase of the Study | Type of Study | Sample Clinical Characteristics | Age* (y) | Sex (n) | |
|-----------------------------|---|---------------------------------------|--------------|---------|----|
| | | | | M | F |
| 1. Item generation | Item generation in five steps | Patients with dysphonia (n = 30) | 48.2 (18–75) | 15 | 15 |
| 2. Reproducibility analysis | Internal consistency | Patients with dysphonia (n = 102) | 54.2 (18–84) | 35 | 67 |
| | Test-retest reliability | Patients with dysphonia (n = 57) | 53.6 (18–84) | 20 | 37 |
| 3. Normative data | Normative data generation | Vocally healthy participants (n = 73) | 47.7 (18–79) | 38 | 35 |
| 4. Validity analysis | Construct validity | Patients with dysphonia (n = 102) | 54.2 (18–84) | 35 | 67 |
| | Criterion validity | Vocally healthy participants (n = 73) | 47.7 (18–79) | 38 | 35 |
| 5. Responsiveness analysis | Comparison of scores before and after speech and language therapy | Patients with dysphonia (n = 45) | 49.7 (20–84) | 14 | 31 |
| | | Patients with dysphonia (n = 30) | 52.7 (18–75) | 11 | 19 |

* Mean age; age range are reported in parentheses.

between the original and the final versions of the I-VTD scale—that is, the ability of the translated version to retain the meaning of the original content—following standard guidelines.²⁴ This process ultimately led to the final version of the Italian VTD scale (I-VTD scale; see [Appendix](#)).

Phase 2: I-VTD scale reproducibility analysis

This phase of the study was designed to assess both internal consistency and test-retest reliability of the I-VTD scale. For this purpose, 102 subjects with dysphonia (35 males and 67 females; mean age 54.2 years; age range 18–84 years) consulting for voice disorders were consecutively recruited. According to the etiology, all patients with dysphonia were assigned to one of the following diagnosis-related groups: (1) functional ($n = 23$; mean age 50.4 years; age range 18–84); (2) organic ($n = 55$; mean age 52.6; age range 24–83); and (3) neurologic ($n = 24$; mean age 61.3; age range 18–84). Gender distribution among diagnosis-related groups is depicted in [Table 2](#).

Internal consistency evaluates the extent to which each item in a factor measures the same underlying construct. We considered Cronbach alpha values between 0.7 and 0.9 as indicators of satisfactory internal consistency.²⁵ For internal consistency analysis, the I-VTD scale scores obtained within the baseline examination of the 102 patients with dysphonia were used. These scores were also used for validity analysis in phase 4 of this study. Out of the 102 individuals with dysphonia involved in this phase of the study, 57 subjects (20 males and 37 females; mean age 53.6 years; age range 18–84 years) were randomly selected to evaluate test-retest reliability. For this purpose, all these individuals filled out the I-VTD scale twice within a 2-week interval because no substantial changes were expected to take place within this period. This time frame was selected in compliance with the patients' needs and basing on the fact that an interval ranging from 2 days to 2 weeks does not affect the results of test-retest reliability analysis, therefore representing a good compromise between recollection bias and unwanted (by the investigators) clinical changes.²⁶ Finally, no access to previous answers was granted to the subjects when fulfilling the I-VTD scale for the second time. Two-way random intraclass correlation coefficient (ICC) was used to assess test-retest reliability; the coefficient was calculated for each of the two subscales (frequency and severity).

TABLE 2.
Gender Distribution Among Diagnosis-related Groups

| Subgroup of Patients | Number of Patients (n) | Sex (n) | | Percentage (%) |
|----------------------|------------------------|---------|----|----------------|
| | | M | F | |
| Functional | 23 | 6 | 17 | 22.55 |
| Organic | 55 | 22 | 33 | 53.93 |
| Neurologic | 24 | 7 | 17 | 23.52 |
| Total | 102 | 35 | 67 | 100 |

Phase 3: I-VTD scale normative data generation

The aim of the third phase of the study was to determine the baseline distribution of I-VTD scores by collecting data from a representative and randomly selected sample of vocally healthy individuals. Therefore, 73 vocally healthy participants (38 males and 35 females; mean age 47.7 years; age range 18–79 years) were consecutively recruited, respecting the following inclusion criteria: no symptoms or history of dysphonia; no previous speech and language therapy nor previous laryngeal surgery; no noticeable deviation on vocal quality (score zero for the Grade parameter of the Grade, Roughness, Breathiness, Asthenia, Strain (GRBAS) scale, assessed by an SLP with more than 5 years of experience in the treatment of voice disorders);^{27,28} no major comorbidities (hearing, swallowing, respiratory, reflux, neurologic, rheumatologic, hematologic, or neoplastic disorders); no history of major head and neck traumas. The cohort included hospital staff members, medical and nursing students, and patients' relatives and visitors who agreed to take part in the study. Each vocally healthy participant completed the I-VTD scale without any help and underwent videolaryngostroboscopy to exclude laryngeal pathologies. The data obtained from this group of participants were also used for clinical validity analysis (phase 4 of the study).

Phase 4: I-VTD scale validity analysis

The purpose of this phase was the assessment of the I-VTD scale validity, which can be defined as the degree to which the instrument measures the construct it purports to measure. In particular, we focused our analysis on construct and criterion validity.

Construct validity assesses the degree to which I-VTD scale scores are consistent with the hypotheses. To evaluate construct validity, I-VTD scale scores obtained from patients with dysphonia ($n = 102$) were compared with those obtained from vocally healthy individuals ($n = 73$). To evaluate the effect of gender on VTD perception, both in dysphonic and in vocally healthy individuals, the I-VTD scores obtained in male and female subjects were also compared. For additional construct validity assessment, the I-VTD scores obtained in the three groups of patients were compared.

Criterion validity defines the degree to which I-VTD scale scores can be considered an adequate reflection of a gold standard. For criterion validity analysis, 45 subjects with dysphonia (14 males and 31 females; mean age 49.7 years; age range 20–84 years) were randomly selected among the three diagnosis-related groups (15 patients per group). All these patients were asked to complete autonomously, at the same time, both the I-VTD scale and the Italian version of the VHI (I-VHI).^{6,29} The I-VHI consists of 30 items, rated on a 5-point Likert scale from 0 (never) to 4 (always), exploring three different domains of voice-related problems: (1) functional, dealing with the influence of dysphonia on common daily routines; (2) emotional, evaluating the psychological impact of the vocal impairment; and (3) physical, dealing with the self-perception of laryngeal and vocal discomfort. The I-VHI has already been used in clinical epidemiology studies, representing an effective instrument to investigate VR-QOL in native Italian speakers affected by dysphonia.^{29–31}

The correlations between the scores of the I-VTD scale and the I-VHI were analyzed using Spearman test. The correlation was considered “strong” for values greater than 0.5, “moderate” for values between 0.3 and 0.5, and “weak” for values lower than 0.3.³²

Phase 5: I-VTD scale responsiveness analysis

Responsiveness consists in the ability of the instrument to detect relevant changes over time in the construct to be measured. To evaluate the I-VTD scale responsiveness, a new cohort of 30 patients affected by muscle tension dysphonia (MTD) was recruited (11 males and 19 females; mean age 52.7 years; age range 18–75 years); each subject autonomously filled out the I-VTD scale and the I-VHI before and after voice therapy (VT). Each patient with dysphonia underwent 45-minute VT sessions twice a week for at least 8 weeks. Therapy was directed toward the progressive development of optimal breathing technique, neck muscles relaxation, and reduction of hyperfunctional vocal behaviors; specifically, abdominal breathing, voiced-sigh technique, flow phonation, and laryngeal self-massage were applied. VT was administered by a senior SLP with at least 20 years of clinical experience in treating patients with dysphonia. The SLP was blind to the I-VTD scale and the I-VHI scores obtained before treatment. Each patient underwent a perceptual voice evaluation right before and after voice rehabilitation; the GRBAS scale^{27,28} was used to assess dysphonia severity. Overall, changes were considered to be statistically significant, and treatment was considered to be successful if the patient showed an improvement in dysphonia severity (at least a 1-point improvement in the G parameter of the GRBAS scale plus either an 8-point improvement in a single VHI subscale score or an 18-point improvement in the VHI total score).

Statistical analysis

Statistical tests were performed using the *SPSS Statistics* 23.0 (SPSS Inc, Chicago, IL). A significance level of $P < 0.05$ was adopted. Kolmogorov-Smirnov test was used to test the normality of the distribution of I-VTD scale scores among the different groups of participants. Because this test demonstrated that the distribution of I-VTD scale scores in all the groups was normal, parametric tests were used when requested. Internal consistency was assessed using Cronbach alpha coefficient. Two-way random ICC was used to evaluate I-VTD scale test-retest reliability by comparing baseline and retesting responses. The comparison of I-VTD scale scores in dysphonic and vocally healthy participants was assessed using Student *t* test. Analysis of variance (ANOVA) with Tukey *post hoc* analysis was used to compare the results obtained in the three groups of patients with dysphonia. The correlations between I-VTD scale and VHI were assessed using Pearson test. The distributions of I-VTD scale scores obtained in pre- and post-treatment conditions were compared using the Student *t* test. The effect size was calculated as the difference between the pretreatment group mean minus the post-treatment group mean, divided by the standard deviation of the initial values. For all statistical comparisons, an $\alpha = 0.05$ and a power of 0.80 were used.³³

RESULTS

Both patients with dysphonia and vocally healthy volunteers filled out the I-VTD scale without any assistance. Moreover, each study participant managed to complete the questionnaire in less than 5 minutes without any missing responses.

Phase 1: Five-step I-VTD scale item generation

The I-VTD scale was developed through a cross-cultural adaptation process performed following standard techniques (I-VTD scale; [Appendix](#)).

Phase 2: I-VTD scale reproducibility analysis

Cronbach alpha scores resulted satisfactory for both the I-VTD subscales: $\alpha = 0.92$ for the Frequency subscale and $\alpha = 0.94$ for the Severity subscale. ICC values for the 57 subjects recruited for test-retest reliability analysis of the I-VTD subscales scores are reported in [Table 3](#). Test-retest reliability was satisfactory for both the subscales: ICC = 0.92 (confidence interval = 0.88–0.95) for the Frequency subscale and ICC = 0.95 (confidence interval = 0.91–0.97) for the Severity subscale.

Phase 3: I-VTD scale normative data generation

The mean age of vocally healthy subjects ($n = 73$) was 47.7 years (range 18–79). Fifty-two percent of subjects were males. The mean score of the Frequency subscale of I-VTD was 3.4 ± 4.3 (range 0–17). The mean score of the Severity subscale of I-VTD was 3.3 ± 4.4 (range 0–18). In vocally healthy subjects, the mean single item score was always lower than 0.80 with the only exception for the item “Sore severity” (mean score 0.808 ± 1.2).

Phase 4: I-VTD scale validity analysis

As far as the construct validity of the I-VTD is concerned, the scores obtained by subjects with dysphonia were compared with the scores obtained by vocally healthy individuals using Student *t* test. The results of this comparison are reported in [Table 4](#). Student *t* test revealed a significant difference between dysphonic and vocally healthy subjects for both the I-VTD subscales. In [Table 5](#), the I-VTD scores obtained by dysphonic and vocally healthy male and female subjects are depicted. Although no differences between vocally healthy male and female subjects were found, significant differences in the I-VTD scores obtained by male and female patients with dysphonia were highlighted. In male patients with dysphonia, a frequency score of 12.7 ± 8.1 and a severity score of 12.9 ± 9.3 were found, whereas female

TABLE 3.
Internal Consistency ($n = 102$) and Test-Retest Reliability ($n = 57$) of the I-VTD Scale

| I-VTD Scale Subscale | Internal Consistency ($n = 102$) | ICC ($n = 57$) |
|----------------------|------------------------------------|------------------|
| Frequency | $\alpha = 0.92$ | 0.92 (0.88–0.95) |
| Severity | $\alpha = 0.94$ | 0.95 (0.91–0.97) |

Cronbach alpha values for internal consistency analysis and ICC with 95% CI values for test-retest reliability analysis are reported.

TABLE 4.
Comparison Between the I-VTD Scale Scores Obtained in Patients With Dysphonia (n = 102) and Vocally Healthy Individuals (n = 73)

| I-VTD Subscale | Dysphonic Patients (n = 102) | Vocally Healthy Subjects (n = 73) | Student <i>t</i> Test |
|----------------|------------------------------|-----------------------------------|-----------------------|
| Frequency | 15.3 ± 8.7 (1–35) | 3.4 ± 4.3 (0–17) | <i>P</i> = 0.014 |
| Severity | 15.5 ± 9.6 (0–36) | 3.3 ± 4.4 (0–18) | <i>P</i> = 0.024 |

Mean, standard deviation, and ranges (in parentheses) are reported. The results of Student *t* test are also reported.

patients with dysphonia scored 16.8 ± 8.7 and 16.9 ± 9.6 for the frequency and the severity subscale respectively.

For additional construct validity assessment, the I-VTD scores obtained in the three groups of patients were compared. The mean score of the Frequency subscale of I-VTD in patients with functional dysphonia was 17.5 ± 4.1; in patients with organic it was 14.8 ± 3.8, whereas in patients with neurologic dysphonia it was 14.3 ± 3.9. These differences were found significant on ANOVA comparison, in particular patients with functional dysphonia scored significantly higher than patients with organic and neurologic dysphonia (*P* = 0.041). The mean score of the Severity subscale

of I-VTD in patients with functional dysphonia was 18.3 ± 4.7; in patients with organic dysphonia it was 15.1 ± 4.1, whereas in patients with neurologic dysphonia it was 13.8 ± 3.2. These differences were found significant on ANOVA comparison, in particular patients with functional dysphonia scored significantly higher than patients with organic and neurological dysphonia (*P* = 0.037).

For the concurrent validity analysis, the correlations between the I-VTD and the I-VHI scores obtained in the group of patients with dysphonia were analyzed; the results are reported in Table 6. Significant correlations between the two scales were found. In particular, the highest correlation was found between the Frequency subscale of I-VTD and the Emotional subscale of I-VHI (*r* = 0.353; *P* < 0.01).

Phase 5: I-VTD responsiveness analysis

All patients with dysphonia demonstrated an improvement in their perceptual voice analysis during the follow-up control after undergoing VT for at least 8 weeks. For the responsiveness analysis, the scores of the I-VTD scale obtained in pretreatment and post-treatment conditions in a group of 30 subjects with dysphonia affected by MTD were analyzed. The differences were statistically significant (*P* = 0.042 at Student *t* test) for both the I-VTD subscale scores. Results of this analysis are depicted in Table 7.

TABLE 5.
Comparison Between the I-VTD Scale Scores Obtained by Male and Female Subjects in Both the Dysphonic (n = 102) and the Vocally Healthy (n = 73) Population

| I-VTD Subscale | Patients With Dysphonia (n = 102) | | | Vocally Healthy Subjects (n = 73) | | |
|----------------|-----------------------------------|-------------------|-----------------------|-----------------------------------|------------------|-----------------------|
| | Males (n = 35) | Females (n = 67) | Student <i>t</i> Test | Males (n = 38) | Females (n = 35) | Student <i>t</i> Test |
| Frequency | 12.7 ± 8.1 (1–30) | 16.8 ± 8.7 (1–35) | <i>P</i> = 0.019 | 2.9 ± 4.3 (0–14) | 4.3 ± 4.4 (0–17) | <i>P</i> = 0.807 |
| Severity | 12.9 ± 9.3 (0–31) | 16.9 ± 9.6 (1–36) | <i>P</i> = 0.037 | 2.8 ± 4.3 (0–18) | 4.5 ± 4.6 (0–15) | <i>P</i> = 0.850 |

Mean, standard deviation, and ranges (in parentheses) are reported. The results of Student *t* test are also reported.

TABLE 6.
Correlation Between the Scores of the I-VTD Scale and of the I-VHI in the Group of Patients With Dysphonia

| I-VTD subscale | | I-VHI | | | |
|----------------|-----------|----------|------------|-----------|--------|
| | | Physical | Functional | Emotional | Total |
| Frequency | Frequency | 0.201* | 0.271† | 0.353† | 0.340† |
| | Severity | 0.205* | 0.280† | 0.313† | 0.327† |

* *P* < 0.05.

† *P* < 0.01.

TABLE 7.
Mean ± Standard Deviation of I-VTD Subscales Scores in the Pre- and Post-Treatment Conditions (8 Weeks) in 30 Patients With Muscle Tension Dysphonia

| I-VTD Subscale | Pre-Treatment Score | Post-Treatment Score | <i>P</i> Score | Cohen <i>d</i> |
|----------------|---------------------|----------------------|----------------|----------------|
| Frequency | 17.5 ± 7.5 (6–33) | 9.5 ± 6.9 (0–20) | 0.001 | 1.110 |
| Severity | 18.4 ± 8.7 (6–36) | 9.8 ± 7.1 (0.17) | 0.001 | 1.083 |

The results of Student *t* test comparison are reported as well as those of Cohen effect size. Ranges are reported in parentheses.

DISCUSSION

The VTD scale is a self-assessment tool originally developed by Mathieson and colleagues¹⁰ to assess VTD in patients affected by MTD before and after laryngeal manual therapy. Since then, the VTD scale has been tested within several cultural and linguistic contexts^{13,16,18–20} and applied to different clinical settings.^{10,13,15} The present study was conducted as a structured translation and adaptation process, following specific methodological guidelines as suggested by Beaton et al²² and by the Consensus-based Standards for the Selection of Health Measurement Instruments checklist manual,²¹ to not alter the measurement capabilities of the original questionnaire and to allow cross-cultural and cross-language comparability of results.³⁴ All the members of the expert committee and the professional translators agreed on a satisfying cross-cultural equivalence (semantic, idiomatic, experimental, and conceptual) between the English and the Italian VTD scale, suggesting that the items of the I-VTD scale had retained the meaning of the original instrument. Specifically, the good level of equivalence between the two versions of the instrument may be because of the single-word formulation of the items of the English VTD scale (except for “Lump in the throat”). Regarding the psychometric properties of the I-VTD scale, the results of the present study highlight good internal consistency, test-retest reliability, clinical validity, and responsiveness. These results further sustain the adoption of the I-VTD as an auxiliary instrument within the QOL evaluation process of dysphonic individuals. Specific findings related to the I-VTD scale are remarkable. In particular, all the study participants managed to autonomously complete the questionnaire in 5 minutes or less and without leaving missing responses; this supports the idea that the I-VTD scale can be considered as an easy self-administrable tool with a satisfying comprehensibility.

The I-VTD scale internal consistency, obtained within the evaluation of 102 patients with dysphonia, appeared satisfactory with a Cronbach α coefficient value of 0.92 for the Frequency subscale and 0.94 for the Severity subscale. These results are not dissimilar to those reported by Mathieson et al ($\alpha = 0.890$ and $\alpha = 0.886$, respectively),¹⁰ Niebudek-Bogusz et al ($\alpha = 0.930$ and $\alpha = 0.936$, respectively),¹⁹ Rodrigues et al ($\alpha > 0.806$ for both the subscales),¹³ and Lukaschyk et al ($\alpha = 0.919$)²⁰; only in the study by Torabi et al¹⁸ Cronbach α values appeared to be slightly lower ($\alpha = 0.77$ for the frequency subscale and $\alpha = 0.73$ for the severity subscale) but equally suggestive of a high internal consistency.

Regarding the I-VTD scale test-retest reliability, the scores obtained in the test-retest analysis suggest excellent stability and reproducibility over time. In fact, the ICC was 0.92 and 0.95 for the frequency subscale and the severity subscale, respectively. These values can be considered highly satisfactory. Also, Torabi and colleagues,¹⁸ who developed a study to test the Persian version of the VTD scale within the assessment of 100 patients affected by MTD at baseline and after an interval of 1 week, highlighted a relevant test-retest reliability for the frequency and severity subscales of this instrument (ICC of 0.93 and 0.91, respectively).

As far as the normative data are concerned, the group of vocally healthy participants scored 3.4 ± 4.3 (range 0–17) for

the frequency subscale and 3.3 ± 4.4 (range 0–18) for the severity subscale of the I-VTD scale. These data appear to be in accordance with those found by Torabi et al,¹⁸ who reported a frequency score of 2.02 ± 2.45 (range 0–10) and a severity score of 2.14 ± 2.58 (range 0–10). Regarding construct validity, in agreement with previous studies, the present research highlighted that dysphonic individuals scored significantly higher values of the I-VTD scale on both frequency and severity subscales, in comparison with vocally healthy participants (15.3 ± 8.7 versus 3.4 ± 4.3 for the frequency subscale, 15.5 ± 9.6 versus 3.3 ± 4.4 for the severity subscale). Similar results were found by Torabi and colleagues.¹⁸ Niebudek-Bogusz et al¹⁹ found slightly higher values among patients affected by occupational dysphonia (a mean score of 23.5 for the frequency subscale and of 24.6 for the severity subscales). A significant difference in the I-VTD scores obtained by male and female patients with dysphonia was found, with higher frequency and severity scores for the female population. This difference might be partially because of the fact that women usually report more intense, more frequent, and more numerous bodily symptoms than men.³⁵ However, these data must be considered preliminary as they are based on small groups of patients with unequal sample sizes (35 male subjects versus 67 female patients with dysphonia). Larger studies with equal sample sizes could better investigate the effect of gender on VTD perception. In addition, significant differences in the scores obtained in patients with different type of dysphonia were found. These results suggest that the I-VTD scale may be a useful tool to identify subjects with dysphonia and help to discriminate among different type of dysphonia. However, in a recent study by Behlau et al,³⁶ aimed at comparing different voice self-assessment instrument for the evaluation of organic and behavioral dysphonia, the VTD scale was considered as the protocol that identified the least number of patients with dysphonia (63 out of 103 subjects). Therefore, further studies will be needed to properly evaluate the actual discriminative power of the VTD scale. Regarding the correlations between the I-VTD scale and the I-VHI, in the present study, moderate correlations between the total I-VHI score and the Frequency and Severity domains of the I-VTD scale ($r = 0.340$ and $r = 0.327$, respectively, $P < 0.01$) were highlighted, demonstrating an acceptable criterion validity of this voice self-assessment instrument. Similar results were reported by Torabi et al¹⁸ within the Persian validation study of the VTD scale ($r = 0.36$ and $r = 0.37$, respectively) and by the team of Behlau et al³⁶ within the evaluation of patients affected by behavioral dysphonia ($r = 0.400$ and $r = 0.435$, respectively, $P < 0.001$). A few authors found higher values, suggesting stronger correlations between the two instruments. For example, Lukaschyk and colleagues²⁰ demonstrated a strong correlation between the total scores of the VTD scale and the VHI ($r = 0.664$), whereas Niebudek-Bogusz et al¹⁹ found a strong correlation between the total VHI and the frequency and severity domains of the VTD scale ($r = 0.577$ and $r = 0.608$, respectively). On the other hand, Woznicka and colleagues¹⁵ found weak correlations between the Frequency and the Severity subscales of the VTD scale and the VHI ($r = 0.221$ and $r = 0.178$, respectively, $P = 0.000$). These differences regarding the correlation between the VTD

scale and the VHI could be partially related to the characteristics of the studied populations. In the present study, patients with different etiologies of dysphonia (organic, functional, neurologic) were recruited. Behlau *et al*³⁶ and Lukaschyk *et al*²⁰ included in their studies both functional and organic dysphonic patients, with most patients being part of the former group (75% and 54%, respectively). Torabi *et al*¹⁸ exclusively enrolled patients affected by MTD, whereas Niebudek-Bogusz *et al*¹⁹ and Woznicka *et al*¹⁵ focused on professional voice users with occupational dysphonia. These differences might be because of the assumption that the two instruments actually deal with slightly different aspects of a vocal disorder. In fact, although the VHI evaluates the self-perception of the handicap deriving from a vocal impairment, the VTD scale measures the degree of sensory discomfort referred to the vocal tract, which may (or may not) be related to the contextual development of dysphonia. Having considered all these aspects, we chose the VHI as the gold standard to allow comparability of results between all the validation studies available to date. However, it is possible that other instruments, such as the VoiSS,^{7,37} would have been a better reference test than the VHI considering that the former includes questions investigating pharyngeal symptoms, which are not addressed in the VHI. Lastly, the relevant differences between the I-VTD scale scores obtained in pre- and post-treatment conditions in a group of 30 patients affected by MTD suggests that the I-VTD scale may be useful also in monitoring the response to treatment in this group of patients. In particular, the magnitude of the effect of VT after at least 8 weeks was around 1.1 for both the frequency and the severity domains, which is considered to be large.³⁸ Similar results were found by Woznicka *et al*¹⁵ who demonstrated that the VTD scale was able to confirm the subjective improvement of the vocal function after treatment in a group of 55 patients with occupational dysphonia.

There are several limitations in the present study. First of all, the number of healthy participants enrolled for normative data generation was quite small and, consequently, the normative data here reported should be considered with caution. Moreover, for validity analysis, groups with unequal numbers of male and female participants were studied; larger studies with equal numbers of male and female subjects will be needed to further investigate the effect of gender on VTD perception. In addition, the responsiveness analysis was performed in a group of only 30 patients with MTD. Even if the results of the I-VTD scale demonstrated an improvement of their symptoms after undergoing VT, a larger group of patients is required to confirm this datum. Finally, for the concurrent validity analysis, only the correlations between the I-VTD and the I-VHI scores were analyzed. It is possible that the analysis of the correlation between the I-VTD scores and the results of other self-assessment questionnaires, such as the Italian versions of the Reflux Symptom Index^{39,40} and of the VoiSS^{7,37} should provide additional information regarding the concurrent validity of the I-VTD scale. Moreover, the comparison between VTD scores and Reflux Symptom Index scores could also provide a better understanding of possible differences in VTD perception between male and female subjects.

CONCLUSIONS

In conclusion, the I-VTD is characterized by high internal consistency, validity, and reproducibility, which make it a reliable and useful instrument in the evaluation of Italian patients with dysphonia. Moreover, considering its ability to provide supplementary and independent information other than the vocal alterations themselves, the introduction of the I-VTD scale in everyday clinical practice as part of the standard voice assessment protocols is highly recommended.

APPENDIX ORIGINAL AND ITALIAN VERSIONS OF THE VTD SCALE

1 English (VTD)

The following are symptoms or sensations that you may feel in your throat, which may occur as part of your voice problem. Please indicate the **frequency** with which they occur and the **severity** of the symptom/sensation, by circling a number in the appropriate column.

| | Frequency of Sensation/Symptom | | | | | | | Severity of Sensation/Symptom | | | | | | |
|-----------------------|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 0 = never; 2 = sometimes; 4 = often; 6 = always | | | | | | | 0 = none; 2 = mild; 4 = moderate; 6 = extreme | | | | | | |
| 1. Burning | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| 2. Tight | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| 3. Dry | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| 4. Aching | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| 5. Tickling | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| 6. Sore | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| 7. Irritable | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| 8. Lump in the throat | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 0 | 1 | 2 | 3 | 4 | 5 | 6 |

Modified from Mathieson et al.¹⁰

2 Italian (I-VTD)

I seguenti sono sintomi o sensazioni che lei può sperimentare a livello della gola e che possono presentarsi come parte del suo problema vocale. Per favore, indichi la **frequenza** con cui si presentano e la **gravità** di ogni sintomo/sensazione, cerchiando un numero nella colonna appropriata.

| | Frequenza Della Sensazione/Sintomo | | | | | | | Severità Della Sensazione/Sintomo | | | | | | |
|---------------------------------|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 0 = mai; 2 = talvolta; 4 = spesso; 6 = sempre | | | | | | | 0 = nulla; 2 = lieve; 4 = moderata; 6 = estrema | | | | | | |
| 1. Bruciore | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| 2. Restringimento | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| 3. Secchezza | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| 4. Dolore | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| 5. Solletico / prurito | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| 6. Infiammazione | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| 7. Irritabilità | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| 8. Sensazione di corpo estraneo | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 0 | 1 | 2 | 3 | 4 | 5 | 6 |

REFERENCES

- World Health Organization Quality of Life. WHOQOL Measuring Quality of Life. *The World Health Organization Quality of Life Instruments, WHO/MSA/MNH/PSF*. 1997;1–15.
- Lassig AA, Yueh B. Outcome research. In: Flint PW, Haughey BH, Lund VJ, et al., eds. *Cummings Otolaryngology Head and Neck Surgery*. Philadelphia: Elsevier Saunders; 2015:3–9.
- Hogikyan ND, Rosen CA. A review of outcome measurements for voice disorders. *Otolaryngol Head Neck Surg*. 2002;126:562–572.
- Benninger MS, Atiuja AS, Gardner G, et al. Assessing outcomes for dysphonic patients. *J Voice*. 1998;12:540–550.
- Phyland DJ, Oates J, Greenwood KM. Self-reported voice problems among three groups of professional singers. *J Voice*. 1999;13:602–611.
- 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.
- Deary IJ, Wilson JA, Carding PN, et al. VoiSS: a patient-derived voice symptom scale. *J Psychosom Res*. 2003;54:483–489.
- Hogikyan ND, Sethuraman G. Validation of an instrument to measure voice-related quality of life (V-RQOL). *J Voice*. 1999;13:557–569.
- Mathieson L. *Greene and Mathieson's The Voice and Its Disorders*. sixth ed. London and Philadelphia: Whurr Publishers Ltd; 2001.
- Mathieson L, Hirani SP, Epstein R, et al. Laryngeal manual therapy: a preliminary study to examine its treatment effects in the management of muscle tension dysphonia. *J Voice*. 2009;23:353–366.
- Roy N, Ford CN, Bless DM. Muscle tension dysphonia and spasmodic dysphonia: the role of manual laryngeal tension reduction in diagnosis and management. *Ann Otol Rhinol Laryngol*. 1996;105:851–856.
- Mathieson L. Vocal tract discomfort in hyperfunctional dysphonia. *J Voice*. 1993;2:40–48.
- Rodrigues G, Zambon F, Mathieson L, et al. Vocal tract discomfort in teachers: its relationship to self-reported voice disorders. *J Voice*. 2013;27:473–480.
- Amaral AC, Zambon F, Moreti F, et al. Vocal tract discomfort in teachers after teaching activity. *Codas*. 2017;29:e20160045. doi:10.1590/23171782/20172016045.
- Woznicka E, Niebudek-Bogusz E, Kwiecień J, et al. Applicability of the vocal tract discomfort (VTD) scale in evaluating the effects of voice therapy on occupational voice disorders. *Med Pr*. 2012;63:141–152.
- Luyten A, Bruneel L, Meerschman I, et al. Prevalence of vocal tract discomfort in the Flemish population without self-perceived voice disorders. *J Voice*. 2016;30:308–314.
- Lopes LW, Cabral FG, Figueiredo de Almeida AA. Vocal tract discomfort symptoms in patients with different voice disorders. *J Voice*. 2014;22:1–7.
- Torabi H, Seyyedeh MK, Nouredin NA, et al. The Vocal Tract Discomfort Scale: validity and reliability of the Persian version in the assessment of patients with muscle tension dysphonia. *J Voice*. 2016;30:711–716.
- Niebudek-Bogusz E, Woznicka E, Wiktorowicz J, et al. Applicability of the Polish Vocal Tract Discomfort (VTD) scale in the diagnostics of occupational dysphonia. *Logoped Phoniatr Vocol*. 2012;37:151–157.

20. Lukaschyk J, Brockmann-Bauser M, Beushausen U. Transcultural adaptation and validation of the German version of the Vocal Tract Discomfort scale. *J Voice*. 2017;31:261e1–261e8. doi:10.1016/j.jvoice.2016.05.006.
21. Mokkink LB, Terwee CB, Patrick DL, et al. International consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes: results of the COSMIN study. *J Clin Epidemiol*. 2010;63:737–745.
22. Beaton DE, Bombardier C, Guillemin F, et al. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*. 2000;25:3186–3191.
23. Boynton PM, Greenhalgh T. Selecting, designing, and developing your questionnaire. *BMJ*. 2004;328:1312–1315.
24. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *J Clin Epidemiol*. 1993;46:1417–1432.
25. Terwee CB, Bot SDM, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*. 2007;60:34–42.
26. Marx RG, Menezes A, Horovitz L, et al. A comparison of two time intervals for test-retest reliability of health status instruments. *J Clin Epidemiol*. 2003;56:730–735.
27. Hirano M. Psycho-acoustic evaluation of voice. In: Hirano M, ed. *Clinical Examination of Voice*. New York: Springer-Verlag; 1981.
28. De Bodt M, Wuyts F, Van de Heyning P, et al. Test B re-test study of GRBAS Scale. *J Voice*. 1997;11:74–80.
29. Schindler A, Ottaviani F, Mozzanica F, et al. Cross-cultural Adaptation and Validation of the Voice Handicap Index into Italian. *J Voice*. 2010;24:708–714.
30. Schindler A, Mozzanica F, Ginocchio D, et al. Vocal improvement after voice therapy in the treatment of benign vocal fold lesions. *Acta Otorhinolaryngol Ital*. 2012;32:304–308.
31. Schindler A, Mozzanica F, Ginocchio D, et al. Voice-related quality of life in patients after total and partial laryngectomy. *Auris Nasus Larynx*. 2012;39:77–83.
32. Vittinghoff E, Glidden DV, Shiboski SC, et al. *Regression Methods in Biostatistics: Linear, Logistic, Survival, and Repeated Measures Models*. New York: Springer; 2005.
33. Sullivan GM, Feinn R. Using effect size—or why the *P* value is not enough. *J Grad Med Educ* 2012;24:279–282.
34. International test commission. *ITC Guidelines for Translating and Adapting Tests*. second ed. 2017 Available at: https://www.intestcom.org/files/guideline_test_adaptation_2ed.pdf. Accessed July 10, 2017.
35. Barsky AJ, Peekna HM, Borus JF. Somatic symptom reporting in women and men. *J Gen Intern Med*. 2001;16:266–275.
36. Behlau M, Zambon F, Moreti F, et al. Voice self-assessment protocols: different trends among organic and behavioural dysphonia. *J Voice*. 2016;31:112e13–112e27. doi:10.1016/j.jvoice.2016.03.014.
37. Mozzanica F, Robotti C, Ginocchio D, et al. Cross-cultural adaptation and validation of the Italian Version of the Voice Symptom Scale (I-VoiSS). *J Voice*. 2017; <https://doi.org/10.1016/j.jvoice.2017.02.001>. In press.
38. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. second ed. Routledge; 1988.
39. Belafsky PC, Postma GN, Koufman JA. Laryngopharyngeal reflux symptoms improve before changes in physical findings. *Laryngoscope*. 2001;111:979–981.
40. Schindler A, Mozzanica F, Ginocchio D, et al. Reliability and clinical validity of the Italian Reflux Symptom Index. *J Voice*. 2010;24:354–358.