Evaluation of Psychometric Properties of Voice Activity and Participation Profile (VAPP): A Spanish Version

*Rosa M. Bermúdez-de-Alvear, †Pablo Gálvez-Ruiz, ‡A. Ginés Martínez-Arquero, §Sara Rando-Márquez, and §Elena Fernández-Contreras, *‡§*Málaga, and* †*Logroño, Spain*

Summary: Objectives: This study aimed to analyze the psychometric properties of the Spanish version of the Voice Activity and Participation Profile (SVAPP) questionnaire.
Study design: A randomized, cross-sectional sampling strategy with controls was used.
Methods: Two samples with a total of 169 participants were analyzed, specifically 61 men (mean age 37.02) and 108 women (mean age 37.78). Of these participants, 112 were patients and 57 were controls. The instrument was submitted to reliability (internal consistency and corrected item-total correlations) and reproducibility analyses. Validation assessment was based on the construct validity, convergent validity, discriminant validity, and concurrent validity.
Results: The global internal consistency was excellent (Cronbach's $\alpha = 0.976$), corrected item-total correlations were satisfactory and ranged 0.63−0.89, and factor loadings were above 0.50. The different subscales showed good internal consistency (alpha coefficients ranged 0.830−0.956) and test-retest values were consistently associated. The exploratory factor analysis evidenced a strongly defined five factors internal structure, with factors loadings ranging 0.51−0.86. Convergent validity demonstrated that all subscales and scores were very strongly correlated (Pearson r above 0.735) and significantly associated. The discriminant validity analysis showed that SVAPP had good specificity to distinguish dysphonic from healthy voice subjects. Concurrent validity with Voice Handicap Index Spanish version (SVHI) showed very strong correlations between total scores, and between SVHI total score and SVAPP Daily and Social Communication subscales; correlations between both tests subscales were strong; only between SVAPP Work and SVHI Physical sections correlations were moderate.
Conclusions: The findings of the present study demonstrated evidence for the SVAPP questionnaire reliability and validity, and provided insightful implications of voice disorders on Spanish patients' quality of life. However, further investigations are required.
Key Words: Voice−Quality of life−Activity limitation−Participation restriction−Reliability.

INTRODUCTION

Since the early 1960s, the traditional biomedical model of health and disease has evolved towards an integrative, nonreductionist clinical and theoretical biopsychosocial model of medicine.[1] In consequence, the health-related quality of life (HRQOL) concept has gained growing interest as a measure to evaluate the impact of illness on an individual's life. Constructs to systematically assess HRQOL have been clearly influenced by the World Health Organization (WHO) successive definitions of health and disease. In 1980, the WHO[2] issued the International Classification of Impairment, Disabilities and Handicaps (ICIDH) and differentiated three levels of impact of illness on HRQOL. Namely, *Impairment* was defined as the effect of disease on bodily functions; *Disability* referred to the repercussion of impairment on the individual's performance; and *Handicap* referred to the consequence of impairment and disability on social, environmental, and economic status. In the late 1990s, the International Classification of Functioning and Disability (ICIDH-2 Beta-1 draft)[3] remodeled this frame of reference and replaced *Disability* by the *Activities Limitation* construct (henceforth AL), which included all those constrains imposed on habitual activities because of a health disorder. The ICIDH-2 Beta-1 draft also substituted *Handicap* by *Participation Restriction* (PR), which came to describe the avoidance of activities and the loss or reduction in economic income and social status derived from AL and impairment. Additionally, ICIDH-2 Beta-1 draft introduced the evaluation of *Contextual Factors* as a measure of the influence of physical and psychosocial environmental conditions on subjects' capacity to deal with illness. On this basis, it was assumed that there is not a lineal progression from impairments to disabilities and handicaps, but rather an interaction among contextual factors and impairments which results in AL and PR. In 2001, the International Classification of Functioning, Disability and Health (ICF)[4] was issued as the WHO framework for measuring health and health-related domains. As it was developed on the basis of ICIDH-2 Beta-1 draft, it defined functioning as a multidimensional construct with four keystones: Body Functions, Body Structures, Activities and Participation, and Contextual Factors.[5]

In response to the evolution of the WHO successive definitions about health and the biopsychosocial factors that have an influence on it, numerous patient-based questionnaires have been developed to analyze voice disorder repercussions on each subject's quality of life.[6−8] In fact, voice HRQOL protocols have become critical tools not

Accepted for publication January 8, 2018.
From the *Departamento de Radiología y Medicina Física, Oftalmología y Otorrinolaringología, Facultad de Medicina, Universidad de Málaga, Málaga, Spain; †Departamento de Educación en Internet, Facultad de Educación, Universidad Internacional de la Rioja, Logroño, Spain; ‡Servicio de Otorrinolaringología, Hospital Regional Universitario de Málaga, Málaga, Spain; and the §Grado de Logopedia, Facultad de Psicología, Universidad de Málaga, Málaga, Spain.
Address correspondence and reprint requests to Rosa M. Bermúdez-de-Alvear, Departamento de Radiología y Medicina Física, Oftalmología y Otorrinolaringología, Facultad de Medicina, Boulevard Louis Pasteur 32, Campus de Teatinos, 29071 Málaga, Spain. E-mails: bermudez@uma.es, bermudezdealvear@gmail.com
Journal of Voice, Vol. 33, No. 4, pp. 582.e15−582.e23
0892-1997
© 2018 The Voice Foundation. Published by Elsevier B.V. All rights reserved.
https://doi.org/10.1016/j.jvoice.2018.01.005

only for research purposes but also in the clinical context. They have proved their suitability for providing clear diagnostic evidence for medical decisions, as well as for promoting patient's adherence to treatment and for testing the efficiency of new therapies.[7] Particularly, the Voice Handicap Index (VHI),[9] Voice-Related Quality of Life (V-RQOL),[10] Voice Outcome Survey (VOS),[11] Voice Symptom Scale (VoiSS),[12] and Voice Activity and Participation Profile (VAPP)[13] are among those voice HRQOL assessment tools that have shown robust psychometric properties.[6] Of them, the VHI[9] and its VHI-10 short form[14] are the most universally used, and they have been translated into several European and Asian languages.[15–19] Because most of these questionnaires were designed on the basis of the ICIDH conceptual model, they focus mainly on the voice HRQOL dimensions such as impairment (ie, physical impact), disability (ie, functional impact), and emotions.

The one exception is the VAPP which was developed[13] according to ICIDH-2 Beta-1 paradigm. Consequently, it does not assess the voice disorder physical impact, but rather the AL and PR derived from it on three specific dimensions, namely Work, Daily Communication, and Social Communication.[5] Validated in three languages,[20–23] VAPP has shown its versatility to test general dysphonic population and professional voice users.[22,24–29] It has also proved its applicability to adapt the ICF construct to the voice assessment,[5] as well as its responsiveness to preventive programs[30] and to treatment-induced changes.[22,23,26–28]

With regard to voice HRQOL appraisal in Spanish-speaking subjects, only VHI[9] and VHI-10[14] have been translated, validated, and used for voice evaluation.[16,31–34] On this basis, this study aims to assess VAPP cultural adaption and psychometric properties in a Spanish group of voice disordered and healthy voice subjects.

METHODS

Participants

The total sample included 169 participants with ages ranging from 18 to 77 years, of whom 61 were men (mean age 37.02; standard deviation [SD] 13.11) and 108 women (mean age 37.78; SD 13.49). The participants were recruited during the period from February 2014 to November 2015 from the appointment lists of Málaga Hospital Regional Universitario (Otorhinolaryngology Service) and from an ENT-Phoniatrics Clinic. The study was approved by the hospital Research Ethics Committee (project code 0375-M1-14). All participants received oral and written information on the objectives and methodology of the study, and they freely agreed to collaborate by signing a Clinical Informed Consent form. According to the 2008 Declaration of Helsinki, patients' confidentiality was guaranteed by using study codes to substitute subjects' names on the completed questionnaires, and a separate document that linked the study codes to subjects' identifying information was secured in a separate location. As a result, the study code was used to identify each subject's responses on the database, and only the primary investigators had access to the secured document.

The sample consisted of an experimental group and a control group. In the experimental group ($N_1 = 112$, age range: 18–77 years), there were 32 men (mean age 37.78; SD 14.37) and 80 women (mean age 37.91; SD 13.65). Patients from the experimental group were diagnosed by two of the authors who are specialized in phoniatrics and otorhinolaryngology, respectively. The following laryngeal impairments were found: hyperfunctional phonation with no evident vocal cord lesions (n = 30); nodules (n = 21); chronic laryngitis (n = 15); serous pseudocyst (n = 14); diffuse Reinke edema (n = 13); polyps (n = 11); recurrent nerve paralysis (n = 3); and congenital lesions (sulcus glottides, vergeture, and dermoid cyst, n = 5). Patients with laryngeal oncologic pathology were excluded from this study. All patients completed two types of surveys, that is, the VAPP[13] Spanish version (SVAPP) and the VHI[9] Spanish version (SVHI[16]). The control group ($N_2 = 57$, age range 19–66 years) was gathered from patients' accompanying persons and from students and teachers of the University of Malaga (UMA), all of them with no history of voice problems. The gender distribution in the control group was 29 men (mean age 36.17; SD 11.76) and 28 women (mean age 37.39; SD 13.27). No significant age or gender differences existed between experimental and control groups.

The test-retest reproducibility was performed by administrating SVAPP on two separate occasions: $SVAPP_0$ (test) and $SVAPP_1$ (retest). For this purpose, a subgroup of 40 randomly selected patients (24% of the experimental group) was studied, with ages ranging from 20 to 63 years. It was composed of 12 men (mean age 43.17; SD 12.73 years) and 28 women (mean age 38.46; SD 12.29 years), without significant differences regarding age or gender. Three weeks before their clinical appointment, they were contacted by phone to request their voluntary participation in this study. The following week, the authors carried out the first SVAPP application ($SVAPP_0$) in the patients' homes. A 15-day interval between test and retest was established, during which no therapeutic interventions were performed. Two weeks later, the second SVAPP administration ($SVAPP_1$) was performed at the hospital before patients entered their consultation. According to a previous study,[19] this 15-day lapse between test and retest is long enough for subjects to forget their first responses and sufficiently short to prevent any consistent change in their voice condition.[19]

Instrumentation

Two types of questionnaires were applied in this study, that is, the VAPP[13] Spanish translation (SVAPP) and the SVHI.[16] The VAPP[13] evaluates the quality of life and treatment outcomes of vocal disorders in dysphonic people. Its 28 items are divided into five QOL dimensions: Self-Perceived Severity of the voice problem (1 question, maximum score 10, item 1); Work (4 questions, maximum score 40, items 2–5); Daily Communication (12 questions, maximum score 120, items 6–17); Social Communication

(4 questions, maximum score 40, items 18−21); and Emotions (7 questions, maximum score 70, items 22−28). On the basis of a previous research,[35] the responses are rated on an equal-appearing interval scale that produces a score for each dimension. The first item (self-perceived severity of the voice disorder) is rated from 0 (normal) to 10 (severe). The rest of the items are measured from never (score 0) to always (score 10). Activity and Participation scores are obtained from Work, Daily Communication, and Social Communication dimensions (AL and PR partial scores). Also, AL and PR totals are provided as a result of adding the respective AL and PR partial scores. The maximum possible total score of this protocol is 280; AL maximum total is 100; PR maximum total is 100. The VHI Spanish version[16] consists of 30 statements and it analyzes the Functional, Physical, and Emotional impact of voice disorders. Each item is individually scored on a Likert scale of 5 points ranging from 0 (never) to 4 (always). As a result, responses produce a total outcome (maximum score 120) and three partial outcomes (each one with a maximum score of 40 points).

Procedure

With the aim of culturally adapting VAPP to the Spanish population, its items were first translated into Spanish by one of the authors of this study. To verify the translation from English to Spanish, the following procedure took place. First, the initial translation was reviewed by a Spanish language specialist who adapted the discourse for greater intelligibility. Then a bilingual professor performed a back translation from Spanish to English. This document was then sent to the VAPP authors who authorized its use. Voice Activity and Participation Profile Spanish title is identified as "Perfil de Voz, Actividad y Participación (VAPP)" (see Appendix S1 for a description of the survey translations).

Statistics

To estimate SVAPP reliability, the internal consistency was assessed using Cronbach's α coefficient and corrected item-total correlation.[36] According to Nunnally and Bernstein,[37] Cronbach's α values greater than 0.8 are considered "good," whereas a value greater than 0.9 is "excellent." The stability of the scale was established by measuring the test-retest reliability, using Cronbach's α and Spearman correlation analyses to compare the results of both $SVAPP_0$ and $SVAPP_1$. Correlations between 0.11 and 0.50 were considered moderate, between 0.51 and 0.75 were strong, between 0.76 and 0.90 were very strong, and higher than 0.91 excellent.[38]

With the aim of validating the findings, the following statistical analyses were used. First an exploratory factor analysis (EFA) using principal component analysis with orthogonal varimax rotation was conducted to assess the dimensionality or construct validity of the questionnaire. To attain the structure and the correct number of factors, the following criteria were used: eigenvalues greater than 1.0 and factor loadings higher than 0.50.[39] Before conducting the factor analysis, Kaiser-Meyer-Olkin (KMO) and Bartlett's test of sphericity were calculated.

As part of the convergent validity analysis, Pearson correlation coefficient was employed to analyze the linear correlation across SVAPP subscales and total scores. This analysis was applied to responses obtained from the total sample of participants (N = 169), including both control and experimental groups.

Next, nonparametric Mann-Whitney *U* test was used to assess the differences between the control and experimental groups, for all subscales and for the three total scores (discriminant validity).

After this and to estimate SVAPP concurrent validity, SVAPP was contrasted with VHI.[16] For this purpose, both questionnaires were applied to the dysphonic group (N = 112) the day of their first medical consultation, and results were compared by means of Spearman rho.

Finally, Statistical Package for Social Sciences (*SPSS*) program, version 20.0 (IBM Corp, Armonk, NY), was used for statistical analyses. The significance level was set at 5% ($P < 0.05$).

RESULTS

Reliability

The internal consistency of the profile was excellent (Cronbach's $\alpha = 0.976$). With the purpose of investigating the extent to which each individual item affected the profile's internal consistency, we examined the α values of the profile after a particular item was deleted. As shown in Table 1, individual α values ranked between $\alpha = 0.97$ and $\alpha = 0.98$, and internal consistency of the whole questionnaire did not increase if any particular item was removed. The analysis of subscales internal consistency showed values between good and excellent ($\alpha > 0.8$): Work $\alpha = 0.870$, Daily Communication $\alpha = 0.956$, Social Communication $\alpha = 0.838$, Emotions $\alpha = 0.946$.

The correlation between an individual item and the total score without that item (corrected item-total correlation) ranged from 0.63 to 0.89, with factor loadings greater than 0.50.

Test-retest consistency

The comparison of test-retest values (Table 2) attained a very good level of reproducibility. Results evidenced adequate consistency for $SVAPP_0$ ($\alpha = 0.952$) and for $SVAPP_1$ ($\alpha = 0.964$), being both correlated at an excellent level ($r = 0.941$). The test-retest reliability analysis rendered homogeneous and consistently associated responses in all subscales and scores, but the Social Communication section showed a tendency to be significantly lower ($P = 0.054$) in the retest compared with the test situation.

Construct validity

The EFA was performed by using orthogonal varimax rotation. The KMO Test for Sampling Adequacy was satisfactory (KMO value = 0.92) and the Bartlett test was significant (χ^2 (378) = 5804.88; $P < 0.01$). Factor loadings of all items were adequate and ranged between 0.51 and 0.86. Construct validity analysis demonstrated a structure

TABLE 1.
Factor Loadings, Corrected Item-Total Correlation, and Coefficient α for the SVAPP When a Particular Item Is Deleted ($N_1 = 169$)

	Factor Loading	Corrected Item-Total Correlation	Cronbach's α if Item Deleted
Item 1	0.64	0.834	0.975
Item 2	0.73	0.781	0.975
Item 3	0.79	0.669	0.976
Item 4	0.66	0.763	0.975
Item 5	0.66	0.690	0.976
Item 6	0.73	0.761	0.975
Item 7	0.58	0.838	0.975
Item 8	0.77	0.712	0.976
Item 9	0.65	0.750	0.975
Item 10	0.81	0.697	0.976
Item 11	0.63	0.696	0.976
Item 12	0.67	0.830	0.975
Item 13	0.65	0.805	0.975
Item 14	0.51	0.812	0.975
Item 15	0.76	0.789	0.975
Item 16	0.72	0.754	0.975
Item 17	0.77	0.785	0.975
Item 18	0.61	0.896	0.974
Item 19	0.51	0.730	0.975
Item 20	0.50	0.638	0.976
Item 21	0.59	0.636	0.976
Item 22	0.62	0.793	0.975
Item 23	0.64	0.774	0.975
Item 24	0.74	0.737	0.975
Item 25	0.71	0.847	0.975
Item 26	0.57	0.854	0.975
Item 27	0.64	0.797	0.975
Item 28	0.86	0.703	0.976

composed of five factors, all of which explained 80.23% of common variance across items. Factor 1 gathered components that indicated the global size of the voice disorder impact, that is, the Self-Perceived Severity (item 1), all Work AL items (items 2 and 4), one of the AL items in Daily Communication (item 12), and one of the Emotions items (item 25). Factor 2 was integrated by PR items in Daily Communication (items 7, 9, 13, 15, and 17), which expressed avoidance of conversations in different contexts such as on the phone or with background noise, and also by items related to PR and AL in Social Communication (items 18, 19, and 20). Factor 3 contained almost all the Emotions related items (items 22–24 and 26–28). Factor 4 included Daily Communication AL constituents (items 6, 8, 10, 14, and 16), which referred to limitations such as speaking to groups, on the phone, and in quiet environments; it also gathered one of the PR components in Daily Communication (item 11), which expressed avoidance of conversations in quiet environments. Factor 5 was composed of the PR components of Work (items 3 and 5) and Social Communication (item 21). Factor 5 thus identified intentions of leaving work placement and eluding communication with friends, family, and co-workers.

TABLE 2.
Reproducibility Outcomes for SVAPP

Sections	Mean	SD	*P*-Value
Self-Perceived Severity			0.816
Test	4.63	2.47	
Retest	4.58	2.43	
Work			0.405
Test	11.85	7.86	
Retest	11.45	8.35	
Daily Communication			0.153
Test	44.73	24.77	
Retest	41.70	28.09	
Social Communication			0.054
Test	8.28	7.92	
Retest	7.58	8.88	
Emotions			0.423
Test	21.18	16.32	
Re-test	20.30	15.37	
Profile Total Score			0.087
Test	90.65	52.15	
Retest	85.60	56.53	
AL Total score			0.100
Test	38.98	20.63	
Retest	35.33	21.57	
PR Total score			0.293
Test	25.88	19.20	
Retest	25.40	21.79	

Values obtained in $SVAPP_0$ (test) and $SVAPP_1$ (re-test) administrations indicated that responses in both applications were very similar. Only results from the Social Communication subscale showed a tendency to be significantly different, with $P = 0.54$. Wilcoxon signed-rank test ($P < 0.05$).

Convergent validity

Linear correlations (Pearson product-moment) across item 1, the four subscales, and the three profile total scores are displayed in Table 3. They were all significantly associated ($P < 0.01$). The Self-Perceived Severity component (item 1) showed a good linear correlation with all subscales: Work ($r = 0.758$), Daily Communication ($r = 0.798$), Social Communication ($r = 0.716$), and Emotions ($r = 0.772$). Linear correlations across the four subscales were good or very good, with Daily Communication and Social Communication being the most correlated sections ($r = 0.855$). Social Communication dimension showed good linear correlations with Emotions ($r = 0.796$) and Work ($r = 0.787$). The Work subscale showed very good correlations with Emotions ($r = 0.827$) and good correlations with sections such as Daily Communication ($r = 0.753$) and Social Communication ($r = 0.787$). The Emotions section showed excellent correlations with Daily Communication ($r = 0.80$) and good correlations with Social Communication ($r = 0.796$). Linear correlations across the three total scores (ie, Profile Total Score, AL Score, and PR Score) were excellent and above $r = 0.866$. Linear correlations between the three total scores and the five sections were also excellent ($r = 0.816$); only the correlation between PR Total Score and item 1 was a bit lower but still good ($r = 0.735$).

TABLE 3.
Pearson's Correlations Between Item 1, Subscales, and Total Scores

Pearson's Correlations	Self-Perceived Severity	Work	Daily Communication	Social Communication	Emotions	Total Score	AL Total Score
Work	0.758*						
Daily Communication	0.798*	0.753*					
Social Communication	0.716*	0.787*	0.855*				
Emotions	0.772*	0.827*	0.800*	0.796*			
Profile Total Score	0.847*	0.879*	0.955*	0.908*	0.926*		
AL Total score	0.849*	0.825*	0.966*	0.837*	0.816*	0.956*	
PR Total score	0.735*	0.838*	0.922*	0.941*	0.834*	0.951*	0.866*

* Statistically significant at $P < 0.001$.

Discriminant validity

Results of Mann-Whitney *U* test comparison between the experimental ($N_1 = 112$) and the control ($N_2 = 57$) groups are shown in Table 4. For all subscales and scores, significant differences were found between dysphonic and healthy subjects' responses ($P < 0.0001$).

Concurrent validity

Results from estimating SVAPP concurrent validity with SVHI[16] are shown in Table 5. All components were positively correlated and significantly associated ($P < 0.0001$). Two of the SVAPP subscales (Daily Communication and Social Communication) and the three total scores (Total, AL Total, and PR Total) correlated very strongly with SVHI[16] Total score ($\rho > 0.79$); and three of SVAPP subscales (Self-Perception, Work, and Emotions) were strongly correlated to SVHI[16] Total score ($\rho > 0.63$). Regarding correlations between both questionnaires individual sections, all subscales correlated strongly ($\rho > 0.55$); it was only the SVAPP Work section that correlated moderately to the SVHI[16] Physical section ($\rho = 0.47$).

DISCUSSION

The VAPP[13] stands out among the voice HRQOL questionnaires with robust psychometric properties[6–8] because it is the only one designed according to the ICDH-2 beta-1 draft.[3] It thus focuses on the limitations of communication, the avoidance of situations, and the loss of social roles that dysphonic subjects' experience in different contexts because of their voice problem. Given that VAPP measures the AL and PR in job, daily communication, and social activities, it allows adapting the ICF[40] framework to the assessment of the impact of dysphonia on subjects' QOL.[5] Additionally, VAPP[13] analyzes the self-perception of the voice disorder severity and its emotional impact. The process followed in this research to culturally adapt and validate SVAPP included two phases: in a first step, VAPP[13] was translated into Spanish; this translation was linguistically revised by a Spanish philologist and back translated into English by an English philologist; this English version was finally forwarded to VAPP[13] authors for their approval. In a second step, a cross-sectional study was performed to explore SVAPP reliability and validity in a sample of 112 dysphonic subjects and 57 controls.

The reliability or stability of SVAPP was assessed by performing Cronbach's alpha to the whole profile and individual subscales, and by measuring its test-retest reproducibility in a subset of 40 patients. Internal consistency reflects the homogeneity of the items in the profile; therefore, testing with homogeneous items is considered to be more reliable than one with heterogeneous items.[41] It is generally considered that α values of at least 0.8 are necessary for adequate

TABLE 4.
Results of Mann-Whitney *U* Test Comparison Between the Control Group and the Experimental Dysphonic Group

	Dysphonic Group			Control Group			
	Min.	Max.	Mean (SD)	Min.	Max.	Mean (SD)	Mann-Whitney U
Subscales							
Self-Perceived Severity	0	10	5.18 (2.57)	0	3	0.44 (.78)	218.50*
Work	0	40	14.89 (10.31)	0	6	0.67 (1.29)	117.00*
Daily Communication	0	111	45.16 (28.28)	0	7	1.21 (1.89)	42.50*
Social Communication	0	40	9.83 (8.96)	0	1	0.05 (.22)	132.00*
Emotions	1	70	26.54 (19.47)	0	4	0.37 (.88)	72.50*
Total scores							
Activity Limitation	7	97	40.21 (22.36)	0	7	1.63 (2.42)	1.50*
Participation Restriction	0	94	29.67 (23.17)	0	4	0.30 (.73)	84.00*
Profile Total	17	255	101.60 (60.88)	0	13	2.74 (3.90)	0.00*

* Statistically significant at $P < 0.001$.
Abbreviations: Min., minimum; Max., maximum.

TABLE 5.
Spearman Coefficients (rho) Between SVAPP and SVHI[16] Scores (N = 112)

	SVHI[16] Sections			
SVAPP Sections	Functional	Physical	Emotional	SVHI[16] Total
Self-perceived severity	0.61*	0.55*	0.65*	0.68*
Emotions	0.58*	0.56*	0.74*	0.70*
Work	0.55*	0.47*	0.66*	0.63*
Daily Communication	0.75*	0.67*	0.60*	0.79*
Social Communication	0.74*	0.64*	0.69*	0.80*
AL Total	0.76*	0.66*	0.67*	0.81*
PR Total	0.73*	0.64*	0.66*	0.79*
SVAPP Total	0.75*	0.68*	0.76*	0.84*

* Statistically significant at $P < 0.0001$.

diagnostic internal consistency, whereas for exploring purposes, $\alpha = 0.7$ can be considered sufficient.[41] In general, α rates between 0.70 and 0.95 are judged acceptable.[37,42] Internal consistency of SVAPP was supported by a very high Cronbach's coefficient ($\alpha = 0.97$) for the whole profile, as in previous research[3,21,23]; and by strong α coefficients (above 0.80) for individual subscales, as corroborated by other authors.[22] Corrected item-total scores correlations were above the cut-off point recommended by other studies,[43] and factor loadings were adequate because they were greater than 0.50.[39]

Regarding SVAPP reproducibility,[13,19,21–23] the test-retest reliability analysis rendered homogeneous and consistently associated responses in all subscales and scores. Mean values obtained in the different sections were slightly lower in the second administration ($SVAPP_1$), similar to the results obtained by previous studies,[21,22] still no significant differences were found. However, the Social Communication mean values were close to being significantly lower in $SVAPP_1$ than in $SVAPP_0$ ($P = 0.054$). As some authors suggested,[21] this slight improvement in the retest may be associated with the subjective relief that patients may experience after receiving external attention to their voice problem, which would support the use of the VAPP questionnaire in the evaluation of voice disorders.

To assess SVAPP construct validity, an EFA was performed with orthogonal varimax rotation because correlations between factors were high.[44] Factor loadings of all items exceeded the minimum 0.50 acceptable value,[39] and in half of the items (ie,14 items), factor loads were higher than 0.70, which is indicative of a strongly defined construct.[43] A five-factor internal structure was evidenced by EFA: factor 1 included issues that resembled the global size of the voice disorder impact; factor 2 was mainly composed by items related to the PR in Daily Communication; factor 3 reflected the emotional impact; factor 4 gathered items largely concerned with AL in Daily Communication; and factor 5 identified PR in Work and Social Communication. This internal structure revealed that SVAPP clearly discriminated between Activity and Participation in Work, Daily Communication, and Social Communication, all of which favors its applicability for clinical purposes, and corroborates other authors' findings regarding the original VAPP.[5] As it is advocated by the ICF[4] and other authors have shown,[5,21,22,24,25,28,30] the differentiated information about AL and PR in voice disorders assessment provides keystones for clinical decision-making, especially regarding therapeutic and preventive measures. When comparing the SVAPP internal structure with that found in the Finnish VAPP,[21] many similarities were found, but some differences appeared regarding the noise-related items. In the Finnish VAPP validation research,[21] items 12 and 13 (which respectively indicate AL and PR due to environmental noise) formed their own separate factor. However, in the SVAPP, item 12 belonged to the global dimension of the voice problem (factor 1), whereas item 13 belonged to the PR in Daily Communication group (factor 2). These internal structure dissimilarities between the Spanish and Finnish versions of VAPP could be attributed to the influence of cross-cultural differences. Other authors[25] found that environmental and cultural influences play a significant role in the perception of the voice disorders impact. In this regard, studies about Spanish[45] and Finnish[46] environmental noise levels have indicated striking differences. For instance, a recent report about outdoor noise levels[47] in Spanish cities has shown that 29.5% of Malaga population is exposed to average noise levels superior to 65 dB(A) during day, evening, and night, well above the WHO recommended limits. However, a document about the City of Tampere environmental policy[48] has stated that there are no residents exposed to daytime average noise above 65 dB(A) or night-time average noise above 60 dB(A). In sum, given the inherent differences in the context, Spanish and Finnish voice patients can logically differ regarding their perception of the environmental noise effects on their QOL. While Spanish voice patients consider the AL generated by environmental noise as part of the global severity of their voice problem, and the PR produced by environmental noise as part of the Daily Communication factor, Finnish voice patients consider the effects of noise as a unique and separate group of agents.

Convergent validity, along with discriminant validity, is a subtype of construct validity. Estimation of SVAPP

convergent validity demonstrated that all subscales were significantly associated and all Pearson correlation coefficients were strong, as in other authors' studies.[13,22] The discriminant validity analysis showed that SVAPP had good specificity to distinguish dysphonic and healthy voice subjects, substantiating former research about VAPP validation in different languages.[13,20–23] In most studies that used VAPP for assessing dysphonia impact,[13,21–26,28,35] high SDs were observed in total scores and subscale mean values, which can be attributed to the sample heterogeneity. In the present cross-sectional study, subjects were randomly selected; therefore, a heterogeneous sample was obtained regarding the size of the voice disorder impact on QOL. It would thus seem interesting for future studies to homogenize subjects under study. For this aim, it could be assumed that the use of item 1 as a classification tool to divide the sample into subgroups of similar self-perceived severity could probably contribute to cluster data about the mean.

According to psychometric normative,[49] the concurrent validity is a subtype of the criterion validity. Concurrent validity is performed to compare the measure in question with an outcome that is obtained simultaneously but with a different instrument. Results from estimating SVAPP concurrent validity with SVHI[16] showed very strong correlations across total scores, and between SVHI[16] total score and SVAPP Daily and Social Communication subscales. Furthermore, strong correlations between the other subscales were obtained. Only in the case of the SVAPP Work section and the SVHI[16] Physical section correlations were moderate; this could be explained because these two sections measure different constructs, that is, while the SVAPP Work dimension estimates AL and PR regarding work, the VHI Physical section explores voice impairments and bodily complaints. Inasmuch as SVAPP and SVHI[16] provide different frameworks to assess voice HRQOL, these concurrent validity outcomes reveal that they can be applied either complementarily or in an isolated fashion, as others have suggested.[26]

In reference to SVAPP cultural and linguistic adaptation, item 1 was not always correctly interpreted by some participants. A few patients believed it was enquiring about the medical interpretation of the voice disorder severity. It may be that the Spanish translation was not sufficiently precise about the "self-perception" concept, and therefore it induced some people to think item 1 referred to the professional's judgment, not to their own subjective perception of severity. In the Finnish version,[21] item 1 was also referred to be unsuccessfully translated. According to Zraick et al's[50] recommendations about readability of HRQOL assessment tools, it seems likely that more attention should be paid to improve the wording of item 1 in future research.

CONCLUSIONS

This cross-sectional study showed that SVAPP is a reliable and valid instrument. Regarding reliability, SVAPP demonstrated very high internal consistency and adequate test-retest values. Validity analyses included four types. First, construct validity analysis evidenced a well-defined internal structure of five factors, clearly discriminating between Activity and Participation in Work, Daily Communication, and Social Communication, all of which proves SVAPP applicability to adapt the ICF framework to the voice disorders assessment. Next, convergent validity analyses evidenced that SVAPP subscales and scores were linearly correlated, whereas discriminant validity proved that SVAPP distinguished dysphonic from healthy voice subjects. And finally, the SVAPP concurrent validity estimate with SVHI[16] showed that both questionnaires were correlated and provided consistently associated information, despite the fact that each one is designed on a different WHO-HRQOL construct.

These outcomes also suggested that expressing item 1 more explicitly would improve SVAPP readability. In future longitudinal studies, further research is warranted to assess SVAPP responsiveness to treatment-induced changes in a Spanish sample.

Acknowledgments

The authors wish to thank Ignacio Moreno Torres, PhD, who kindly helped to improve the Spanish language adaptation of SVAPP.

SUPPLEMENTARY DATA

Supplementary data related to this article can be found online at doi:10.1016/j.jvoice.2018.01.005.

REFERENCES

1. Benning TB. Limitations of the biopsychosocial model in psychiatry. *Adv Med Educ Pract*. 2015;6:347–352. https://doi.org/10.2147/AMEP.S82937.
2. World Health Organization. International classification of impairments, disabilities, and handicaps (ICIDH). 1980.
3. World Health Organization. International classification of functioning and disability. ICIDH-2, Beta-1 draft. 1997.
4. Organización Mundial de la Salud. Clasificación internacional del funcionamiento, de la discapacidad y de la salud. Ministerio del trabajo y asuntos sociales. Secretaría general de asuntos sociales. Instituto de migraciones y servicios sociales (IMSERSO); 2001.
5. Ma EP, Yiu EM, Verdolini-Abbott K. Application of the ICF in voice disorders. *Semin Speech Lang*. 2007;28:343–350. https://doi.org/10.1055/s-2007-986531.
6. Franic DM, Bramlett RE, Bothe AC. Psychometric evaluation of disease specific quality of life instruments in voice disorders. *J Voice*. 2005;19:300–315. https://doi.org/10.1016/j.jvoice.2004.03.003.
7. Branski RC, Cukier-Blaj S, Pusic A, et al. Measuring quality of life in dysphonic patients: a systematic review of content development in patient-reported outcomes measures. *J Voice*. 2010;24:193–198. https://doi.org/10.1016/j.jvoice.2008.05.006.
8. Zraick RI, Risner BY. Assessment of quality of life in persons with voice disorders. *Curr Opin Otolaryngol Head Neck Surg*. 2008;16:188–193. https://doi.org/10.1097/MOO.0b013e3282febd10.
9. Jacobson BH, Johnson A, Grywalski C, et al. The voice handicap index: development and validation. *Am J Speech Lang Pathol*. 1997;6:66–70.
10. Hogikyan ND, Sethuraman G. Validation of an instrument to measure voice-related quality of life (V-RQOL). *J Voice*. 1999;13:557–569.

11. Giklich RE, Glovsky RM, Montgomery WW. Validation of a voice outcome survey for unilateral vocal cord paralysis. *Otolaryngol Head Neck Surg*. 1999;120:153–158. https://doi.org/10.1016/S0194-5998(99)70399-2.
12. Dreary IJ, Wilson JA, Carding PN, et al. VoiSS: a patient-derived Voice Symptom Scale. *J Psychosom Res*. 2003;54:483–489.
13. Ma EP, Yiu EM. Voice activity and participation profile: assessing the impact of voice disorders on daily activities. *J Speech Lang Hear Res*. 2001;44:511–524.
14. Rosen CA, Lee AS, Osborne J, et al. Development and validation of the voice handicap index-10. *Laryngoscope*. 2004;114:1549–1556. https://doi.org/10.1097/00005537-200409000-00009.
15. Lam PKY, Chan KM, Ho WK, et al. 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.
16. Nuñez-Batalla F, Corte-Santos P, Senaris-Gonzalez B, et al. Adaptación y validación del índice de incapacidad vocal (VHI-30) y su versión abreviada (VHI-10) al español [Adaptation and validation to the Spanish of the Voice Handicap Index (VHI-30) and its shortened version (VHI-10)]. *Acta Otorrinolaringol Esp*. 2007;58:386–392.
17. Verdonck-de Leeuw IM, Kuik DJ, De Bodt M, et al. Validation of the voice handicap index by assessing equivalence of European translations. *Folia Phoniatr Logop*. 2008;60:173–178. https://doi.org/10.1159/000127836.
18. 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. https://doi.org/10.1016/j.jvoice.2009.05.006.
19. Helidoni ME, Murry T, Moschandreas J, et al. Cross-cultural adaptation and validation of the voice handicap index into Greek. *J Voice*. 2010;24:221–227. https://doi.org/10.1016/j.jvoice.2008.06.005.
20. Sukanen O, Sihvo M, Rorarius E, et al. Voice Activity and Participation Profile (VAPP) in assessing the effects of voice disorders on patients' quality of life: validity and reliability of the Finnish version of VAPP. *Logoped Phoniatr Vocol*. 2007;32:3–8.
21. Kleemola L, Helminen M, Rorarius E, et al. Voice activity and participation profile in assessing the effects of voice disorders on quality of life: estimation of the validity, reliability and responsiveness of the Finnish version. *Folia Phoniatr Logop*. 2011;63:113–121. https://doi.org/10.1159/000316311.
22. Ricarte A, Oliveira G, Behlau M. Validation of the voice activity and participation profile protocol in Brazil. *Codas*. 2013;25:242–249.
23. Fava G, Paolillo NP, Oliveira G, et al. Cross-cultural adaptation, validation, and cutoff point of the Italian version of the voice activity and participation profile: Profilo di Attività e Partecipazione Vocale. *J Voice*. 2015;29:130,e11–e19. https://doi.org/10.1016/j.jvoice.2014.04.008.
24. Yiu EM-L, Ma EP-M. Voice activity limitation and participation restriction in the teaching profession: the need for preventive voice care. *J Med Speech Lang Pathol*. 2002;10:51–60.
25. Yiu EM-L, Ho EM, Ma EP-M, 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.
26. Bassi IB, Assunção AÁ, de Medeiros AM, et al. Quality of life, self-perceived dysphonia, and diagnosed dysphonia through clinical tests in teachers. *J Voice*. 2011;25:192–201. https://doi.org/10.1016/j.jvoice.2009.10.013.
27. Kleemola L, Helminen M, Rorarius E, et al. Twelve-month clinical follow-up study of voice patients' recovery using the Voice Activity and Participation Profile (VAPP). *J Voice*. 2011;25:e245–e254. https://doi.org/10.1016/j.jvoice.2010.05.003.
28. Duarte de Almeida L, Santos LR, Bassi IB, et al. Relationship between adherence to speech therapy in patients with dysphonia and quality of life. *J Voice*. 2013;27:617–621. https://doi.org/10.1016/j.jvoice.2013.02.003.
29. Cantor Cutiva LC, Burdorf A. Factors associated with voice-related quality of life among teachers with voice complaints. *J Commun Disord*. 2014;52:134–142. https://doi.org/10.1016/j.jcomdis.2014.05.003.
30. Dragone MLS. Voice activity and participation profile presenting coordinates for readjustment of preventive action of educators. *Folia Phoniatr Logop*. 2011;63:49–54. https://doi.org/10.1159/000316324.
31. García-López I, Núñez-Batalla F, Gavilán Bouzas J, et al. Validación de la versión en español del índice de incapacidad vocal (S-VHI) para el canto [Validation of the Spanish Version of the Voice Handicap Index for Vocal Singing (S-VHI)]. *Acta Otorrinolaringol Esp*. 2010;61:247–254.
32. Sanz L, Bau P, Arribas I, et al. Adaptation and validation of Spanish version of the pediatric Voice Handicap Index (P-VHI). *Int J Pediatr Otorhinolaryngol*. 2015;79:1439–1443. https://doi.org/10.1016/j.ijporl.2015.06.021.
33. Sanz L, Bau P, Arribas I, et al. Development and validation of a short version of the Spanish pediatric voice handicap index (P-VHI-10). *Int J Pediatr Otorhinolaryngol*. 2016;88:113–116. https://doi.org/10.1016/j.ijporl.2016.06.053.
34. Señaris-González B, Núñez Batalla F, Corte-Santos P, et al. Índice de Incapacidad Vocal: factores predictivos [Factors predicting voice handicap index]. *Acta Otorrinolaringol Esp*. 2006;57:101–108.
35. Ma EP, Yiu EM. Scaling voice activity limitation and participation restriction in dysphonic individuals. *Folia Phoniatr Logop*. 2007;59:74–82. https://doi.org/10.1159/000098340.
36. Groth-Marnat G. *Handbook of Psychological Assessment*. Hoboken, NJ: John Wiley & Sons; 2009.
37. Nunnally JC, Bernstein AH. *Psychometric Theory*. third ed. New York, NY: McGraw Hill Series in Psychology; 1994.
38. Mondragón Barrera M. Uso de la correlación de Spearman en un estudio de intervención en fisioterapia. *Mov Científico*. 2014;8:98–104.
39. Kahn J. Factor analysis in counseling psychology. Research, training, and practice: Principles, advances and applications. *Couns Psychol*. 2006;34:684–718.
40. World Health Organization. International Classification of Functioning, Disability and Health; 2001. Available at: http://www.icfillustration.com/top_e.html. Accessed August 15, 2017.
41. Bryman A, Cramer D. *Concepts and Their Measurement: Quantitative Data Analysis, with SPSS for Windows*. London: Routledge; 1997.
42. Fernández-Domínguez JC, Sesé-Abad A, Morales-Asencio JM, et al. Validity and reliability of instruments aimed at measuring Evidence-Based Practice in Physical Therapy: a systematic review of the literature. *J Eval Clin Pract*. 2014;20:767–778. https://doi.org/10.1111/jep.12180.
43. Hair J, Black W, Babin B, et al. *Multivariate Data Analysis*. 6th ed. New Jersey: Pearson Educational, Inc.; 2006.
44. Tabachnick B, Fidell L. *Using Multivariate Statistics*. New York: Harper & Row; 2001.
45. *Europapress.es*. Mapa del ruido: este es el nivel de contaminación acústica que soportan los españoles; 2017. Published January 12; Available at: http://www.europapress.es/sociedad/noticia-mapa-ruido-nivel-contaminacion-acustica-soportan-espanoles-20170112101349.html. Accessed July 1, 2017.
46. Halonen JI, Lanki T, Yli-Tuomi T, et al. Associations of traffic noise with self-rated health and psychotropic medication use. *Scand J Work Environ Health*. 2014;40:235–243. https://doi.org/10.5271/sjweh.3408.
47. Ministerio de Agricultura, Alimentación y Medioambiente. Sistema de Información sobre Contaminación Acústica SICA; 2017. Updated; Available at: http://sicaweb.cedex.es/docs/mapas/fase2/aglomeracion/MALAGA/Ag_AND_Malaga_C_Lden.pdf. Accessed November 3, 2017.
48. City of Tampere Environmental Policy 2020. Proactivity and responsible operating methods. Available at: https://www.tampere.fi/liitteet/e/6G5lrE7uX/Environmental_Policy_2020_English.pdf. Accessed November 1, 2017.
49. American Educational Research Association, American Psychological Association, National Council on Measurement Education. *Standards for Educational and Psychological Testing*. Washington, DC: American Psychological Association; 1985.
50. Zraick RI, Atcherson SR, Ham BK. Readability of patient-reported outcome questionnaires for use with persons with swallowing disorders. *Dysphagia*. 2012;27:346–352. https://doi.org/10.1007/s00455-011-9373-x.