



Translation, Validity, and Reliability of Singing Voice Handicap Index in Persian Traditional Singers: A New Revised Version

Nassim Ahmadi ¹, Fatemeh Rajati ², Behnoosh Vasaghi-Ghara Malaki ³, Abbas Ebadi ⁴, Esmail Ebrahimi Takamjani ⁵, Katherine Verdolini Abbott⁶ and Farhad Torabinezhad ^{1,*}

¹Department of Speech and Language Pathology, School of Rehabilitation Sciences, Iran University of Medical Sciences, Tehran, Iran

²Research Center Environmental Determinants of Health, Health Institute, Kermanshah University of Medical Sciences, Kermanshah, Iran

³Department of Basic Sciences, School of Rehabilitation Sciences, Iran University of Medical Sciences, Tehran, Iran

⁴Behavioral Sciences Research Center, Life Style Institute, Baqiyatallah University of Medical Sciences, Tehran, Iran

⁵Department of Physiotherapy, School of Rehabilitation Sciences, Iran University of Medical Sciences, Tehran, Iran

⁶College of Health Sciences, University of Delaware, Newark, Delaware, USA

*Corresponding author: Department of Speech and Language Pathology, School of Rehabilitation Sciences, Iran University of Medical Sciences, Tehran, Iran. Email: torabinezhad.f@iums.ac.ir

Received 2021 November 14; Revised 2022 January 08; Accepted 2022 January 09.

Abstract

Background: The singing voice handicap index (SVHI) was developed to assess voice-related handicaps in singers.

Objectives: The purpose of this study was to evaluate the validity and reliability of a Persian translation of the SVHI (P-SVHI) for traditional Persian singers.

Methods: A cross-sectional study was performed to translate and validate the P-SVHI using a "forward-backward" procedure. The questionnaire was administered to 275 traditional Persian singers (168 with a normal voice and 107 with dysphonia) ranging from 16 to 83 years. Content, construct, discriminant, convergent, and face validity, test-retest reliability, internal consistency, and floor and ceiling effects were assessed.

Results: The 36-item SVHI, translated from the original English by two translators, was completed by 21 traditional Persian singers. After the face validity was approved, construct validity assessment using factor analysis identified four factors in the P-SVHI with the exclusion of seven items. The mean score of the 29-item P-SVHI on a 0-100 scale was 16.9 for healthy subjects and 40.9 for dysphonic subjects. Discriminant validity confirmed that P-SVHI scores were significantly different between normal and dysphonic singers, using an independent sample *t*-test ($P < 0.001$). Convergent validity was proven by strong correlations between P-SVHI scores and the Iowa Patient's Voice Index (IPV1, IPV2, and "IPV1 effort" scores of $r = 0.58, 0.69,$ and $0.42,$ respectively, $P < 0.001$). Internal consistency and test-retest reliability were also excellent (Cronbach $\alpha = 0.95$, average inter-item correlation = 0.40, and intra-class correlation coefficients using two-way random model with absolute agreement, $ICC_{\text{two-way random, absolute agreement}} = 0.69 - 0.95$). No floor or ceiling effects were observed.

Conclusions: The revised 29-item P-SVHI is a valid and reliable tool in differentiating normal and dysphonic Persian traditional singers.

Keywords: Disability Evaluation, Iran, Psychometrics, Singing, Voice Disorder

1. Background

It is acknowledged that voice disorders can cause poor quality of life for professional voice users such as singers. Singers constitute 11.5% of clients of some voice clinics while representing only 0.02% of the general population (1). As a working population that relies on voice in their profession, singers have higher vocal demands and are more at risk of developing vocal disorders than the general population (2).

According to the World Health Organization (WHO), the handicap resulting from any disability may exert oc-

cupational, social, or economic effects (3). Accordingly, a valid and reliable vocal disorder scale for singers needs to evaluate all aspects of handicap and be sensitive to all levels of function. A valid and reliable vocal disorder scale for singers could guide voice service providers to capture different aspects of singers' voice problems, in particular, to address specific problems in the treatment process, and evaluate the effect of management on a voice index over the therapeutic process.

Persian traditional singing is the original, "popular" singing style in Iran (4), where "popular" is under-

stood as “of the people,” perhaps best translated as “folk.” “Tahrir/Tahreer” is musical ornamentation similar to vibrato, produced in this singing style (4). Production of diverse Tahrir/Tahreer styles is a delicate and crucial vocal skill that may be affected by vocal impairment (5). Moreover, since Tahrir/Tahreer production may create uncomfortable laryngeal muscle tension, traditional singers may be more exposed to vocal impairment than many others (6). Along similar lines, traditional Persian singing codifies songs according to melodic figures called “Radif.” These figures are transmitted orally across generations. Iran’s traditional music is based on Radif. The acquisition of these figures typically requires years of practice (see, for example, [here](#), making the learning process different from the process involved in most other singing styles (6). This learning process may place traditional Persian singers at greater risk of voice disorders than others (6).

The singing voice handicap index (SVHI) is a commonly used measure for clinical purposes, developed and validated by Cohen et al. (7). The Persian translation of the SVHI (P-SVHI) assesses symptoms commonly reported to laryngologists and speech-language pathologists by singers.

The SVHI has been translated into various languages (3, 8-16). The validation of the P-SVHI was previously investigated by Ghaemi et al. (15) in all singing styles. However, their study did not provide evidence regarding the validity and reliability of the SVHI for traditional Persian singers specifically, who comprise a vast majority of singers in Iran. Moreover, the study did not assess the construct validity to explore relevant dimensions of the P-SVHI.

2. Objectives

Although Ghaemi et al. study represented an initial step in assessing singing disruptions in Persian singers (15), considering traditional singers’ particular vocal demands, the current study aimed to investigate the psychometric properties of SVHI in traditional Persian singers.

3. Methods

3.1. Study Design and Participants

A methodological study was conducted from September 2019 to February 2020 after obtaining approval from the Research Ethics Committee. A convenience sample of 285 traditional Persian singers (222 males and 63 females) from 10 music and singing centers in Tehran was recruited over six months.

Inclusion criteria for all participants were as follows: having at least four years of singing performance experience, the age range of 16 to 85 years, and no history of hearing loss or smoking. Participants with the following criteria were excluded from the study: having any forms of rhinitis or common cold at the time of participation.

3.2. Procedure

3.2.1. Translation

After obtaining permission from the scale developer (Cohen), cross-cultural adaptation using a “forward-backward” procedure was applied to translate the original version of SVHI into Persian according to Beaton et al. guidelines (17). Two translators, who had an M.Sc. degree in speech-language pathology, synthesized two initial translations and resolved any inconsistencies. The synthesized Persian version of the SVHI was then back-translated into English by two bilingual translators and a native English speaker unfamiliar with the original English version. The two back translations were assessed by our expert committee, including three speech-language pathologists who compared them to the original version, approved by the developer. The final P-SVHI was developed after approval. As the P-SVHI was available, it was unnecessary to evaluate the face validity. However, the understanding of the questionnaire was done during the translation process in the current study.

As the face validity of P-SVHI was previously approved using the impact score by multiplying frequency and importance (score 4 or 5) ratings for each item in the previous P-SVHI, first, we compared our version with the previous Persian version. All items were approved by the previous version after some changes in wording. Face validity evaluation was considered a step of cross-cultural adaptation in ensuring the quality of the P-SVHI. Therefore, in this section, we applied our P-SVHI to assess the items’ clarity and unambiguity using a qualitative process by 21 traditional singers who were asked to determine ambiguous items and suggest other sentences in a pilot study if needed. This section resulted in no change in any of the items.

3.2.2. Data Collection

Subjects completed a demographic questionnaire and P-SVHI. To evaluate the P-SVHI test-retest reliability, 35 healthy and 10 dysphonic singers completed it again in a seven-day interval. Two clinicians were randomly interviewed with 96 subjects with 0.87 Kappa agreement, and they rated voice quality separately using the grade, roughness, breathiness, asthenia, and strain (GRBAS) Scale on three tasks of vowel prolongation, counting to 10, and reading the rainbow text. The GRBAS Scale was previously

used with Persian speakers by Hosseinifar et al. (18). Subjects were also asked about their vocal health status within the past month on a binary scale (with voice problem, without voice problem).

3.3. Measures

3.3.1. Singing Voice Handicap Index

The SVHI comprises 36 questions rated on a five-point scale from 0 (never) to 4 (always), with Cronbach alpha of 0.97 of the first article. The total score ranges between 0 and 144, with higher scores indicating more voice disability (3, 7). Sobol et al. reported the cutoff of the original SVHI as 20.35 (95% CI: 10.6 - 30.1) (19). In a previous version of the P-SVHI, the internal consistency was confirmed for Iranian singers generally (Cronbach $\alpha = 0.83$) (15).

3.3.2. Iowa Patient's Voice Index

The Iowa Patient's voice index (IPVI) was developed by Verdolini et al. cited in Karnell et al. with three times (20). The first two items, "IPVI1" and "IPVI2," are rated on a 0 - 6 scale (0 for "no impact" and 6 for "great impact"). The third item, "IPVI effort," rates relative vocal effort. On this scale, 100 indicates comfortable effort, 200 indicates twice as much effort as comfortable, and so on (20).

3.3.3. Grade, Roughness, Breathiness, Asthenia, and Strain Measurers

The GRBAS is an auditory-perceptual voice assessment scale that addresses (overall) grade (of vocal dysphonia), roughness, breathiness, asthenia, and strain (21, 22). Clinicians rate these subscales from 0 to 3, in which zero indicates normal, and 3 implies severely impaired voice quality (21). The grade score was used to determine the clinicians' evaluations of subjects' voices in this study. Samples were further partitioned into normal (0 scores), slightly, moderately, and severely dysphonic groups (1, 2, and 3 scores, respectively) based on the GRBAS Scale (23).

3.4. Statistical Analysis

Descriptive statistics were obtained for all variables with SPSS software ver.16 (SPSS Inc., Chicago, IL). We used descriptive analysis to calculate the mean \pm SD and linear regression to assess the correlation between quantitative variables. There were no missing data. Univariate distribution of data was used to evaluate normality. Skewness (± 3) and kurtosis (± 8) were applied to evaluate the univariate normal distribution. The independent-sample *t*-test was used to assess the differences between healthy singers and individuals with a vocal disorder. A one-way analysis of variance (ANOVA) with Tukey post hoc test was used to assess the difference between different levels of P-SVHI based

on the GRBAS diagnostic tool. The statistical significance level was $P < 0.05$. We followed COSMIN recommendations to report this study (24).

3.5. Psychometric Analysis

3.5.1. Validity

We assessed face, construct, discriminant, and convergent validity, as follows.

3.5.1.1. Construct Validity

An exploratory factor analysis was employed using the maximum likelihood method with Promax rotation (oblique). The extraction of each factor was determined based on a factor loading of almost 0.31, which was estimated using the following formula:

$$CV = 5.152 \div \sqrt{(n - 2)}$$

Where CV is the number of extractable factors and n is the sample size (25). The Eigenvalue was considered ≥ 1 (26). The sample size should be more than 200 participants for factor analysis, according to MacCallum et al. (27). However, we recruited 285 cases.

Items with communalities less than 0.2 were excluded from the analysis. Bartlett's test of sphericity was performed to evaluate the appropriateness of the study sample. The Kaiser-Meyer-Olkin (KMO) test was also used to assess the data for factor analysis with a cutoff of 0.9 as an excellent result.

3.5.1.2. Scoring

To facilitate the interpretation of the results and compare them to other studies, we used a linear transformation to homogenize the scoring of original SVHI scores with the new P-SVHI using the following formula:

$$\text{Score in } 0 - 100 = ((\text{The obtained raw score} - \text{The lowest possible raw score}) / (\text{The highest possible score} - \text{The lowest possible score})) \times 100$$

The final scoring of this transformation is 0 - 100 (28).

3.5.1.3. Discriminant Validity

We used known-group analysis to investigate the clinical validity of the tool and differences between healthy and dysphonic singers. Then, we assumed that singers with a voice disorder would have higher scores on the P-SVHI than healthy singers. Since the P-SVHI scores showed normal distribution, an independent sample *t*-test was used.

The GRBAS was used as a diagnostic tool among 96 participants. No one received the three scores of GRBAS. Thus, we compared P-SVHI scores to GRBAS scores. A one-way analysis of variance (ANOVA) assessed statistical differences in P-SVHI scores across the GRBAS levels.

3.5.1.4. Convergent Validity

We expected that if the questionnaire was valid, there would be a correlation between P-SVHI and IPVI. It was assessed randomly among 179 participants. A Spearman correlation was calculated between the total score and subscale scores of P-SVHI and IPVI test items. A moderate correlation coefficient ($r = 0.3$) between the relevant scales was acceptable (29).

3.5.2. Reliability

3.5.2.1. Test-retest Reliability and Internal Consistency Analysis

The reliability of the P-SVHI was determined using (a) internal consistency with the average inter-item correlation (AIC) and alpha (30) and (b) test-retest analysis using the intra-class correlation coefficient ($ICC_{\text{two-way random, absolute agreement: } ICC_{2,k}}$) (31) using two-way random model, with interaction for the absolute agreement between scores ($ICC_{2,k}$) (32). A cutoff of 0.70 was considered acceptable, a value < 0.8 was considered good, and a value > 0.9 was considered the evidence of excellent internal consistency (25). Item internal consistency was assessed using item and subscale correlation corrected for overlap, and the Spearman rank correlation coefficient was used (33); a value of 0.40 was considered acceptable, and ≥ 0.80 indicated high internal consistency (34). The $ICC_{2,k}$ values < 0.4 showed poor agreement, between 0.4 and 0.75 fair to good agreement, and values > 0.75 excellent agreement (35).

3.5.3. Floor and Ceiling Effects

The distribution of responses was evaluated by assessing potential floor and ceiling effects. These effects were considered at the lowest (0) and highest (4) scores. A floor or ceiling effect was present if more than 15% of the respondents indicated the lowest or the highest possible score on the P-SVHI (36).

4. Results

4.1. Demographic Data

Of 300 traditional singers, 15 were excluded due to either rhinitis comorbidity or smoking status. Of all 285 participants, we selected 275 traditional singers because 10 persons had outlier data. Therefore, the final analysis was done among 168 healthy subjects and 107 subjects who experienced a voice disorder in the last month. The mean \pm SD of age was 35.43 ± 10.7 , and the singing experience was 12.06 ± 13.27 years. The majority of the subjects (77.9%) were males. All characteristics are shown in Table 1.

Table 1. Demographic Characteristics of the Sample (N = 275)

Characteristics	No. (%)
Age (y)	
Range 16 - 83	35.45
Gender	
Male	216 (78.5)
Female	59 (21.5)
Singing experience (y)	
	12.06 \pm 13.27
Voice disorder (GRBAS) (n = 96)	
Yes	
Male	38 (17.6)
Female	11 (18.7)
No	
Male	22 (10.2)
Female	25 (42.4)
Education	
Primary school	13 (4.6)
Diploma	58 (20)
Associate degree	28 (9.8)
Bachelors	111 (42)
Master	59 (21.5)
Doctorate	6 (2.1)

Abbreviations: GRBAS, grade, roughness, breathiness, asthenia, and strain.

4.2. Validity

4.2.1. Construct Validity

All Skewness values were between ± 3 , and kurtosis values were between ± 7 . Bartlett's test of sphericity ($\chi^2 = 6410.07$, P-value < 0.001) and Kaiser-Meyer-Olkin test (KMO = 0.91) demonstrated the appropriateness of sampling. The exploratory factor analysis extracted four factors with eigenvalues > 1.0 (Table 2). None of the items had communalities less than 0.2. Items 3, 7, 8, 13, 15, 19, 20, and 27 were removed from analysis according to the following method: first, we removed items that loaded on two factors from EFA one by one if their factor loadings had less than 0.2 differences and re-run factor analysis to identify the next item should be removed. Therefore, because the cross-loadings persisted, other items became candidates for deletion. Therefore, these items were excluded based on the EFA results (37). The introducing factors of the final 29-item P-SVHI (Table 2) followed factor 1 including 10 items (5, 21, 10, 6, 23, 31, 14, 26, 17, and 24), factor 2 including seven items (35, 34, 36, 32, 28, and 33), factor 3 with seven items (2, 4, 18, 9, 1, and 16), and factor 4 including three items (29, 25, and 30). In the final structure, item 25 was cross-loaded on

factors 1 and 2, but the difference between factor loadings was more than 0.2. Therefore, item 25 belonged to factor 4. A similar pattern was also seen in factor loadings of item 31. Therefore, it was considered an item of factor 1.

4.2.2. Scoring

We used the linear transformation to score the new P-SVHI. Since 7 items were eliminated after factor analysis, the range of the total score of P-SVHI decreased from 0 - 144 in the original 36-item version to 0 - 100 in the 29-item P-SVHI. Moreover, the mean \pm SD score was 16.00 ± 14.08 for healthy singers and 40.00 ± 17.05 for singers with a voice disorder. After performing the linear transformation, the scoring of the 29-item P-SVHI was transferred from 0 - 144 to 0 - 100. The mean \pm SD total score was 26.26 ± 19.87 . Therefore, in a 0 - 100 scoring, healthy subjects obtained a lower mean score than the dysphonic participants (16.9 vs. 40.9).

4.2.3. Discriminant Validity

The means \pm SD of P-SVHI scores were statistically different between healthy (16.90 ± 14.9) and dysphonic (40.9 ± 17.60) singers (P -value < 0.001) (Figure 1).

In a further approach, we randomly used the GRBAS Scale among 96 participants to discriminate between the total P-SVHI and GRBAS scores. A significant difference between the P-SVHI total and three levels of GRBAS scores was observed using one-way ANOVA ($F = 92.2$, $P < 0.000$). The P-SVHI total score was 18.3 ± 10.1 in singers with normal voice ($GRBAS = 0$), 39.2 ± 7.2 in singers with slight voice problems ($GRBAS = 1$), and 58.8 ± 7.3 in singers with moderate voice problems ($GRBAS = 2$). The Tukey post hoc test revealed that these significant differences were related to all levels of the GRBAS Scale (P -values $0 < 0.001$) (Figure 2).

4.2.4. Convergent Validity

The correlations between the P-SVHI total score and IPVI1, IPVI2, and "IPVI effort" scores were $r = 0.52$, 0.48 , and 0.33 , respectively. The correlation between IPVI items and all P-SVHI subscales ranged from $r = 0.30$ to 0.62 , as shown in Table 3.

4.3. Reliability

4.3.1. Test-retest Reliability and Internal Consistency Analysis

The average inter-item correlation (AIC) was 0.4 , and Cronbach's alpha was 0.95 for the total scale. Test-retest analysis using the intra-class correlation coefficient ($ICC_{2,k}$) ranged from low (0.69) to high (0.95) (Table 4).

The $ICC_{2,K}$ value for the total score was 0.97 ($P < 0.001$). The lowest $ICC_{2,K}$ was 0.69 , and the greatest was 0.95 . Corrected total item correlation was calculated for each item, ranging from 0.69 to 0.95 . Therefore, all remaining 29

items met the acceptance criteria ($ICC_{2,K} > 0.4$) and were retained in the new P-SVHI.

4.4. Floor and Ceiling Effects

The total floor and ceiling factor effects were 7% and 0% , respectively.

5. Discussion

Because of particular voice demands, specific tools should be used to evaluate potential occupational handicaps in singers (7). In the current study, the psychometric properties of the P-SVHI, with a focus on factor analysis, were investigated in traditional Persian singers. We found that the 29-item P-SVHI is valid and reliable for traditional Persian singers.

Concerning the demographic features of subjects, there was a significant difference between the number of male and female subjects. One reason may be that in Iran, female singers are only allowed to perform in public in choir groups, and they are not allowed to perform as solo singers unless for female audiences.

Investigating the construct validity of the original SVHI in English, Cohen et al. (2007 and 2009) performed Principal Component Analysis (PCA) by the quartimax and varimax methods to assess the potential subscales of the SVHI and then develop the short form of the SVHI Scale (7, 38). It yielded a different number of constructs. One of the significant issues of Cohen et al. studies is the rotation method applied in EPA (7, 38). By considering the essential association between given factors and an inter-correlation between factors extracted, oblique rotation is preferred over varimax and quatrimax. Accordingly, we used the Promax method as one of the most potent ways to rotate X and Y axes with a non-orthogonal angle to get a proper structure of the items. Another issue that needs to be addressed in Cohen et al.'s study is that they used principal component analysis. Once the concept underlying a questionnaire such as SVHI is extracted into various items, the maximum likelihood method seems to be the proper method of performing explanatory factor analysis (EFA). At the same time, PCA is used when some factors define the central concept of an index (34). However, these differences may be related to either the method of EFA running or applied rotation. In addition, based on the formula presented in the construct validity section, we used the ≥ 0.3 value as a cut-off point to extract factors through the EFA, while Cohen et al. study considered a threshold of 0.4 for factor loading instead of 0.30 . Thus, extracting only one factor by Cohen et al. (7, 38) compared to four factors extracted by the current study might be related to the cutoff they chose. Moreover, Cohen et al. omitted 26 items of the original 36 items,

Table 2. Factor Structure of P-SVHI Using Maximum Likelihood Method with Promax Rotation (Oblique)^a

Items	Factors			
	1	2	3	4
22	I have difficulty staying on the pitch when I sing.	0.79		
20	I have trouble controlling the raspiness in my voice.	0.76		
5	My ability to sing varies day to day.	0.75		
21	I have trouble singing loudly.	0.68		
10	I am unable to use my "high voice."	0.63		
6	My voice "gives out" on me while I am singing.	0.61		
23	I feel anxious about my singing.	0.58		
31	I am unsure of what will come out when I sing.	0.56	0.31	
14	I have had to eliminate certain songs from my singing/performances.	0.52		
26	My voice quality is inconsistent.	0.50		
17	I have trouble making my voice do what I want it to.	0.50		
24	My singing sounds forced.	0.36		
35	My singing makes me feel incompetent.		0.98	
34	I feel left out of the music scene because of my voice.		0.88	
36	I have to cancel performances, singing engagements, rehearsals, or practices.	0.75		
32	I feel something is missing in my life because of my inability to sing.		0.62	
28	My singing makes me feel handicapped.		0.52	
33	I am worried my singing problems will cause me to lose money.		0.50	
11	I get nervous before I sing because of my singing problems.		0.36	
2	My voice cracks and breaks.			0.84
4	People ask, "what is wrong with your voice?" when I sing.			0.79
12	My speaking voice is not normal.			0.58
18	I have to "push it" to produce my voice when singing.			0.55
9	I am embarrassed by my singing.			0.53
1	It takes a lot of effort to sing.			0.38
16	My singing voice is never normal.			0.37
29	My singing voice tires easily.			0.72
25	My speaking voice is hoarse after I sing.	0.30		0.55
30	I feel pain, tickling, or choking when I sing.			0.40
% of variance	44.5	0.06	0.05	0.04
Eigenvalue	12.93	1.85	1.40	1.08

^a Extraction method: Maximum likelihood. Rotation method: Promax with Kaiser normalization. Items 25 and 31 had more than 0.2 differences between their factor loadings. So, it is accepted that item 25 belongs to factor 4, and item 31 belongs to factor 1.

including items 2, 3, 4, 5, 8, 9, 11, 12, 13, 14, 16, 19, 20 - 28, 30, and 33 - 36 considering consensus comments, items being loaded on two factors, presence of low item-total correlation, presence of low factor loading, and floor effects lower than 0.5. In the current study, items 3, 7, 8, 13, 15, 19, and 27 were omitted after Promax oblique rotation, five of which were also deleted by Cohen et al. in their next arti-

cle (38). An oblique rotation allows a degree of correlation between the factors to improve the inter-item correlations within the factors. Cross-loading of seven items on two factors reduced the items from 36 to 29 items. Therefore, we obtained a new version of P-SVHI with 29 items that seemed appropriate to traditional singers.

We converted the total score of all these studies into a

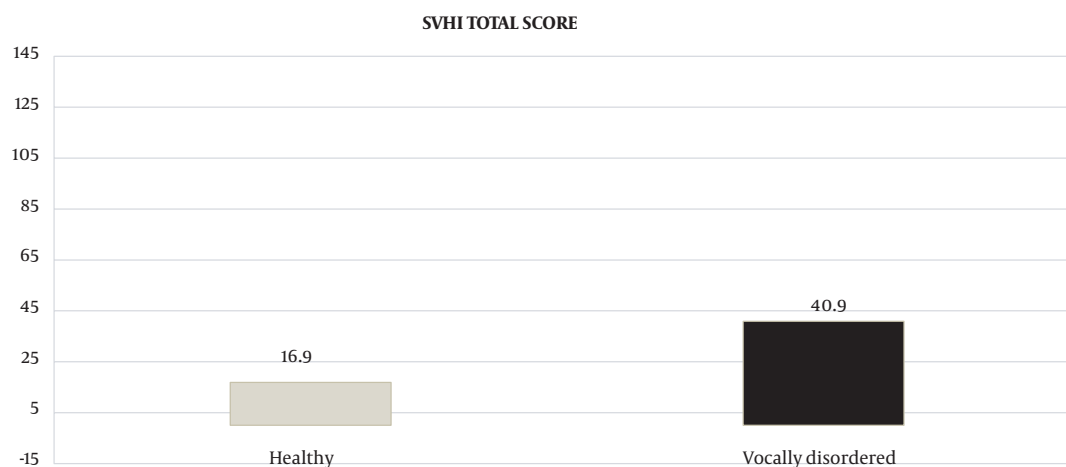


Figure 1. Difference between SVHI scores of healthy and dysphonic singers

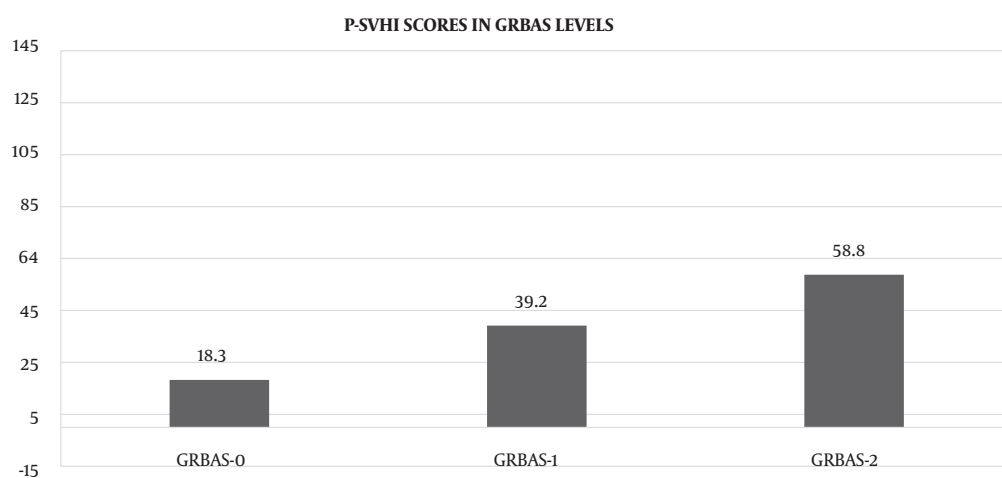


Figure 2. Average P-SVHI score according to different levels of GRBAS score

0 - 100 score so the results would be comparable to ours. The mean score for healthy singers in the current study was 16.90 on the P-SVHI score, range 0 - 100, which is similar to the results of healthy singers in Sobol et al., in which 729 healthy singers pooled from different languages had an average score of 14.10 on a 0 - 100 scale (19). The mean P-SVHI score, range 0 - 100, in healthy singers in the current study (16.9) also is similar to the results previously reported by Ghaemi et al. (15). The mean \pm SD scores of the SVHI were reported in Spanish (voice disorder patients vs. healthy group: 61.46 ± 30.50 vs. 28.43 ± 18.58) (13), Korean (dysphonic vocalist group vs. normal vocalist group: 70.63 ± 33.90 vs. 19.84 ± 12.84) (16), Kannada (voice problem group vs. normal group: 61.05 ± 20.57 vs. 19.81 ± 10.74) (9), Per-

sian (voice disorder patients vs. healthy group: 68.75 ± 10.02 vs. 18.0 ± 2.65) (15), and Turkish (pathology group vs. healthy group: 53.6 ± 28.9 vs. 21.8 ± 18.5) (11). Also, the P-SVHI score (0 - 100) for dysphonic subjects in our study was 40.9, which is more than the results for an Italian version in which dysphonic signers scored 31.68 (converted to 0-100), suggesting that voice disorder in our traditional singers is more than that in other style singers. In the Italian version, the mean \pm SD scores of SVHI in the normal group of singers and vocal fold abnormality group were 29.26 ± 25.72 and 45.62 ± 27.95 , respectively.

There was a significant and moderate correlation between P-SVHI and IPVI scores, implying that singers with higher IPVI scores had higher P-SVHI scores. Similar results

Table 3. Correlations Between P-SVHI Total and Subscale Scores and IPVI Items^a

IPVI	Factor 1	Factor 2	Factor 3	Factor 4	P-SVHI Total Score
IPVI1	0.52	0.54	0.45	0.43	0.56
IPVI2	0.48	0.62	0.51	0.40	0.57
IPVI effort	0.33	0.44	0.43	0.30	0.40

^a All correlation coefficients were significant at a level $P < 0.001$.

Table 4. Test-retest Reliability and Internal Consistency of P-SVHI Subscales

Subscales of P-SVHI	Corrected Total-Item Correlation	Cronbach α	Cronbach α If Item Deleted	ICC _{2,k}
Factor 1	0.51 - 0.69	0.90	0.89 - 0.90	0.83 - 0.94
Factor 2	0.55 - 0.78	0.86	0.82 - 0.85	0.82 - 0.86
Factor 3	0.38 - 0.66	0.82	0.76 - 0.77	0.78 - 0.84
Factor 4	0.58 - 0.71	0.79	0.55 - 0.62	0.63 - 0.77
Total	0.45 - 0.76	0.95	0.95 - 0.95	0.69 - 0.95

were shown in the Korean version of SVHI, in which Lee et al. reported that the correlation coefficient between the SVHI score and IPVI's first item was 0.52 for normal vocalists and 0.78 for dysphonic vocalists (16). As the IPVI reflects the singers' perception of the quality of their singing voice, we logically expected to observe the relationship between the voice quality and all the subscales of the SVHI.

Robust test-retest reliability was also shown for the 29-item P-SVHI with a high value of ICC_{2,k} (0.63 - 0.95). Internal consistency was confirmed through a Cronbach alpha coefficient and average inter-item correlation, similar to the previous version in the Ghaemi et al. study (Cronbach $\alpha = 0.83$) (15).

Our study found no floor or ceiling effects for the P-SVHI total score. If floor and ceiling effects are present for a questionnaire, they may influence the tool's responsiveness; as a result, it cannot identify singers' improvement and deterioration beyond the extreme ends of the scales. Therefore, the lack of floor or ceiling effects in our study reflects the reliability of the P-SVHI.

5.1. Limitations

One of the limitations of the current study was that only an auditory perceptual examination was performed to distinguish between healthy and dysphonic singers. Therefore, to better distinguish between healthy and disordered people, we suggest a comprehensive laryngeal examination performed by a laryngologist. Besides, another limitation of the current study is that only traditional Persian singers were included. Thus, to better understand vocal handicaps in Persian singing styles, investigating the psychometric features of the P-SVHI in different singing styles is proposed. Moreover, to better understand the ef-

fect of various vocal pathologies on vocal handicap, the current study researchers suggest that participants be categorized based on different voice disabilities.

5.2. Conclusions

This study showed that the revised 29-item P-SVHI is a valid, reliable, and sensitive tool for measuring voice-induced handicaps for traditional Persian singers with four subscales. Also, factor analysis showed that this questionnaire could capture different aspects of vocal handicaps with acceptable divergent and convergent validity. This P-SVHI version is an excellent questionnaire to discriminate between healthy and dysphonic traditional singers. This scale also showed high internal consistency and good test-retest reliability.

Acknowledgments

We would like to thank the cooperation of singers and staff of music centers.

Footnotes

Authors' Contribution: Study concept and design: F.T, N.A, F.R, and I.E.T; Acquisition of data: N.A and B.V.M; Analysis and interpretation of data: N.A, F.R, and I.E.T. Drafting of the manuscript: F.T, N.A, F.R, and K.V.A. Critical revision of the manuscript for important intellectual content: F.T, F.R, A.E, and K.V.A; Statistical analysis: F.R and A.E; Administrative, technical, and material support: F.T, F.R, and B.V.M; Study supervision: F.T.

Conflict of Interests: The authors declare no conflict of interest.

Data Reproducibility: The data presented in this study are openly available in one of the repositories or will be available on request from the corresponding author by this journal representative at any time during submission or after publication. Otherwise, all consequences of possible withdrawal or future retraction will be with the corresponding author.

Ethical Approval: The Research Ethics Committee of Iran University of Medical Sciences approved this study (Registration code: IR.IUMS.REC.1398.808, link: ethics.research.ac.ir/ProposalCertificateEn.php?id=95180).

Funding/Support: The Research Council of Iran University of Medical Sciences funded the study.

Informed Consent: Informed consent was obtained.

References

1. Titze IR, Lemke J, Montequin D. Populations in the U.S. workforce who rely on voice as a primary tool of trade: a preliminary report. *J Voice*. 1997;11(3):254-9. doi: [10.1016/s0892-1997\(97\)80002-1](https://doi.org/10.1016/s0892-1997(97)80002-1).
2. Kwok M, Eslick GD. The Impact of Vocal and Laryngeal Pathologies Among Professional Singers: A Meta-analysis. *J Voice*. 2019;33(1):58-65. doi: [10.1016/j.jvoice.2017.09.002](https://doi.org/10.1016/j.jvoice.2017.09.002). [PubMed: 29523383].
3. Baracca G, Cantarella G, Forti S, Pignataro L, Fussi F. Validation of the Italian version of the singing voice handicap index. *Eur Arch Otorhinolaryngol*. 2014;271(4):817-23. doi: [10.1007/s00405-013-2658-y](https://doi.org/10.1007/s00405-013-2658-y). [PubMed: 23974331].
4. Miller L. *Music and Song in Persia (RLE Iran B): The Art of Avaz*. London, England: Routledge; 1999.
5. Tahamtan M, Scherer RC. Tahrir Patterns and Acoustical Analysis of Tekyeh in a Professional Iranian Classical Singer. *J Voice*. 2020;34(6):830-7. doi: [10.1016/j.jvoice.2019.04.018](https://doi.org/10.1016/j.jvoice.2019.04.018). [PubMed: 31221483].
6. Salehi A, Barkmeier-Kraemer J. Laryngeal manual therapy as a treatment for impaired production of tahrir vibrato in traditional Iranian singers. *Folia Phoniatr Logop*. 2014;66(6):265-72. doi: [10.1159/000369063](https://doi.org/10.1159/000369063). [PubMed: 25676737].
7. Cohen SM, Jacobson BH, Garrett CG, Noordzij JP, Stewart MG, Attia A, et al. Creation and validation of the Singing Voice Handicap Index. *Ann Otol Rhinol Laryngol*. 2007;116(6):402-6. doi: [10.1177/000348940711600602](https://doi.org/10.1177/000348940711600602). [PubMed: 17672240].
8. Abou-Elsaad T, Baz H, Afsah O, Abo-Elhoud H. Validation and Adaptation of the Singing Voice Handicap Index for Egyptian Singing Voice. *J Voice*. 2017;31(1):1-6. doi: [10.1016/j.jvoice.2015.12.002](https://doi.org/10.1016/j.jvoice.2015.12.002). [PubMed: 26774846].
9. Gunjawate DR, Aithal VU, Guddattu V, Bellur R. Adaptation and Validation of the Kannada Version of the Singing Voice Handicap Index. *J Voice*. 2017;31(4):7-11. doi: [10.1016/j.jvoice.2016.10.012](https://doi.org/10.1016/j.jvoice.2016.10.012). [PubMed: 27913091].
10. Capucho MC, Ruah C, da Silva JM, Escada PA. Validation of the SVHI into Portuguese from Portugal. *Otolaryngol Head Neck Surg*. 2011;145(2_suppl):P200. doi: [10.1177/0194599811415823a216](https://doi.org/10.1177/0194599811415823a216).
11. Denizoglu J, Sahin K, Kazancioglu A, Dagdelen Z, Akdeniz S, Oguz H, et al. [Validation and reliability of Turkish Singing Voice Handicap index]. *Kulak Burun Bogaz Ihtis Derg*. 2016;26(1):1-6. doi: [10.5606/kbbihtisas.2016.34682](https://doi.org/10.5606/kbbihtisas.2016.34682). [PubMed: 26794328].
12. Sielska-Badurek EM, Sobol M, Cioch A, Osuch-Wojcikiewicz E, Rzepakowska AM, Niemczyk K. Adaptation and validation of the Singing Voice Handicap Index into Polish. *Clin Otolaryngol*. 2017;42(5):994-9. doi: [10.1111/coa.12833](https://doi.org/10.1111/coa.12833). [PubMed: 28107611].
13. Garcia-Lopez I, Nunez-Batalla F, Gavilan Bouzas J, Gorriz-Gil C. [Validation of the Spanish version of the voice handicap index for vocal singing (SVHI)]. *Acta Otorrinolaringol Esp*. 2010;61(4):247-54. doi: [10.1016/j.otorri.2010.01.012](https://doi.org/10.1016/j.otorri.2010.01.012). [PubMed: 20236622].
14. Lorenz A, Kleber B, Buttner M, Fuchs M, Murbe D, Richter B, et al. [Validation of the German version of the Singing Voice Handicap Index]. *HNO*. 2013;61(8):699-706. doi: [10.1007/s00106-013-2721-4](https://doi.org/10.1007/s00106-013-2721-4). [PubMed: 23868652].
15. Ghaemi H, Dehqan A, Mahmoodi-Bakhtiari B, Sobhani-Rad D. The Singing Voice Handicap Index (SVHI): Validation and Adaptation Into Persian. *J Voice*. 2020;34(1):17-23. doi: [10.1016/j.jvoice.2018.07.025](https://doi.org/10.1016/j.jvoice.2018.07.025). [PubMed: 30170911].
16. Lee AR, Sim HS. The Korean Version of the Singing Voice Handicap Index. *Commun Sci Disord*. 2013;18(2):194-202. doi: [10.12963/csd.13019](https://doi.org/10.12963/csd.13019).
17. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine (Phila Pa 1976)*. 2000;25(24):3186-91. doi: [10.1097/00007632-200012150-00014](https://doi.org/10.1097/00007632-200012150-00014). [PubMed: 11124735].
18. Hosseiniifar S, Torabinezhad F, Ghelichi L, Roudbari M, Silverman EP, Faham M. How Do Voice Perceptual Changes Predict Acoustic Parameters in Persian Voice Patients? *J Voice*. 2018;32(6):705-9. doi: [10.1016/j.jvoice.2017.08.015](https://doi.org/10.1016/j.jvoice.2017.08.015). [PubMed: 29032355].
19. Sobol M, Sielska-Badurek EM, Osuch-Wojcikiewicz E. Normative values for singing voice handicap index - systematic review and meta-analysis. *Braz J Otorhinolaryngol*. 2020;86(4):497-501. doi: [10.1016/j.bjorl.2018.12.004](https://doi.org/10.1016/j.bjorl.2018.12.004). [PubMed: 30846419].
20. Karnell MP, Melton SD, Childes JM, Coleman TC, Dailey SA, Hoffman HT. Reliability of clinician-based (GRBAS and CAPE-V) and patient-based (V-RQOL and IPV) documentation of voice disorders. *J Voice*. 2007;21(5):576-90. doi: [10.1016/j.jvoice.2006.05.001](https://doi.org/10.1016/j.jvoice.2006.05.001). [PubMed: 16822648].
21. Lee M, Drinnan M, Carding P. The reliability and validity of patient self-rating of their own voice quality. *Clin Otolaryngol*. 2005;30(4):357-61. doi: [10.1111/j.1365-2273.2005.01022.x](https://doi.org/10.1111/j.1365-2273.2005.01022.x). [PubMed: 16209679].
22. Hirano M, McCormick KR. Clinical Examination of Voice. *J Acoust Soc Am*. 1986;80(4):1273. doi: [10.1121/1.393788](https://doi.org/10.1121/1.393788).
23. De Bodd MS, Wuyts FL, Van de Heyning PH, Croux C. Test-retest study of the GRBAS scale: Influence of experience and professional background on perceptual rating of voice quality. *J Voice*. 1997;11(1):74-80. doi: [10.1016/s0892-1997\(97\)80026-4](https://doi.org/10.1016/s0892-1997(97)80026-4).
24. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res*. 2010;19(4):539-49. doi: [10.1007/s11136-010-9606-8](https://doi.org/10.1007/s11136-010-9606-8). [PubMed: 20169472]. [PubMed Central: PMC2852520].
25. Hahs-Vaughn DL. *Applied Multivariate Statistical Concepts*. 1st ed. Routledge; 2016. doi: [10.4324/9781315816685](https://doi.org/10.4324/9781315816685).
26. Stevens JP. *Applied Multivariate Statistics for the Social Sciences*. Routledge; 2012. doi: [10.4324/9780203843130](https://doi.org/10.4324/9780203843130).
27. MacCallum RC, Widaman KF, Preacher KJ, Hong S. Sample Size in Factor Analysis: The Role of Model Error. *Multivariate Behav Res*. 2001;36(4):611-37. doi: [10.1207/S15327906MBR3604_06](https://doi.org/10.1207/S15327906MBR3604_06). [PubMed: 26822184].
28. Nikpour M, Tirgar A, Ebadi A, Ghaffari F, Firouzbakht M, Hajiahmadi M. Development and psychometric evaluation of a women shift workers' reproductive health questionnaire: study protocol for a sequential exploratory mixed-method study. *Reprod Health*. 2018;15(1):22. doi: [10.1186/s12978-018-0456-0](https://doi.org/10.1186/s12978-018-0456-0). [PubMed: 29409501]. [PubMed Central: PMC5801669].
29. Portney LG, Watkins MP. *Foundations of clinical research: applications to practice*. 892. 2nd ed. Upper Saddle River, New Jersey: Prentice Hall; 2000.
30. Bujang MA, Omar ED, Baharum NA. A Review on Sample Size Determination for Cronbach's Alpha Test: A Simple Guide for Researchers.

- Malays J Med Sci.* 2018;**25**(6):85–99. doi: [10.21315/mjms2018.25.6.9](https://doi.org/10.21315/mjms2018.25.6.9). [PubMed: [30914882](https://pubmed.ncbi.nlm.nih.gov/30914882/)]. [PubMed Central: [PMC6422571](https://pubmed.ncbi.nlm.nih.gov/PMC6422571/)].
31. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J Chiropr Med.* 2016;**15**(2):155–63. doi: [10.1016/j.jcm.2016.02.012](https://doi.org/10.1016/j.jcm.2016.02.012). [PubMed: [27330520](https://pubmed.ncbi.nlm.nih.gov/27330520/)]. [PubMed Central: [PMC4913118](https://pubmed.ncbi.nlm.nih.gov/PMC4913118/)].
 32. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull.* 1979;**86**(2):420–8. doi: [10.1037//0033-2909.86.2.420](https://doi.org/10.1037//0033-2909.86.2.420). [PubMed: [18839484](https://pubmed.ncbi.nlm.nih.gov/18839484/)].
 33. Fayers PM, Machin D. *Quality of Life: The Assessment, Analysis and Reporting of Patient-reported Outcomes.* 3rd ed. Wiley; 2016. doi: [10.1002/9781118758991](https://doi.org/10.1002/9781118758991).
 34. Nunnally JC. *Psychometric theory* 3E. Tata McGraw-Hill Education; 1994.
 35. Rosner B. *Fundamentals of biostatistics: Nelson Education.* Nelson Education; 2015.
 36. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol.* 2007;**60**(1):34–42. doi: [10.1016/j.jclinepi.2006.03.012](https://doi.org/10.1016/j.jclinepi.2006.03.012). [PubMed: [17161752](https://pubmed.ncbi.nlm.nih.gov/17161752/)].
 37. Le TC, Cheong F. Perceptions of Risk and Risk Management in Vietnamese Catfish Farming: An Empirical Study. *Aquac Econ Manag.* 2010;**14**(4):282–314. doi: [10.1080/13657305.2010.526019](https://doi.org/10.1080/13657305.2010.526019).
 38. Cohen SM, Statham M, Rosen CA, Zullo T. Development and validation of the Singing Voice Handicap-10. *Laryngoscope.* 2009;**119**(9):1864–9. doi: [10.1002/lary.20580](https://doi.org/10.1002/lary.20580). [PubMed: [19572269](https://pubmed.ncbi.nlm.nih.gov/19572269/)].