

The Interoception in Voice Scale¹

^{#1}Alexander Jones, ^{#,1}Jonathan Silas, [†]Debbie Bethlehem, ^{‡,§}Sylvia Taylor-Goh, [¶]Michael Eggart, ^{||}Jennifer Todd, and ^{**,††}Sebastian Brand, ^{#†‡}London, [§]Manchester, ^{||}Cambridge, UK, [¶]Rüdersdorf, ^{**}Mainz, and ^{††}Heidelberg, Germany

SUMMARY: Emerging evidence suggests that interoception—the perception and interpretation of internal bodily signals—may play a role in voice perception and voice disorders; however, no self-report measure currently assesses interoception specifically in relation to voice processing. The present research aimed to develop and validate the Interoception in Voice Scale (IVS), a self-report measure designed to capture awareness of bodily cues related to voice production and perception. In total, 33 novel items were developed by drawing on existing interoceptive measures and expert input from speech and language specialists. In Study 1 ($N = 584$), exploratory factor analysis indicated a 19-item scale with four interpretable factors: Emotional Voice Awareness, Physical Voice Awareness, Voice-related Concern, and Voice-related Attention. These factors demonstrated good-to-excellent internal consistency ($\omega = .84-.91$). In Study 2 ($N = 320$), confirmatory factor analysis supported the four-factor structure, demonstrating acceptable model fit. Correlational analyses supported construct validity through meaningful associations with dimensions of the Multidimensional Assessment of Interoceptive Awareness (MAIA-2), and Voice-related Concern was consistently associated with self-perceived voice handicap. Together, the findings support the composite reliability and construct validity of the IVS. The scale provides a tool for investigating the role of interoception in voice perception and may be useful for future research and for speech and language therapists engaged in clinical work involving voice disorders. The survey and scoring instructions are available here: <https://jswlab.co.uk/IVS>

Key Words: Interoception–Voice–Speech and language–Psychometric–Scale development.

INTRODUCTION

A dependable voice is fundamental to everyday communication, but its production relies on a complex interplay of physiological, biological, and psychosocial systems. While there is some debate regarding the definition and classification of voice disorders,^{1,2} the voice may be considered disordered when the quality, pitch, loudness, flexibility, and vocal effort are perceived to be different from others with similar demographic characteristics (eg, gender, age, or cultural group).³ Voice disorders are relatively common, affecting approximately 7% of children and 17% of adults^{4,5} and with one in five adults reporting having a voice disorder at some point in their lifetime.⁶ Moreover,

there is an increased prevalence in occupations that are reliant on one's voice, such as teaching.⁷

While voice disorders are often defined in terms of observable vocal characteristics, they are also shaped by how individuals internally perceive and experience their own voice production. Interoception refers to the perception of internal bodily states. More specifically, interoception refers to the representation of the internal body, including the sensing, interpretation, integration, and regulation of internal body signals.⁸ Interoception is multidimensional, with components that span both conscious and non-conscious levels,⁹ and aspects that can be measured via neuroimaging, objective tasks, and self-report measures.^{9,10} It is acknowledged as an important component of numerous physical and mental health conditions, including anxiety disorders, eating disorders, addictive disorders, and somatic symptom disorders.^{11,12}

Vocal production depends on the monitoring and regulation of internal sensations such as respiratory effort, laryngeal activity, resonance within the vocal tract, and (dis)comfort in the throat, but the role of interoception in voice is not straightforward. As highlighted by Chen et al⁸ (see also Desmedt et al¹³), there is ambiguity in the boundaries between interoception and related sensory modalities, and vocal production also involves proprioceptive elements, including posture and positioning of the tongue and lips, and exteroceptive elements (auditory feedback and environmental acoustics) as well as the sensation of vibrotactile signals through bone conduction.¹⁴ Vocal production involves a complex relationship between different sensory inputs, outputs, and monitoring, and research on interoception has only recently surged in

Accepted for publication July 10, 2026.

This work was conducted without funding.

¹ The dataset and analysis script for both studies reported in this manuscript are available at the Open Science Framework. The URL pointing to this is: <https://osf.io/cnqw4/overview> DOI: <https://doi.org/10.17605/OSF.IO/CNQW4>

From the #Department of Psychology, School of Science and Technology, Middlesex University, London, UK; †Speech and Language Therapy, Independent Consultant, London, UK; ‡Speech and Language Therapy, Relational Communication Practice Ltd, London, UK; §Sensory Integration Education, Manchester, UK; ¶Center for Mental Health, Immanuel Hospital Rüdersdorf, Brandenburg Medical School Theodor Fontane, Rüdersdorf, Germany; ||School of Psychology, Sports and Sensory Sciences, Anglia Ruskin University, Cambridge, UK; **Department of Clinical Psychology and Psychotherapy of Childhood and Adolescence, Johannes Gutenberg-University Mainz, Mainz, Germany; and the ††Department of Clinical Psychology of Childhood and Adolescence, Heidelberg University, Heidelberg, Germany

¹ Share first authorship.

Address correspondence and reprint requests to Alexander Jones, Department of Psychology, Middlesex University, The Burroughs NW4 4BT, London, UK. E-mail: a.j.jones@mdx.ac.uk

Journal of Voice, Vol xx, No xx, pp. xxx–xxx
0892-1997

© 2026 The Author(s). Published by Elsevier Inc. on behalf of The Voice Foundation. This is an open access article under the CC BY-NC license (<http://creativecommons.org/licenses/by-nc/4.0/>).
<https://doi.org/10.1016/j.jvoice.2026.07.035>

popularity in the last decade. Together, this is perhaps the reason why studies on vocal production and interoception are only now starting to emerge.

Indeed, to date, only a handful of studies have examined associations between interoception and voice. In terms of clinical work, the findings of Smeltzer et al.¹⁵ suggest there may be differences in interoception in people with voice disorders. Specifically, they found associations between the self-perceived ability to sustain attention toward internal body signals and voice-related handicap and fatigue indices. This finding was in a combined cohort of participants with functional voice disorders and typical voice users. Typical voice users here were defined as having no current or prior voice disorder diagnoses, any related symptomatic complaints associated with their voice, or history or diagnosis of somatoform disorders. In the same study,¹⁵ Smeltzer et al. also found that participants with primary muscle tension dysphonia (a functional voice disorder thought to be the result of excessive muscle tension and hyperfunction in and around the larynx) reported significantly lower emotional awareness compared with typical voice users.

In terms of nonclinical work, Crow et al.¹⁶ examined the associations between interoception and vocal congruence, which they define as the extent to which one's voice is in alignment with one's sense of self. Here, they found that vocal congruence was significantly associated with interoceptive beliefs (ie, participants' degree of confidence in their performance on a heartbeat perception task), and interoceptive insight (the congruence between performance on the heartbeat perception task and confidence judgments). They also found group differences in vocal congruence when dividing participants into groups that are high or low in interoceptive insight, with the high insight group reporting higher vocal congruence. Other nonclinical studies have found associations between cardiac interoceptive accuracy and singing ability.^{17,18} Zamorano et al.¹⁸ found that accumulated singing practice predicted accuracy and that emotional awareness partly mediated this relationship. Additionally, among nonsingers, higher cardiac interoceptive accuracy was linked to better singing accuracy.

The present studies

In addition to the challenges in isolating the interoceptive aspects of voice, another factor that may have inhibited research in this domain is the lack of a measure designed for this purpose. While numerous self-report measures of interoception exist (for discussions, see^{19,20}), none specifically capture voice-related sensations or abilities. To address this issue, the overarching aim of the present studies was to develop and psychometrically evaluate a novel measure of bodily awareness in vocal production and perception. To do so, we drew upon the vast literature which documents that the self-reported component of interoception is multidimensional,^{19–21} with (variously termed) components such as interoceptive accuracy (the ability to accurately detect internal bodily sensations),²² attention (the degree to which interoceptive signals are the object of attention¹⁰), appreciation (valuing and appreciating the role of body sensations²³), attribution (interpretation of the causes of interoceptive sensations⁹), and appraisal (how signals are

interpreted, including affective experience and ascribed meaning in relation to the self^{24,25}). Accordingly, we broadly aimed to develop an item pool that can capture individual differences in the noticing and appraisal of changes in one's voice and related bodily sensations, including vocal quality, emotional state, and throat-related (dis)comfort, focusing on interoceptive sensations while acknowledging their interaction with proprioceptive and exteroceptive auditory processes.

In Study 1, we report on the process of item generation and refinement and an examination of content validity, as well as the results of an initial examination of factorial validity, composite reliability, and construct validity. Here, we used exploratory factor analysis (EFA) that allowed us to determine the best-fitting model in our sample, without any *a priori* modeling limitations. In Study 2, we cross-validated the factor structure using confirmatory factor analysis (CFA) and provide further evidence of composite reliability and construct validity. Although largely exploratory, we expected to extract a multidimensional structure reflecting distinct aspects of interoceptive vocal perception. We also expected that the measure would evidence appropriate composite reliability and construct validity via moderate correlations with aspects of self-reported interoception (indexed via the Multidimensional Assessment of Interoceptive Awareness-2 [MAIA-2]²⁶), and self-perceived voice-related impairment (assessed via the Voice Handicap Index-10²⁷) across both studies. In particular, based on the findings of Smeltzer et al.,¹⁵ we expected that our measure would evidence significant associations with the Emotional Awareness scale of the MAIA-2.

STUDY 1

Method

Ethical approval for both studies was obtained from the Middlesex Psychology departmental ethics committee, and all participants provided informed consent prior to data collection.

Scale development

Development of the items included in Study 1 was an iterative process. An initial pool of 40 items was developed by AJ, JS, DB, and ST-G by modifying the items of the MAIA-2²⁶ and MAIA-youth.²⁸ AJ and JS have experience of interoception scale development with the publications of MAIA-2 (Mehling, Acree, Stewart, and Silas & Jones²⁶) and MAIA-youth (Jones et al.²⁸). DB and ST-G are both highly experienced practicing speech and language therapists with backgrounds in academia as well as clinical practice. Their specialties include voice disorders, acquired neurological disorders, and sensory integration.

Through discussion, several items were developed by modification of the MAIA-2, for example, by adapting the item "I can maintain awareness of my whole body even when a part of me is in pain or discomfort" to "I can still

pay attention to my body even when my voice feels uncomfortable.” Additionally, laryngeal-specific items were developed by including voice-specific terminology and conditions such as the voice being husky, clear, breathy, strained, and weak. For example, “I notice when my voice feels strained” and “I notice when my voice feels weak”. We opted not to include any negatively worded items nor any reverse-scored items as these have suggested to lower internal consistency and contaminate a construct’s dimensionality by loading on a distinct factor.^{29,30} After several rounds of discussion by the team and external colleagues, a pool of 33 items was selected (see Appendix A). The response scale was also developed as a five-point Likert scale ranging from Never (0), Occasionally (1), Sometimes (2), Often (3), and Always (4).

Content validity and item revision

The 33 items were sent out to four independent expert voice speech and language therapists/pathologists, three based in the United Kingdom and a fourth in the United States.

The experts assessed the items based on the following five questions: 1). “Do these statements make sense to you? If any statements do not, please identify the statement(s) and explain why this is.” 2). “Do you think we need to make any changes to these statements to make them clearer and/or more suitable for you? If yes, please provide us with any suggestions you may have and why.” 3). “Were you able to respond to these statements based on the available options (i.e., all of the time, none of the time)? If not, could you explain why this is and suggest any alternative responses you think may be more suitable?” 4). “Do any of the statements listed above refer to offensive or ableist concepts/language? If yes, please indicate which statement(s) and explain why.” 5). “Do you have any other concerns you wish to share?” The experts also provided free-text feedback on how well the items capture the breadth of the construct. Based on the expert feedback, seven items were revised. For example, Item 10 was revised from “I find it difficult to tell if my voice will start easily” to “Before I speak, I can tell if my voice will start easily.” At the end of this process, we retained 33 items (see Appendix A) for EFA.

Participants

Data were collected between April 2024 and June 2025 using the online research platform Prolific (<https://www.prolific.co/>). In total, $N = 659$ individuals accessed the questionnaire, of whom $N = 595$ provided data. *A priori* determination of the required sample size followed widely accepted guidelines, which suggest a minimum of 200 participants (Howard, 2015), ideally a participant-to-item ratio between 10:1 and 20:1³¹ and consider sample sizes higher than $N = 500$ as very good.³² Accordingly, we aimed to recruit a sample of ~600 participants.

Eligibility criteria included being 18 years of age or older and fluent in English. Participants provided digital informed consent prior to participation. Completion of all

questions was required in order to proceed with the study. Upon finishing the survey, participants were compensated approximately £1.00 for their time, which is commensurate with Prolific guidelines for a survey of this duration.

After excluding participants who completed the questionnaire too quickly, measured via the relative-speed-index, RSI ($n = 11$ ³³), the final sample consisted of $N = 584$ adults. Of these, 248 (42.5%) identified as male, 327 (56.0%) as female, 7 (1.2%) as nonbinary, and 2 (0.3%) chose not to disclose their gender. Participants ranged in age from 18 to 77 years ($M = 35.0$, $SD = 13.0$).

In terms of education, 40 participants (6.8%) held a doctoral degree, 180 (30.8%) a postgraduate degree, 267 (45.7%) an undergraduate degree, and 97 (16.6%) had completed secondary education or lower. Additionally, 70 participants (12.0%) reported having a clinical voice condition. To increase statistical power and as no further details on the nature of voice condition were provided, these 70 participants were included in all analyses, but analyses without them are also reported in Appendix B (no differences in results have been observed).

Materials

Interoceptive in Voice Scale (IVS): The IVS included 33 items (see Appendix A).

Voice Handicap: The Voice Handicap Index-10 (VHI-10²⁷) is a widely used screening tool that assesses the perceived psychosocial impact of voice problems. It comprises 10 items rated on a 5-point Likert scale ranging from 0 (never) to 4 (always), with higher scores indicating higher perceived voice-related handicap.

Self-reported Interoception: The Multidimensional Assessment of Interoceptive Awareness, Version 2 (MAIA-2²⁶) was included to assess multiple dimensions of self-reported interoception across 37 items, grouped into eight subscales: Noticing (4 items), Not-Distracting (6 items), Not-Worrying (5 items), Attention Regulation (7 items), Emotional Awareness (5 items), Self-Regulation (4 items), Body Listening (3 items), and Trusting (3 items). Each item is rated on a 6-point Likert scale ranging from 0 (never) to 5 (always).

Demographic Items: Participants were requested to report demographic data comprising gender, age, educational attainment, and present/historical diagnosis of a voice condition (“Do you have, or have you ever been diagnosed with, a voice problem condition?” Response options were Yes or No). Participants also reported their occupation and ethnicity as free-text answers, but these demographics were not analyzed further.

Data analysis

Prior to analysis, data were inspected for completeness. Due to the use of forced response settings in the survey software, there were no missing responses at the item level. However, $n = 3$ participants exited the study just before the final page and did not complete the MAIA-2 questionnaire.

These participants were excluded from analyses involving MAIA-2 data but retained for all other analyses.

Formal analysis was conducted using R version 4.5.1³⁴ with the packages *psych*, *dplyr*, *EFAtools*, *EFA.dimensions*, and *lavaan*. We conducted a principal axis exploratory factor analysis (EFA). For factor extraction, an oblimin (oblique) rotation was used, reflecting the assumption of multiple, potentially interrelated dimensions. Data factorability was evaluated using the Kaiser-Meyer-Olkin (KMO³⁵) measure of sampling adequacy (which should ideally be $\geq .80$) and Bartlett's test of sphericity (which should be significant).

Item-level analyses were conducted to evaluate item quality before factor extraction. This included examining item-total correlations, item difficulty, and distributional properties. In line with psychometric recommendations, items were considered for exclusion if they showed item-total correlations near zero ($r_{it(i)} < .10$), extreme difficulty ($P_i < .05$ or $> .95$), or distributional irregularities such as excessive skewness or kurtosis.³⁶

As the commonly used Kaiser criterion (eigenvalues > 1) has frequently been criticized for its tendency to overestimate the number of factors, the number of factors to retain was determined using a combination of Velicer's Minimum Average Partial test (MAP³⁷), parallel analysis (PA³⁸), and inspection of the scree plot (visual scree test). Items were retained in the final solution if they showed primary factor loadings $\geq .40$, no substantial cross-loadings $\geq .30$, and a minimum difference of $.20$ between primary and secondary loadings.³⁹

Construct validity was assessed through bivariate correlations between IVS scores, MAIA-2 subscales, and VHI-10 scores. Based on Cohen,⁴⁰ values $\leq .10$ were considered weak, $\sim .30$ were considered moderate, and $\sim .50$ were considered strong correlations. Composite reliability was estimated using McDonald's⁴¹ omega with values $\geq .70$ reflecting adequate internal reliability.⁴²

The VHI-10 showed an excellent internal consistency, with $\omega = .93$.² In the MAIA-2, all except for one (Not-Worrying) subscales demonstrated an acceptable-to-excellent internal consistency: Noticing ($\omega = .80$), Not-Distracting ($\omega = .84$), Not-Worrying ($\omega = .67$), Attention Regulation ($\omega = .89$), Emotional Awareness ($\omega = .87$), Self-Regulation ($\omega = .87$), Body Listening ($\omega = .88$), and Trusting ($\omega = .87$).

RESULTS

Exploratory factor analysis

An item-level analysis (Table 1) indicated that item difficulty ranged from .32 to .68, with most items clustering around moderate values. Additionally, corrected item-total correlations were generally strong and above the predefined cutoff (range: .38-.70). Furthermore, skewness and kurtosis values were within an unproblematic range. Thus, based on

TABLE 1.
Item Analysis

Item	M	SD	P_i	$r_{it(i)}$	Skewness	Kurtosis
Item 1	2.3	1.3	.58	.52	-0.26	-1.02
Item 2	2.4	1.3	.59	.62	-0.32	-0.96
Item 3	2.2	1.3	.55	.69	-0.12	-1.06
Item 4	2.2	1.3	.56	.65	-0.18	-1.16
Item 5	2.2	1.4	.54	.62	-0.18	-1.17
Item 6	2.4	1.3	.59	.62	-0.25	-1.00
Item 7	2.4	1.3	.61	.62	-0.45	-0.98
Item 8	2.3	1.2	.58	.70	-0.18	-0.94
Item 9	2.3	1.2	.58	.67	-0.23	-0.97
Item 10	2.1	1.3	.51	.49	-0.11	-1.05
Item 11	2.0	1.3	.49	.59	0.04	-1.15
Item 12	2.5	1.3	.62	.64	-0.37	-1.04
Item 13	2.5	1.2	.61	.68	-0.39	-0.85
Item 14	2.6	1.2	.64	.66	-0.47	-0.65
Item 15	1.8	1.3	.46	.63	0.14	-1.04
Item 16	2.4	1.1	.60	.42	-0.24	-0.75
Item 17	1.9	1.4	.48	.54	0.06	-1.21
Item 18	2.1	1.3	.52	.60	-0.04	-1.09
Item 19	1.9	1.2	.47	.58	0.09	-0.93
Item 20	2.2	1.2	.56	.57	-0.16	-0.84
Item 21	2.5	1.2	.61	.64	-0.39	-0.78
Item 22	2.5	1.1	.63	.63	-0.42	-0.55
Item 23	2.7	1.2	.67	.61	-0.61	-0.58
Item 24	2.7	1.2	.68	.62	-0.63	-0.54
Item 25	2.5	1.2	.63	.63	-0.33	-0.78
Item 26	1.8	1.3	.44	.62	0.10	-1.15
Item 27	2.7	1.1	.68	.38	-0.55	-0.41
Item 28	2.2	1.2	.55	.68	-0.15	-0.89
Item 29	2.0	1.2	.50	.62	0.05	-0.95
Item 30	1.3	1.3	.32	.52	0.57	-0.85
Item 31	2.0	1.2	.49	.61	0.07	-1.02
Item 32	1.9	1.3	.48	.56	0.10	-1.02
Item 33	2.3	1.2	.58	.60	-0.30	-0.86

Notes: $r_{it(i)}$ = item-total correlation; P_i = item difficulty; scale from 0 to 4.

our criteria, none of the items were eliminated prior to factor analysis.

The Kaiser-Meyer-Olkin statistic for calculating sample suitability, KMO = 0.95, indicated that the sample was suitable for factor extraction. Additionally, Bartlett's test of sphericity was significant, $\chi^2(528) = 11662.7$, $p < .001$.

Based on the results of the MAP test, parallel analysis, and the scree plot, a four-factor solution was retained. Both the original ($r^2_p = .012$) and revised MAP tests ($r^4_p < .001$) indicated four factors, as this solution yielded the lowest average partial correlations. Parallel analysis supported this conclusion, with the first four eigenvalues from the actual data exceeding those from the simulated data ($\lambda_1 = 13.30 > 1.42$, $\lambda_2 = 2.42 > 1.37$, $\lambda_3 = 1.79 > 1.33$, and $\lambda_4 = 1.51 > 1.30$), while the fifth eigenvalue fell below the respective criterion ($\lambda_5 = 1.13 < 1.26$). Additionally, the scree plot revealed a steep decline in eigenvalues after the first factor, followed by a flattening of the curve after the fourth factor. Together, the four factors accounted for 52.6% of the common variance.

² Values $\geq .70$ were interpreted as an acceptable internal consistency (c.f.⁴²).

Factor interpretation and item reduction

We inspected the oblimin-rotated factor loadings (Table 2) using our *a priori* item-retention rule. Thirteen items did not meet the selection criteria: Eight items (Items 9, 10, 13, 16, 20, 27, 28, and 33) failed to reach the primary loading threshold ($\lambda < 1.40$), three items (Items 5, 7, and 12) did not show sufficient separation between primary and secondary loadings (difference < 1.20), and two items (Items 8 and 30) showed substantial cross-loadings (≥ 1.30). These items were excluded from further analysis. Although Item 21 met the primary loading threshold, its conceptual fit was weak (ie, “I notice if my stomach is tense” as an indicator of “Emotional voice awareness”—see below), and the difference between primary and secondary loadings was low and close to the cutoff ($\Delta = .259$). Therefore, this item was also excluded from the final set of items.

The remaining 19 items exhibited strong and distinct primary loadings. Based on content, the following four interpretable factors were identified: emotional voice awareness (EA), physical voice awareness (PA), voice-related concern (VC), and voice-related attention (VA).

TABLE 2.
Factor Loadings After Oblimin Rotation

Item	Factor 1	Factor 2	Factor 3	Factor 4
Item 1	-.096	.902	.012	-.103
Item 2	-.004	.870	.044	-.076
Item 3	-.021	.735	.043	.167
Item 4	.017	.816	.006	.030
Item 5	.151	.258	.000	.408
Item 6	.061	.728	-.020	.059
Item 7	.288	.122	-.013	.412
Item 8	.196	.447	-.041	.311
Item 9	.292	.375	-.017	.216
Item 10	.182	.228	-.036	.266
Item 11	.042	.122	.640	.044
Item 12	.231	.430	.196	-.004
Item 13	.297	.395	.189	.027
Item 14	.519	.242	.122	-.026
Item 15	.062	.183	.539	.108
Item 16	.354	.105	-.169	.228
Item 17	.045	.010	.785	-.026
Item 18	.126	-.001	.806	-.039
Item 19	.162	.068	.025	.513
Item 20	.248	.183	-.035	.340
Item 21	.450	.131	.056	.191
Item 22	.617	.084	.058	.058
Item 23	.749	.022	.064	-.045
Item 24	.813	-.114	.075	.027
Item 25	.585	-.005	.111	.128
Item 26	.121	-.051	.234	.546
Item 27	.385	.000	-.092	.188
Item 28	.216	.179	.198	.311
Item 29	.126	.110	.097	.502
Item 30	-.245	-.024	.466	.570
Item 31	.070	.029	-.042	.752
Item 32	.022	-.038	-.024	.804
Item 33	.378	.085	.109	.214

EA—Items 14, 22, 23, 24, and 25—captures the awareness of emotional modulation in one’s voice. PA—Items 1, 2, 3, 4, and 6—reflects the awareness and evaluation of one’s voice characteristics. VC—Items 11, 15, 17, and 18—represents the tendency to worry about sensations of pain or discomfort related to the voice production. VA—Items 19, 26, 29, 31, and 32—describes how individuals attend to posture, breathing, and vocal cues to stay aware of voice-related cues. The resulting 19-item IVS subscales were used in all subsequent analyses (see Appendix B).

Composite reliability and construct validity

The final four subscales exhibited good-to-excellent composite reliability, as $\omega_{EA} = .86$, $\omega_{PA} = .91$, $\omega_{VC} = .86$, and $\omega_{VA} = .84$.

Across all four IVS subscales, significant associations emerged with multiple MAIA-2 dimensions (Table 3). Both EA and PA showed consistent positive correlations with noticing, attention regulation, emotional awareness, self-regulation, body listening, and trusting, which were moderate-to-strong in magnitude. In addition, EA showed a small negative correlation with Not-Worrying. VC demonstrated a similar profile of positive associations but was further characterized by a moderate-to-strong negative association with not-worrying, indicating that individuals who tend to worry about bodily sensations also report higher voice-related concern. VA was strongly associated with noticing, attention regulation, emotional awareness, self-regulation, and body listening and trusting.

Regarding self-perceived voice impediment, VC and VA showed small-to-moderate positive associations with VHI scores, whereas EA and PA were not significantly associated with VHI. Thus, participants who reported higher voice-related concern or heightened attention to voice-related signals were also more likely to report voice impairment.

STUDY 2

Methods

Participants

Data collection for Study 2 took place in November 2025, again using the online research platform Prolific. A total of $N = 320$ individuals accessed the questionnaire, and all provided complete data.

A priori sample size determination for Study 2 was conducted using Monte Carlo simulation⁴³ using the *simsem* package in R.⁴⁴ Data were generated from the hypothesized four-factor measurement model using the respective loadings, 1000 replications, and robust maximum likelihood estimation, consistent with the modeling approach planned for Study 2 (see below). Simulations showed that at $N = 200$, the CFA model already achieved excellent fit with good parameter recovery. However, to increase estimate precision, improve stability under potential model misspecification, and to align with widely used item-to-participant guidelines (cf.³¹), we targeted a sample size of approximately $N \approx 300$ for Study 2.

TABLE 3.
Exploratory Correlation Analysis Across Constructs

Variable	M	1	2	3	4	5	6	7	8	9	10	11	12
	(SD)	[95% CI]	[95% CI]	[95% CI]	[95% CI]	[95% CI]	[95% CI]	[95% CI]	[95% CI]	[95% CI]	[95% CI]	[95% CI]	[95% CI]
1. EA	2.6 (0.9)												
2. PA	2.3 (1.1)	.51*** [.44, .56]											
3. VC	2.0 (1.1)	.56*** [.50, .61]	.49*** [.42, .54]										
4. VA	1.9 (1.0)	.60*** [.54, .65]	.42*** [.35, .48]	.51*** [0.45, 0.57]									
5. Noticing	3.5 (1.0)	.47*** [.40, .53]	.31*** [.23, .38]	.29*** [.21, .36]	.47*** [.40, .53]								
6. Not-distracting	2.1 (1.0)	-.01 [.09, .07]	-.03 [.11, .05]	-.01 [.08, .08]	.04 [-.04, .12]	-.03 [.11, .05]							
7. Not-worrying	2.4 (0.9)	-.18*** [.26, .01]	-.07 [-.15, .01]	-.43*** [.50, .36]	-.09* [.17, .01]	-.10* [.18, .02]	-.09* [.17, .01]						
8. Attention regulation	3.1 (1.0)	.38*** [.31, .45]	.25*** [.17, .32]	.20*** [.12, .27]	.54*** [.48, .59]	.52*** [.46, .58]	-.08 [.16, .01]	.8 [.00, .16]					
9. Emotional awareness	3.6 (1.0)	.56*** [.50, .61]	.35*** [.27, .42]	.35*** [.27, .42]	.46*** [.39, .52]	.62*** [.57, .67]	-.15*** [.23, .07]	-.19*** [.27, .11]	.55*** [.49, .60]				
10. Self-regulation	3.3 (1.1)	.42*** [.35, .48]	.22*** [.14, .30]	.30*** [.23, .38]	.53*** [.47, .59]	.52*** [.46, .58]	-.09* [.17, .01]	-.03 [.11, .06]	.60*** [.54, .65]				
11. Body listening	2.8 (1.3)	.42*** [.35, .48]	.21*** [.13, .29]	.37*** [.29, .43]	.59*** [.54, .64]	.55*** [.49, .61]	.01 [.08, .08]	-.14*** [.22, .06]	.62*** [.56, .67]	.69*** [.65, .73]			
12. Trusting	3.6 (1.1)	.43*** [.36, .49]	.23*** [.15, .30]	.25*** [.17, .32]	.44*** [.37, .50]	.43*** [.36, .50]	.06 [.02, .14]	-.01 [.09, .08]	.49*** [.43, .55]	.59*** [.54, .64]	.55*** [.49, .60]		
13. VHI	8.1 (7.5)	-.03 [.12, .05]	.02 [.06, .11]	.27*** [.19, .34]	.14*** [.06, .22]	-.04 [.12, .04]	-.07 [.15, .01]	-.17*** [.25, .09]	-.05 [.13, .03]	-.09* [.17, .01]	-.04 [.12, .04]	.06 [.02, .14]	-.17*** [.25, .09]

Notes: * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$ two-tailed. EA, Emotional Voice Awareness; PA, Physical Voice Awareness; VC, Voice-related Concern; VA, Voice-related Attention; VHI, Voice Handicap Index.

Eligibility criteria for Study 2 were identical to those in Study 1 (ie, being at least 18 years old and fluent in English). Participants provided digital informed consent prior to participation, and completion of all questionnaire items was required to proceed through the study. Upon finishing the survey, participants were compensated £1.00 for their time, which is commensurate with Prolific guidelines.

No participants in Study 2 had to be excluded for completing the questionnaire too quickly; thus, the final sample consisted of $N = 320$ individuals. Of these, 170 (53.1%) identified as male, 149 (46.6%) as female, and 1 (0.3%) as nonbinary. Participants ranged in age from 19 to 77 years ($M = 44.4$, $SD = 12.8$). Regarding education, 11 participants (3.4%) held a doctoral degree, 55 (17.2%) a postgraduate degree, 143 (44.7%) an undergraduate degree, and 111 (34.6%) had completed secondary school or lower education. Only four participants (1.2%) reported having a voice condition, again without further details.

Materials

The measures used in Study 2 were identical to those administered in Study 1, with the exception that we only administered the IVS items that were retained in the EFA.

Data analysis

All analyses for Study 2 were conducted in R version 4.5.1³⁴ with the packages *psych*, *dplyr*, *EFAtools*, *EFA.dimensions*, *lavaan*, and *semPlot*. The aim of Study 2 was to confirm the factor structure identified in Study 1. To this end, a CFA was performed on the 19 retained IVS items, specifying the four-factor model derived from EFA. Each item was constrained to load on only one factor, and all four latent factors were allowed to correlate. Factor loadings were freely estimated, and all residual covariances were fixed to zero.

The model was estimated using robust maximum likelihood estimation (MLR), to cover potential violation of multivariate normality assumptions. Model fit was evaluated using the Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), Root Mean Square Error of Approximation (RMSEA), Standardized Root Mean Square Residual (SRMR), and the model χ^2 statistic. Following established guidelines,⁴⁵ acceptable fit was defined as $CFI/TLI \geq .90$ (preferably $\geq .95$), $RMSEA \leq .08$ (preferably $\leq .05$), and $SRMR \leq .10$ (preferably $\leq .05$).

Construct validity was assessed through bivariate Pearson correlations between IVS scores, MAIA-2 subscales, and VHI-10 scores. Based on Cohen,⁴⁰ values $\leq .10$ were considered weak, $\sim .30$ were considered moderate, and $\sim .50$ were considered strong correlations. Composite reliability was estimated using McDonald's omega⁴¹ with values $\geq .70$ reflecting adequate internal reliability.⁴²

The VHI-10 again demonstrated excellent internal consistency ($\omega = .94$). The MAIA-2 subscales showed acceptable-to-excellent internal consistency across all scales: Noticing ($\omega = .78$), Not-Distracting ($\omega = .85$), Not-Worrying ($\omega = .82$), Attention Regulation ($\omega = .90$), Emotional Awareness

($\omega = .85$), Self-Regulation ($\omega = .87$), Body Listening ($\omega = .88$), and Trusting ($\omega = .90$).

RESULTS

Confirmatory factor analysis

The initial CFA indicated insufficient global fit, $\chi^2(146) = 564.28$, $P < 0.001$, $RMSEA = .10$ (90% CI [.09, .11]), $CFI = .88$, $TLI = .86$, and $SRMR = .10$. However, all items loaded strongly and significantly on their intended latent factors ($\lambda = .53-.91$), and the four IVS factors showed significant latent intercorrelations ($r = .18-.64$).

To explore potential sources of misfit, modification indices were inspected to identify theoretically interpretable model adjustments.⁴⁶ To preserve the factor structure derived from the EFA, only within-factor residual covariances among closely related items were considered, and only when indices indicated substantial misspecification ($MI > 100$). The largest modification indices suggested correlated residuals between item 31 and item 32 (both asking about paying attention to voice-related signals when other things are happening; $MI = 147.4$) and between item 1 and item 2 (asking about husky or rough voice; $MI = 128.0$).

Including these two within-factor residual covariances improved overall model fit to an acceptable level, $\chi^2(144) = 360.7$, $P < .001$, $RMSEA = .07$ (90% CI [.06, .08]), $CFI = .94$, $TLI = .93$, and $SRMR = .08$. In this modified model, standardized factor loadings remained high ($\lambda = .53-.91$), and the latent IVS factors continued to show moderate-to-strong positive intercorrelations ($r = .32-.79$). The final model is shown in Figure 1.

Internal consistency and construct validity

The confirmatory correlations with the MAIA-2 and VHI-10 scales are presented in Table 4. Across the two samples, the overall correlational patterns between the IVS subscales and the MAIA-2 dimensions were highly consistent. EA and PA once again showed a uniform pattern of moderate-to-strong positive associations with Noticing, Attention Regulation, Emotional Awareness, Self-Regulation, Body Listening, and Trusting. EA also demonstrated a small negative association with Not-Worrying. VC exhibited a partially similar correlational profile, with positive associations with Noticing, Emotional Awareness, and Body Listening, and a strong negative association with Not-Worrying. VA again demonstrated moderate-to-strong associations with Noticing, Attention Regulation, Emotional Awareness, Self-Regulation, Body Listening, and Trusting.

Regarding voice-related impairment, VC again demonstrated a small-to-moderate positive association with VHI, and PA also showed a small positive association. However, EA and VA were not related to perceived voice handicap.

DISCUSSION

The present study aimed to develop and validate a self-report measure that can capture individual differences in

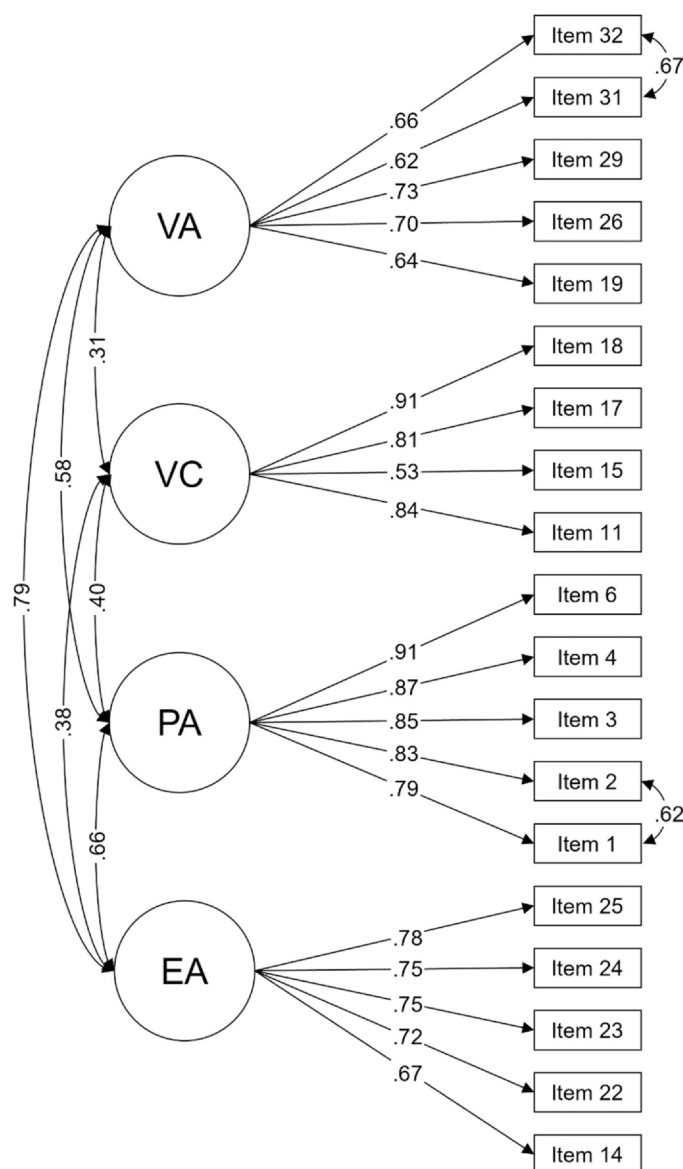


FIGURE 1. Four-factor model of the interoception in voice scale.

Note: Circles represent latent variables, squares refer to manifest variables, single-headed arrows represent factor loadings, and double-headed arrows between manifest variables represent standardized residual correlation paths; all factor loadings and correlation coefficients are significant at $P < .05$; error terms of manifest variables are not shown. VA, Voice-related attention; VC, Voice-related concern; PA, Physical voice awareness; EA, Emotional voice awareness.

the noticing and appraisal of changes in voice and associated bodily sensations, with a particular focus on interoceptive processes and their interaction with proprioceptive and auditory inputs. Across two studies, we introduced the Interoception in Voice Scale (IVS), a 19-item instrument capturing four dimensions of voice-related bodily experience: emotional voice awareness, physical voice awareness, voice-related concern, and voice-related attention. Exploratory and confirmatory factor analyses supported this four-factor structure, albeit following estimation of

modification indices for two pairs of items, and the scale demonstrated appropriate indices of composite reliability and construct validity via moderate-to-large associations with facets of self-reported interoception and self-perceived voice-related impairment.

The four-factor structure identified in this study reflects related but theoretically distinct components of voice-related interoceptive experience. Accordingly, future researchers should use the four scales as separate indices—that is, researchers should not compute a global score across all 19 items. The first scale, Emotional Voice Awareness, reflects the ability to recognize emotional states through changes in vocal characteristics. This dimension aligns with theoretical accounts suggesting that vocal expression provides an important channel through which emotional states are both communicated to others and perceived by the self.⁴⁷ Across both studies, emotional voice awareness showed strong positive correlations with MAIA-2 dimensions such as emotional awareness, body listening, and self-regulation. This pattern is consistent with evidence that adaptive forms of self-reported interoception (reflecting confidence in detecting and interpreting internal states) are associated with greater emotional differentiation and well-being.²¹

The second factor, Physical Voice Awareness, captures awareness of acoustic and physiological properties of the voice, such as strain, breathiness, or weakness. This dimension reflects the monitoring of physical vocal sensations that arise during voice production. Prior work has found a relationship with vocal experience and expertise in the manipulation of voice properties.¹⁸ For example, professional singers have been shown to exhibit enhanced interoceptive accuracy compared with nonsingers,¹⁷ suggesting that awareness of internal physiological cues may support skilled vocal control. The Physical Voice Awareness factor identified here may therefore represent an aspect of interoceptive attention related to vocal self-monitoring. Within this factor, model fit was improved in the CFA analyses by allowing the residuals of two items (“I notice when my voice sounds husky” and “I notice when my voice sounds rough”) to correlate, likely reflecting their semantic similarity and shared emphasis on closely related vocal qualities, and indicating a small degree of local dependence beyond the latent construct.

The third factor, Voice-related Concern, reflects the tendency to experience worry or emotional reactions in response to vocal discomfort or throat sensations, aligning with theorized construct of interoceptive appraisal (encompassing the affective evaluation and personal meaning ascribed to bodily signals^{24,25}). This dimension showed a strong negative relationship with the MAIA-2 Not-Worrying scale across both studies, indicating that individuals who tend to worry about bodily sensations also report greater concern about voice-related sensations. Such findings are consistent with broader interoception research demonstrating that maladaptive interpretations of bodily signals can contribute to heightened symptom perception

TABLE 4.
Confirmatory Correlation Analysis Across Constructs

Variable	M	1	2	3	4	5	6	7	8	9	10	11	12
(SD)		[95% CI]	[95% CI]	[95% CI]	[95% CI]	[95% CI]	[95% CI]	[95% CI]	[95% CI]	[95% CI]	[95% CI]	[95% CI]	[95% CI]
1. EA	2.4 (0.9)												
2. PA	2.3 (1.1)	.59*** [.51,.66]											
3. VC	1.6 (1.0)	.39*** [.29,.48]	.44*** [.34,.52]										
4. VA	1.5 (0.9)	.63*** [.56,.69]	.49*** [.40,.56]	.27*** [.17,.37]									
5. Noticing	3.4 (0.9)	.48*** [.39,.56]	.33*** [.23,.42]	.22*** [.11,.32]	.53*** [.45,.61]								
6. Not-distracting	2.0 (0.9)	.05 [−.06,.15]	.01 [−.10,.12]	.06 [−.05,.17]	.09 [−.02,.20]	−.03 [−.14,.08]							
7. Not-worrying	2.7 (1.0)	−.16** [−.26, −.05]	−.07 [−.18,.04]	−.51*** [−.59, −.42]	−.04 [−.15,.07]	−.06 [−.17,.05]	−.17** [−.27, −.06]						
8. Attention regulation	3.0 (1.0)	.42*** [.32,.50]	.26*** [.16,.36]	−.01 [−.12,.10]	.59*** [.52,.66]	.60*** [.53,.67]	.01 [−.10,.12]	.19*** [.08,.29]					
9. Emotional awareness	3.3 (1.1)	.53*** [.45,.60]	.30*** [.20,.40]	.22*** [.12,.32]	.55*** [.47,.63]	.64*** [.57,.70]	−.11 [−.21,.00]	−.12* [−.23, −.01]	.56*** [.48,.63]				
10. Self-regulation	3.0 (1.1)	.36*** [.26,.45]	.31*** [.21,.41]	.05 [−.06,.16]	.54*** [.46,.61]	.50*** [.42,.58]	−.02 [−.13,.09]	.17** [.06,.27]	.70*** [.64,.75]	.60*** [.53,.67]			
11. Body listening	2.5 (1.2)	.40*** [.30,.49]	.29*** [.18,.38]	.23*** [.12,.33]	.56*** [.48,.63]	.54*** [.46,.61]	.04 [−.07,.15]	−.06 [−.17,.05]	.57*** [.50,.64]	.66*** [.60,.72]	.66*** [.60,.72]		
12. Trusting	3.5 (1.1)	.29*** [.19,.39]	.17** [.06,.28]	−.01 [−.12,.10]	.36*** [.26,.45]	.32*** [.22,.41]	.11 [−.00,.22]	.15** [.04,.26]	.50*** [.41,.58]	.36*** [.26,.45]	.57*** [.49,.64]	.43*** [.34,.51]	
13. VHI	5.6 (6.4)	.01 [−.10,.12]	.12* [.01,.23]	.21*** [.11,.32]	.04 [−.07,.15]	.03 [−.08,.14]	−.14* [−.25, −.03]	−.17** [−.27, −.06]	−.14* [−.24, −.03]	.06 [−.05,.16]	−.09 [−.20,.02]	.01 [−.09,.12]	−.23*** [−.33, −.12]

Note. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$ two-tailed. EA, Emotional Voice Awareness; PA, Physical Voice Awareness; VC, Voice-related Concern; VA, Voice-related Attention; VHI, Voice Handicap Index.

and distress.^{48,49} In the context of voice perception, this pattern suggests that not all forms of voice awareness are necessarily adaptive; rather, excessive concern about vocal sensations may reflect a more anxiety-related component of interoceptive processing.

Notably, the Voice Concern factor also demonstrated consistent positive correlations with self-perceived voice-related impairment, as indexed by VHI-10 scores, across both samples. This finding may have particular relevance for understanding voice disorders. More specifically, we found that individuals who report greater concern about vocal sensations or who monitor these sensations more closely tend to experience greater perceived voice impairment. The convergence provides preliminary support for the construct validity of the factor, suggesting that it captures a dimension of voice-related body awareness that is meaningfully linked to perceived functional impact. Our findings here align with previous work by Smeltzer et al,¹⁵ who reported associations between interoceptive processes (particularly the ability to sustain attention toward internal bodily signals) and voice-related handicap and fatigue indices in both clinical and nonclinical adults.

The fourth factor, Voice-related Attention, reflects the capacity to monitor voice-related cues such as posture, breathing, and vocal signals during communication. This dimension showed particularly strong associations with MAIA-2 scales related to attentional processes, including Attention Regulation and Body Listening. These findings align with theoretical accounts emphasizing the role of attentional processes in interoception.^{50,51} Monitoring posture and breathing during speech may represent a form of deliberate attention toward internal vocal cues, which could be relevant for both typical vocal functioning and therapeutic interventions. Within this factor, it was necessary to correlate the error terms between Items #31 ("I can continue to pay attention to my voice even when there is a lot happening around me") and #32 ("I can pay attention to my voice even when my mind is racing"), likely reflecting shared variance associated with sustained attentional control under conditions of external and internal distraction.

The IVS makes a significant novel contribution to a nascent area of research, and may therefore have several applications. From a research perspective, it provides a tool for investigating how interoceptive attention and appraisal contribute to vocal perception, emotional expression, and communication. For example, emerging research suggests that interoceptive processes may play an important role in how individuals experience and interpret their own voice, and the importance of this to the sense of self.¹⁶ In clinical contexts, the scale may help identify patterns of voice-related interoceptive processing that could be relevant for the assessment and treatment of voice disorders. For example, interventions targeting body awareness or attentional processes have been proposed as potential approaches for improving interoceptive functioning.⁵² Understanding how individuals attend to and interpret vocal sensations may therefore inform voice therapy strategies aimed at improving vocal self-monitoring and reducing maladaptive concern about vocal symptoms.

Limitations and future directions

Nevertheless, several limitations should be considered when interpreting the present findings. First, both studies relied on self-report measures, collected from online samples and drawn from the general population, and only a small proportion of participants reported having a voice condition, which prevented more detailed analyses of clinical subgroups. Future research should therefore examine the psychometric properties of the IVS in clinical populations with diagnosed voice disorders and should also seek to explore whether the measure has clinical utility, for example, by examining sensitivity to change and minimal clinically important difference indices (eg, see Eggart et al⁵³). Additionally, further work could investigate relationships between IVS scores and objective measures of vocal function or interoceptive accuracy, such as physiological or behavioral assessments of voice production.

Future work should also be conducted to cross-validate the factor structure identified in the present study. The estimation of modification indices to respecify less-than-adequately fitting models is commonly recommended,⁴⁶ but it should also be acknowledged that this represents a post hoc, data-driven modification of the hypothesized model. Future examinations of factorial validity will help determine the extent to which modification indices represent true or idiosyncratic characteristics of the CFA dataset.^{54,55} This work could also examine a broader set of psychometric properties, which were beyond the scope of the present work, including test-retest reliability and discriminant validity.

CONCLUSION

Overall, the present research introduces the first scale specifically designed to measure interoception related to voice processing. The IVS demonstrated promising psychometric properties across two independent samples and provides a novel tool for studying the intersection between interoception and vocal perception. By enabling researchers and clinicians to assess how individuals perceive and interpret internal vocal signals, the scale may contribute to a deeper understanding of both typical voice perception and voice-related disorders.

CRedit Author Statement

None.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would like to thank the experts who provided feedback during the development of items.

Appendix A

Items included for data collection in study 1

Please read the sentences below and circle your answer. There is no right or wrong!

Item	Never	Occasionally	Sometimes	Often	Always
1 I notice when my voice sounds husky.	0	1	2	3	4
2 I notice when my voice sounds rough.	0	1	2	3	4
3 I notice when my voice sounds breathy.	0	1	2	3	4
4 I notice when my voice sounds weak.	0	1	2	3	4
5 I notice when my voice sounds clearer than usual.	0	1	2	3	4
6 I notice when my voice feels strained.	0	1	2	3	4
7 I notice when my voice feels comfortable.	0	1	2	3	4
8 I notice when my voice changes, even when there is a lot going on around me.	0	1	2	3	4
9 I notice when my voice becomes tired.	0	1	2	3	4
10 Before I speak, I can tell if my voice will start easily.	0	1	2	3	4
11 I get worried if I feel pain/discomfort in my throat.	0	1	2	3	4
12 I notice when I do not breathe freely.	0	1	2	3	4
13 I notice the first signs of discomfort in my throat before they become very strong.	0	1	2	3	4
14 I notice if my voice changes with different emotions.	0	1	2	3	4
15 When my voice feels uncomfortable, it dominates my attention.	0	1	2	3	4
16 I can still pay attention to my body even when my voice feels uncomfortable.	0	1	2	3	4
17 I get upset when I feel pain/discomfort in my throat.	0	1	2	3	4
18 I worry about my throat when it feels uncomfortable.	0	1	2	3	4
19 I notice my posture when I am talking.	0	1	2	3	4
20 I notice if there is tension in my shoulders.	0	1	2	3	4
21 I notice if my stomach is tense.	0	1	2	3	4
22 I notice how my voice changes when I speak in different situations.	0	1	2	3	4
23 I can tell I am angry by the sound of my voice.	0	1	2	3	4
24 I can tell if I am happy by the sound of my voice.	0	1	2	3	4
25 I can tell I am stressed/anxious by the sound of my voice.	0	1	2	3	4
26 I listen for clues from my voice to let me know what emotion I am feeling.	0	1	2	3	4
27 I feel comfortable in my body.	0	1	2	3	4
28 I can tell if my throat changes in how tight it feels.	0	1	2	3	4
29 I notice changes in my breathing over the day.	0	1	2	3	4
30 I worry about my voice even when I am doing other activities.	0	1	2	3	4
31 I can continue to pay attention to my voice even when there is a lot happening around me.	0	1	2	3	4
32 I can pay attention to my voice even when my mind is racing.	0	1	2	3	4
33 I can tell that I am tired by how my voice sounds.	0	1	2	3	4

Appendix B

TABLE A1.

Item Analysis Without Participants Who Reported a Voice Condition

Item	<i>M</i>	<i>SD</i>	<i>P_i</i>	<i>r_{it}</i>	Skewness	Kurtosis
Item 1	2.3	1.3	.57	.52	−0.22	−1.11
Item 2	2.4	1.3	.59	.63	−0.32	−1.01
Item 3	2.2	1.3	.54	.70	−0.08	−1.12
Item 4	2.2	1.4	.55	.65	−0.14	−1.21
Item 5	2.1	1.4	.53	.62	−0.12	−1.25
Item 6	2.3	1.3	.58	.62	−0.22	−1.05
Item 7	2.4	1.4	.60	.63	−0.42	−1.07
Item 8	2.3	1.2	.57	.71	−0.14	−1.01
Item 9	2.3	1.3	.58	.66	−0.22	−1.03
Item 10	2.0	1.3	.51	.51	−0.09	−1.09
Item 11	2.0	1.3	.47	.58	0.13	−1.15
Item 12	2.5	1.3	.62	.63	−0.39	−1.05

TABLE A1 (Continued)

Item	<i>M</i>	<i>SD</i>	<i>P_i</i>	<i>r_{it}</i>	Skewness	Kurtosis
Item 13	2.4	1.3	.61	.69	−0.38	−0.89
Item 14	2.6	1.2	.64	.66	−0.48	−0.67
Item 15	1.8	1.3	.44	.62	0.23	−1.02
Item 16	2.4	1.1	.60	.42	−0.25	−0.76
Item 17	1.8	1.4	.46	.54	0.13	−1.21
Item 18	2.0	1.3	.50	.59	0.02	−1.10
Item 19	1.8	1.2	.45	.56	0.15	−0.93
Item 20	2.2	1.2	.55	.55	−0.11	−0.86
Item 21	2.4	1.2	.60	.63	−0.35	−0.84
Item 22	2.5	1.1	.62	.64	−0.41	−0.60
Item 23	2.7	1.2	.68	.61	−0.63	−0.55
Item 24	2.7	1.2	.68	.64	−0.67	−0.52
Item 25	2.5	1.2	.62	.64	−0.31	−0.81
Item 26	1.7	1.3	.42	.61	0.19	−1.13
Item 27	2.8	1.1	.69	.37	−0.59	−0.38
Item 28	2.2	1.2	.54	.68	−0.10	−0.97
Item 29	2.0	1.2	.49	.61	0.12	−0.93
Item 30	1.2	1.3	.29	.50	0.76	−0.55
Item 31	1.9	1.2	.48	.59	0.12	−1.02
Item 32	1.8	1.3	.45	.55	0.19	−0.99
Item 33	2.3	1.2	.57	.60	−0.24	−0.93

Note. $r_{it(i)}$ = item–total correlation; P_i = item difficulty; scale from 0 to 4.

TABLE A2.

Factor Loadings Without Participants Who Reported a Voice Condition

Item	Factor 1	Factor 2	Factor 3	Factor 4
Item 1	−.091	.901	.009	−.107
Item 2	−.013	.874	.045	−.063
Item 3	−.023	.738	.055	.166
Item 4	.028	.827	.004	.002
Item 5	.159	.248	.026	.383
Item 6	.053	.750	−.032	.049
Item 7	.290	.117	.018	.398
Item 8	.216	.452	−.025	.279
Item 9	.302	.382	−.039	.204
Item 10	.174	.238	−.031	.274
Item 11	.048	.122	.627	.049
Item 12	.174	.463	.177	.029
Item 13	.265	.405	.162	.076
Item 14	.517	.250	.116	−.025
Item 15	.057	.207	.55	.072
Item 16	.316	.145	−.194	.244
Item 17	.046	.011	.778	−.018
Item 18	.127	−.009	.809	−.044
Item 19	.151	.075	.020	.496
Item 20	.255	.189	−.050	.313
Item 21	.460	.137	.040	.179
Item 22	.666	.070	.058	.019
Item 23	.748	.024	.060	−.054
Item 24	.806	−.109	.082	.039
Item 25	.618	−.011	.105	.116
Item 26	.132	−.051	.247	.516
Item 27	.319	.014	−.098	.226
Item 28	.192	.183	.204	.330
Item 29	.114	.112	.082	.514

TABLE A2 (*Continued*)

Item	Factor 1	Factor 2	Factor 3	Factor 4
Item 30	-.249	-.011	.439	.577
Item 31	.056	.017	-.037	.765
Item 32	.037	-.044	-.050	.810
Item 33	.399	.088	.071	.223

Note. Based on the MAP tests, parallel analysis, and the scree plot, a four-factor solution was retained. Both the original ($r_p^2 = .013$) and revised MAP tests ($r_p^4 < .001$) reached their minima at four factors. Parallel analysis likewise supported this structure, with four eigenvalues exceeding the corresponding random eigenvalues ($\lambda_1 = 13.21 > 1.45$, $\lambda_2 = 2.50 > 1.39$, $\lambda_3 = 1.81 > 1.35$, and $\lambda_4 = 1.49 > 1.31$). Together, the four factors accounted for 52.5% of the common variance.

Appendix C

Final Version of the Interoception in Voice Scale

Please read the sentences below and select your answer. There is no right or wrong.

	Never	Occasionally	Sometimes	Often	Always
1 I notice if my voice changes with different emotions	0	1	2	3	4
2 I notice how my voice changes when I speak in different situations	0	1	2	3	4
3 I can tell I am angry by the sound of my voice	0	1	2	3	4
4 I can tell if I am happy by the sound of my voice	0	1	2	3	4
5 I can tell I am stressed/anxious by the sound of my voice	0	1	2	3	4
6 I notice when my voice sounds husky	0	1	2	3	4
7 I notice when my voice sounds rough	0	1	2	3	4
8 I notice when my voice sounds breathy	0	1	2	3	4
9 I notice when my voice sounds weak	0	1	2	3	4
10 I notice when my voice feels strained	0	1	2	3	4
11 I get worried if I feel pain/discomfort in my throat	0	1	2	3	4
12 When my voice feels uncomfortable, it dominates my attention	0	1	2	3	4
13 I get upset when I feel pain/discomfort in my throat	0	1	2	3	4
14 I worry about my throat when it feels uncomfortable	0	1	2	3	4
15 I notice my posture when I am talking	0	1	2	3	4
16 I listen for clues from my voice to let me know what emotion I am feeling	0	1	2	3	4
17 I notice changes in my breathing over the day	0	1	2	3	4
18 I can continue to pay attention to my voice even when there is a lot happening around	0	1	2	3	4
19 I can pay attention to my voice even when my mind is racing	0	1	2	3	4

Note. Items 1-5 = Emotional voice awareness (EA), Items 6-10 = Physical voice awareness (PA), Items 11-14 = Voice-related Concern (VC), Items 15-19 = Voice-related Attention (VA); Scale: 0 = Never, 4 = Always.

Scoring instructions: Compute the average score for the items on each scale.

- Emotional voice awareness (EA):** awareness of emotional modulation in one's voice (Item 1 ____ + Item 2 ____ + Item 3 ____ + Item 4 ____ + Item 5 ____) $\div 5 =$ ____
- Physical voice awareness (PA):** awareness of acoustic and physiological properties of the voice (Item 6 ____ + Item 7 ____ + Item 8 ____ + Item 9 ____ + Item 10 ____) $\div 5 =$ ____
- Voice-related concern (VC):** tendency to experience worry or emotional reactions in response to vocal discomfort or throat sensations (Item 11 ____ + Item 12 ____ + Item 13 ____ + Item 14 ____) $\div 4 =$ ____
- Voice-related attention (VA):** capacity to monitor voice-related cues such as posture, breathing, and vocal signals during communication (Item 15 ____ + Item 16 ____ + Item 17 ____ + Item 18 ____ + Item 19 ____) $\div 5 =$ ____

References

- Constantini AC, Ribeiro VV, Behlau M. Voice disorder classifications: a scoping review – part A. *J Voice*. 2025;39:676–684. <https://doi.org/10.1016/j.jvoice.2022.11.016>.
- Payten CL, Chiapello G, Weir KA, Madill CJ. Frameworks, terminology and definitions used for the classification of voice disorders: a scoping review. *J Voice*. 2024;38:1070–1087. <https://doi.org/10.1016/j.jvoice.2022.02.009>.
- Merrill RM, Roy N, Lowe J. Voice-related symptoms and their effects on quality of life. *Ann Otol Rhinol Laryngol*. 2013;122:404–411. <https://doi.org/10.1177/000348941312200610>.
- Fujiki RB, Thibeault SL. Voice disorder prevalence and vocal health characteristics in children. *JAMA Otolaryngol-Head & Neck Surg*. 2024;150:677. <https://doi.org/10.1001/jamaoto.2024.1516>.
- Lyberg-Åhlender V, Rydell R, Fredlund P, Magnusson C, Wilén S. Prevalence of voice disorders in the general population, based on the Stockholm public health cohort. *J Voice*. 2019;33:900–905. <https://doi.org/10.1016/j.jvoice.2018.07.007>.
- Huston MN, Puka I, Naunheim MR. Prevalence of voice disorders in the United States: a national survey. *Laryngoscope*. 2024;134:347–352. <https://doi.org/10.1002/lary.30929>.
- Feng S, Weng C, Cai S, Yang Z, Wu M, Kang N. The prevalence and risk factors for perceived voice disorders in public school teachers. *Laryngosc Investig Otolaryngol*. 2022;7:790–798. <https://doi.org/10.1002/lio2.803>.
- Chen WG, Schloesser D, Arensdorf AM, et al. The emerging science of interoception: sensing, integrating, interpreting, and regulating signals within the self. *Trends Neurosci*. 2021;44:3–16. <https://doi.org/10.1016/j.tins.2020.10.007>.
- Suksasilp C, Garfinkel SN. Towards a comprehensive assessment of interoception in a multi-dimensional framework. *Biol Psychol*. 2022;168:108262. <https://doi.org/10.1016/j.biopsycho.2022.108262>.
- Murphy J, Catmur C, Bird G. Classifying individual differences in interoception: Implications for the measurement of interoceptive awareness. *Psychon Bull Rev*. 2019;26:1467–1471. <https://doi.org/10.3758/s13423-019-01632-7>.
- Khalsa SS, Adolphs R, Cameron OG, et al. Interoception and mental health: a roadmap. *Biol Psychiatry Cogn Neurosci Neuroimag*. 2018;3:501–513. <https://doi.org/10.1016/j.bpsc.2017.12.004>.
- Nord CL, Garfinkel SN. Interoceptive pathways to understand and treat mental health conditions. *Trends Cogn Sci*. 2022;26:499–513. <https://doi.org/10.1016/j.tics.2022.03.004>.
- Desmedt O, Luminet O, Maurage P, Corneille O. Discrepancies in the definition and measurement of human interoception: a comprehensive discussion and suggested ways forward. *Perspect Psychol Sci*. 2025;20:76–98. <https://doi.org/10.1177/17456916231191537>.
- Orepic P, Kannape OA, Faivre N, Blanke O. Bone conduction facilitates self-other voice discrimination. *R Soc Open Sci*. 2023;10:1–17. <https://doi.org/10.1098/rsos.221561>.
- Smeltzer JC, Chiou SH, Shembel AC. Interoception, voice symptom reporting, and voice disorders. *J Voice*. 2025;39:1236–1245. <https://doi.org/10.1016/j.jvoice.2023.03.002>.
- Crow KM, van Mersbergen M, Payne AE. Vocal congruence: the voice and the self measured by interoceptive awareness. *J Voice*. 2021;35:324.e15–324.e28. <https://doi.org/10.1016/j.jvoice.2019.08.027>.
- Schirmer-Mokwa KL, Fard PR, Zamorano AM, Finkel S, Birbaumer N, Kleber BA. Evidence for enhanced interoceptive accuracy in professional musicians. *Front Behav Neurosci*. 2015;9:1–13. <https://doi.org/10.3389/fnbeh.2015.00349>.
- Zamorano AM, Haumann NT, Ligato D, et al. Embodied singing: dual role of interoception in vocal expertise and musical competence. *Ann N Y Acad Sci*. 2025;1554:187–202. <https://doi.org/10.1111/nyas.70107>.
- Desmedt O, Heeren A, Corneille O, Luminet O. What do measures of self-report interoception measure? Insights from a systematic review, latent factor analysis, and network approach. *Biol Psychol*. 2022;169:108289. <https://doi.org/10.1016/j.biopsycho.2022.108289>.
- Todd J, Swami V, Aspell JE, Furnham A, Horne G, Stieger S. Are some interoceptive sensibility components more central than others? Using item pool visualisation to understand the psychometric representation of interoception. *PLoS One*. 2022;17:e0277894. <https://doi.org/10.1371/journal.pone.0277894>.
- Ventura-Bort C, Wendt J, Weymar M. The role of interoceptive sensibility and emotional conceptualization for the experience of emotions. *Front Psychol*. 2021;12:1–15. <https://doi.org/10.3389/fpsyg.2021.712418>.
- Garfinkel SN, Seth AK, Barrett AB, Suzuki K, Critchley HD. Knowing your own heart: distinguishing interoceptive accuracy from interoceptive awareness. *Biol Psychol*. 2015;104:65–74. <https://doi.org/10.1016/j.biopsycho.2014.11.004>.
- Mehling WE, Todd J, Schuman-Olivier Z. Multidisciplinary assessment of interoceptive awareness, Version 2 (MAIA-2). *Handbook of Assessment in Mindfulness Research*. Cham, Switzerland: Springer Nature Switzerland,; 2025:1509–1530. https://doi.org/10.1007/978-3-031-47219-0_40.
- Joshi V, Graziani P, Del-Monte J. The role of interoceptive attention and appraisal in interoceptive regulation. *Front Psychol*. 2021;12:1–7. <https://doi.org/10.3389/fpsyg.2021.714641>.
- Savage HS, Garfinkel S. Making sense of sensation: a model of interoceptive attribution and appraisal with clinical applications. Preprint posted online July 11, 2025. doi:10.31234/osf.io/bsma9_v1.
- Mehling WE, Acree M, Stewart A, Silas J, Jones A. The multidimensional assessment of interoceptive awareness, Version 2 (MAIA-2). *PLoS One*. 2018;13. <https://doi.org/10.1371/journal.pone.0208034>. (Published online).
- Rosen CA, Lee AS, Osborne J, Zullo T, Murry T. Development and validation of the Voice Handicap Index-1000. *Laryngoscope*. 2004;114:1549–1556. <https://doi.org/10.1097/00005537-200409000-00009>.
- Jones A, Silas J, Todd J, et al. Exploring the multidimensional assessment of interoceptive awareness in youth aged 7–17 years. *J Clin Psychol*. 2021;77:661–682. <https://doi.org/10.1002/jclp.23067>.
- Gnambs T, Schroeders U. Cognitive abilities explain wording effects in the rosenberg self-esteem scale. *Assessment*. 2020;27:404–418. <https://doi.org/10.1177/1073191117746503>.
- Weijters B, Baumgartner H. Misresponse to reversed and negated items in surveys: a review. *J Market Res*. 2012;49:737–747. <https://doi.org/10.1509/jmr.11.0368>.
- Costello AB, Osborne J. Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. *Pract Assess Res Eval*. 2005;10:1–9.
- Comrey AL, Lee HB. *A First Course in Factor Analysis*. 2nd ed., Mahwah, NJ: Lawrence Erlbaum Associates,; 1992.
- Leiner DJ. Too fast, too straight, too weird: non-reactive indicators for meaningless data in internet surveys. *Direct Open Access J*. 2019;13:229–248.
- R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Preprint posted online 2021.
- Kaiser HF. A second generation little jiffy. *Psychometrika*. 1970;35:401–415. <https://doi.org/10.1007/BF02291817>.
- Kelava A, Moosbrugger H. Deskriptivstatistische evaluation von items (Itemanalyse) und Testwertverteilungen. In: 2012:75–102. doi:10.1007/978-3-642-20072-4_4.
- Velicer WF. Determining the number of components from the matrix of partial correlations. *Psychometrika*. 1976;41:321–327. <https://doi.org/10.1007/BF02293557>.
- Horn JL. A rationale and test for the number of factors in factor analysis. *Psychometrika*. 1965;30:179–185. <https://doi.org/10.1007/BF02289447>.
- Howard MC. A review of exploratory factor analysis decisions and overview of current practices: What we are doing and how can we improve? *Int J Hum Comput Interact*. 2016;32:51–62. <https://doi.org/10.1080/10447318.2015.1087664>.

40. Cohen J. Statistical power analysis. *Curr Dir Psychol Sci*. 1992;1:98–101. <https://doi.org/10.1111/1467-8721.ep10768783>.
41. McDonald RP. The theoretical foundations of principal factor analysis, canonical factor analysis, and alpha factor analysis. *Br J Math Stat Psychol*. 1970;23:1–21. <https://doi.org/10.1111/j.2044-8317.1970.tb00432.x>.
42. Madadzadeh F, Bahariniya S. Tutorial on internal consistency assessment by Cronbach's alpha and McDonald's omega. *Perioperative Care and Operating Room Management*. 2025;41:100568. <https://doi.org/10.1016/j.pccorm.2025.100568>.
43. Muthén LK, Muthén BO. How to use a Monte Carlo study to decide on sample size and determine power. *Struct Equ Model*. 2002;9:599–620. https://doi.org/10.1207/S15328007SEM0904_8.
44. Pornprasertmanit S, Miller P, Schoemann AM, Jorgensen TD, & Quick C. SIMulated Structural Equation Modeling (Version 0.5–17). CRAN. 2025.
45. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equ Model*. 1999;6:1–55. <https://doi.org/10.1080/10705519909540118>.
46. Whittaker TA. Using the modification index and standardized expected parameter change for model modification. *J Exp Educ*. 2012;80:26–44. <https://doi.org/10.1080/00220973.2010.531299>.
47. Pinheiro AP. Behind a voice there is a speaker: why vocal emotion research needs to become 'personal'. *Affect Sci*. 2025;6:562–574. <https://doi.org/10.1007/s42761-025-00317-w>.
48. Barsky AJ, Peekna HM, Borus JF. Somatic symptom reporting in women and men. *J Gen Intern Med*. 2001;16:266–275. <https://doi.org/10.1046/j.1525-1497.2001.016004266.x>.
49. Bogaerts K, Walentynowicz M, Van Den Houte M, Constantinou E, Van den Bergh O. The interoceptive sensitivity and attention questionnaire: evaluating aspects of self-reported interoception in patients with persistent somatic symptoms, stress-related syndromes, and healthy controls. *Psychosom Med*. 2022;84:251–260. <https://doi.org/10.1097/PSY.0000000000001038>.
50. Gabriele E, Spooner R, Brewer R, Murphy J. Dissociations between self-reported interoceptive accuracy and attention: evidence from the Interoceptive Attention Scale. *Biol Psychol*. 2022;168:108243. <https://doi.org/10.1016/j.biopsycho.2021.108243>.
51. Murphy J, Brewer R, Plans D, Khalsa SS, Catmur C, Bird G. Testing the independence of self-reported interoceptive accuracy and attention. *Quarterly J Exp Psychol*. 2020;73:115–133. <https://doi.org/10.1177/1747021819879826>.
52. Heim N, Bobou M, Tanzer M, Jenkinson PM, Steinert C, Fotopoulou A. Psychological interventions for interoception in mental health disorders: a systematic review of randomized-controlled trials. *Psychiatry Clin Neurosci*. 2023;77:530–540. <https://doi.org/10.1111/pcn.13576>.
53. Eggart M, Todd J, Valdés-Stauber J. Validation of the multi-dimensional assessment of interoceptive awareness (MAIA-2) questionnaire in hospitalized patients with major depressive disorder. *PLoS One*. 2021;16:e0253913. <https://doi.org/10.1371/journal.pone.0253913>.
54. Worthington RL, Whittaker TA. Scale development research: a content analysis and recommendations for best practices. *Couns Psychol*. 2006;34:806–838.
55. Peterson C. Accommodation, prediction and replication: model selection in scale construction. *Synthese*. 2019;196:4329–4350. <https://doi.org/10.1007/s11229-017-1660-0>.