



**HAL**  
open science

# A new Speech, Spatial, and Qualities of Hearing Scale Short-Form: Factor, Cluster and Comparative Analyses

Annie Moulin, Judith Vergne, Stéphane Gallégo, Christophe Micheyl

## ► To cite this version:

Annie Moulin, Judith Vergne, Stéphane Gallégo, Christophe Micheyl. A new Speech, Spatial, and Qualities of Hearing Scale Short-Form: Factor, Cluster and Comparative Analyses. *Ear and Hearing*, 2019, 40 (4), pp.938-950. 10.1097/AUD.0000000000000675 . hal-02367082v1

**HAL Id: hal-02367082**

**<https://hal.science/hal-02367082v1>**

Submitted on 17 Nov 2019 (v1), last revised 27 Aug 2023 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**A new Speech, Spatial, and Qualities of Hearing Scale (SSQ) Short-Form:  
Factor, Cluster, and Comparative Analyses**

Annie Moulin<sup>1,2\*</sup>, Judith Vergne<sup>1,2</sup>, Stéphane Gallego<sup>3,4,5</sup>, Christophe Micheyl<sup>6,7,8</sup>

<sup>1</sup>. INSERM U1028, CNRS UMR 5292, Lyon Neuroscience Research Center, Brain Dynamics and Cognition Team, F-69000 Lyon, France.

<sup>2</sup>. Université C. Bernard, Lyon 1, Lyon Neuroscience Research Center, Brain Dynamics and Cognition Team, University of Lyon, F-69 000 Lyon, France.

<sup>3</sup>. Institut des sciences et technologies de réadaptation (ISTR), University of Lyon, Lyon, France

<sup>4</sup>. Audition Conseil, France

<sup>5</sup>. Laboratory of Integrative and Adaptive Neurosciences (LNIA) UMR 7260, Aix-Marseille University-CNRS, Marseille, France.

<sup>6</sup>. Starkey Hearing Technologies, Créteil, France

<sup>7</sup>. Starkey Hearing Research Center, Berkeley, CA, USA

<sup>8</sup>. INSERM U1028, CNRS UMR 5292, Cognition and Auditory Perception Team (CAP), Lyon Neuroscience Research Center, University of Lyon, F-69000 Lyon, France.

\*Corresponding author: [annie.moulin@cnrs.fr](mailto:annie.moulin@cnrs.fr)

**This is not the final published version.**

**This is a pre-print, pre-refereeing version.**

The paper is published in Ear and hearing, ISSN: 0196-0202, ESN: 1538-4667

**Full reference of the final version:**

Moulin Annie, Vergne Judith, Gallego Stéphane, Micheyl Christophe (2018) A new Speech, Spatial, and Qualities of Hearing Scale (SSQ) Short-Form: Factor, Cluster, and Comparative Analyses. Ear Hear. 2019 Jul/Aug;40(4):938-950. doi: 10.1097/AUD.0000000000000675.

<https://insights.ovid.com/crossref?an=00003446-201907000-00016>

**A new Speech, Spatial, and Qualities of Hearing Scale (SSQ) Short-Form:  
Factor, Cluster, and Comparative Analyses**

Annie Moulin<sup>1,2\*</sup>, Judith Vergne<sup>1,2</sup>, Stéphane Gallego<sup>3,4,5</sup>, Christophe Micheyl<sup>6,7,8</sup>

<sup>1</sup>. INSERM U1028, CNRS UMR 5292, Lyon Neuroscience Research Center, Brain Dynamics and Cognition Team, F-69000 Lyon, France.

<sup>2</sup>. Université C. Bernard, Lyon 1, Lyon Neuroscience Research Center, Brain Dynamics and Cognition Team, University of Lyon, F-69 000 Lyon, France.

<sup>3</sup>. Institut des sciences et technologies de réadaptation (ISTR), University of Lyon, Lyon, France

<sup>4</sup>. Audition Conseil, France

<sup>5</sup>. Laboratory of Integrative and Adaptive Neurosciences (LNIA) UMR 7260, Aix-Marseille University-CNRS, Marseille, France.

<sup>6</sup>. Starkey Hearing Technologies, Créteil, France

<sup>7</sup>. Starkey Hearing Research Center, Berkeley, CA, USA

<sup>8</sup>. INSERM U1028, CNRS UMR 5292, Cognition and Auditory Perception Team (CAP), Lyon Neuroscience Research Center, University of Lyon, F-69000 Lyon, France.

\*Corresponding author: [annie.moulin@cnr.fr](mailto:annie.moulin@cnr.fr)

Postal address : DYCOG - Dynamique cérébrale et Cognition, Centre de Recherche en Neurosciences de Lyon, CRNL, Inserm U1028 - CNRS UMR5292, CH Le Vinatier, Bâtiment 452, 95 Bd Pinel, 69675 BRON Cedex France.

**Conflicts of interest and sources of funding**

Author CM is supported by Starkey Hearing Technologies, a private entity and manufacturer of hearing technology. Author SG is supported by Audition Conseil, a private company and group of audiology clinics. Other than through funding of these two coauthors' salaries, the sponsors for this study had no involvement in the design of the study, the data analysis, or the writing of the manuscript.

This work was supported in part by the "Fondation de l'Avenir" and "Visaudio" (research program ET4-738-VI4-001); by the LABEX CELYA (ANR-11-LABX-0060) of Université de Lyon, France and the LABEX CORTEX (ANR-11-LABX-0042) of Université de Lyon, within the program "Investissements d'Avenir" (ANR-11-IDEX-0007) operated by the French National Research Agency (ANR).

**A new Speech, Spatial, and Qualities of Hearing Scale (SSQ) Short-Form:  
Factor, Cluster, and Comparative Analyses**

Annie Moulin, Judith Vergne, Stéphane Gallego, Christophe Michey

**Acronyms and abbreviations**

**SSQ5:** 5-item SSQ short-form

**SSQ12:** 12-item SSQ short-form

**SSQ15:** 15-item SSQ short-form

**15iSSQ:** new 15-item SSQ short-form

**15iSSQs:** new 15-item SSQ short-form, obtained as a subset of the SSQ

**SSQ:** Speech Spatial and Qualities of hearing scale

**Speech:** Speech subscale of the SSQ

**Spatial:** Spatial-hearing subscale of the SSQ

**Qualities:** Qualities of hearing subscale of the SSQ

**HI:** Hearing-impaired

**HIHA:** Hearing-impaired, hearing-aid wearers

**HIN:** Hearing-impaired, non-hearing-aid wearers

**NH:** Normal-hearing subjects

**PTA:** Pure-tone threshold average (0.5, 1, 2 and 4 kHz)

**SD:** standard deviation

**NA:** Non-applicable answers

**F-ANOVA:** Friedman's analysis of variance

**R-ANOVA:** repeated-measures analysis of variance

## **ABSTRACT**

**Objectives:** (a) To build a 15-item short-form of the Speech Spatial and Qualities of Hearing Scale (SSQ) that maintains the three-factor structure of the full form, using a data-driven approach consistent with internationally recognized procedures for short-form building. (b) To validate the new short-form on an independent sample. (c) To perform an in-depth, comparative analysis of all existing, full and short SSQ forms.

**Design:** Data from a previous study involving 98 normal-hearing (NH) individuals and 196 hearing-impaired, non-hearing-aid wearers (HIN), along with results from several other published SSQ studies were used for developing the short-form. Data from a new and independent sample of 35 NH and 88 HI hearing-aid wearers (HIHA) were used to validate the new short-form. Factor and hierarchical cluster analyses were used to check the factor structure and internal consistency of the new short-form. In addition, the new short-form was compared to all other SSQ forms, including the full SSQ, the German SSQ15, the SSQ12, and the SSQ5. Construct validity was further assessed by testing statistical relationships between scores and audiometric factors, including pure-tone threshold averages (PTAs) and left/right PTA asymmetry. Receiver-operating characteristic (ROC) analyses were used to compare the ability of different SSQ forms to discriminate between NH and HI (HIN and HIHA) individuals.

**Results:** Compared all other SSQ forms, including the full SSQ, the new short-form showed negligible cross-loading across the three main subscales, and greater discriminatory power between NH and HI subjects (as indicated by a larger area under the ROC curve), as well as between the main subscales (especially Speech and Qualities). Moreover, the new, 5-item Spatial subscale showed increased sensitivity to left/right PTA asymmetry. Very good internal consistency and homogeneity, and high correlations with the SSQ were obtained for all short forms.

**Conclusion:** While maintaining the three-factor structure of the full SSQ, and exceeding the latter in terms of construct validity, and sensitivity to audiometric variables, the new 15-item SSQ, the 15iSSQ, affords a substantial reduction in the number of items and thus, in test time. Based on overall scores, Speech subscores, or Spatial subscores, but not Qualities subscores, the 15iSSQ appears to be more sensitive to differences in self-evaluated hearing abilities between NH and HI subjects than the full SSQ.

**Key words:** Self-report Measure; Short form building; Hearing Disability; Hearing Loss; Spatial Hearing; Speech; Speech Spatial and Qualities of Hearing Scale.

## **INTRODUCTION**

Self-report outcome measures have become an essential component in the evaluation of rehabilitation benefits for patients. In the field of audiology, several scales have been developed for the self-evaluation of various aspects of hearing, such as speech perception, binaural hearing, or hearing in challenging situations. Among these, the Speech Spatial and Qualities of Hearing scale (SSQ) designed by Gatehouse and Noble (2004) is a widely used self-report measure of hearing, and has already been translated into several languages (Dutch, Korean, German, and French).

In its full form, the SSQ includes 50 items. One advantage of this relatively large number of items is that it makes it possible to explore some very specific aspects of hearing via three main subscales (Speech, Spatial and Qualities) and 10 pragmatic subscales (Gatehouse & Akeroyd 2006). One disadvantage relates to the substantial amount of time and cognitive effort required for completion. Using a French version of the SSQ, Moulin et al. (2015) found that the self-assessed time to complete the scale in a hearing-impaired (HI) group varied from 10 minutes to 1 hour, with more than 25% of respondents reporting completion times above 25 minutes. This makes it difficult to use the SSQ in routine clinical use, or for swift assessments of hearing. Moreover, a tendency for the number of missing responses to increase over the last third of the scale was observed, with more missing responses for the most difficult-to-read items (Moulin et al., 2015). Indeed, with over a thousand words, the SSQ can be quite taxing for respondents.

In order to address this problem, short-forms of the SSQ have been developed. Demeester et al. (2012) created a short-form with five items (SSQ5) out of the Dutch version of the SSQ, specifically for screening purposes. Using the UK version of the SSQ, Noble et al. (2013) created a short-form with 12 items (SSQ12) based on a large multi-center dataset. Neither of these short forms include subscales. Validating a German version of the SSQ, Kiessling et al. (2011) proposed a short-form containing 15 items, with five items per subscale. The resulting short-form has the potential to retain the information present in the three main subscales of the SSQ. However, each of these short forms was constructed based on results from a single study, using a single language and, to our knowledge, none of these short-forms has been validated on a different sample of participants from the one used to build it. Lastly, to our knowledge, Kiessling et al.'s SSQ15 has only been published in German and has not been disseminated internationally.

In this context, the main goals of the present study were as follows. Firstly, to create a short-form of the SSQ using a data-driven approach, taking into account SSQ results from multiples studies, spanning different languages. Secondly, to validate the newly created short-form on a new, independent sample of participants. To this aim, we tried to follow the recommendations of Smith et al. (2000) for avoiding the “sins of short-form development”. Hence, the short form was built using a data-driven approach and strategies specific to short-form building (Stanton et al. 2002), using published data from SSQ studies involving four different languages (Dutch, English, French, and German). First, responses to items of the newly built short-form were extracted from existing data from a previous study in which the

full SSQ was administered to 196 hearing-impaired (HI) participants (non-wearers of hearing-aids: HIN) and 100 normal-hearing (NH) participants (Moulin et al., 2015). Subsequently, both the full SSQ and the newly built short-form were administered to a new sample of 88 HI subjects (all hearing-aid wearers: HIHA) and 35 NH subjects. In-depth analyses of all existing, full and short SSQ forms were performed. In particular, factor and hierarchical cluster analyses were used to check the factor structure and internal consistency of the new short-form. In addition, the new short-form was compared to all other SSQ forms, including the full SSQ, the German SSQ15, the SSQ12, and the SSQ5. Construct validity was further assessed by testing statistical relationships between scores and audiometric factors, including pure-tone threshold averages (PTAs) and left/right PTA asymmetry. Receiver-operating characteristic (ROC) analyses were used to compare the ability of different SSQ forms to discriminate between NH and HI (HIN and HIHA) individuals.

## MATERIAL AND METHODS

### 1. Short-form creation

To create the new short-form, we considered the following sources of information:

- Missing-response rates across items for the German SSQ (Kiessling et al. 2011), the English SSQ (Akeroyd et al. 2014), and the French SSQ (Moulin et al. 2015; Moulin & Richard 2016b).

- Results and normative values obtained on NH subjects for the Dutch SSQ (Demeester et al. 2012), the English SSQ (Banh et al. 2012) and the French SSQ (Moulin et al. 2015).

- Factor analyses of the English SSQ (Akeroyd et al. 2014) and French SSQ (Moulin et al. 2015).

One of the recommended methods suggested by Widaman et al. (2011) for building short-forms, is to identify a subset of items in the full-length questionnaire, which maintains the factorial integrity of the main scale. For the SSQ, factor analyses have shown that the main factors correspond to the three main subscales, Speech, Spatial and Qualities (Akeroyd et al. 2014; Moulin et al. 2015). The recommended minimum number of items per subscale ranges between three and five (Loewenthal 2001). We opted to create a short-form with five items per subscale (15 items in total), thus keeping a possibility of further scaling down the questionnaire to three or four items per subscale. To distinguish this new short-form from other SSQ short-forms, namely, Noble et al.'s (2013) SSQ12 and Kiessling et al.'s (2011) SSQ15, we hereafter refer to it as the "15-item SSQ" (15iSSQ).

The selection of items for inclusion in the short-form used the three categories of "item quality" proposed by Stanton et al. (2002): (1) *Internal items qualities*, which refer to scale properties such as inter-item and total-to-item correlations, internal-consistency measures (e.g. Cronbach's alpha), factor-analysis results, and subscale internal structure; (2) *External item qualities*, which refer to the way the items interact with external factors (construct validity), such as correlations between scores and hearing loss, ear asymmetry, age, or number of years of education. (3) *Judgmental item qualities*, which refers to the subjective evaluation by the scale user, of the relevance and adaptation to the patient's level of understanding. Because these indices can yield contradictory information regarding the "quality" of an item, item selection often reflects a compromise between different evaluation criteria. To deal with this, Stanton et al. (2002) proposed to code the quality indices for each item and to sort the items by their level of quality, starting from *external items quality* such as item-level validity, followed by *judgmental quality and face validity*, and finally, *internal-consistency item qualities*. As we wanted to eliminate, first, the most unreliable items, we adapted this approach by devising a system of "penalty points", awarded for each item according to the extent of its departure from several numeric criteria related to the quality categories. The scale for "penalty points" was 0, 1, and 2; a penalty of 2 meant that the corresponding item was eliminated. In a first phase, the following criteria were used:



1.1 Elimination of items with the greatest percentage of missing answers and NA responses (Missing >10% in English, French and German SSQ).

1.2 Elimination of items with less than 20% of variance explained by hearing loss, ear asymmetry, age, gender, and number of years of education (based on an analysis of French SSQ data from Moulin et al. 2016).

1.3 Elimination of all items with the lowest communalities (<50%) in factor analyses of the French SSQ and the English SSQ.

1.4 Elimination of items having a main-factor load lower than 0.60 and/or cross-loadings larger than 0.20 (based on factor analyses of the French SSQ and the English SSQ).

1.5 Elimination of items whose cutoff scores (defined as the mean – 2 SD), in young NH subjects, were below 3.5. This was based on 98 subjects for the French SQ, 103 subjects for the Dutch SSQ (Demeester et al. 2012), and 48 subjects for the English SSQ (Banh et al. 2012).

This first selection phase left a total of 23 items remaining, with 8 from Speech, 10 from Spatial and 5 from Qualities.

In a second phase, the criteria listed above were used more coarsely, in that the “penalty points” were summed together rather than considered separately; in addition; four other criteria were added:

2.1 Favoring items with scores not significantly predicted by unwanted factors, such as the number of years of education or gender.

2.2 Favoring, for the spatial scale, items showing a strong prediction by left/right asymmetry in pure-tone thresholds, and eliminating items predicted strongly by ear asymmetry for the other subscales.

2.3 Favoring, whenever possible, items already present in the SSQ12, the SSQ15, or the SSQ5.

2.4. Favoring items with high item-to-total correlations.

This second phase led to a 15-item short-form (#1.1, #1.4, #1.5, #1.6, #1.11, #2.2, #2.6, #2.7, #2.11, #2.17, #3.4, #3.5, #3.6, #3.8, #3.9) having one item in common with the SSQ5, five items in common with the SSQ12, and 7 items in common with the SSQ15. As an aside, the SSQ12 and SSQ15 have 5 items in common, while the SSQ5 and SSQ12 have 3 items in common (Appendix, Table 1). Importantly, the 15-item short-form thus created conforms to the three main subscales of the SSQ.

## **2. Short-form validation**

Two samples of subjects were used to validate the short-forms. Sample A contained 98 young NH subjects (aged 18 to 27 years; mean age = 20.8 years; SD = 2.2 years) and 196 HI subjects (aged 18 to 88 years; age = 53.4 years; SD = 17.1 years), who did not use hearing-aids

(HIN). For the latter, the four-frequency (0.5, 1, 2, and 4 kHz) PTA was equal to 27.2 dB HL (SD = 14.4 dB) on the best ear, and 41 dB HL (SD = 24 dB) on the worse ear, on average (Fig. 1). The HIN subjects formed a subset of an original sample of 216 HI subjects, which was used in two previous publications, along with the 98 NH subjects (Moulin & Richard 2016a; Moulin & Richard 2016b).

Sample B was a new sample, completely independent from the first one. It comprised 35 NH subjects aged 18 to 30 years (mean = 21.6 years, SD = 2.7 years), and 88 HI hearing-aid users (HIHA) aged 19 to 94 years (mean = 73.6 years, SD = 16.2 years). The NH subjects were recruited mostly from a pool of university undergraduates. They were free from any medical condition, history of otologic pathology, medication, and heavy noise exposure. They underwent pure-tone audiometry in half-octave steps from 125 to 8000 Hz using an Interacoustic AC40 audiometer in a sound-proof booth. The HIHA subjects were recruited from an audiology clinic. Their unaided pure-tone thresholds were measured at 250, 500 Hz, then at half-octave steps from 500 to 8000 Hz, using an Aurical Astera. Aided thresholds were measured using warble tones at these same frequencies, through speakers. Those patients had been wearing their hearing-aids for an average of 3.4 years (SD = 3.7 years) and an average of 10.4 hours per day (SD = 5.1 hours). The average PTA for the best ear was 43.9 dB HL (SD = 13.4 dB) and 53.8 dB HL (SD = 19.5 dB) for the worse ear (Fig. 1).

Subjects from sample B were administered the validated full French language version SSQ (Moulin et al., 2015) and the new short-form version at two different time points, separated by at least two hours and up to a few days, in randomized testing order; the short-form was presented first for half of the subjects.

The research was conducted in agreement with the World Medical Association declaration of Helsinki and in agreement with the French law pertaining to biomedical research (Agreement number A-11-385, ‘CPP Sud-Est IV’).

### **3. Data and statistical analyses**

#### ***3.1. Score calculations***

PTA was computed as the average hearing threshold (in dB HL) across 0.5, 1, 2, and 4 kHz. Ear asymmetry was calculated as the difference in PTAs (in dB) between the left and right ears. SSQ scores were calculated over the full scale and over the different short-forms that were available in the literature: the SSQ5 (#1.8, #2.3 and #2.9, #3.9 and #3.14; its weighted version (SSQ5w) (Demeester et al. 2012)), the SSQ12 (#1.1, #1.4 and #1.10 to #1.12; #2.6, #2.9 and #2.13; #3.2, #3.7, #3.9 and #3.14; (Noble et al. 2013) and the SSQ15 (#1.4, #1.5, #1.7, #1.9, #1.10; #2.5 to #2.7, #2.9 and #2.12; #3.3, #3.4) (Kiessling et al. 2011). Scores corresponding to the new 15-item short-form, 15iSSQ, (#1.1, #1.4, #1.5, #1.6, #1.11, #2.2, #2.6, #2.7, #2.11, #2.17, #3.4, #3.5, #3.6, #3.8, #3.9) were also calculated.

Scores were computed individually for every NH and HI subjects in samples A and B. For sample B, two scores were computed for the 15iSSQ: one score was obtained by scoring a

subset of items from the SSQ questionnaire, hereafter referred to as the 15iSSQs, for ‘15-item SSQ subset’; the second score was obtained by scoring the actual 15-item short-form, which was administered separately to the subjects rather than as a part of the full SSQ.

For the SSQ15 and the 15iSSQ, both of which were designed to maintain the three-subscale structure of the SSQ (Speech, Spatial *and* Qualities), three “differential” sub-scores were calculated by subtracting sub-scores for the three subscales, pairwise (Qualities – Speech; Qualities – Spatial and Speech – Spatial).

Similarly to Demeester et al. (2012), we calculated the SSQ-disability cutoffs, defined as the mean scores obtained in normal hearing population, minus 2 standard deviations, for the total scores obtained from SSQ and all short-forms, and the Speech, Spatial and Qualities subscales.

### **3.2 Internal-structure analysis**

Data from the various short-forms were analyzed with common factor analysis, using the same parameters as for the full SSQ (see Moulin et al. 2015). Briefly, these analyses were performed using the R package “Psych”, with the following parameters: correlation matrix of 15x15 (SSQ15i), maximum likelihood method for factors extraction, parallel analysis for number of factors extraction, and oblique (oblimin) factor rotation.

Secondly, to determine whether factors of the short-forms reflect the same structure as the full scale, hierarchical cluster analyses (Revelle’s ICLUST algorithm, (Revelle 1978; Revelle 1979) were performed. Revelle’s method is designed specifically to visualize questionnaire scales and subscales. It relies on two indices: the alpha coefficient (mean split-half reliability), a measure of internal consistency, and Revelle’s beta coefficient (Zinbarg et al. 2005), a measure of factorial homogeneity; specifically, beta is the worst split-half reliability of a scale, and hence is lower than alpha. In short, the scale structure is built starting from two item clusters that are most similar to each other; an item is added to the initial two item clusters only if this addition improves the internal consistency (measured by alpha) and/or the factorial homogeneity (measured by beta) of the cluster. The results are shown using a hierarchical tree diagram of clusters that displays the internal sub-structure of the scale, allowing the definition of homogeneous subscales. The tree diagram connects increasingly less similar items and/or clusters from left to right: the most similar items are combined first, and the most dissimilar items are added last. Alpha and beta coefficients are provided for each cluster, and correlations coefficients are given for each connection between clusters and/or items. An alpha above 0.8 and a beta above 0.7, with a difference between these two coefficients lower than 0.1, are indicative of good homogeneity and good internal consistency (Cooksey & Soutar 2006). A goodness-of-fit index was used to compare the quality of the structures. Specifically, the fits of the different models to the data were assessed using root-mean-square residuals (RMSRs). An RMSR lower than 0.05 indicates a good fit; RMSRs comprised between 0.05 and 0.08 indicate a fair fit (Cooksey & Soutar 2006; Fabrigar et al. 1999). This technique is complementary to the more classical approach of factor analysis and is less method-dependent (Cooksey & Soutar

2006). The ICLUST algorithm was applied to the whole sample. All statistical analyses were performed using the “Psych” package within the R statistical package version 3.1.0 (2014-04-10).

### ***3.3. Comparisons between different short-forms***

Short-form and subscale scores were compared to each other, and to SSQ scores. Since the data of the NH participants deviated from normality (Shapiro-Wilk test,  $p < 0.001$ , kurtosis between 2 and 4 for some variables), these data were compared using non-parametric tests: Wilcoxon’s sign test; Friedman’s analysis of variance (F-ANOVA); Spearman’s correlation coefficient. The data of the HI subjects, and those used for whole-group analyses (including both NH and HI subjects), met the normality assumptions ( $|\text{skewness}| < 2.6$ ;  $|\text{kurtosis}| < 2.1$ ). Accordingly, for these data, parametric tests were used: analysis of variance for repeated measures (R-ANOVA); paired t-tests; Pearson’s correlation coefficients. For both NH and HI groups, internal validity of each short-form was assessed using Cronbach’s alpha coefficient. Correlations between the SSQ and the different short-forms were calculated. To correct for the hyper-inflation of those correlation coefficients due to the presence of common items in the SSQ and the short-forms, Girard and Christensen’s correction for overlapping error variance was used (Girard & Christensen 2008). In order to obtain reciprocal equations allowing to transform the SSQ scores to a short-form score and vice versa, taking into account measurement error of both scores, orthogonal distance regressions were performed on the total population.

Multiple-regression analysis of the SSQ, the short-forms, and the subscale scores of the HIN and HIHA subjects were performed using three predictors: Better Ear PTA (in dB HL), Ear Asymmetry (in dB), and Age (in years). The normality of the residuals was checked using the Shapiro-Wilk test, and the assumptions of non-multicollinearity were checked using Durbin-Watson test and VIF statistics. Correlation coefficients were compared using Revelle’s two-tailed test for correlated coefficients (R psych package).

### ***3.4. ROC analysis***

Receiver-operating-characteristic (ROC) curves were calculated for HI and NH subjects, for the SSQ, each of the short-forms, and each of the three subscales, using Robin et al. (2011) R package, pROC. The ROC curves were compared using Venkatraman’s test for paired ROCs. The areas under the ROC curve (AUC) and partial AUC across regions of high sensitivity (90 to 100%) were compared using Robin et al.’s bootstrap test, based on the percentile bootstrap method, using 10000 replicates. The Z statistic and two-tailed p value associated with Robin et al.’s bootstrap are given. As the SSQ is designed to assess a difference across two conditions (e.g. HI versus NH, or hearing-aid versus no hearing-aid), the region of high sensitivity appeared to be most relevant.

## **RESULTS**

### **1. Factor and cluster analysis of the short-forms**

To check the underlying factor structure of the various short-forms and their similarity with the SSQ, common factor analysis with maximum likelihood factor extraction was performed on all of the SSQ short-forms except for the SSQ5 and the SSQ12. This was done separately for the 196 HIN and the 88 HIHA subjects.

#### ***1.1. Factor analysis***

All the results from the factor analyses are in table 1. The Kaiser-Meyer-Olkin (KMO) indexes of Sampling adequacy were above 0.84 (which corresponds to “very good” (Field et al. 2012; Fabrigar et al. 1999), with values for individual items all higher than 0.71. Bartlett’s tests of sphericity were all highly significant. The analysis showed systematically three factors with eigenvalues greater than or equal to 1 (8.8, 1.5 and 1.0), with cumulative variance explained between 69% and 75%. The three-factor structure was systematically confirmed by Catell’s scree test, parallel analysis, and Velicer’s MAP criterion. The mean communalities ranged from 0.69 to 0.75. The RMSR ranged between 0.03 and 0.04. Hence, a three-factor extraction was always chosen for the final analysis, and an oblique rotation (oblimin) was applied, as the scores of all items were intercorrelated. The three rotated factors explained at least 21% of the variance, and each factor loaded primarily on items corresponding to the main subscales of the SSQ. The three rotated factors correlated with each other, with correlation coefficients ranging from 0.44 to 0.67. The different indices showed that this three factor solution was adequate, with a Tucker-Lewis index above 0.9 (one exception at 0.87), a CFI (comparative fit index) above 0.92, and root mean square error of approximation (RMSEA) between 0.08 and 0.13.

For the 15iSSQ, the items loadings on each factor showed clear separation between each factor, with no consistent cross-loading across the 3 factor analysis. A small degree of cross-loading was obtained for #1.5 especially for the 15iSSQs on the 88 HIHA, and for #2.2 but only for 15iSSQ (Fig. 2).

For the SSQ15, the separation between the different loads is less clear: for both samples, item #3.4 showed low communality (0.56) and cross-loading on factor 3 and factor 2 and both #3.3 and #2.12 showed a minor degree of cross-loadings on factors 2 and 3. Items #1.5 showed some cross-loadings on factors 1 and 2 for the 88 HIHA (Fig. 2). Hence, in the two different populations, items #3.3, #3.4 and #2.12 underperformed.

#### ***1.2. Cluster analysis***

Hierarchical cluster analysis (ICLUST) corroborated the three-factor structure, with three distinct clusters corresponding the three subscales (Speech, Spatial and Qualities), for both the 15iSSQ and the SSQ15 forms.

For the 15iSSQ (Fig. 3), Alpha and beta coefficients were higher than 0.82, which is above their usual criterion values (0.8 for alpha and 0.7 for beta), and the difference between these coefficients was small (<0.1), indicating high homogeneity and consistency within each cluster (Cooksey & Soutar 2006). The goodness-of-fit measure for this three-subscale solution

was 0.98, with an RMSR  $< 0.05$ , which corresponds to an excellent fit. The alpha coefficient above 0.91 for the 3 main sub-clusters showed their high reliability.

The spatial items are represented in a close knit cluster (C8), with  $\alpha=0.94$ , in which subcluster C6 can be identified and represents of the items pertaining to localization (locate vehicle, dog, person).

Speech items show a close association between items 1.4, 1.11 and 1.6, as a subcluster (C5), all pertaining to auditory perception in conditions of several talkers. Items 1.1 and 1.5, concerning talking with one person with one source of noise, are associated in subcluster C9.

Within the quality items, the two items about voice pitch and familiar music are associated in cluster C3, whereas the other 3 items, about clarity and naturalness of sounds are associated in subcluster C11.

For the SSQ15, the hierarchical cluster analysis showed a three-factor structure, with each factor corresponding to the expected subscale, with an acceptable fit (0.97, RMSR = 0.05), a minimum beta coefficient at 0.77, and a minimum alpha coefficient at 0.81. For the 196 HIN subjects, the subcluster “quality” showed a lower reliability ( $\alpha= 0.89$ ,  $\beta=0.77$ ), probably due to a poorly functioning item (#3.3: Music and voice as separate items), whose addition to the cluster led to a decrease of 0.1 in the beta coefficient, with a difference between alpha and beta coefficient of 0.12 for this factor, slightly higher than the expected  $<0.1$  (Appendix, Figure 1).

## **2. Comparison between SSQ and short-forms scores**

### ***2.1. Missing responses***

The number of missing responses was analysed per item for the HIHA subjects. It ranged from 0 to 14.8%, with an average of 3.1% (SD = 3.5) across the 49 items of the SSQ. Those percentages were significantly correlated with the percentages obtained in the HIN group ( $r = 0.67$ ,  $p < 0.0001$ ), with the same items (i.e., items 2.14 to 2.16 and 3.16) yielding the greatest number of missing responses for these two groups. The 15iSSQ gave an average of 2% (SD = 1.4) and 2% (SD = 2%) for the 15iSSQs, with a significant correlation between the two ( $r = 0.72$ ,  $p < 0.003$ ).

### ***2.2. Global scores***

F-ANOVA showed highly significant differences between the different scores (full version and short-forms) for the NH subjects ( $\text{Chi}^2 = 58.3$ ,  $\text{df} = 7$ ,  $p < 0.00001$ ) and for the HIHA subjects ( $\text{Chi}^2 = 68.4$ ,  $\text{df} = 7$ ,  $p < 0.00001$ ) (Fig. 4)

For NH subjects, the newly developed short-forms gave significantly greater scores than the full scale, with  $W = 163$ ,  $p < 0.02$  for the 15iSSQ, and  $W = 113$ ,  $p < 0.0006$  for the 15iSSQs. SSQ12 scores were significantly lower than the SSQ scores ( $W=563$ ,  $p < 0.0001$ ). No significant differences were obtained for the SSQ15s or the SSQ5s.

For HIHA subjects, all of the short-forms except SSQ5s gave significantly lower scores than the full-form, with  $W = 2517$ ,  $p < 0.03$  for the 15iSSQ,  $W = 2985$ ,  $p < 0.0001$  for the SSQ15s, and  $W = 3698$ ,  $p < 0.00001$  for the SSQ12s.

The 15iSSQ gave greater cut-off scores than the SSQ in both NH samples (6.13 vs 6.00 for 35 NH), the SSQ15 (5.96) and the SSQ12 (5.31).

### **2.3. Subscale scores**

For NH subjects, no statistically significant difference was obtained between the full form and the different short-forms for the Speech ( $\text{Chi}^2 = 11.6$ ,  $\text{df} = 5$ ,  $p < 0.05$ ) and Spatial subscales ( $\text{Chi}^2 = 6.3$ ,  $\text{df} = 5$ ,  $p > 0.20$ ). For HIN subjects, no statistically significant difference was obtained for the Spatial subscale (F-ANOVA,  $\text{Chi}^2 = 10$ ,  $\text{df} = 5$ ,  $p = 0.07$ ).

However, for the Speech subscale and the HIN subjects, all the different short-forms gave significantly lower scores than the full subscale (F-ANOVA  $\text{chi}^2 = 81.5$ ,  $\text{df} = 5$ ,  $p < 0.00001$ ), with  $W = 2728$ ,  $p < 0.0006$  for the 15iSSQ, and  $W = 3606$ ,  $p < 0.00001$  for the SSQ15s.

For the Qualities subscale, highly significant differences between full form and the different short-forms were observed for NH subjects and for HIHA subjects (F-ANOVA,  $\text{Chi}^2 = 68$ ,  $\text{df} = 5$ ,  $p < 0.00001$  for NH and  $\text{Chi}^2 = 73$ ,  $\text{df} = 5$ ,  $p < 0.00001$  for HIHA), albeit in opposite directions: for NH subjects, all of the different short-forms yielded significantly greater Qualities-subscale scores than the full form, with  $W = 4$ ,  $p < 0.00001$  for the 15iSSQ, and  $W = 56$ ,  $p < 0.0001$  for the SSQ15s; for HIHA subjects, the short-forms yielded significantly lower Qualities sub-scores than the full form, with  $W = 1071$ ,  $p < 0.0006$  for the 15iSSQ, and  $W = 741$ ,  $p < 0.00001$  for the SSQ15s.

Although the mean scores of the 15iSSQ Speech and Spatial subscales were not lower than the SSQ sub-scores, the cut-off scores of the 15iSSQ were lower (6.07 versus 6.15 for Speech and 4.13 versus 4.77 for Spatial) than the SSQ Speech and Spatial sub-scores, due to larger standard deviations. The SSQ15 gave the lowest cut-off scores (5.9 and 3.85 for Speech and Spatial respectively). However, the Qualities cut-off scores were all greater for all short-forms (15iSSQ, SSQ15) than for the SSQ (Appendix, table 2).

### **2.4. Internal validity and correlations between the different SSQ forms.**

Cronbach's alpha was higher than 0.83 for SSQ5, 0.93 for the SSQ12, and 0.94 for the SSQ15 and 15iSSQ. Good consistency was also shown within each subscale, with Cronbach's alphas higher than 0.92 for the Speech and Spatial subscales, and comprised between 0.85 and 0.91 for the Quality subscale (Appendix, Table 3a). Item-to-total correlations were above 0.67 for the 196 HIN, and above 0.61 for the HIHA (Appendix, Table 3b).

The different short-forms correlated very highly with the SSQ, with the lowest correlation coefficients obtained for the SSQ5 ( $r = 0.88$ , corrected Pearson) and coefficients between 0.94 and 0.95 for the other short-forms. The scores for the three subscales correlated highly with the corresponding SSQ subscale scores ( $r$  above 0.91 for the Speech and Spatial subscales, and  $r$  between 0.78 and 0.85 for the Quality subscale) (Appendix, Table 4a). Similar results were obtained for the non-aided population (Appendix, Table 4b).

As the relationship between the SSQ and the SSQ12 (and to a lesser degree the SSQ15, see Moulin & Richard 2016a) follows a power function rather than a linear one (Noble et al. 2013), we fitted the data for each of the short-forms tested using power functions, according to Noble's formulae:

$$\text{SSQ Short forms} = 10 (\text{SSQ}/10)^b$$

$$\text{SSQ} = 10(\text{SSQ Short forms}/10)^c$$

Non-linear relationships between scores for the SSQ and the SSQ12, on the one hand, and the SSQ and the SSQ15, on the other hand, were confirmed ( $b = 1.215$  [1.180, 1.250] and  $b = 1.069$  [1.041, 1.097], respectively). For the new short-form, the relationship was linear, as shown by a 95% confidence interval for  $b$  overlapping 1 (Appendix, Table 5). Similar analyses performed on the larger sample of hearing-aid non-wearers (NH and HIN) yielded similar results, with a linear relationship between scores of the SSQ and of the new short-form.

Reciprocal equations obtained by orthogonal distance regression between SSQ and the new short form, using all subjects (i.e., 417) are:

$$15i\text{SSQs} = 0.931 * \text{SSQ} + 0.492$$

$$\text{SSQ} = 1.074 * 15i\text{SSQs} - 0.529$$

### **3. Construct validity**

#### ***3.1. Influence of hearing-loss***

The scores of all the different forms decreased highly significantly with increasing hearing-loss in the better ear, with a slope ranging from 0.56 (SSQ5) to 0.81 (15iSSQ) (0.65 for the SSQ) scale points per 10-dB HL. These slopes did not differ significantly from each other. However, the different short forms differed from each other in the strength of the correlation between score and HL, with a significantly higher correlation coefficient for the newly developed short-form than for the full form ( $r = 0.51$  versus  $r = 0.58$ ,  $t = 2.5$ ,  $p < 0.02$ ), and a significantly lower correlation for the SSQ5 than for the full form ( $r = 0.41$ ,  $t = 3.6$ ,  $p < 0.0005$ ).

When analysing subscale scores, correlations between scores and HL were not found to differ significantly across the different SSQ forms (full versus short).

Linear regressions performed on the scores of the NH and HIHA subjects showed a decrease in short-form scores with hearing loss, with a greater slope for the short-forms in general, except for the SSQ5.

$$\text{SSQ} = -0.065 * \text{HL} + 8.23 \quad (r = 0.51, 26\% \text{ variance explained})$$

$$15i\text{SSQ} = -0.081 * \text{HL} + 8.47 \quad (r = 0.58, 34\% \text{ variance explained})$$

All the details of the relationships between scores (for all the different forms and their subscales) and hearing loss are in the Appendix, Table 6.



### 3.2. Influence of patients' characteristics

This analysis sought to investigate predictors of the full- and short-form scores. To this aim, multiple-regression analyses were performed on the scores for each form, using the following explanatory variables: Gender, Age, Better-ear PTA, and (PTA) Asymmetry. The latter two audiometric variables were measured with, and without, hearing-aids, and both measurements were entered as potential predictors in the model. For these analyses, the data of the 88 HIHA subjects were used. Gender and Age were never found to be statistically significant predictors. The best models included hearing thresholds measured with hearing-aids, and PTA Asymmetry measured without hearing-aids. Therefore, multiple-regression results are presented using the better-ear PTA measured with hearing-aids, and PTA asymmetry measured without hearing-aids. The size of the effect ( $r^2$ ), and the beta coefficients that allow the comparison of the relative influence of each statistically significant predictor on the SSQ scores are summarized in Table 7 in the Appendix. The different short-forms yielded results similar to the full SSQ, with 32% to 39 % of variance explained by the two predictors: Better-ear PTA (beta ranging from -0.28 to -0.36) and PTA Asymmetry (beta ranging from -0.22 to -0.19) for the seven different SSQ forms. Ear asymmetry was not a significant predictor of SSQ12 scores, but was the only significant predictor of SSQ5 scores.

The three subscales of the short-forms gave similar results to the three subscales of the SSQ: Better-ear PTA and Asymmetry were significant predictors. For all forms, the greatest dependencies were between the Asymmetry predictor and scores of the Spatial subscale, and between Better-ear PTA and scores of the Quality subscale. The differential scores between the main subscales (Qualities – Spatial; Qualities – Speech) did not show any significant dependency on the two predictors (Appendix, Table 7a).

Similar analyses were performed on the 196 non-hearing-aid wearers. Results showed that the different short-forms yielded results similar to the SSQ, with 37% to 38 % of variance explained by the two predictors: Better Ear PTA (beta from -0.54 to -0.58) and Ear Asymmetry (Beta = -0.29 to -0.32) for the five short-forms. The three subscales of the SSQ15 and 15iSSQ gave results similar to the three subscales of the SSQ, with the Better-ear PTA and PTA Asymmetry as significant predictors. The relative importance of PTA Asymmetry for the Spatial subscale was higher for the 15iSSQ, with an even greater beta coefficient (-0.47) than for hearing loss (-0.44). The differential scores (Qualities – Spatial) had PTA Asymmetry as the main predictor, with 21% of variance explained for the 15iSSQ, which is substantially greater than the same score using the SSQ (11%) or the SSQ15 (10%) (Appendix, Table 7b).

### 3.3. ROC analysis

For the SSQ, the area under the ROC curve (AUC) was equal to 77.3 for the HIN group, and to 80.1 for the HIHA group. For the HIHA group, significantly larger AUCs (Fig 5) were obtained for the 15iSSQ (AUC=84) and the 15iSSQs (AUC=83) than for the SSQ ( $Z = 1.75$ ,  $p < 0.08$ , and  $Z = 2.02$ ,  $p < 0.05$  for the 15iSSQ and 15iSSQs, respectively). By contrast, AUCs for the SSQ12 and the SSQ5 were not significantly larger than the SSQ AUC. For the HIN

group, differences in AUC between the SSQ and the different short-forms did not reach statistical significance (Appendix, table 8a, b).

Similar analyses were performed for the three subscales (Fig. 5). For the Speech subscale, the different short-forms all showed a larger AUC ( $> 86.4$ ) than the full-form AUC (84.2), with significant differences for the 15iSSQs ( $Z = 2.9$ ,  $p < 0.004$  for the HIHA group;  $Z = 2.4$ ,  $p < 0.02$  for the HIN group). For the Spatial subscale, no significant difference between the different short-forms and the full-form were obtained for the HIN group; however, for the HIHA group, the short-forms all produced a smaller AUC (71.9 for the 15iSSQ and 69.3 for the SSQ15) than the SSQ (72.1), with a significant difference for the 15iSSQs ( $Z = -2.2$ ,  $p < 0.03$ ). For the Qualities subscale, the AUC for the 15iSSQs ( $Z = -4.4$ ,  $p < 0.0001$ ) was significantly smaller than the AUC for the full-form for the HIN group, but not for HIHA group. (Appendix, Tables 8a and 8b).

AUCs for the spatial subscale were systematically smaller than AUCs for the other subscales, while the AUCs for the Speech subscale were systematically larger than AUCs for the other subscales: this pattern was seen for the SSQ, as well as for the different short-forms.

A similar analysis was performed for the high-sensitivity region (sensitivity above 90%). For the HIHA group, the newly developed short-forms showed a significantly and systematically larger partial-AUC than the SSQ ( $Z = 2.5$ ,  $p < 0.02$  for the 15iSSQ, and  $Z = 2.4$ ,  $p < 0.02$  for the 15iSSQs). No significant differences were obtained between the SSQ and the SSQ15, the SSQ12, or the SSQ5. No significant differences were obtained for the HIN group either.

## DISCUSSION

The main goals of this study were the creation of a new SSQ short-form using a rigorous data-driven approach, in agreement with the criteria suggested by Stanton et al. (2002) and the factorial-integrity check suggested by Widaman et al. (2011) and Smith et al. (2000). By using data from multiple studies, we tried to avoid the caveat of building a short-form specific to a single data set (Widaman et al. 2011). By using various types of criteria for selecting items, we tried to overcome the caveats of a single-criterion approach. For instance, by using only internal-consistency criteria, or by over-using such criteria, one can end up with too narrowly focused a short-form, which explores only some aspects of the full form (Stanton et al. 2002). Following Stanton et al. (2002), we considered three different categories of criteria: “judgmental” (such as the percentage of missing answers, the belonging of the items to another short form, or the readability), “external qualities” (percentage of variance explained by the five main predictors), and “internal qualities” (communalities and factor loadings of factor analysis obtained in two different studies; normative data obtained in NH subjects; inter-item and item-to total correlations). The internal structure of the short-form was cross-checked using factor and cluster analyses. The results revealed three clusters for the SSQ15 and the 15iSSQ, each corresponding to a main subscale of the SSQ (Speech, Spatial, and Qualities). Lastly, we validated the new 15iSSQ using two independent and different samples of participants, including both NH and HIHA, hence avoiding the last two of the nine “sins” of short form building described by Smith et al. (2000): not administering the short and full forms independently, and not using independent samples for validation.

The results provide evidence for the construct validity of the new 15-items short-form thus obtained, the 15iSSQ. In particular, factor and hierarchical cluster analyses confirmed that the 15iSSQ has the same three-factor structure as the full SSQ, with no consistent cross-loading. The results further indicate that the new 15iSSQ is superior to the other existing 15-item SSQ short-form, the SSQ15 (Kiessling et al. 2011). In particular, from a construct-validity perspective, factor and hierarchical cluster analyses indicated less cross-loading, and higher fits to the three-factor structure of the original SSQ, for the 15iSSQ than for the SSQ15.

Regarding external validity, consistent with the full SSQ, scores for the 15iSSQ decreased with increasing hearing loss, with a higher slope for the 15iSSQ (0.81 points per 10 dB HL) than for the full SSQ (0.65 points per 10 dB HL). This outcome can be explained by one of our criteria for item selection: the presence of relatively high scores and lower variability in NH subjects. Although SSQ scores have been initially interpreted as if young NH subjects should have perfect scores (e.g., House et al. 2010), in reality, actual SSQ scores for NH are often lower than 10 (Demeester et al. 2012; Moulin et al. 2015). Indeed, in the present study, the average score of NH respondents for the Spatial subscale was lower than 8 (see Fig. 4). The 15iSSQ scores were found to be greater than the full-SSQ scores for both NH samples, and smaller for the HIN and HIHA subjects, thus reinforcing the contrast between NH and HI subjects. In addition, the expected correlation between PTA and score was significantly higher for the 15iSSQ (21% of variance explained) than for the SSQ (11% of variance explained). Moreover, for both HIHA and HIN groups, the AUC, a measure of the difference between the

scores of the NH and HI groups was significantly larger for the 15iSSQ than for the SSQ, whereas AUCs obtained using the other short-forms were not (SSQ5, SSQ12 and SSQ15). One of the caveats of reducing the number of items in a questionnaire, is the potential increase of inter-subjects variability, the scores being less “smoothed out” by an averaging procedure across a great number of items than across a smaller number of items. This has potential detrimental consequences on the disability SSQ cut-off scores. This is not the case here, as for the total score, the cut-off scores obtained with the new short-form were consistently larger than the scores obtained with (1) the SSQ and (2) all the other short-forms.

A look at the three main subscale scores gives a more complex picture: the cut-off scores for the Speech and Spatial subscales tend to be lower for all short-forms than for the SSQ. For the qualities subscale, all short-form gave greater cut-off scores than the SSQ qualities. Indeed, the 15iSSQ (and to a lesser degree, the SSQ15) gave larger scores for the Qualities subscale in all samples, especially in NH. For the Spatial and Speech subscales, 15iSSQ scores were similar to SSQ scores for both NH samples, but were significantly lower than SSQ scores for HIN and HIHA. As a result, the contrast between Qualities and Speech sub-scores and, to a lesser degree, between Speech and Spatial sub-scores, was higher for HI with the 15iSSQ than with the SSQ. The lower Spatial sub-scores for the HI participants may be attributed to our choice of favoring items that were more sensitive to Ear Asymmetry. This can be observed in the measured dependency of the Spatial scores to Best Ear PTA and Ear Asymmetry, where both predictors contributed almost equally to the scores; by contrast, for the SSQ15 and the SSQ, the contribution of the Best Ear PTA was always substantially greater than the contribution of Ear Asymmetry. The statistical relationship between PTA asymmetry and the differential sub-score, Qualities – Spatial, was substantially stronger for the 15iSSQ (21% of variance explained) than for the SSQ15 (10% of variance explained) or the SSQ (11% of variance explained). Hence, although our selection of items for the new short-form was guided primarily by the imperative to maintain the three main subscales of the SSQ, two advantageous by-products of this selection are a better contrast between NH and HI, and a Spatial subscale that is more sensitive to Ear Asymmetry. This greater sensitivity to ear asymmetry can be highly advantageous in the analysis of self-reported hearing disabilities linked to ear asymmetry (Vannson et al. 2015), unilateral hearing loss (Olsen et al. 2012; Dwyer et al. 2014; Douglas et al. 2007), and the evaluation of the benefits of hearing rehabilitation strategies of those asymmetrical losses (Pai et al. 2012; Dumper et al. 2009).

It is worth noting that we found a linear relationship between scores obtained with the new short-form and scores for the full SSQ. This differs from the SSQ12 for which the relationship is nonlinear (Moulin & Richard 2016b; Noble et al. 2013) and, to a lesser degree, from the SSQ15, and makes it straightforward to infer scores for the full SSQ based on the new short-form scores, or *vice versa*. One limitation of this study stems from the fact that test-retest reliability of the new short-form was not verified. To the best of our knowledge, this shortcoming applies also to the other SSQ short-forms. However, it appears highly unlikely that the 15iSSQ is any less reliable than its existing 15- and 12-item counterparts, the SSQ15 and the SSQ12. First, the items in this new short-form come from the same set of items that was

used to create the SSQ15 and the SSQ12, namely the complete set of SSQ questions, which has shown a high degree of correlation between several administration modes and at different time points (Singh & Kathleen Pichora-Fuller 2010). Second, we obtained very good correlations between the 15iSSQs and the 15iSSQ ( $r=0.93$ ) and no statistically significant differences between 15iSSQs and 15iSSQ in the same samples (NH and HIHA), whether the full score or scores per subscales were considered. Although the delay between the two administrations was too low (less than a week) to really assess long-term reproducibility, and responses to the subset version (15iSSQs) were, to some degree, influenced by the other questions of the full SSQ, it does give a clear indication of the good test/retest reproducibility of the 15iSSQ.

Even if it were merely equivalent to the full SSQ for differentiating between NH and HI subjects, the 15iSSQ could be advantageously used for this purpose in research studies or clinical work, given its shorter length. In fact, the results showed that, in some respects, the short-form affords *better* discrimination between NH and HI subjects than the full SSQ, and a better sensitivity to ear asymmetry thanks to the Spatial subscale. One qualification to this conclusion stems from the finding of smaller AUCs for the 15iSSQ than for the SSQ, for one specific combinations of HI subgroups and subscales: the Qualities subscale with the HIN group. For this reason, we cannot recommend using the 15iSSQ in lieu of the full SSQ in all situations, including those in which longer length of the latter is not a major limitation.

### **Acknowledgements**

The authors wish to thank Jeremy Montagnat-Misson for technical help during the data collection.

## REFERENCES

- Akeroyd, M.A., Guy, F.H., Harrison, D.L., et al. (2014). A factor analysis of the SSQ (Speech, Spatial, and Qualities of Hearing Scale). *Int. J. Audiol.*, 53, 101–114.
- Banh, J., Singh, G., Pichora-Fuller, M.K. (2012). Age affects responses on the Speech, Spatial, and Qualities of Hearing Scale (SSQ) by adults with minimal audiometric loss. *J. Am. Acad. Audiol.*, 23, 81-91; quiz 139-140.
- Cooksey, R.W., Soutar, G.N. (2006). Coefficient Beta and Hierarchical Item Clustering An Analytical Procedure for Establishing and Displaying the Dimensionality and Homogeneity of Summated Scales. *Organ. Res. Methods*, 9, 78–98.
- Demeester, K., Topsakal, V., Hendrickx, J.-J., et al. (2012). Hearing disability measured by the speech, spatial, and qualities of hearing scale in clinically normal-hearing and hearing-impaired middle-aged persons, and disability screening by means of a reduced SSQ (the SSQ5). *Ear Hear.*, 33, 615–616.
- Douglas, S.A., Yeung, P., Daudia, A., et al. (2007). Spatial hearing disability after acoustic neuroma removal. *The Laryngoscope*, 117, 1648–1651.
- Dumper, J., Hodgetts, B., Liu, R., et al. (2009). Indications for bone-anchored hearing AIDS: a functional outcomes study. *J. Otolaryngol. - Head Neck Surg. J. Oto-Rhino-Laryngol. Chir. Cervico-Faciale*, 38, 96–105.
- Dwyer, N.Y., Firszt, J.B., Reeder, R.M. (2014). Effects of unilateral input and mode of hearing in the better ear: self-reported performance using the speech, spatial and qualities of hearing scale. *Ear Hear.*, 35, 126–136.
- Fabrigar, L.R., Wegener, D.T., MacCallum, R.C., et al. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychol. Methods*, 4, 272–299.
- Field, A., Miles, J., Field, Z. (2012). *Discovering Statistics Using R* 1 edition., London ; Thousand Oaks, Calif: SAGE Publications Ltd.
- Gatehouse, S., Akeroyd, M. (2006). Two-eared listening in dynamic situations: Audición con dos oídos en situaciones dinámicas. *Int. J. Audiol.*, 45, 120–124.
- Girard, T.A., Christensen, B.K. (2008). Clarifying problems and offering solutions for correlated error when assessing the validity of selected-subtest short forms. *Psychol. Assess.*, 20, 76–80.
- House, J.W., Kutz, J.W., Chung, J., et al. (2010). Bone-anchored hearing aid subjective benefit for unilateral deafness. *The Laryngoscope*, 120, 601–607.
- Kiessling, J., Grugel, L., Meister, I.G., et al. (2011). Übertragung der Fragebögen SADL, ECHO und SSQ ins Deutsche und deren Evaluation. German translations of questionnaires SADL, ECHO and SSQ and their evaluation. *Z Audiol.*, 6–16.
- Loewenthal, K.M. (2001). *An Introduction to Psychological Tests and Scales*, Psychology Press.

Moulin, A., Pauzie, A., Richard, C. (2015). Validation of a French translation of the speech, spatial, and qualities of hearing scale (SSQ) and comparison with other language versions. *Int. J. Audiol.*, 54, 889–898.

Moulin, A., Richard, C. (2016a). Sources of variability of speech, spatial, and qualities of hearing scale (SSQ) scores in normal-hearing and hearing-impaired populations. *Int. J. Audiol.*, 55, 101–109.

Moulin, A., Richard, C. (2016b). Validation of a French-Language Version of the Spatial Hearing Questionnaire, Cluster Analysis and Comparison with the Speech, Spatial, and Qualities of Hearing Scale. *Ear Hear.*, 37, 412–423.

Noble, W., Jensen, N.S., Naylor, G., et al. (2013). A short form of the Speech, Spatial and Qualities of Hearing scale suitable for clinical use: The SSQ12. *Int. J. Audiol.*, 52, 409–412.

Olsen, S.Ø., Hernvig, L.H., Nielsen, L.H. (2012). Self-reported hearing performance among subjects with unilateral sensorineural hearing loss. *Audiol. Med.*, 10, 83–92.

Pai, I., Kelleher, C., Nunn, T., et al. (2012). Outcome of bone-anchored hearing aids for single-sided deafness: A prospective study. *Acta Otolaryngol. (Stockh.)*, 132, 751–755.

Revelle, W. (1979). Hierarchical Cluster Analysis And The Internal Structure Of Tests. *Multivar. Behav. Res.*, 14, 57–74.

Revelle, W. (1978). ICLUST: A cluster analytic approach to exploratory and confirmatory scale construction. *Behav. Res. Methods Instrum.*, 10, 739–742.

Robin, X., Turck, N., Hainard, A., et al. (2011). pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*, 12, 77.

Singh, G., Kathleen Pichora-Fuller, M. (2010). Older adults' performance on the speech, spatial, and qualities of hearing scale (SSQ): Test-retest reliability and a comparison of interview and self-administration methods. *Int. J. Audiol.*, 49, 733–740.

Smith, G.T., McCarthy, D.M., Anderson, K.G. (2000). On the sins of short-form development. *Psychol. Assess.*, 12, 102–111.

Stanton, J.M., Sinar, E.F., Balzer, W.K., et al. (2002). Issues and Strategies for Reducing the Length of Self-Report Scales. *Pers. Psychol.*, 55, 167–194.

Vannson, N., James, C., Fraysse, B., et al. (2015). Quality of Life and Auditory Performance in Adults with Asymmetric Hearing Loss. *Audiol. Neurotol.*, 20, 38–43.

Widaman, K.F., Little, T.D., Preacher, K.J., et al. (2011). On creating and using short forms of scales in secondary research. In K. H. Trzesniewski, M. B. Donnellan, & R. E. Lucas, eds. *Secondary data analysis: An introduction for psychologists*. (pp. 39–61). Washington, DC, US: American Psychological Association.

Zinbarg, R.E., Revelle, W., Yovel, I., et al. (2005). Cronbach's  $\alpha$ , Revelle's  $\beta$ , and McDonald's  $\omega^2$ : their relations with each other and two alternative conceptualizations of reliability. *Psychometrika*, 70, 123–133.



FIGURES AND FIGURE LEGENDS

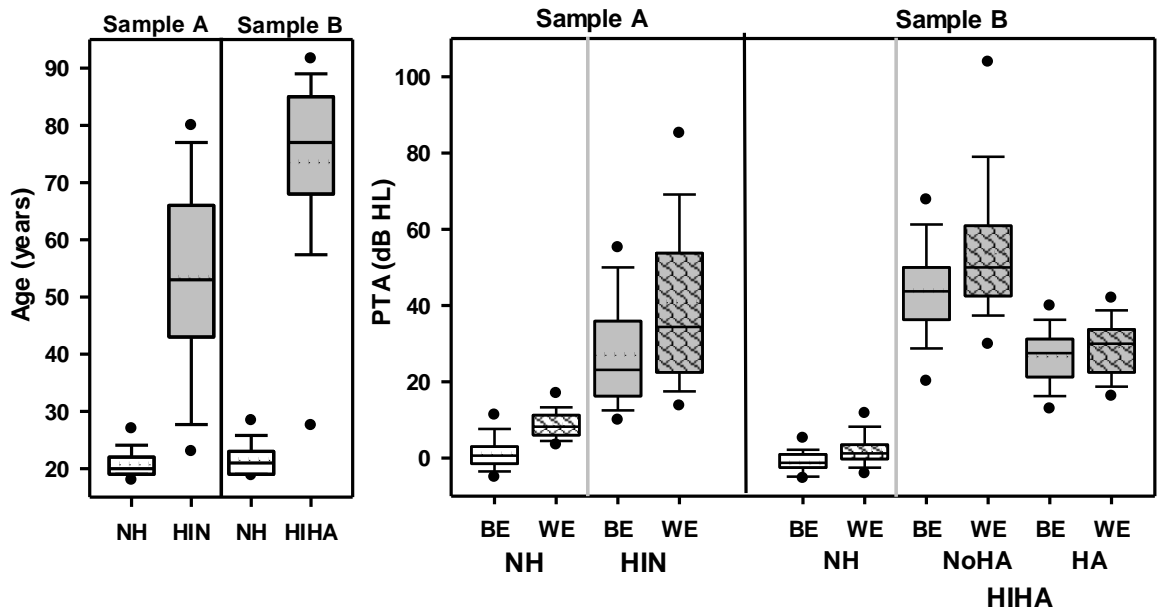


Figure 1

Statistical characteristics of the different subject samples. Left panel: age. Middle and right panels: best-ear (BE) and worst-ear (WE) PTA. For Sample B, all HI subjects were hearing-aid wearers, and PTAs measured with HAs and without HAs ('No HA') are reported.

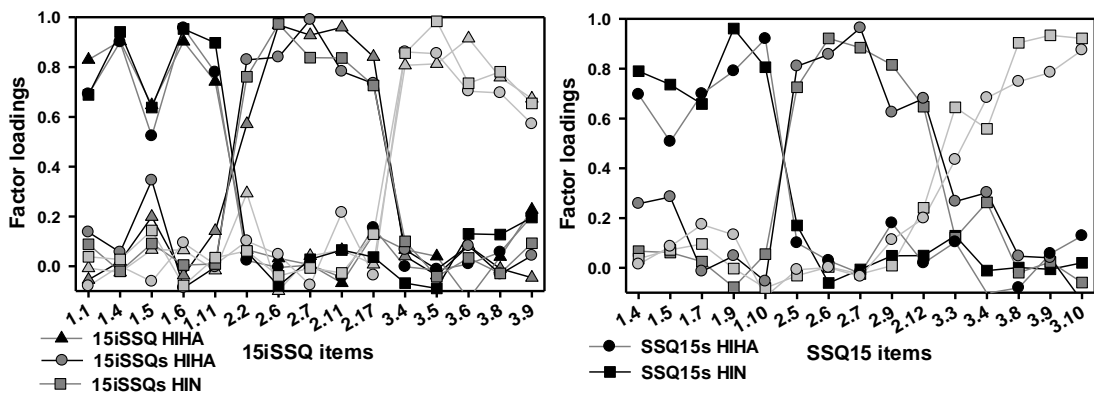


Figure 2

Factor loadings computed using factor analysis applied on the responses to 15-item subsets from the full SSQ corresponding to the 15iSSQ (left panel) or to the SSQ15 (right panel). The three factors are indicated by different colors (black: factor 1; dark-grey: factor 2; light-grey: factor 3). The different subject samples are shown using different symbols, as indicated on the figure. '15iSSQs' refers to data collected separately in 88 HIHA subjects using the 15iSSQ.

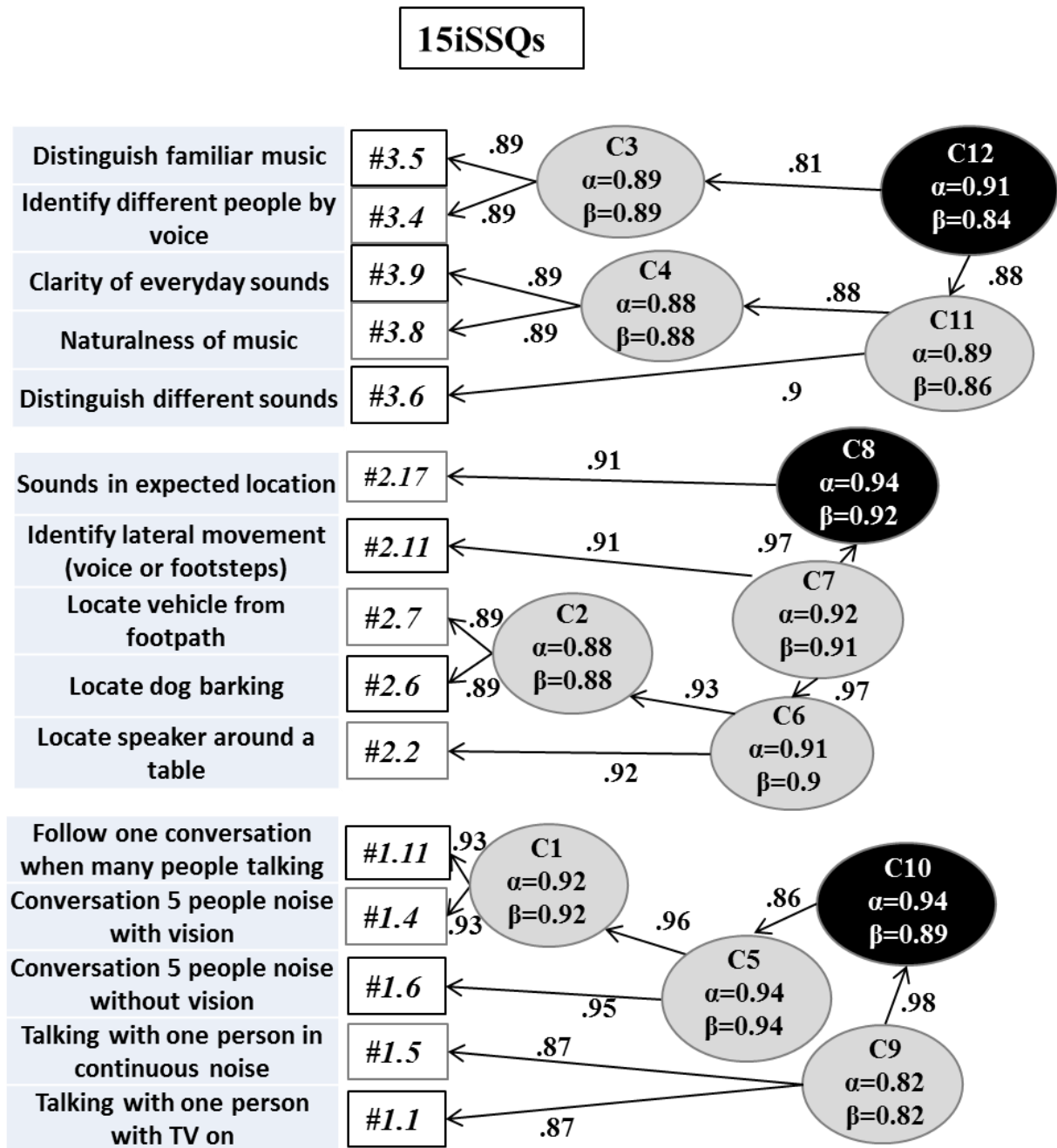


Figure 3

Tree diagram obtained using hierarchical cluster analysis (ICLUST) of the 15iSSQs data from 196 HI subjects. Shortened verbal descriptions (after Bahn et al., 2012) of the 15 items are listed on the left, grouped by subscale. The most similar items are combined first, and increasingly less similar clusters are represented from left to right. For each cluster, the alpha coefficient and Revelle's B (worst-split-half reliability) are provided. Three main clusters (in black: C8, C10 and C12), corresponding to a main subscale, are identified.

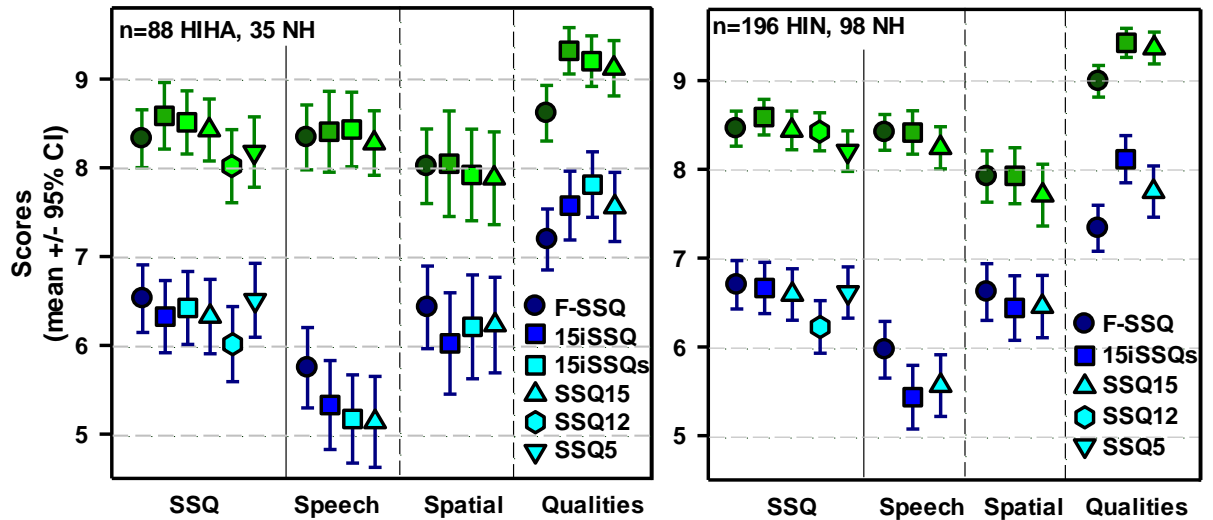


Figure 4

Left panel: full and short-form SSQ scores for the 88 HIHA subjects (blue) and 35 young NH subjects (green) composing Sample B. The short-form scores were either, obtained using the 15iSSQ or, computed by tallying scores across five, 12, or 15-item subsets of the (full) SSQ corresponding to the SSQ5, the SSQ12, the SSQ15, or the 15iSSQ; the latter subscores are referred to as ‘15iSSQs’, to distinguish them from scores obtained using the 15iSSQ. Right panel: full and short-form SSQ scores obtained using subsets of the (full) SSQ in the 196 HI subjects (blue) and 98 young NH subjects (green) composing Sample A. Error bars show 95% confidence intervals.

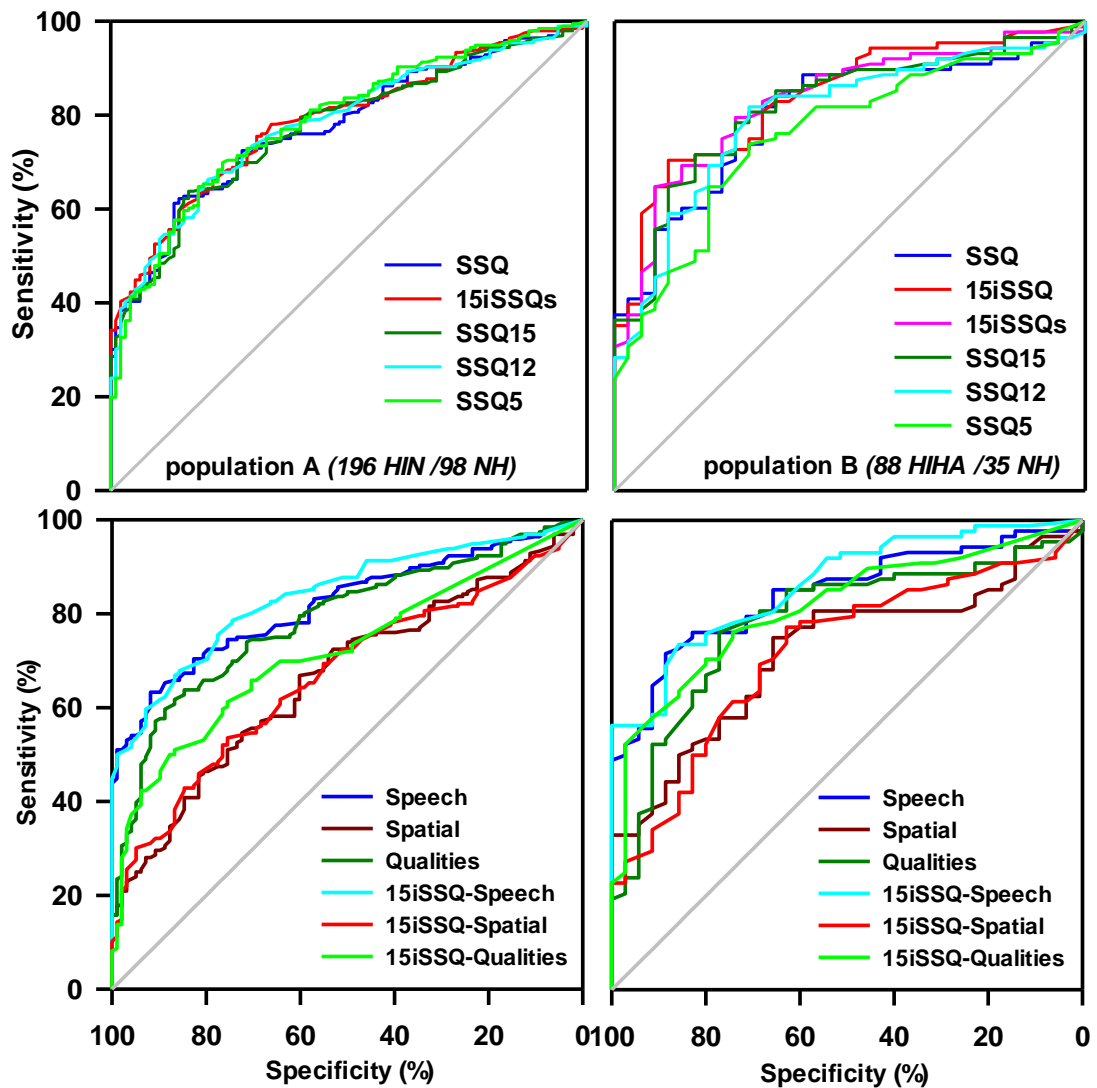


Figure 5

ROCs for each subscale (Speech, Spatial, Qualities). Left: ROCs computed based on Sample-A data. Right: ROCs computed based on Sample-B data. Top row: ROCs computed on scores for the (full) SSQ and for various SSQ short-forms. Bottom row: ROC curves computed on subscale scores for either the (full) SSQ or the 15iSSQ.

Table 1

Population analyzed	88 HIHA patients			196 HIN patients	
Form analyzed	15iSSQ	15iSSQs	SSQ15s	15iSSQs	SSQ15s
Kaiser-Meyer-Olkin (kmo) index	0.92	0.84	0.9	0.93	0.93
KMO values for individual items	0.86	0.71	0.73	0.88	0.86
Bartell's test of sphericity (df=105, p<0.00001) (Chi2)	1300	1220	1069	2838	2656
Eigenvalues of first three factors (F1, F2 and F3)	8.2, 1.7, 1.4	8.2, 1.3, 1.1	8.0, 1.3, 1.0	8.8, 1.5, 1.0	8.6, 1.2, 1.0
% variance explained by the three factors	55, 11, 9	55, 9, 8	54, 9, 7	58, 10, 7	58, 8, 7
Cumulative variance explained (%)	75	71	69	75	72
Mean communality (SD)	0.75 (0.12)	0.71, (0.15)	0.69, (0.10)	0.75 (0.08)	0.71 (0.10)
Minimum communality	0.59	0.49	0.55	0.6	0.55
% variance explained by the three rotated factors	28, 24, 24	27, 23, 21	27, 21, 21	26, 25, 24	25, 24, 23
Correlation coefficient between the three factors	0.44 to 0.63	0.52 to 0.67	0.45 to 0.59	0.54 to 0.66	0.65 to 0.66
Items on which F1 loads, with minimum load	Spatial, 0.57	Spatial, 0.73	Spatial, 0.62	Speech, 0.64	Spatial, 0.65
Items on which F2 loads, with minimum load	Speech, 0.65	Speech, 0.52	Speech, 0.51	Quality, 0.65	Quality, 0.56
Items on which F3 loads, with minimum load	Quality, 0.67	Quality, 0.57	Quality, 0.43	Spatial, 0.73	Speech, 0.66
Tucker-Lewis index	0.96	0.87	0.9	0.93	0.92
CFI (Comparative fit index)	0.98	0.92	0.94	0.96	0.95
RMSR	0.03	0.04	0.04	0.03	0.03
RMSEA (root mean square error of approximation)	0.08	0.13	0.11	0.1	0.1

Table 1: Results of factor analyses performed on 15iSSQ, 15SSQs, and SSQ15 data from the HIHA and HIN subjects.

**List of SDC:**

Appendix : SupplementaryData.xlsx, available upon request ([annie.moulin@cnrs.fr](mailto:annie.moulin@cnrs.fr))

