



HAL
open science

Validation of a French-Language Version of the Spatial Hearing Questionnaire, Cluster Analysis and Comparison with the Speech, Spatial, and Qualities of Hearing Scale

Annie Moulin, Celine Richard

► To cite this version:

Annie Moulin, Celine Richard. 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 and Hearing*, 2016, 37 (4), pp.412-423. <10.1097/AUD.0000000000000269>. <hal-04189635>

HAL Id: hal-04189635

<https://hal.science/hal-04189635v1>

Submitted on 28 Aug 2023

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.



Distributed under a Creative Commons CC BY-NC 4.0 - Attribution - Non-commercial use - International License

**Validation of a French-language Version of the
Spatial Hearing Questionnaire,
Cluster Analysis and Comparison with the Speech, Spatial, and
Qualities of Hearing Scale**

Annie Moulin^{1,2,3*} and Celine Richard⁴

¹INSERM U1028, Lyon Neuroscience Research Center, Brain Dynamics and Cognition Team, Lyon, F-69000, France

²CNRS UMR5292, Lyon Neuroscience Research Center, Brain Dynamics and Cognition Team, Lyon, F-69000, France

³University of Lyon, Lyon 1, F-69000, France

⁴Department of Oto-Rhino-Laryngology, Head and Neck Surgery, University Hospital (CHUV), Lausanne, Switzerland.

*Corresponding author: Annie Moulin annie.moulin@cnrs.fr

This is the final version of the manuscript accepted for publication in “*Ear and Hearing*”, the official journal of the American Auditory Society.

ISSN: 0196-0202, ESN: 1538-4667.

Publisher: Lippincott <https://lww.com/pages/default.aspx>

Supplemental material:

The first supplemental material (the French language version of the SHQ) is available at the end of the present file. The second supplemental material is available page 49 onwards of the present file.

The publisher’s version is available at:

https://journals.lww.com/ear-hearing/abstract/2016/07000/validation_of_a_french_language_version_of_the.5.aspx

Full reference of the final version:

Moulin, A., & Richard, C. (2016). 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 and Hearing*, 37(4), 412-423.

<https://doi.org/10.1097/AUD.0000000000000269>

1 **Validation of a French-language Version of the Spatial Hearing Questionnaire, Cluster**
2 **Analysis and Comparison with the Speech, Spatial, and Qualities of Hearing Scale**

3 Annie Moulin^{1,2,3*} and Celine Richard⁴

4 ¹INSERM U1028, Lyon Neuroscience Research Center, Brain Dynamics and Cognition Team,
5 Lyon, F-69000, France

6 ²CNRS UMR5292, Lyon Neuroscience Research Center, Brain Dynamics and Cognition Team,
7 Lyon, F-69000, France

8 ³University of Lyon, Lyon 1, F-69000, France

9 ⁴Department of Oto-Rhino-Laryngology, Head and Neck Surgery, University Hospital
10 (CHUV), Lausanne, Switzerland.

11

12 *Corresponding author: Annie Moulin annie.moulin@cnrs.fr

13 Address : INSERM U1028, CNRS UMR5292, Centre de Recherche en Neurosciences de
14 Lyon, DYCOG, CH le Vinatier, Bâtiment 452, 95 Bd Pinel, Bron, F-69500, France.

15 Phone: 33 (0)4 72 13 89 37 ; Fax: 33 (0)4 72 13 89 01.

16

17

18 **Conflicts of Interest and Source of Funding**

19 Part of this work was supported by la “Fondation de l’Avenir”, research program ET2-
20 652 and by the LABEX CELYA (ANR-11-LABX-0060), the LABEX CORTEX (ANR-11-
21 LABX-0042) of Université de Lyon, within the program "Investissements d’Avenir" (ANR-11-
22 IDEX-0007) operated by the French National Research Agency. Those institutions had no
23 involvement on the study design, data collection and analysis, nor results reporting or
24 manuscript writing. The authors declare no conflict of interest.

25 **Validation of a French-language Version of the Spatial Hearing Questionnaire, Cluster**
26 **Analysis and Comparison with the Speech, Spatial, and Qualities of Hearing Scale**

27

28 **Annie Moulin and Celine Richard**

29

30 **Short summary**

31

32 The objective of this work was to validate a French-language version of the Spatial
33 Hearing Questionnaire (SHQ), including investigating its internal structure using cluster
34 analysis and exploring its construct validity on a large population of hearing-impaired and
35 normal-hearing subjects, and to compare the SHQ with the Speech, Spatial and Qualities of
36 Hearing Scale (SSQ) in the same population. The SHQ showed good conceptual equivalence
37 across languages, suggesting its potential utility for non-cochlear-implanted hearing-impaired
38 patients and in international settings. SHQ and SSQ scores showed remarkable similarities,
39 suggesting the possibility of extrapolating results from one questionnaire to the other.

40 **Validation of a French-language Version of the Spatial Hearing Questionnaire, Cluster**
41 **Analysis and Comparison with the Speech, Spatial, and Qualities of Hearing Scale**

42

43 **Annie Moulin and Celine Richard**

44

45 **ABSTRACT**

46

47 **Objectives:** To validate a French-language version of the Spatial Hearing Questionnaire
48 (SHQ), including investigating its internal structure using cluster analysis and exploring its
49 construct validity on a large population of hearing-impaired and normal-hearing subjects, and
50 to compare the SHQ with the Speech, Spatial and Qualities of Hearing Scale (SSQ) in the same
51 population.

52

53 **Design:** The SHQ was translated in accordance with the principles of the Universalist Model
54 of cross-cultural adaptation of patient-reported outcome instruments. The SSQ and SHQ were
55 then presented in a counterbalanced order, in a self-report mode, in a population of 230 hearing-
56 impaired (HI) subjects (mean age = 54 years and pure-tone audiometry (PTA) on the better ear
57 = 28 dB HL) and 100 normal-hearing (NH) subjects (mean age = 21 years). The SHQ feasibility,
58 readability and psychometric properties were systematically investigated using reliability
59 indices, cluster and factor analyses and multiregression analyses. SHQ characteristics were
60 compared both to different literature data obtained with different language versions and to the
61 SSQ scores obtained in the same population.

62

63 **Results:** Internal validity was high and very good reproducibility of scores and intersubject
64 variability were obtained across the 24 items between the English and French SHQ for NH

65 subjects. Factor and cluster analyses concurred in identifying 5 correlated factors,
66 corresponding to several SHQ subscales: (1) Speech in noise (corresponding to SHQ subscales
67 7 and 8), (2) Localization of voice sounds from behind, (3) Speech in quiet (corresponding to
68 SHQ subscale 1), (4) Localization of everyday sounds and (5) localization of voices and music
69 (corresponding to parts of the SHQ localization subscale). Correlations between SSQ subscales
70 and SHQ factors identified the greatest correlations between SHQ factors 2, 4 and 5 and SSQ
71 spatial subscales, whereas SHQ factor 1 had the greatest correlation with SSQ_speech. SHQ
72 and SSQ scores were similar, whether in NH subjects (8.5 versus 8.4) or in HI subjects (6.6 for
73 both), sharing more than 80% of variance. The SHQ localization subscale gave similar scores
74 as the SSQ spatial subscale, sharing more than 75% of variance. Construct validity identified
75 better ear PTA and PTA asymmetry as the two main predictors of SHQ scores, to a degree
76 similar to that seen for the SSQ. The SHQ was shorter, easier to read and less sensitive to the
77 number of years of formal education than the SSQ, but this came at a cost of ecological validity,
78 which was rated higher for the SSQ than for the SHQ.

79

80 **Conclusions:** A comparison of factor analysis outcomes among the English, Dutch and French
81 versions of the SHQ confirmed good conceptual equivalence across languages and robustness
82 of the SHQ for use in international settings. In addition, SHQ and SSQ scores showed
83 remarkable similarities, suggesting the possibility of extrapolating the results from one
84 questionnaire to the other. Although the SHQ was originally designed in a population of
85 cochlear implant patients, the present results show that its usefulness could easily be extended
86 to non-cochlear-implanted, hearing-impaired subjects.

87

INTRODUCTION

88
89 To assess a person's hearing impairment, audiologists have a variety of hearing
90 assessment measures ranging from standard pure-tone audiometry to speech-perception tests.
91 However, those tests can fail to quantify the effects of hearing impairment on the patient's daily
92 life (Stephens & Héту 1991). Because health policies require a precise and exhaustive
93 evaluation of the benefits of rehabilitation strategies, a growing interest in patient-centered
94 outcome measurements has led to the development of numerous questionnaires designed to
95 assess the patient's appraisal of his or her own hearing disability in a wide range of contexts
96 (Bentler & Kramer 2000; Cox et al. 2000; Noble 2013). Two recent questionnaires focus
97 specifically on spatial hearing and difficult listening situations. The Spatial Hearing
98 Questionnaire (SHQ) was designed by Tyler et al. (2009) and primarily addresses spatial
99 hearing. The Speech Spatial Qualities scale (SSQ) encompasses broader aspects of hearing
100 (including speech, spatial and qualities of hearing) and uses an interview mode (Gatehouse &
101 Noble 2004). The availability of two independently designed questionnaires, tested in different
102 populations, triggers two questions: "Which questionnaire should I use for my patients?" and
103 "If I use one questionnaire, will I be able to compare its results to the results obtained with the
104 other questionnaire?" The first question involves obtaining information about the
105 characteristics and behavior of a questionnaire in the same population as the one intended for
106 its use and the answer to the second question involves a direct comparison between
107 questionnaires in the same population. But before we could answer either, as a non-English
108 speaking country, we were confronted with a third problem: "Is there a validated version of a
109 French-language SHQ and SSQ?"

110 There are few validated health-related questionnaires available in languages other than
111 English, as validation procedures involve much more than a simple translation (Arlinger, 2000).
112 Indeed, the Universalist model of cross-cultural adaptation of health questionnaires (Herdman

113 et al. 1998) identifies six types of equivalence that should be fulfilled for a valid adaptation.
114 Item equivalence and Semantic equivalence relate to the translation of the questionnaire,
115 whereas Operational equivalence addresses how the questionnaire can be administered and
116 scored. Conceptual equivalence relates to how the different concepts included in the
117 questionnaire are equivalent from one country to another, and can be measured by comparing
118 the internal structure of the questionnaire in different languages and populations. Psychometric
119 properties, internal consistency measures and construct validity are reflected in Measurement
120 equivalence. Functional equivalence assesses the validity of the adapted version of the
121 questionnaire and is part of the general psychometric properties that must be reported to validate
122 the questionnaire (Hyde 2000). In the first part of this report, we will present a validation of a
123 French-language version of the SHQ (F-SHQ), following the Universalist model guidelines.
124 Our validation includes reliability measures and a comparison to results obtained in the
125 literature, including factor analysis of the SHQ.

126 In a second part, we largely extend our analysis beyond a language translation validation
127 procedure, by analyzing the SHQ characteristics using data obtained from a population of 230
128 hearing-impaired and 100 normal-hearing subjects. The internal structure of the SHQ will be
129 explored using Revelles' clustering method (Revelle, 1978), which is specifically designed to
130 enable visualization of questionnaire scales and subscales and their hierarchical
131 interdependency, allowing the definition of homogeneous subscales and the identification of
132 similar items (Cooksey & Soutar 2006). Different sources of variability of the SHQ scores are
133 explored: the expected ones (hearing loss, ear asymmetry, age) and the less desirable ones
134 (gender or number of years of education).

135 In the last part of this work, we compare the SHQ and SSQ questionnaires in the same
136 population, using the same presentation format (visual analogue scale) and mode of
137 administration (self-report). To our knowledge, the only direct comparison of these two

138 questionnaires in the same population was performed by Tyler et al. (2009) in a population of
139 cochlear implant (CI) subjects. Because Tyler et al.'s study was aimed primarily at the
140 design of the SHQ, comparisons between the SSQ and SHQ were performed at the scale level,
141 not at the subscale or item levels. Here, SHQ and SSQ scores are compared at the subscale and
142 item levels and their dependency on hearing thresholds, ear asymmetry, and age are explored.
143 Finally, feasibility, including response missing rate, item readability, time to complete the
144 questionnaire and the potential influence of unwanted factors (such as the number of years of
145 education, gender, etc.) are compared.

146

MATERIALS AND METHODS

Cross-cultural adaptation of the SHQ

The cross-cultural adaptation and translation of both questionnaires followed the same procedures, which have already been fully described elsewhere (Moulin et al., 2015). In brief, we used a multistep translation procedure recommended by the European Regulatory Issues on Quality of Life Assessment (ERIQQA) group (Acquadro et al. 2008) involving a double translation, a translation committee, a focus group and a pilot study in 26 hearing-impaired subjects (see the French version of the SHQ (F-SHQ), Supplemental Digital Content 1).

Because our subsequent aim was to compare the SHQ and SSQ questionnaires, we used the same visual analogue scale with a range of 0 to 10 for both¹ (Streiner et al. 2014). The subjects had the possibility to tick an answer box “non applicable” (NA) if they felt the question could not be applied to their personal experience, similarly to Gatehouse and Noble ’s (2004) design for the SSQ. Thus, both the SHQ and SSQ questionnaires included the same presentation and instructions, with an easy sample question (“*You can hear somebody talking to you, facing you, in a quiet room*”) with its visual analogue scale as an example. At the end of the questionnaire, two items were added, using the same visual analogue scale format, asking subjects their opinion about the relevance of the questions to the situations they experience every day (termed “ecological validity”: “*How well do the questions reflect your hearing experiences in daily life? (very badly/ very well)*”) and the difficulty of the questions (termed “difficulty”: “*Did the questions seem easy to answer? (not at all/ yes, no problem)*”).

Procedures

- *Audiological assessment*

¹ Because a numerical scale from 0 to 100 is unusual in France, where marks usually range from 0 to 10, the use of the same scale as the SSQ was needed to obtain a strict equivalence between the presentation formats of the SHQ and SSQ for comparison and to achieve Operational equivalence (Herdman et al. 1998).

170 All of the hearing-impaired (HI) subjects were recruited from a pool of patients who had
171 come for a routine otology consultation with a specialist ENT surgeon. Subjects underwent
172 otoscopic and audiological evaluation, including air- and bone-conduction hearing thresholds
173 and tympanometry performed using an Interacoustic AA222 tympanometer. Pure-tone
174 audiometry was performed using an Interacoustic AC 33 clinical audiometer in a sound-proof
175 booth, between 250 and 8000 Hz in octave steps in each ear. Medical history and general
176 information such as the number of years of education/current professional activity were
177 gathered. Normal-hearing (NH) subjects were recruited mostly from university undergraduates
178 and were selected based on an absence of any medical condition, history of otologic pathology
179 and/or traumatic noise exposure. They underwent pure-tone audiometry in half-octave steps
180 from 125 Hz to 8000 Hz using an interacoustic AC40 audiometer in a sound-proof booth.
181 Written informed consent was obtained from each subject, within the framework of French laws
182 related to non-invasive biomedical research on humans (Agreement number A-11-385, “CPP
183 sud-Est IV”).

184

185 • *Questionnaire data collection*

186 The SHQ questionnaire was administered with the SSQ (Noble & Gatehouse, 2004) and
187 the order of questionnaire completion was counterbalanced. The subjects were asked to measure
188 the time (in minutes) they needed to complete each questionnaire. The questionnaires were
189 explained and given to the HI subjects in a pre-stamped addressed A4 format envelope at the
190 end of the clinical examination. Most of the NH subjects completed the questionnaires on-site.

191

192

193

194

195 **Population characteristics**

196 • *Normal-hearing subjects*

197 The questionnaires were given to 100 normal-hearing subjects. Two of them
198 accidentally missed a page (5 questions) and so the analysis was performed on 98 sets of
199 responses (mean age = 21 years, SD = 2.2, 69 women). The mean audiometric threshold (taken
200 as the mean of the 500 Hz, 1, 2 and 4 kHz frequencies) in their worse ear was 8.7 dB HL (SD
201 = 3.6) and in their better ear was 1.1 dB HL (SD = 4.5). The difference in the mean audiometric
202 threshold between ears was 8 dB (SD = 2.6).

203

204 • *Hearing-impaired subjects*

205 A total of 230 HI subjects took part in the project, but seven of them failed to complete
206 some of the SSQ or SHQ questionnaire due to technical problems and several subjects did not
207 provide a numerical answer to some questions (either by using “NA” or by not answering the
208 question). As a result, we restricted the numerical analysis to the 205 sets of responses that
209 contained responses to more than 80% of the questions. Amongst these subjects (105 women),
210 18 wore hearing aids in at least one ear. Sixteen percent of HI subjects suffered from conductive
211 hearing-loss, 46% of HI subjects came for presbycusis, 11% suffered from noise-induced
212 hearing-loss, and 12% suffered from congenital or early idiopathic sensorineural hearing-loss.
213 The mean age was 54 years (SD = 17) and the number of years of formal education was 13 (SD
214 = 4.2). The PTA on the better ear was 28 dB HL (SD = 15) and 43 dB HL (SD = 27) on the
215 worse ear, with 15 dB (SD = 24) of asymmetry.

216

217

218

219

220 **Data analysis**

221 • *Readability analysis (Item equivalence)*

222 One of the concerns when developing and translating questionnaires is to ensure that the
223 questionnaire remains easily readable by the target population (Calderon et al. 2006; Paz et al.
224 2009; Atcherson et al. 2011). Several readability indexes have been defined, most of them based
225 on the number of words per sentence and the number of syllables per word (Reck & Reck 2007).
226 Although those indexes cannot reflect all of the processes involved in reading comprehension,
227 and most of them have been developed for the English language, they still represent a standard
228 by which to measure readability. We analyzed the readability of both the English language SHQ
229 (E-SHQ) and the F-SHQ using the Gunning-Fog index (Gunning, 1952) and the Flesch Kincaid
230 index (Flesch 1948; Kandel & Moles 1958). Each index was calculated for the entire SHQ, per
231 subscale and for each item, as in Calderon et al. (2006) and Paz et al. (2009).

232

233 • *Scores and subscales*

234 For each subject, the mean score was calculated over the 24 SHQ items, and for each of
235 the eight subscales defined by Tyler et al. (2009) (perception of different voices with Man's
236 voice (#1, #5, #9, #13 and #17), Woman's voice (#2, #6, #10, #14 and #18) and Child's voice
237 (#3, #7, #11, #15 and #19), Music perception (#4, #8, #12, #16 and #20), Source localization
238 (#13 to #24), Understanding speech in quiet (#1, #2, #3 and #4), Understanding speech in noise
239 with target and noise sources from the front (#5 to #8), and Understanding speech in noise with
240 spatially separate target and noise sources (#9 to #12)). In addition, mean scores were calculated
241 over the items belonging to the five different factors obtained from a factor analysis of the SHQ
242 (F1 to F5). Finally, SHQ scores were compared to SSQ scores obtained from the same subjects
243 and the average scores over the 49 SSQ items for each subject.

244 A comparison of the F-SHQ scores was then performed with Perreau et al.'s data
245 (Perreau et al. 2014b) obtained from NH subjects, Potvin et al.'s data (Potvin et al. 2011) on HI
246 subjects and data on cochlear implant patients (Tyler et al. 2009; Perreau et al. 2014a) to
247 evaluate the Item equivalence and the Measurement equivalence of the Universal model.

248

249 **Statistical analysis**

250 • *General procedures*

251 The mean skewness across items was -0.74 for HI subjects and -1.10 for NH subjects,
252 showing a skew toward high scores (as would be expected for NH subjects that tend to give
253 scores above 6 and often close to 10). The kurtosis values were -0.26 and 0.97, respectively.
254 We applied a logit transform transformation as in Perreau et al. (2014b), but it did not yield
255 normal data and the correction was too high for HI subjects, with the skewness and kurtosis
256 worsening after data transformation (kurtosis increasing to 1.61). Therefore, instead of
257 transforming the data, we checked the reliability of our results with bootstrap procedures and
258 used non-parametric statistics, especially for NH subjects.

259 For the analysis of the time needed to complete the questionnaires, the reported time
260 was log-transformed, which reduced the skewness of the data (from 1.8 to -0.1) and the kurtosis
261 (from 4 to -0.3). Regression analysis was performed between questionnaire scores and better
262 ear PTA in order to define the relationship between scores, and non-parametric bootstrap
263 confidence intervals have been calculated for the parameters of the equations, using 10000
264 bootstrap replicates. Relationships between SSQ and SHQ scores were established using
265 orthogonal distance regression (Petras and Podlubny, 2010), in order to take into account
266 measurement errors both in the SSQ and SHQ scores.

267 The multi-regression analyses of scores (and time to complete the questionnaires) as a
268 function of different predictors were restricted to the HI subjects, as including NH subjects

269 would introduce a bias due to the NH subjects having a normal hearing threshold, younger age
270 and greater number of years of formal education than the HI subjects. In addition, investigating
271 predictors in questionnaire scores yields more relevant information when performed in a
272 population similar to the one for which the questionnaire is intended, i.e., HI subjects. Thus,
273 multi-regression analyses were performed to attempt to explain the variation of both SHQ and
274 SSQ scores for the 205 HI subjects using the following explanatory variables: Age, PTA on the
275 better ear, number of years of education, PTA asymmetry and gender.

276 Internal consistency was assessed by calculating both Cronbach's alpha (Cronbach et
277 al. 1951) and item-total correlations for each subject group and for each factor, as determined
278 by factor and cluster analyses. The reliability estimates were compared with the internal
279 consistency indices reported by Tyler et al. (2009) and Perreau et al. (2014b).

280 With any exploratory research involving multi-item questionnaires, multicomparison
281 procedures involve the risk of labeling a result as statistically significant (i.e., $p < 0.05$) when it
282 is only the result of random variation. However, classical multicomparison p adjustments (such
283 as Bonferroni) involve dividing the significance level so that the frequency of incorrect
284 statements about the absence of a significant association between variables increases (type II
285 error inflation) (Rothman, 1990). To avoid this, we chose to follow the guidelines reported by
286 Asendorpf et al. (2013) for psychological studies and by Glickman et al. (2014) for clinical
287 studies, i.e., systematically reporting the exact p -value and using a false discovery rate approach
288 for the problem of multiple comparisons (Benjamini & Hochberg 1995), with p considered as
289 significant for $p < 0.01$.

290

291 • *Cluster and factor analysis (conceptual and functional equivalence)*

292 To explore the underlying structure of the SHQ and the similarity between the F-SHQ
293 and the E-SHQ (i.e., Conceptual equivalence), we performed common factor analyses

294 separately in the group of NH subjects (to compare with Perreau et al.'s (2014b) data), the group
295 of HI subjects (to compare with Tyler et al. (2009)'s data and with Potvin et al. (2011)'s Dutch-
296 SHQ) and all subjects together (NHHI subjects). The analysis involving all subjects might show
297 a more stable structure, because the ratio of the number of subjects to items is well above 10:1
298 (Costello and Osborne, 2005). This ratio was 8:1 for the HI subjects and only 4:1 for the NH
299 subjects, but the Kaiser-Meyer-Olkin (KMO) index of sampling adequacy (Kaiser, 1970) was
300 at least 0.68 (minimum observed for a single item in NH subjects), with an overall KMO of
301 0.86 for NH subjects, 0.87 for HI subjects and 0.9 for all subjects, suggesting that the samples
302 were sufficient (a KMO index is considered good when above 0.70 and excellent above 0.90
303 (Kaiser and Rice, 1974)). For factor extraction, we used the principal axis factoring method,
304 because it is free of distribution assumptions (Fabrigar et al. 1999). For the number of factors
305 to extract, the commonly used subjective Cattell's scree test (Cattell 1966) gave a number of
306 factors of 3 for HI subjects, and between 3 and 5 for NH subjects and the whole sample. The
307 objective parallel analysis (Horn 1965), which is based on a comparison between the
308 eigenvalues of the sample data and the eigenvalues taken from a random sample of uncorrelated
309 data, gave 4 factors for NH subjects (as in Perreau's data), 3 factors for HI subjects (as in Tyler
310 et al.'s data), and 5 factors for all. However, Velicer's minimum average partial method (MAP)
311 (Velicer 1976) gave 5 factors for both HI subjects and all samples, but 8 factors for NH subjects,
312 which seemed to be clearly overfactoring. Thus, we decided to test systematically the 3-, 4- and
313 5-factor options for each group, following the guidelines of Fabrigar et al. (1999) and
314 Hutcheson and Sofroniou (1999). In addition, this method provided a better comparison with
315 Tyler et al. (2009), Perreau et al. (2014b) and Potvin et al. (2011)'s study, which suggested 3-
316 and 4-factor solutions. Because the different dimensions explored by the SHQ are very likely
317 to be correlated with each other, an oblique factor rotation method (direct oblimin) was chosen
318 (Hutcheson & Sofroniou 1999), similar to Akeroyd et al. (2014) in their factor analysis of the

319 SSQ. The percentages of variance explained by the different factors will be given before
320 rotation, for easier comparison with literature data, and after rotation.

321 The fit of the different models to the data was checked by RMSR (root mean square
322 residuals), where a value below 0.05 is considered a good fit, and values between 0.05 and 0.08
323 are considered a fair fit (Field et al. 2012; Fabrigar et al. 1999).

324 Next, we used a hierarchical clustering method (Revelle's ICLUST algorithm, Revelle 1978,
325 1979), which has the advantages, over more classical factor analysis, of exploring similarities
326 and hierarchical interdependencies between items and clusters of items of a questionnaire. The
327 ICLUST approach is less method dependent, thus allowing us to explore the Universalist
328 model's Functional equivalence and to further describe the SHQ characteristics and internal
329 structure. The ICLUST method is based on two indices: the coefficient alpha measure of
330 internal consistency and Revelle's beta coefficient (Revelle, 1978) that reflects factorial
331 homogeneity (specifically, beta is the worst split-half reliability of a scale and thus is lower
332 than alpha). In short, the scale structure is built starting from the two item clusters with the
333 greatest similarity. An item is then added to the initial two item clusters only if this addition
334 improves the cluster's internal consistency (measured by alpha) and/or the factorial
335 homogeneity (measured by beta). This technique is considered complementary to the more
336 classical approach of factor analysis and is less dependent on the subjects-to-variables ratio and
337 the method (Cooksey & Soutar 2006). In this case, this method has a definite advantage over
338 exploratory factor analyses for the number of factors extraction, because the number of factors
339 to be extracted varies between 3 and 5. Thus, the ICLUST algorithm (Revelle, 1978) was
340 applied to the NH subjects and HI subjects separately and to the entire sample, without any
341 assumption about the latent structure of the SHQ.

342 To analyze functional equivalence, the SHQ factor structures obtained with factor
343 analysis and cluster analyses were compared, in addition to a comparison with literature data

344 (E-SHQ data from Tyler et al. (2009) and Perreau et al. (2014b) and the D-SHQ data from
345 Potvin et al. (2011)). Factor congruence coefficients were computed between the cluster and
346 the factor analyses and between the present F-SHQ and the E-SHQ results of Tyler et al. (2009)
347 and Perreau et al. (2014).

348 Finally, the SHQ item scores were averaged according to the 5-factor structure obtained
349 by factor and cluster analyses using an item weight of 1, instead of the specific loading of each
350 item on the factor, i.e., using an equally weighted mean of item scores. Provided the item
351 loadings on each factor are high, this procedure does not result in a loss of accuracy (Akeroyd
352 et al. 2014).

353 All of the statistics were calculated using the “Psych” package (Revelle, 2014) within
354 the R statistical package version 3.1.0 (2014-04-10).

RESULTS

F-SHQ CROSS-CULTURAL VALIDATION AND CHARACTERISTICS

Feasibility and difficulty (Item equivalence, Semantic equivalence)

- *Percentage of valid responses (Item equivalence)*

No items were missing in the NH subjects' data, and only a few "NA" responses were given. However, the HI subjects showed between 0.9% and 4% (#20) missing answers, and when the "NA" answers are taken into account, between 1.3% and 9% (#11) of item responses were missing (mean 4.4% (SD = 2.2)). The number of missing + "NA" responses was higher for items 9, 10, 11 and 12 (all items pertaining to perception with a "loud fan off to one side," which were the items queried the most by the patients during the pilot study). This could reflect a problem of item equivalence, which is specific to the F-SHQ: perhaps the experience of fan noise is less common in France than in the US.

There were no significant correlations between the percentage of missing responses and age, years of education, or PTA on the better ear, but there was a significant small increase with PTA on the worse ear ($r = 0.19$, $p < 0.005$), explaining only 3.8% of the variance.

- *Readability analysis*

The number of words per item varied from 7 to 21, with a mean of 13.1 words. The total length of the questionnaire was 512 words. The readability analysis showed a Gunning-Fog index of 5.6 and a Flesch-Kincaid index of 7.3, i.e., very easy to easy. Per-item analysis revealed large inter-item variability, with readability ranging from 3.1 and 5.3 (very easy) to 10.6 and 13.2 (difficult) using the Gunning-Fox and Flesch-Kincaid indexes, respectively. There were excellent correlations between readability indices obtained for the F-SHQ and the E-SHQ ($r > 0.84$ for Gunning fox, $r = 0.78$ for Flesch-Kincaid, $p < 0.0001$). No significant relationship was

379 found between the rate of missing responses per item and readability indices, nor between
380 scores and readability indices.

381

382 • *Time needed to complete the SHQ*

383 For HI subjects, the median time needed to complete the questionnaire was 10 minutes,
384 and 75% of HI subjects took less than 15 minutes to complete² (Figure 1). This duration
385 increased significantly as HI subjects showed a greater PTA on the better ear ($\beta = 0.25$, $p <$
386 0.003), a lower number of years of formal education ($\beta = -0.20$, $p < 0.005$) and a greater age
387 ($\beta = 0.16$, $p < 0.05$), with 24% of the variance explained by those 3 predictors ($F(3, 189) =$
388 19.7 , $p < 0.0001$).

389 • *SHQ difficulty and ecological validity*

390 The significant correlations obtained between the SHQ difficulty item and (1) the PTA
391 on the better ear (Spearman's $r = -0.20$ $p < 0.006$) and (2) the overall SHQ scores ($r = 0.23$ $p <$
392 0.001) suggest that the HI subjects' answers were at least partly based on their own hearing
393 difficulties. There was no relationship between SHQ difficulty score and number of years of
394 formal education (Spearman's $r = -0.02$, $p = ns$).

395 No systematic relationship was observed between the SHQ Ecological Validity and
396 subject characteristics (HI subjects versus NH subjects, with or without hearing aid, Better Ear
397 PTA (Spearman's $r = -0.11$, $p = ns$) or number of educational years (Spearman's $r = -0.12$, p
398 $= ns$)).

399

400

401

² Although the time needed to complete the questionnaire seems trivial, it gives an insight about its feasibility and is often the first question asked before using questionnaires in a clinical setting.

402 **Overall SHQ reliability (Measurement equivalence)**

403 The value of Cronbach's alpha was 0.96 for NH subjects and 0.98 for HI subjects,
404 indicating very good internal consistency. Cronbach's alpha was then calculated for each of the
405 five factors and each of the eight subscales for NH subjects and HI subjects (alpha in italics and
406 brackets), and ranged from 0.90 (*0.94*) for F4 to 0.98 (*0.98*) for F2, and from 0.77 (*0.90*) for
407 subscales 1 and 2 to 0.97 (*0.98*) (subscale 8).

408 The item-total correlation (see table in Supplemental Digital Content 2) was lower for
409 the first 4 items in NH subjects (from 0.22 to 0.32) and in HI subjects (from 0.63 to 0.67) than
410 for the other items. Except for those first 4 items, the item-total correlations were at least 0.67
411 for NH subjects and at least 0.74 for HI subjects. The mean inter-item correlation was 0.5 for
412 NH subjects and 0.7 for HI subjects.

413

414 **Scores in NH subjects (Measurement equivalence)**

415 The mean score for the NH subjects was 8.4 (SD = 1.2), ranging from 6.2 (SD = 2.6)
416 for #21 to 9.9 (SD = 0.4) for #1 (Figure 2). The highest scores were obtained for the "*Speech*
417 *in quiet*" subscale with 9.9 (SD = 0.4), whereas the lowest scores were obtained for the
418 "*localization*" subscale (7.9, SD = 1.6) (Figure 3).

419 F-SHQ scores tended to be lower than Perreau et al. (2014b)'s E-SHQ data (Figure 2),
420 with significant differences at the 0.05 level for 5 items (out of 24), and two of the eight
421 subscales (*localization* and *noise separate*). However, those degrees of significance did not
422 survive p-correction for multicomparison procedures, except for item #21, which showed a
423 significantly lower score on the F-SHQ. There was very good agreement with scores across the
424 different items between the F-SHQ and the E-SHQ (Figure 2): the scores obtained in NH
425 subjects across the 24 items correlated highly significantly with the scores reported in the NH
426 subjects by Perreau et al. (2014b): Spearman $r = 0.78$ ($p < 0.0001$) for the mean scores and $r =$

427 0.91 for the standard deviations ($p < 0.0001$). The table in Supplemental Digital Content 2
428 provides mean scores and the cutoff points at two standard deviations lower than the mean
429 scores, in accordance with international guidelines for the identification of diseases (WHO
430 1980).

431

432 **Cluster and Factor analysis (Conceptual equivalence)**

- 433 • *Factor analysis*

434 For each of the NH, HI and NHHI subject groups, the extraction of 3, 4 and 5 factors
435 gave a similar but not identical distribution of items on which factors were loading (Figure 4).
436 Whether 3, 4 or 5 factors were extracted, the communalities per item were all at least 0.53 for
437 NH subjects and at least 0.67 for HI subjects. The Spearman correlation between communalities
438 obtained across the 24 items in NH subjects and in HI subjects was 0.84 ($p < 0.0001$), and the
439 correlation between Tyler's data and our HI subjects data was above 0.71 ($p < 0.0001$). No
440 significant correlations were obtained between our data and Potvin et al.'s (2011) data ($r = -$
441 0.09 , $p = ns$), nor between Tyler et al.'s (2009) data and Potvin et al.'s (2011) data ($r = -0.12$,
442 $p = ns$). The details regarding the different number of factor extractions in the different groups
443 are available in the Supplemental Digital Content 2 and Figure 4 provides a summary of the
444 outcomes of the different factor analysis reported here and in the literature.

445 For the total sample, both parallel analysis and Velicer's MAP index (Velicer 1976)
446 gave a 5-factor extraction that explained 89% of the total variance (70%, 8%, 6%, 2% and 2%
447 for the 5 factors, and 34%, 18%, 16%, 17% and 15% after oblique rotation). The RMSR was
448 below 0.02. The 5-factor structure matched the one presented in Figure 4 and the communalities
449 and the pattern matrix, which show the load of each factor on each item, are provided in a figure
450 in the Supplemental Digital Content 2.

451

- *Cluster analysis*

452

453 Cluster analysis gave a 19-cluster hierarchical tree (labelled C1 to C19), with a 5-cluster

454 solution at the top-most level, with identical items within each cluster for all samples considered

455 (NH, HI or NHHI subjects), as shown for NHHI subjects in Figure 5 (clusters labelled C12,

456 C13, C15, C18, C19). The goodness of fit of the three structures was 0.98 for NH subjects and

457 1 for HI and NHHI subjects, with the RMSR below 0.03 for NH subjects and below 0.02 for

458 HI and NHHI subjects, showing near perfect fits. The eigenvalues of the 5 clusters were 6.9, 4,

459 3.6, 3.5 and 3.5 (NHHI subjects). Thus, for the three samples, the SHQ is not unidimensional,

460 but rather comprises five distinct scales, which are reliable and homogeneous, as shown by

461 alpha coefficients consistently above 0.89 and beta coefficients above 0.78. Those five scales

462 are strongly intercorrelated (the lowest correlation was 0.52 between C19 and C15, whereas the

463 correlation between C12 and C13 was 0.87), and correspond to the five factors obtained by

464 factor analysis. There was one 8-item cluster and there were four 4-item clusters. Each cluster

465 corresponds to a successive part of the SHQ, grouped by type of question (Figure 5). In addition,

466 by looking at the tree for cluster C18 (Figure 5), 3 subscales can be identified: a subscale (C11)

467 pertaining to speech in noise when the noise is at the front, a subscale (C9) pertaining to speech

468 in noise when the noise is at the side, and a subscale (C7) pertaining to music perception. The

469 addition of C7 to C9 yielded a decrease in homogeneity (the beta coefficient decreased from

470 0.97 to 0.95), suggesting that #8 and #12 reflect a slightly different aspect than the other C9

471 items. Revelle's (1978) cluster analysis shows items associated together by degree of similarity

472 from the left to the right side of the tree, which enables the identification of a systematic

473 grouping of male voice items with female voice items (clusters C1 to C4, C6). Then, items

474 pertaining to a child voice are added, with the music items always added last to the structure,

475 thus showing their greater degree of dissimilarity from the other items.

476 In summary, a similar 5 factor structure was obtained with factor analysis and cluster
477 analysis, with the different factors (or clusters) defined as follows:(1) Speech in quiet (F3)
478 corresponding to the Speech in quiet subscale; (2) Speech in noise (F1) corresponding to
479 subscales 7 and 8; (3) Localization of voice sounds from behind (F2); (4) Localization of
480 everyday sounds (F4); and (5) Localization of voices and music (F5).

481

482

COMPARISON BETWEEN SHQ AND SSQ

483

Feasibility and difficulty

484 There was no significant difference between the SHQ missing rate of 4.4% (SD = 2.2)
485 and the SSQ missing rate of 3.8% (SD = 2.8). The number of missing responses per subject
486 correlated between the SHQ and SSQ at $r = 0.37$ ($p < 0.00001$).

487 Although there was a significant correlation between the time to complete the SSQ and
488 the time to complete the SHQ ($r = 0.68$, $p < 0.00001$), the same analysis as the one performed
489 for the SHQ yielded different predictors for the SSQ (24% variance explained, $F(4, 192) = 15$,
490 $p < 0.00001$): The main predictor was the number of years of formal education ($\beta = -0.22$, p
491 < 0.003), followed by increasing age ($\beta = 0.18$, $p < 0.03$), hearing loss ($\beta = 0.17$, $p < 0.04$)
492 and having not filled out a similar questionnaire before ($\beta = 0.17$, $p < 0.01$).

493 The “difficulty” and “ecological validity” scores showed that the SHQ was perceived as
494 significantly easier than the SSQ for both NH and HI subjects (SSQ = 7.1, SD = 2.6 versus
495 SHQ = 7.5, SD = 2.5; $t = 3$, $p < 0.004$ for HI subjects) but less ecologically valid than the SSQ
496 for HI subjects (SSQ = 8.15, SD = 2.1 versus SHQ = 7.7, SD = 2.6).

497

498

Comparison of SHQ and SSQ scores

499 As detailed results concerning the SSQ have been presented elsewhere (Moulin and
 500 Richard, 2015), only the main SSQ results will be presented for comparison purposes. However,
 501 the results regarding the SHQ will be fully detailed.

502 There was no significant difference between SHQ and SSQ scores in the NH subjects :
 503 8.5 versus 8.4 (Wilcoxon $z = 1.5$, $p = ns$) or in HI subjects: 6.6 versus 6.6 (Wilcoxon $z = 0.75$,
 504 $p = ns$). Similarly, the SHQ *Localization* subscale and the SSQ *Spatial* subscales scores were
 505 not different for NH or HI subjects (Figures 3 and 6). There was a highly significant correlation
 506 between SSQ and SHQ scores ($r = 0.90$ $p < 0.0001$), with a linear relationship³ shown in
 507 equations (1) and (2) (Figure 6).

$$508 \quad SHQ = 1.083 * SSQ - 0.624 \quad (1)$$

$$509 \quad SSQ = 0.923 * SHQ + 0.576 \quad (2)$$

$$510 \quad SHQ \text{ Localization} = 1.107 * SSQ \text{ Spatial} - 0.680 \quad (3)$$

$$511 \quad SSQ \text{ Spatial} = 0.904 * SHQ \text{ Localization} + 0.615 \quad (4)$$

512

513 All of the correlations between SHQ subscales and the main SSQ subscales were highly
 514 significant ($p < 0.00001$) (table 1), with some expected patterns: the SHQ localization and
 515 factors 2, 4 and 5 showed a significantly greater correlation with SSQ Spatial (equations 3 and
 516 4, Figure 6) than with the other SSQ subscales ($Z > 2.6$, $p < 0.01$), whereas SHQ factor 1,
 517 Speech in Noise and Speech in noise (Separate) showed a significantly greater correlation with
 518 SSQ Speech than with the other 2 SSQ subscales ($Z > 3.8$, $p < 0.0001$).

519 Linear regression between SHQ score and Better Ear PTA ($r=0.56$, $p<0.00001$) in NH
 520 and HI subjects showed a 0.66 point decrease in SHQ score per 10 dB HL increase, according

³ Similarly to the relationship between the SSQ and one of its short form (the SSQ12), reported to follow a power function (Noble et al., 2013), we fitted a power function to the SHQ/SSQ relationship, with: $SHQ=10*(SSQ/10)^{1.010}$ [0.966, 1.055]. The 95% confidence interval for the exponent (in brackets) overlaps 1, suggesting that the SHQ:SSQ relationship is linear.

521 to equation 5, with HL for Better ear PTA in dB HL and non-parametric bootstrap 95%
 522 confidence intervals specified in brackets. This finding is not significantly different from the
 523 result obtained with the SSQ in the same population (equation 6).

$$524 \quad \text{SHQ score} = -0.066 [-0.078, -0.054] * \text{HL} + 8.45 [8.28, 8.78] \quad (5)$$

$$525 \quad \text{SSQ score} = -0.069 [-0.079, -0.058] * \text{HL} + 8.55 [8.39, 8.82] \quad (6)$$

526

527 Multiregression analysis on HI subjects, showed that SHQ score variances were
 528 explained for 31% by the Better Ear PTA (beta = -0.53) and PTA asymmetry (beta = -0.20).
 529 The other predictors (age, gender, number of years of formal education) were not statistically
 530 significant. A subscale analysis explained between 20% and 34% of the subscales variances
 531 with two predictors, Better Ear PTA and PTA asymmetry. Figure 7a shows the Beta coefficients
 532 obtained for each predictor, in the same population of HI subjects, for each SHQ and SSQ
 533 subscale. Better Ear PTA was the most significant predictor of scores, and PTA asymmetry was
 534 present in F2, F4 and F5 and the first five subscales of the SHQ, with a large beta coefficient (-
 535 0.26) for the localization subscale that was very similar to the SSQ spatial subscale (beta = -
 536 0.27). The three factors with PTA asymmetry as predictors represent the three parts of the
 537 localization SHQ subscale. Surprisingly, the eighth SHQ subscale (Speech in noise with
 538 spatially separate noise and speech sources) did not have PTA asymmetry as a significant
 539 predictor.

540 **Individual item analysis**

541 One of the concerns that can arise when designing a questionnaire is that items may
 542 have other significant predictors (i.e. explanatory variables) besides those that are expected. For
 543 instance, PTA on the better Ear and PTA asymmetry are expected predictors, whereas other
 544 variables, such as gender or number of years of education, are less welcome and may represent
 545 confounds. To identify such predictors, multi-regression analyses were performed for each

546 item, with gender, number of years of education, Better Ear PTA and PTA asymmetry as
547 predictors (age was never a significant predictor, so it was removed from the model) on the
548 group of 169 HI subjects who had no missing responses.

549 The percentage of variance explained across the 24 items ranged from 16% (#21) to
550 32% (#3), with both the mean and median at 24%. The beta coefficients allowed us to compare
551 the relative influence of each statistically significant predictor on the scores (Figure 7b). Better
552 Ear PTA was a significant predictor for all items, with the highest beta coefficient. PTA
553 asymmetry was a significant predictor from items #13 to #24, but beta coefficients were
554 substantially lower than Better Ear PTA. For items #17, #18 and #20 (and for SHQ factor 2 that
555 regroups items #17 to #20), the number of years of education tended to be a significant predictor
556 ($0.01 < p < 0.05$), with lower scores for HI subjects with a higher number of years of education
557 ($p = 0.05$, $\beta = -0.14$).

558 For item #23, related to the capacity of knowing from which direction a car is coming
559 when standing beside a road with eyes closed, gender showed a significant relationship, with
560 lower scores for women versus men. However, the influence of both number of years of
561 education and gender remained weak, and did not survive p-correction for multiple comparisons
562 ($0.01 < p < 0.05$).

563

564

DISCUSSION

565

Cross-language comparison of the SHQ

566

567

568

569

570

571

572

573

574

575

Internal structure of the SHQ

576

577

578

579

580

581

582

583

584

585

586

587

588

One of the most widely reported elements, to validate a translation of a questionnaire, is a measure of internal consistency, such as Cronbach's alpha. The internal consistency of the F-SHQ was very high for both NH and HI subjects, similarly to the Dutch (Potvin et al. 2011) and the English versions (Tyler et al. 2009, Perreau et al. 2014a, b) of the SHQ. The excellent correlation across items between Perreau's data and the F-SHQ for NH subjects, both in scores and standard deviations, is another strong argument in favor of the validation of the F-SHQ and the validity of the SHQ across different languages. It suggests a good Universalist Model Functional equivalence of the F-SHQ.

Investigating the internal structure of a questionnaire is essential to ascertain its robustness across different language versions for cross-cultural validation and potential international use. Such an analysis allows us to assess Conceptual equivalence within the Universalist model of cross-cultural adaptation (Regnault & Herdman 2014). In addition, it enables exploration of the robustness of the internal structure of a questionnaire across different populations, using data driven methods, which can reveal a different structure from that intended at the design stage of a questionnaire. Indeed, factor analysis was used to assess the F-SHQ internal structure, and gave slightly different results according to the dataset used (NH or HI subjects). Previous analyses of the SHQ have reported 3 factors (Tyler et al., 2009; Potvin et al. 2011) or 4 factors (Perreau et al. 2014b). This can be expected from factor analyses with a relatively low ratio of number of subjects to number of variables (less than 2:1 in Perreau et al. (2014b) and less than 4:1 in Potvin et al. (2011)) and different populations (cochlear implant (CI) subjects in Tyler et al. (2009) versus non-cochlear implantees with a 6:1 ratio). The lack

589 of a significant correlation between the communalities obtained across the 24 items by Potvin
590 et al. (2011) and both our dataset and Tyler's dataset could be attributed to the very low ratio
591 used by Potvin (below 2:1). For all of the different number of factor extractions (3, 4 or 5)
592 performed, we obtained very high communalities. Although different structures were obtained
593 between the different literature data and our different data sets, there were great similarities in
594 the item groupings: specifically, all of the structures identified items #5 to #12 (addressing
595 speech in noise, regardless of whether noise was at the front (subscale 7) or at the side (subscale
596 8)) as a single factor (factor 2). The first four items, corresponding to the subscale Speech in
597 quiet, were grouped in the same factor, with a low percentage of variance explained. This
598 suggests that those items, which describe relatively easy situations, were categorized together
599 (the exception represented by item #4 in Tyler's data obtained in CI subjects is explained
600 below). The differences in structures, from 3 to 5 factors, lies entirely in how items 13 to 24
601 (addressing localization) are grouped: in one single group of 12, down to 3 groups of 4. This
602 suggests a difference in the degree of detail in the structures obtained, rather than true
603 differences in the SHQ structure. If cross-loading is to be kept to a minimum, five factors seem
604 the most appropriate for all datasets and correspond to the 5 clusters obtained using Revelle's
605 method (Revelle, 1978 & 2014).

606 The difference in the percentage of variance explained in the five factor structure
607 between NH and HI subjects suggests that for NH subjects, the most difficult situations were
608 in the localization of sounds, which is in agreement with the fact that their lowest scores were
609 obtained for localization items.

610 For HI subjects, the presence of background noise was the most difficult situation, with
611 the first factor (8 items pertaining to noise) explaining most of the variance. This difference
612 between NH and HI subjects could be attributed either to hearing threshold differences between
613 the groups, or to the age difference. Indeed, auditory perception in noise has been shown to

614 deteriorate with age independent of hearing loss (Pichora-Fuller, 2003), and a significant age
615 effect on SSQ items pertaining to hearing speech in noise has been reported in a larger group
616 of 216 HI subjects, including the 202 HI subjects of the present study (Moulin and Richard,
617 2015).

618 Although we did not obtain statistically significant correlations between the
619 communalities obtained here and Potvin's data, the lower communalities on the music items
620 (#4, #8, #12, #16, #20), which were particularly obvious in NH subjects and (to a lesser degree)
621 in HI subjects, were also observed by Potvin. This could be attributed to the fact that NH
622 subjects tend both to categorize different voices (male, female, children) together and to put
623 music hearing in a different category. Indeed, cluster analysis revealed a systematic association,
624 and hence the highest degree of similarity, between the items pertaining to the male voice and
625 the corresponding items pertaining to the female voice, resulting in four subclusters. To those
626 subclusters, the third item to be added was systematically the child voice item, and then the
627 music item was added the last, showing some degree of dissimilarity from the others. In
628 addition, items #8 and #12 are associated in a subcluster (subcluster C7 in Figure 5) that
629 represents a subscale of cluster C18 (corresponding to F2), emphasizing the dissimilarity of
630 music items from the other items. Except for item #4, such a pattern is much less obvious in
631 Tyler et al.'s data in CI subjects. Indeed, perception of speech in quiet can be quite good in CI
632 subjects, but their difficulty with music perception is notorious (Kang et al. 2009; Singh et al.
633 2009). This could also explain why #4 has very low communality and is part of the second
634 factor in Tyler's data and not part of the third factor as in all other data sets. For all of the other
635 items, whether speech in noise or music in noise, CI subjects would have substantial difficulties,
636 and music items would not stand out as much as in a quiet situation. Perhaps Revelle's (1978)
637 method on CI subjects' data could highlight some items similarities obtained specifically in CI
638 subjects.

639 Regardless of whether factor or cluster analysis is used, the latent internal structure of
640 the SHQ revealed by the data is quite different from the initial 8 subscales (Tyler et al. 2009)
641 intended at the design stage: The first 3 subscales (male, female and child voice) seem to be
642 redundant and do not appear in the internal structure of the SHQ, whereas the localization
643 subscale appears as a composite of 3 smaller subscales.

644

645 **Comparison between the SHQ and SSQ**

646 The direct comparison between the SHQ and the SSQ in the same population, using the
647 same administration mode and presentation format, allowed us to assess the similarities and
648 differences between the two questionnaires and to define a precise relationship between the two
649 scores.

650 The availability of data on a homogeneous population of NH subjects seems essential
651 to the interpretation of scores obtained in HI subjects. Indeed, the maximum score of 10 is not
652 necessarily obtained for each item, and some items show high intersubject variability so that
653 the cutoff point for hearing disability, defined as the mean minus 2 standard deviations, can be
654 too low, rendering some items inefficient when considered individually. This is the case for
655 SHQ items #21 and #22, with cutoffs of 1.0 and 2.2, respectively. Item 21 (*location of an*
656 *airplane*) also had a low score and high SD in Perreau et al.'s data. The last four items pertaining
657 to localization had the smallest scores in NH subjects, the lowest prediction by Better Ear PTA
658 but, contrary to what could be expected, a relatively lower prediction by PTA asymmetry
659 (compared to the other items in the *localization* subscale). The same low cutoff point is
660 observed for several items of the SSQ as well, mostly items pertaining to the spatial subscale
661 (Moulin et al. 2015). For both questionnaires, the lowest scores in NH subjects is observed for
662 items pertaining to localization, not only in the present study, but in other language versions of

663 both questionnaires (Demeester et al. 2012 and Bahn et al. 2012 for the SSQ and Perreau et al.
664 2014b for the SHQ).

665 The similarity between the main scores (6.6 versus 6.6 in HI subjects and 8.5 versus 8.4
666 in NH subjects) and their standard deviations, the very good correlation between the two (more
667 than 81% variance explained) in two different populations (NH and HI subjects), and the very
668 good correlation between the SSQ spatial subscale and SHQ localization are extremely
669 encouraging, enabling comparisons between studies using the different questionnaires.

670 Both questionnaires depended on Better Ear PTA and PTA asymmetry. The SSQ total
671 score depended more on Better Ear PTA than did the SHQ total score, but the SSQ *Spatial* and
672 SHQ *Localization* subscales depended on PTA asymmetry to a similar degree. The SHQ
673 “*speech in quiet*” behaved in a very similar manner to the SSQ *Speech*, depending only on
674 Better Ear PTA. Age was not identified as a statistically significant predictor of either SSQ or
675 SHQ scores in the present data. Surprisingly, the SHQ *Speech in Noise* and *Speech in Noise*
676 (*Separate*) did not show PTA asymmetry as a significant predictor. The present results show
677 that SSQ and SHQ scores are very similar and behave in the same way as a function of Better
678 Ear PTA and PTA asymmetry in HI subjects, allowing comparison of SSQ scores and SHQ
679 scores in different studies. This similarity enables a follow-up of the scores from the SSQ to the
680 SHQ, for instance, in long-term hearing-impaired patients who will eventually need a cochlear
681 implant and perhaps will be more specifically assessed by the SHQ. The similarities between
682 both questionnaires could also provide the basis for building a composite questionnaire, taking
683 advantage of the good points of both the SSQ and the SHQ.

684 Although numerous studies have pointed out the need for health-related questionnaires
685 to meet patients’ reading abilities (Paz et al. 2009; Calderon et al. 2006), such aspects are rarely
686 reported during questionnaire design. Although we used readability indexes on the French
687 version of the questionnaires, the excellent correlations between the E-SHQ and F-SHQ

688 readability indices indicate that the results are likely extendable to the English-language
689 versions. The readability indices are clearly easier for the SHQ than for the SSQ, with much
690 shorter questions (13 words versus 21 for the SSQ in the French version). This relative easiness
691 of the SHQ was confirmed by the lack of a significant influence of the number of years of
692 education on SHQ scores. A tendency for a decrease in scores as the years of education
693 increased was observed for only a few items. In contrast, in the SSQ, a significant increase in
694 scores for several questions was obtained when years of education increased (Moulin and
695 Richard, 2015). The time needed to complete the SSQ appeared to be more sensitive to the
696 number of years of education than the SHQ. The number of years of education was the most
697 important predictor of time to complete the SSQ, but it ranked only third in importance for the
698 SHQ. This probably reflects the higher degree of difficulty of the SSQ compared to the SHQ,
699 due to the former's greater length and more-difficult questions as shown by both the readability
700 analysis and the self-report scores of the items pertaining to items difficulty, which were greater
701 for the SSQ. To allow an unbiased comparison between the SHQ and the SSQ, we used a similar
702 visual analog scale for both questionnaires, including adding a "non-applicable" answer option
703 to the SHQ. Both of the questionnaires showed a similar rate of missing answers, but in contrast
704 to the SSQ, the SHQ showed no correlation between missing response rate and readability
705 indices when tested on the same population as the SSQ. However, the relative simplicity of
706 SHQ items compared to SSQ comes at a cost of ecological validity: subjects tended to think
707 that SSQ items were more relevant to everyday communication situations. Indeed, some SHQ
708 items appeared to be redundant, such as the repeated items for male and female voices, as shown
709 by the cluster analysis. Finally, the SSQ was designed to be administered in interview mode,
710 where the observer could better explain and facilitate the reading of the scale (Singh & Pichora-
711 Fuller 2010), whereas the SHQ was designed to be used in self-report mode. Using both of the

712 questionnaires in self-report mode could explain why in this study, the SSQ seemed to be more
713 sensitive to undesired factors such as the patient's number of years of education.

714 In summary, the SHQ is easier, quicker and seems to be more specific for cochlear
715 implant subjects due to its inclusion of music items, but some of the SHQ items seem redundant.
716 In contrast, the SSQ is longer, more difficult and could perform as well without some
717 particularly difficult items, but it reflects the hearing difficulties of hearing impaired subjects
718 in a more realistic manner, with a robust internal structure of 3 subscales. However and more
719 importantly, the similarities of scores and behaviors between the SSQ and the SHQ in the same
720 population show that the use of different questionnaires does not preclude the possibility of a
721 direct comparison between scores, allowing comparison between international studies, even
722 when they use a different questionnaire.

723

724

725

CONCLUSION

726
727
728
729
730
731
732
733
734
735
736
737
738
739
740

This study validated a French-language version of the SHQ within the framework of the Universalist model of cross-cultural adaptation of health-related questionnaires and showed that the F-SHQ version was similar to other language versions. Thus, the psychometric properties and the internal structure (based on cluster analysis) of the F-SHQ obtained in this study from a large sample of normal-hearing and hearing-impaired subjects are easily extendable to other language versions.

Secondly, both the SHQ and SSQ were administered to the same population of normal-hearing and hearing-impaired subjects, allowing a direct comparison between SHQ and SSQ scores and their behavior as a function of the main predictors: Better Ear PTA and PTA asymmetry. The questionnaires showed remarkable similarities, demonstrating the possibility of extrapolation of results from one questionnaire to the other. The SHQ was perceived as both quicker and easier to complete by the patients in a self-report mode but was somewhat less ecologically valid than the SSQ.

ACKNOWLEDGEMENTS

741

742

743 This work was supported in part by the “Fondation de l’Avenir” research program ET2-
744 652 and by the LABEX CELYA (ANR-11-LABX-0060), the LABEX CORTEX (ANR-11-
745 LABX-0042) of Université de Lyon, within the program "Investissements d'Avenir" (ANR-11-
746 IDEX-0007) operated by the French National Research Agency (ANR). The authors wish to
747 thank Prof. Michael Akeroyd for his very valuable suggestions.

748

REFERENCES

- 749
750
- 751 Acquadro, C., Conway, K., Hareendran, A., et al. (2008). Literature Review of Methods to
752 Translate Health-Related Quality of Life Questionnaires for Use in Multinational
753 Clinical Trials. *Value Health*, 11, 509–521.
- 754 Akeroyd, M.A., Guy, F.H., Harrison, D.L., et al. (2014). A factor analysis of the SSQ (Speech,
755 Spatial, and Qualities of Hearing Scale). *Int. J. Audiol.*, 53, 101–114.
- 756 Arlinger, S. (2000). Can we establish internationally equivalent outcome measures in
757 audiological rehabilitation? *Ear Hear.*, 21, 97S–99S.
- 758 Asendorpf, J.B., Conner, M., De Fruyt, F., et al. (2013). Recommendations for Increasing
759 Replicability in Psychology: Recommendations for increasing replicability. *Eur. J.*
760 *Personal.*, 27, 108–119.
- 761 Atcherson, S.R., Zraick, R.I., Brasseur, R.E. (2011). Readability of patient-reported outcome
762 questionnaires for use with persons with tinnitus. *Ear Hear.*, 32, 671–673.
- 763 Banh, J., Singh, G., Pichora-Fuller, M.K. (2012). Age Affects Responses on the Speech, Spatial,
764 and Qualities of Hearing Scale (SSQ) by Adults with Minimal Audiometric Loss. *J. Am.*
765 *Acad. Audiol.*, 23, 81–91.
- 766 Benjamini, Y., Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and
767 Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B Methodol.*, 57, 289–300.
- 768 Bentler, R.A., Kramer, S.E. (2000). Guidelines for choosing a self-report outcome measure.
769 *Ear Hear.*, 21, 37S–49S.
- 770 Calderon, J.L., Morales, L.S., Liu, H., et al. (2006). Variation in the readability of items within
771 surveys. *Am. J. Med. Qual. Off. J. Am. Coll. Med. Qual.*, 21, 49–56.
- 772 Cattell, R. (1966). The Scree test for the number of factors. *Multivar. Behav. Res.*, 1, 245–276.

- 773 Cooksey, R.W., Soutar, G.N. (2006). Coefficient Beta and hierarchical item clustering: An
774 analytical procedure for establishing and displaying the dimensionality and
775 homogeneity of summated scales. *Organ. Res. Methods*, 9, 78–98.
- 776 Costello, A. B., Osborne, J. (2005). Best practices in exploratory factor analysis: four
777 recommendations for getting the most from your analysis. *Practical Assessment
778 Research & Evaluation*, 10, 173-178.
- 779 Cox, R., Hyde, M., Gatehouse, S., et al. (2000). Optimal outcome measures, research priorities,
780 and international cooperation. *Ear Hear.*, 21, 106S–115S.
- 781 Cronbach, L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16,
782 297–334.
- 783 Demeester, K., Topsakal, V., Hendrickx, J.-J., et al. (2012). Hearing disability measured by the
784 speech, spatial, and qualities of hearing scale in clinically normal-hearing and hearing-
785 impaired middle-aged persons, and disability screening by means of a reduced SSQ (the
786 SSQ5). *Ear Hear.*, 33, 615–616.
- 787 Fabrigar, L.R., Wegener, D.T., MacCallum, R.C., et al. (1999). Evaluating the use of
788 exploratory factor analysis in psychological research. *Psychol. Methods*, 4, 272–299.
- 789 Field, A., Miles, J., Field, Z. (2012). *Discovering Statistics Using R*, London ; Thousand Oaks,
790 Calif: SAGE Publications Ltd.
- 791 Flesch, R. (1948). A new readability yardstick. *J. Appl. Psychol.*, 32, 221–233.
- 792 Gatehouse, S., Noble, W. (2004). The Speech, Spatial and Qualities of Hearing Scale (SSQ).
793 *Int. J. Audiol.*, 43, 85–99.
- 794 Glickman, M.E., Rao, S.R., Schultz, M.R. (2014). False discovery rate control is a
795 recommended alternative to Bonferroni-type adjustments in health studies. *J. Clin.
796 Epidemiol.*, 67, 850–857.
- 797 Gunning, R., 1952. *The Technique of Clear Writing*. McGraw-Hill. pp. 36–37.

- 798 Herdman, M., Fox-Rushby, J., Badia, X. (1998). A model of equivalence in the cultural
799 adaptation of HRQoL instruments: the universalist approach. *Qual. Life Res. Int. J.*
800 *Qual. Life Asp. Treat. Care Rehabil.*, 7, 323–335.
- 801 Horn, J.L. (1965). A rationale and test for the number of factors in factor analysis.
802 *Psychometrika*, 30, 179–185.
- 803 Hutcheson, G.D., Sofroniou, N. (1999). Factor Analysis, Chapter 6 in *The Multivariate Social*
804 *Scientist: Introductory Statistics Using Generalized Linear Models*, London; Thousand
805 Oaks, Calif: SAGE Publications Ltd.
- 806 Hyde, M.L. (2000). Reasonable psychometric standards for self-report outcome measures in
807 audiological rehabilitation. *Ear Hear.*, 21, 24S–36S.
- 808 Kaiser, H.F. (1970). A second generation little jiffy. *Psychometrika*, 35, 401–415.
- 809 Kaiser, H.F., Rice, J. (1974). Little Jiffy, Mark IV. *J. Educ. Psychol. Meas.*, 34, 111-117.
- 810 Kandel, L., Moles, A. (1958). Application de l'Indice de Flesch à la langue française. *Cah.*
811 *Etudes Radio-Telev.*, 253–274.
- 812 Kang, R., Nimmons, G.L., Drennan, W., et al. (2009). Development and validation of the
813 University of Washington Clinical Assessment of Music Perception test. *Ear Hear.*, 30,
814 411–418.
- 815 Moulin, A., Pauzie, A., Richard, C., 2015. Validation of a French translation of the Speech,
816 Spatial, and Qualities of Hearing Scale (SSQ) and Comparison with other language
817 versions. *Int. J. Audiol.*, 54, 889-98.
- 818 Moulin, A., Richard, C., 2016. Sources of Variability of Speech, Spatial, and Qualities of
819 Hearing Scale (SSQ) Scores in Normal-Hearing and Hearing-Impaired Populations.
820 *Int. J. Audiol.* 55, 101-109.
- 821 Noble, W. (2013). *Self-Assessment of Hearing* Édition : 2nd Revised edition., San Diego, CA:
822 Plural Publishing Inc.

- 823 Noble W., Jensen N.S., Naylor G., Bhullar N. & Akeroyd M.A. (2013). A short form of the
824 Speech, Spatial, and Qualities of Hearing scale suitable for clinical use: The SSQ12.
825 *Int. J. Audiol.*, 52, 409–412.
- 826 Paz, S.H., Liu, H., Fongwa, M.N., et al. (2009). Readability estimates for commonly used
827 health-related quality of life surveys. *Qual. Life Res. Int. J. Qual. Life Asp. Treat. Care*
828 *Rehabil.*, 18, 889–900.
- 829 Perreau, A.E., Ou, H., Tyler, R., et al. (2014a). Self-reported spatial hearing abilities across
830 different cochlear implant profiles. *Am. J. Audiol.*, 23, 374–384.
- 831 Perreau, A.E., Spejcher, B., Ou, H., et al. (2014b). The spatial hearing questionnaire: data from
832 individuals with normal hearing. *Am. J. Audiol.*, 23, 173–181.
- 833 Petras, I., and Podlubny, I. (2010). Least Squares or Least Circles? *CHANCE* 23, 38–42.
- 834 Pichora-Fuller, M.K., (2003). Processing speed and timing in aging adults: psychoacoustics,
835 speech perception, and comprehension. *Int. J. Audiol.*, 42, 59–67.
- 836 Potvin, J., Punte, A.K., Van de Heyning, P. (2011). Validation of the Dutch version of the
837 Spatial Hearing Questionnaire. *B-ENT*, 7, 235–244.
- 838 Reck, R.P., Reck, R.A. (2007). Generating and rendering readability scores for Project
839 Gutenberg texts. *Corpus Linguist. Conf. Birm. UK*, #120.
- 840 Regnault, A., Herdman, M. (2014). Using quantitative methods within the Universalist model
841 framework to explore the cross-cultural equivalence of patient-reported outcome
842 instruments. *Qual. Life Res.* Available at: [http://link.springer.com/10.1007/s11136-014-](http://link.springer.com/10.1007/s11136-014-0722-8)
843 [0722-8](http://link.springer.com/10.1007/s11136-014-0722-8) [Accessed July 14, 2014].
- 844 Revelle, W. (1978). ICLUST: A cluster analytic approach to exploratory and confirmatory scale
845 construction. *Behav. Res. Methods Instrum.*, 10, 739–742.
- 846 Revelle, W. (1979). Hierarchical cluster-analysis and the internal structure of tests. *Multivariate*
847 *Behav. Res.*, 14, 57–74.

- 848 Revelle, W. (2014). *psych: Procedures for Personality and Psychological Research*,
849 Northwestern University, Evanston, Illinois, USA, [http://CRAN.R-](http://CRAN.R-project.org/package=psych)
850 [project.org/package=psych](http://CRAN.R-project.org/package=psych) Version = 1.4.8.
- 851 Rothman, K.J. (1990). No Adjustments Are Needed for Multiple Comparisons. *Epidemiology*,
852 1, 43–46.
- 853 Singh, S., Kong, Y.-Y., Zeng, F.-G. (2009). Cochlear implant melody recognition as a function
854 of melody frequency range, harmonicity, and number of electrodes. *Ear Hear.*, 30, 160–
855 168.
- 856 Singh, G., Pichora-Fuller, K.M. (2010). Older adults' performance on the speech, spatial, and
857 qualities of hearing scale (SSQ): Test-retest reliability and a comparison of interview
858 and self-administration methods. *Int. J. Audiol.*, 49, 733–740.
- 859 Stephens, D., Héту, R. (1991). Impairment, Disability and Handicap in Audiology: Towards a
860 Consensus. *Int. J. Audiol.*, 30, 185–200.
- 861 Streiner, D.L., Norman, G.R., Cairney, J. (2014). *Health Measurement Scales: A practical*
862 *guide to their development and use*, Oxford University Press.
- 863 Tyler, R.S., Perreau, A.E., Ji, H. (2009). Validation of the Spatial Hearing Questionnaire. *Ear*
864 *Hear.*, 30, 466–474.
- 865 Velicer, W.F. (1976). Determining the number of components from the matrix of partial
866 correlations. *Psychometrika*, 41, 321–327.
- 867 World Health Organisation. (1980). International Classification of Impairments, Disabilities, and
868 Handicaps. A manual of classification relating to the consequences of disease. World Health
869 Organisation. Geneva. Available at:
870 http://apps.who.int/iris/bitstream/10665/41003/1/9241541261_eng.pdf?ua=1

FIGURES

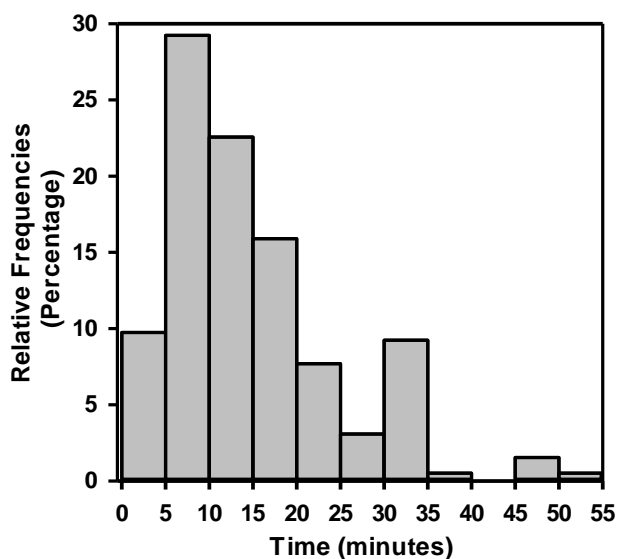


Figure 1. Relative frequencies (percentages) of the time, per 5-minute bins, that 202 hearing-impaired subjects reported taking to complete the Spatial Hearing Questionnaire.

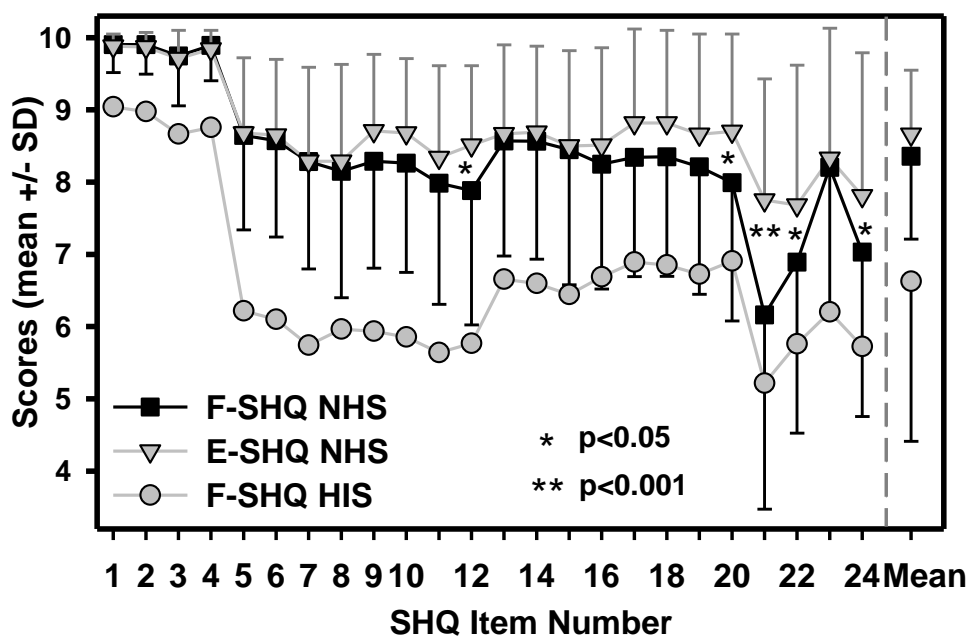


Figure 2. Mean scores (\pm standard deviation) per SHQ item for 98 normal-hearing subjects (NHS, black squares) and for 205 hearing-impaired subjects (HIS, gray circles). For comparison, the scores obtained by Perreau et al. (2014) with the E-SHQ in normal-hearing subjects are shown as gray triangles.

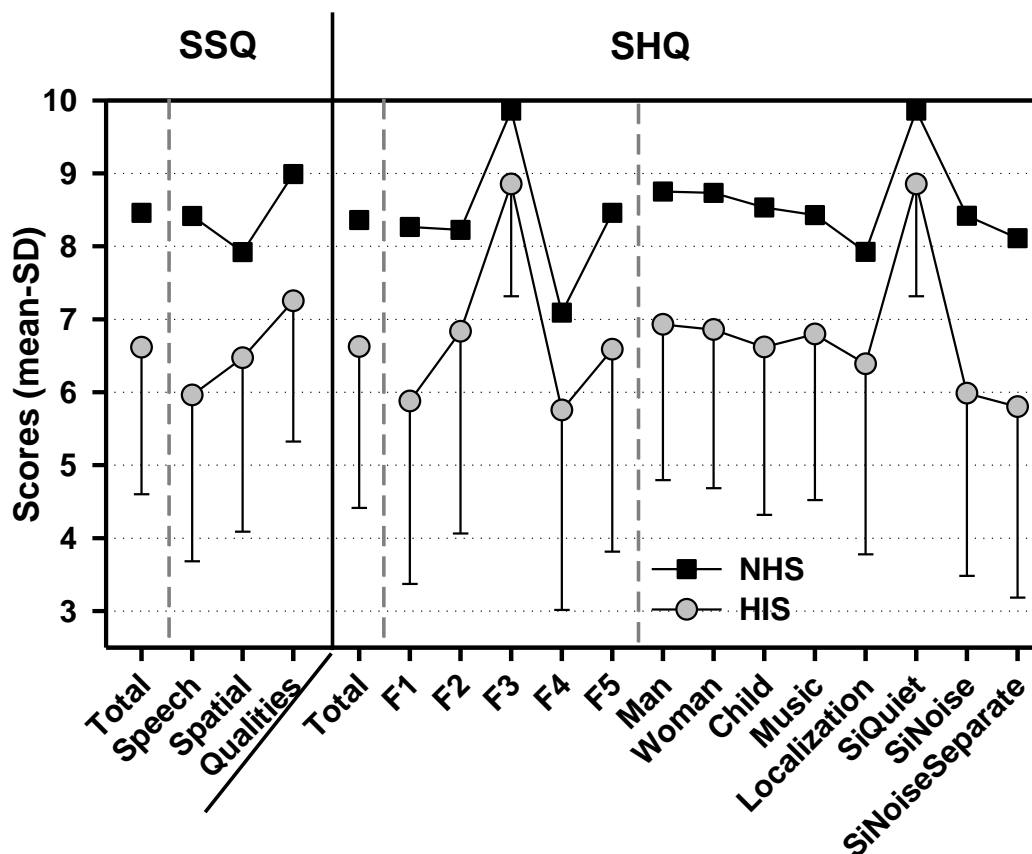


Figure 3. Spatial Hearing Questionnaire (SHQ) mean scores obtained in normal-hearing subjects (NHS, black squares, $n = 98$) and in hearing-impaired subjects (HIS, gray dots, $n = 205$). The 8 subscales and the 5 factors (F1 to F5) identified by factor analysis are shown. The mean scores of the Speech Spatial Hearing (SSQ) questionnaire and its 3 main subscales (Speech, Spatial and Qualities) obtained in the same subjects are also shown. The standard deviations are displayed as error bars for the hearing-impaired subjects only. Although they are identical, both factor F3 and the subscale “Speech in quiet” are shown.

Subscales	Question elements	Item	3 factors		4 factors		5 factors		Cluster		
			NHHIS	CIS *	NHS **	HIS ***	NHHIS	NHS	NHHIS		
			303	146 *	51 **	71 ***	303	98	303	303	
6. Speech in quiet	Speech in Quiet	man	F3 (6%)	F3 (5%)	F3 (12%)	F4 (5%)	F3 (6%)	F3 (11%)	F3 (6%)	C19 (3.6)	
		women									2
		child									3
		music									4
7. Speech in noise (Front)	Speech in Noise	man	F2 (8%)	F2 (13%)	F2 (15%)	F1 (68%)	F2 (8%)	F2 (13%)	F1 (70%)	C11	
		women									6
		child									7
		music									8
8. Speech in noise (separate)	noise at the side	man	F2 (8%)	F2 (13%)	F2 (15%)	F1 (68%)	F2 (8%)	F2 (13%)	F1 (70%)	C18 (6.9)	
		women									10
		child									11
		music									12
5. Locali- zation	location of (can't see)	man	F1 (70)	F1 (65)	F1 (48%)	F2 (11%)	F1 (70%)	F4 (5%)	F5 (2%)	C13 (3.5)	
		women									14
		child									15
		music									16
	location of (when behind you)	man			17						
		women			18						
		child			19						
		music			20						
	Locali- zation	Air plane			21						
		car coming			22						
car passing by		23									
loud noise		24									
					F4 (7%)	F3 (6%)	F1 (70)	F5 (3%)	F4 (2%)	C15 (4)	

Figure 4. Distribution of the different SHQ items on the factors obtained by factor analyses on 98 normal-hearing subjects (NHS), 205 hearing-impaired subjects (HIS) and the total sample (NHHIS). The leftmost columns describe the 24 items of the SHQ, showing the questions wording and the SHQ subscales defined by Tyler et al. (2009). The distribution of factors corresponding to 3-, 4- and 5-factor extractions are shown: the factors are sorted by the % of variance they explain (with gray background intensity decreasing as % of variance decreases). For comparison, factor analyses obtained in the literature for different populations are shown with asterisks, * for Tyler et al. (2009) for 146 cochlear implant subjects (CIS), ** for Perreau et al. (2014) for 51 NHS and *** for Potvin et al. (2011) for 71 HIS. The right-most column describes the cluster analysis, with the eigenvalues of each cluster shown.

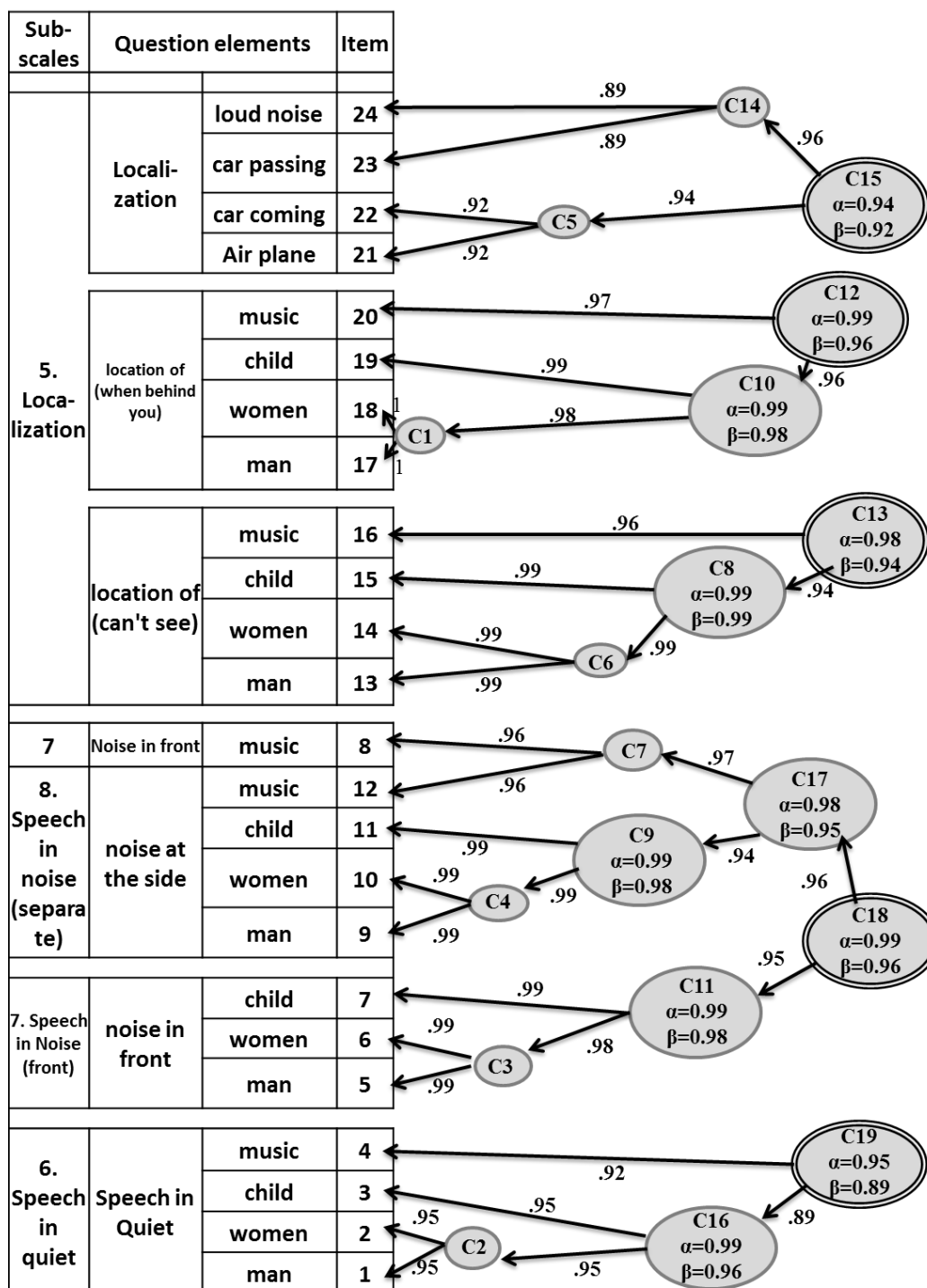


Figure 5. Tree-cluster diagram from the Revelle's cluster analysis of the F-SHQ data obtained in normal-hearing and hearing-impaired subjects ($n = 316$) and correspondence with each item and subscale of the SHQ. Highly similar items/clusters are combined first (clusters C1 to C7) and other items/clusters are added so that the internal consistency (measured by alpha) and/or the factorial homogeneity (measured by beta) of the cluster does not decrease. This procedure results in a hierarchical tree diagram of 5 distinct main clusters, shown by a double line (C12, C13, C15, C18 and C19) that displays the internal structure of the scale. Correlations coefficients between each cluster and sub-clusters are shown beside the arrows and alpha and beta coefficients are specified in the main clusters and sub-clusters.

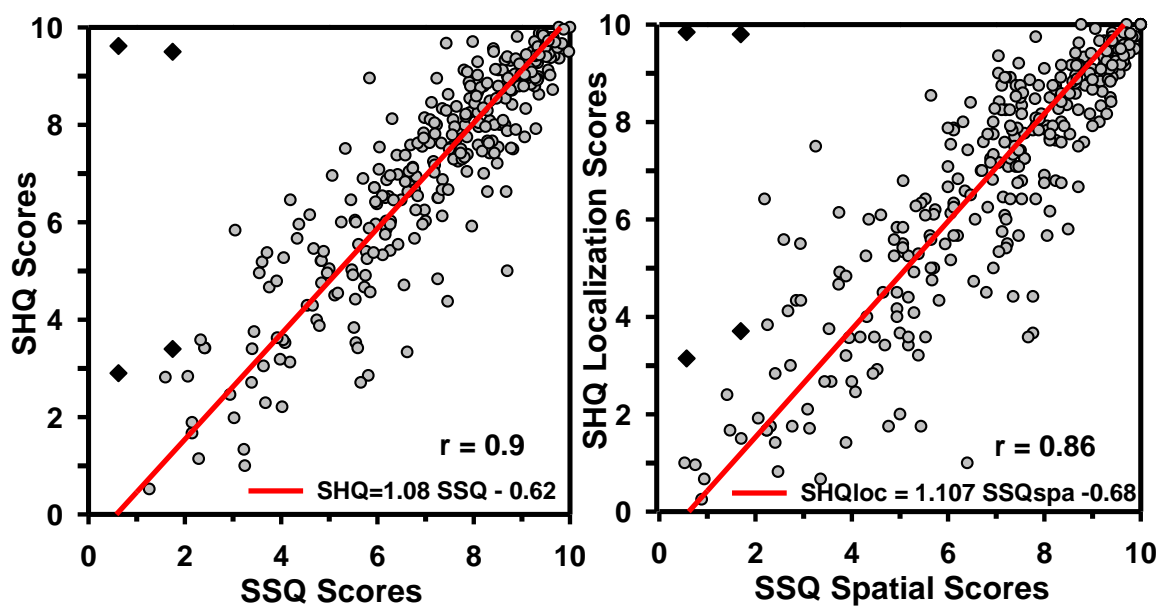


Figure 6: Scatterplots of the SHQ scores as a function of the SSQ scores, for the total score (left panel) and the SHQ localization subscale as a function of SSQ spatial scores, for 202 hearing-impaired and 98 normal-hearing subjects. Orthogonal distance regression lines are shown in red, along with their equations. Box plots in the top left corner of each panel shows the distribution of the SHQ (white boxes) and SSQ scores (grey boxes) with the 95th percentiles shown as black diamonds, the median as a solid line and the mean as a dashed line.

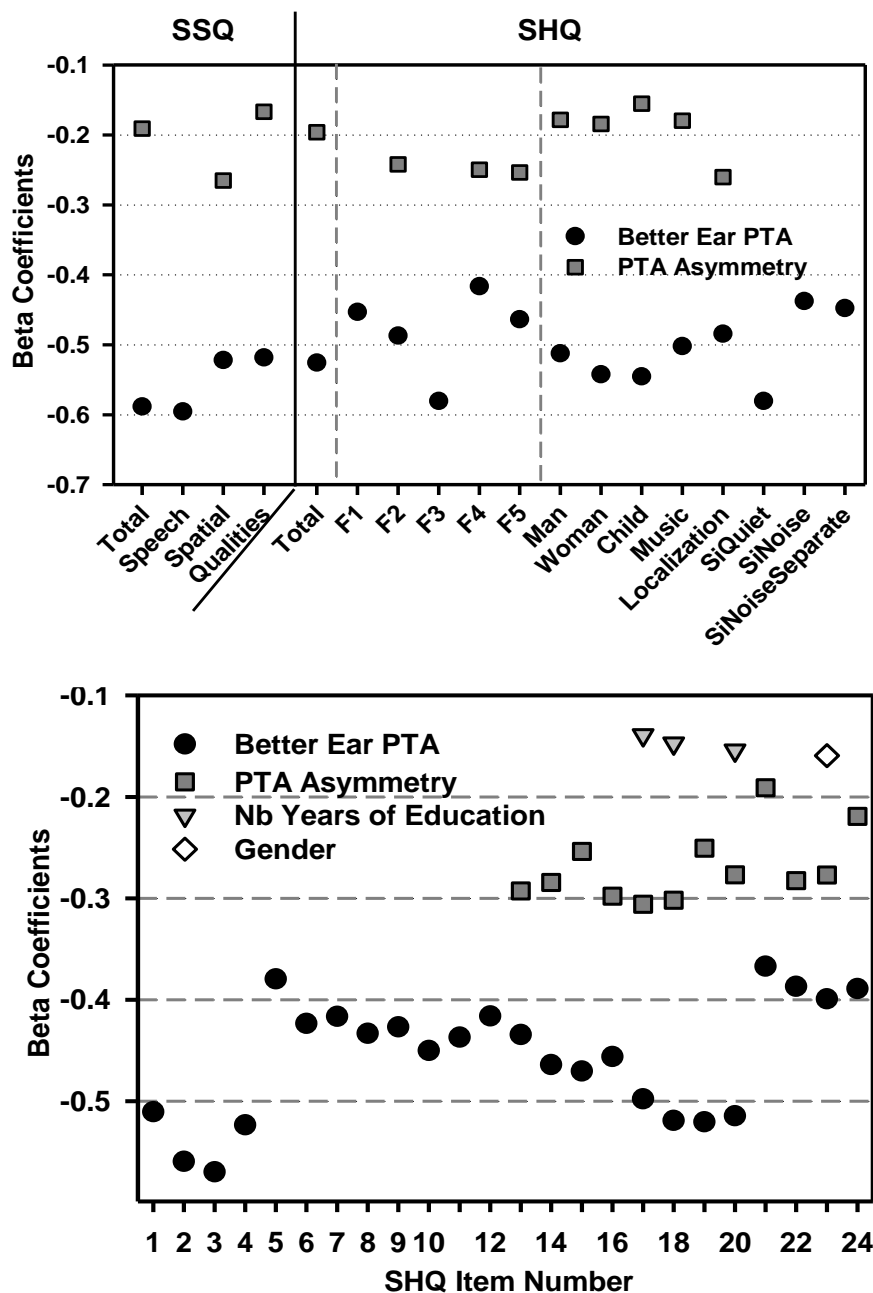


Figure 7a. Beta coefficients of multi-regression analysis of the Spatial Hearing Questionnaire (SHQ) scores, for 2 significant predictors: Better Ear PTA (black dots) and PTA Asymmetry (gray squares), obtained in hearing-impaired subjects ($n = 205$). Beta coefficients are shown for the total score (Total), each of the 5 factors (from F1 to F5) and the 8 SHQ subscales. The same analysis on the same population is shown for the Speech Spatial Hearing (SSQ) questionnaire total score (total) and the 3 main SSQ subscales (speech, Spatial and Qualities) for comparison.

Figure 7b. Beta coefficients of multi-regression analysis of the Spatial Hearing Questionnaire (SHQ) scores, for 4 significant predictors: Better Ear PTA (black dots), PTA asymmetry (gray squares), number of years of education (gray triangles) and gender (white diamonds) obtained in hearing-impaired subjects ($n = 169$). Beta coefficients are shown only for the significant predictors.

TABLES

Table 1. Pearson correlation coefficients obtained between SSQ scores and SHQ scores on 98 normal-hearing and 205 hearing-impaired subjects. Correlations between each SHQ subscale and SHQ_factors with each of the 3 main subscales of the SSQ (Speech, Spatial and Qualities) are compared. Correlation coefficients that were significantly stronger with one SSQ subscale than the other two SSQ subscales are on a gray background (Z scores, $p < 0.01$).

	SSQ	Speech	Spatial	Qualities
SHQ	0.90	0.83	0.85	0.82
SHQ_Man	0.87	0.82	0.80	0.80
SHQ_Woman	0.88	0.84	0.81	0.81
SHQ_Child	0.87	0.83	0.80	0.79
SHQ_Music	0.86	0.79	0.81	0.81
SHQ_Localization	0.86	0.74	0.86	0.77
SHQ_SiNoise	0.79	0.81	0.68	0.73
SHQ_SiNoiseSeparate	0.79	0.80	0.69	0.72
SHQ_F1	0.80	0.81	0.69	0.73
SHQ_F2	0.78	0.68	0.78	0.72
SHQ_F3 (SiQuiet)	0.73	0.71	0.61	0.73
SHQ_F4	0.81	0.68	0.85	0.72
SHQ_F5	0.82	0.71	0.82	0.74

Supplemental Digital Contents

Supplemental Digital Contents 1: French-language version of the Spatial Hearing Questionnaire (SHQ). F_SHQ2015.pdf.

Supplemental Digital Contents 2: Validation of the French language version of the spatial hearing questionnaire: normative data obtained in 98 normal-hearing subjects, item-to-total correlations and details of the different factor analyses. F_SHQ_VALIDATION.pdf.

Supplemental material 2, related to the paper:

Validation of a French-Language Version of the Spatial Hearing Questionnaire, Cluster Analysis and Comparison with the Speech, Spatial, and Qualities of Hearing Scale, by A. Moulin & C. Richard. *Ear and Hearing*, 2016.

1. Normative data and item-total correlation

Item	Normative values (n=98 NHS)			Item - total correlation	
	Mean	SD	- 2 SD	NHS	HIS
1	9.90	0.39	9.13	0.23	0.63
2	9.90	0.41	9.09	0.24	0.64
3	9.75	0.69	8.36	0.32	0.67
4	9.89	0.49	8.91	0.29	0.67
5	8.65	1.31	6.03	0.72	0.82
6	8.57	1.33	5.91	0.74	0.82
7	8.28	1.49	5.31	0.74	0.83
8	8.15	1.55	5.04	0.70	0.85
9	8.29	1.48	5.33	0.75	0.84
10	8.26	1.51	5.24	0.71	0.84
11	7.99	1.68	4.63	0.74	0.85
12	7.88	1.69	4.50	0.68	0.83
13	8.57	1.59	5.38	0.72	0.86
14	8.57	1.63	5.30	0.73	0.87
15	8.45	1.67	5.10	0.74	0.88
16	8.25	1.73	4.79	0.78	0.86
17	8.34	1.65	5.04	0.81	0.87
18	8.35	1.66	5.04	0.81	0.88
19	8.21	1.77	4.68	0.82	0.87
20	7.99	1.91	4.16	0.80	0.85
21	6.16	2.57	1.02	0.73	0.74
22	6.89	2.37	2.16	0.78	0.84
23	8.21	2.10	4.01	0.67	0.80
24	7.03	2.17	2.69	0.73	0.78

Table 1.

Mean scores (+/- standard deviation) obtained in normally hearing subjects (NHS) with item-total correlations for the NHS and the hearing-impaired subjects (HIS).

2. Factor analysis results of the Spatial Hearing Questionnaire.

Methods are briefly summarized here (detailed in the manuscript)

For factor extraction, we used the principal axis factoring method. For the number of factors to extract, the commonly used subjective Cattell's scree test (Cattell 1966) was used, with the more objective parallel analysis (Horn 1965) and Velicer's minimum average partial method (MAP) (Velicer 1976). An oblique factor rotation method (direct oblimin) was chosen and the percentages of variance explained by the different factors will be given before and after rotation.

NH subjects:

For NH subjects, the 3-factors option gave an underfactoring result, with a cumulative variance explained at 74% and RMSR (root mean square residuals) at least 0.06. The 3-factors option gave F1, a 12-item factor (#13 to #24) explaining 50% of variance; F2, an 8 item factor (#5 to #12) explaining 13% of variance and F3, a 4-item factor (#1 to 4) explaining 11% of variance (32%, 27% and 13% after rotation). Just one Item (#2) loaded on two factors: F1 (0.53) and F2 (0.33).

A 4-factor extraction was suggested by the parallel analysis and an eigenvalue above 1.0. The 4 factors explained 51%, 13%, 11% and 5% (cumulative 80% of variance explained) of the variance (35, 29, 20 and 16% after rotation), and the mean communalities were 0.8, with the 3 lowest values for #23 (0.46), #21 and #24 (0.61). The RMSR was 0.04. The second and third factors were identical to the 3-factor solution (Figure 2). The factor congruence coefficients with Perreau et al.'s data for F2 and F3 were 0.97 and 0.96, respectively. Item #23 loaded equally on factors 1 (0.27), 2 (0.27) and 4 (0.30).

As 5-factor extraction was recommended in the HIS and NHHIS group, we tried it for NHS. The cumulative variance explained was above 83% and the fifth factor explained 3% of the variance (32 %, 21 %, 18 %, 16 % and 13 % after rotation). The RMSR was below 0.03. The two lowest communalities were 0.57 (#23) and 0.68 (#8), with an average of 0.83. #23 loaded mostly on factor 5 (0.46), and on factor 3 to a lesser degree (0.36) (Figure 1)

HI subjects:

For HI subjects, the parallel analysis gave 3 factors which explained 70%, 8% and 6% of the variance (cumulative variance explained 84%) (48%, 37% and 15% after rotation). The lowest communalities were 0.63 for #21 and 0.66 for #4, with a mean of 0.84. The RMSR was 0.03. The distribution of the 3 main factors was identical to that observed for NH subjects, without any cross-loading. The congruence factors with Tyler et al.'s data were 95 for F3, 87 for F2 and 83 for F1 (for HI subjects).

A 4-factor extraction gave heavy cross-loadings, with items 13 to 16 loading on F1 and F4, and was abandoned.

The extraction of 5 factors, as determined by Velicer's MAP criterion (Velicer 1976), explained 89% of variance (with 70%, 8%, 7%, 2% and 2% for the 5 factors and 34 %, 19 %, 18 %, 15% and 14% after rotation). The lowest communalities were 67% and 73%, with a mean of 89%. The distribution of factor loadings on the 24 items was identical to the distribution mentioned for the total sample, with high loads for all of the items without cross-loadings (Figure 1).

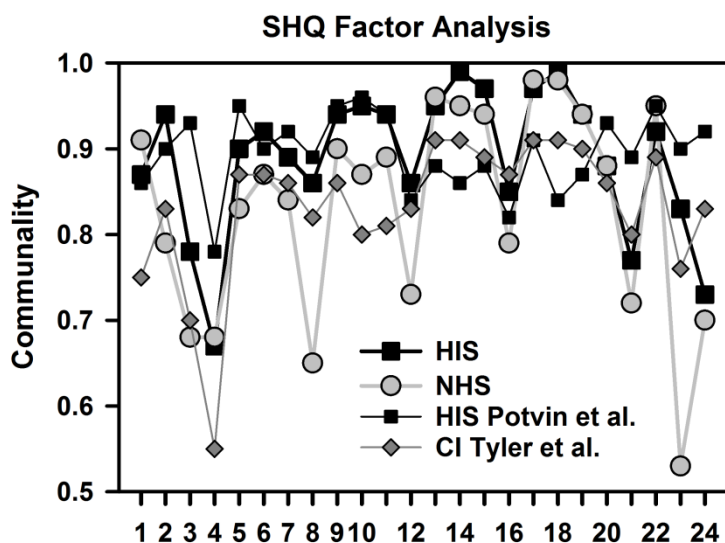


Figure 1 Communalities are shown, representing the amount of variance in each of the 24 items that is accounted for by the five retained factors, for the normal-hearing subjects (NHS, $n = 98$, black squares), the hearing-impaired subjects (205 HIS, gray dots). For comparison, literature data obtained in cochlear implant subjects (CIS) (small gray diamonds) by Tyler et al. (2009) and in 71 HIS (small black squares) by Potvin et al. (2011) are provided.

Total sample:

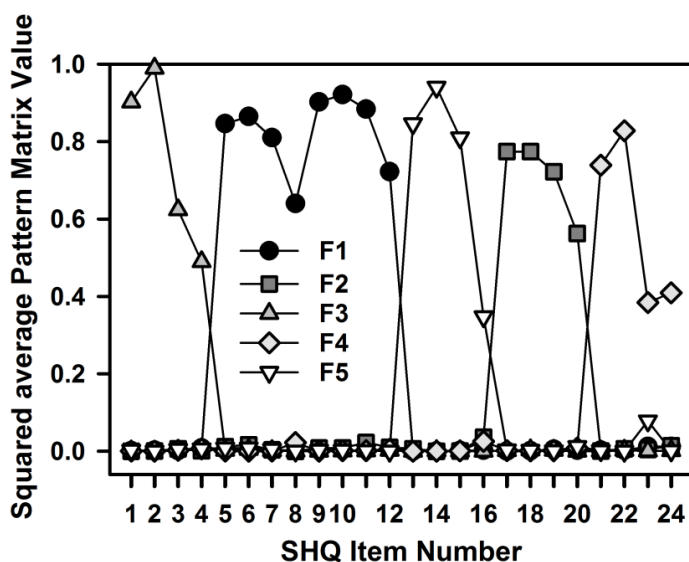


Figure 2. The square of the mean factor weights for each of the 24 SHQ items (i.e., the proportion of variance accounted for by each factor) obtained by factor analysis using a 5-factor extraction and oblique rotation on a group of normal-hearing and hearing-impaired patients ($n = 303$). The symbols mark the five rotated factors, by order of % of variance explained (from F1 to F5).

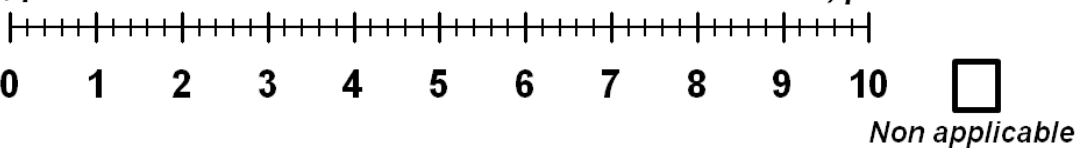
Questionnaire d'écoute spatiale

Conseils pour répondre aux questions

Les questions suivantes concernent vos capacités et votre expérience en matière d'audition et d'écoute dans le cadre de situations diverses.

Pour chacune des questions, vous devez mettre la réponse sur l'échelle située à droite, à l'aide d'une croix, à l'endroit choisi entre 0 et 10. Une croix sur la valeur 10 signifie que vous êtes parfaitement capable de faire ce qui est décrit dans la question correspondante. Une croix sur la valeur 0 indique que vous ne pouvez pas faire ce qui est décrit.

Exemple :

0. Vous pouvez suivre une conversation avec une personne en face de vous, dans le calme.	<p><i>Non, pas du tout</i> <i>Oui, parfaitement</i></p>  <p>0 1 2 3 4 5 6 7 8 9 10 <input type="checkbox"/></p> <p style="text-align: right;"><i>Non applicable</i></p>
--	--

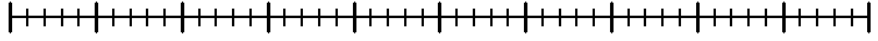
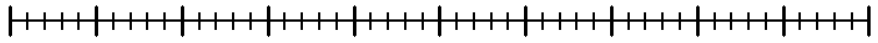
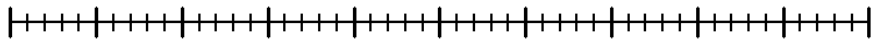
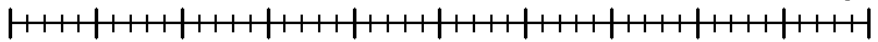
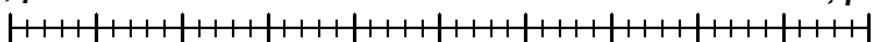
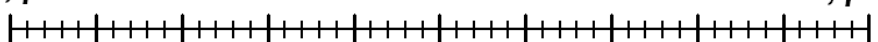
La question 0 se rapporte à votre capacité de suivre une conversation. Si vous êtes tout à fait capable de le faire, placez une croix sur l'extrémité droite de l'échelle. Si vous êtes capable de suivre environ la moitié de la conversation dans une telle situation, placez la marque vers le milieu de l'échelle (le 5) et ainsi de suite. Si vous n'arrivez pas du tout à suivre une conversation, placez la croix sur le 0, à gauche de l'échelle.

Nous espérons que toutes les questions sont pertinentes au regard de votre vie quotidienne. Si ce n'est pas le cas pour certaines questions, veuillez cocher la case « non applicable ». Merci de bien vouloir expliquer en quelques mots à côté de la question la raison pour laquelle elle n'est pas pertinente dans votre cas.

Date: _____

Heure de début :.....

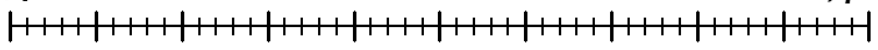
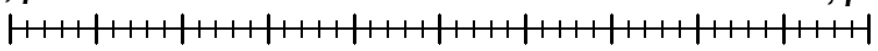
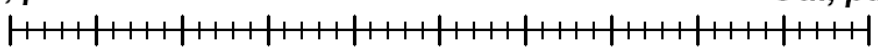
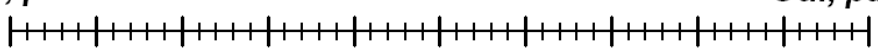
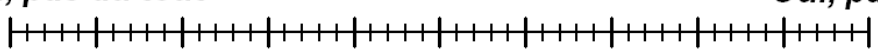
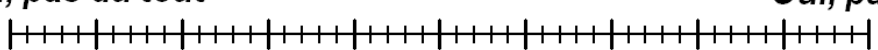
Questionnaire d'audition spatiale (SHQ) version 1.0

<p>1. Un homme, face à vous, vous parle. La pièce dans laquelle vous êtes est très calme. Est-il facile de le comprendre ?</p>	<p><i>Non, pas du tout</i> <i>Oui, parfaitement</i></p> <p>  </p> <p>0 1 2 3 4 5 6 7 8 9 10 <input type="checkbox"/></p> <p style="text-align: right;"><i>Non applicable</i></p>
<p>2. Une femme, face à vous, vous parle. La pièce dans laquelle vous êtes est très calme. Est-il facile de la comprendre ?</p>	<p><i>Non, pas du tout</i> <i>Oui, parfaitement</i></p> <p>  </p> <p>0 1 2 3 4 5 6 7 8 9 10 <input type="checkbox"/></p> <p style="text-align: right;"><i>Non applicable</i></p>
<p>3. Un enfant, face à vous, vous parle. La pièce dans laquelle vous êtes est très calme. Est-il facile de le comprendre ?</p>	<p><i>Non, pas du tout</i> <i>Oui, parfaitement</i></p> <p>  </p> <p>0 1 2 3 4 5 6 7 8 9 10 <input type="checkbox"/></p> <p style="text-align: right;"><i>Non applicable</i></p>
<p>4. Vous écoutez une musique d'intensité confortable, venant en face de vous. La pièce dans laquelle vous êtes, est très calme. Pouvez-vous entendre clairement la musique ?</p>	<p><i>Non, pas du tout</i> <i>Oui, parfaitement</i></p> <p>  </p> <p>0 1 2 3 4 5 6 7 8 9 10 <input type="checkbox"/></p> <p style="text-align: right;"><i>Non applicable</i></p>
<p>5. Un homme, face à vous, vous parle. Il y a un ventilateur bruyant juste derrière lui. Est-il facile de le comprendre ?</p>	<p><i>Non, pas du tout</i> <i>Oui, parfaitement</i></p> <p>  </p> <p>0 1 2 3 4 5 6 7 8 9 10 <input type="checkbox"/></p> <p style="text-align: right;"><i>Non applicable</i></p>
<p>6. Une femme, face à vous, vous parle. Il y a un ventilateur bruyant juste derrière elle. Est-il facile de la comprendre ?</p>	<p><i>Non, pas du tout</i> <i>Oui, parfaitement</i></p> <p>  </p> <p>0 1 2 3 4 5 6 7 8 9 10 <input type="checkbox"/></p> <p style="text-align: right;"><i>Non applicable</i></p>

Questionnaire d'audition spatiale (SHQ) version 1.0

<p>7. Un enfant, face à vous, vous parle. Il y a un ventilateur bruyant juste derrière lui. Est-il facile de le comprendre ?</p>	<p><i>Non, pas du tout</i> <i>Oui, parfaitement</i></p> <p> ----- ----- ----- ----- ----- ----- ----- ----- ----- ----- </p> <p>0 1 2 3 4 5 6 7 8 9 10 <input type="checkbox"/></p> <p style="text-align: right;"><i>Non applicable</i></p>
<p>8. Vous écoutez une musique d'intensité confortable, venant en face de vous. Il y a également un ventilateur bruyant en face de vous. Pouvez-vous entendre clairement la musique ?</p>	<p><i>Non, pas du tout</i> <i>Oui, parfaitement</i></p> <p> ----- ----- ----- ----- ----- ----- ----- ----- ----- ----- </p> <p>0 1 2 3 4 5 6 7 8 9 10 <input type="checkbox"/></p> <p style="text-align: right;"><i>Non applicable</i></p>
<p>9. Un homme, face à vous, vous parle. Il y a un ventilateur bruyant sur le côté. Est-il facile de le comprendre ?</p>	<p><i>Non, pas du tout</i> <i>Oui, parfaitement</i></p> <p> ----- ----- ----- ----- ----- ----- ----- ----- ----- ----- </p> <p>0 1 2 3 4 5 6 7 8 9 10 <input type="checkbox"/></p> <p style="text-align: right;"><i>Non applicable</i></p>
<p>10. Une femme, face à vous, vous parle. Il y a un ventilateur bruyant sur le côté. Est-il facile de la comprendre ?</p>	<p><i>Non, pas du tout</i> <i>Oui, parfaitement</i></p> <p> ----- ----- ----- ----- ----- ----- ----- ----- ----- ----- </p> <p>0 1 2 3 4 5 6 7 8 9 10 <input type="checkbox"/></p> <p style="text-align: right;"><i>Non applicable</i></p>
<p>11. Un enfant, face à vous, vous parle. Il y a un ventilateur bruyant sur le côté. Est-il facile de le comprendre ?</p>	<p><i>Non, pas du tout</i> <i>Oui, parfaitement</i></p> <p> ----- ----- ----- ----- ----- ----- ----- ----- ----- ----- </p> <p>0 1 2 3 4 5 6 7 8 9 10 <input type="checkbox"/></p> <p style="text-align: right;"><i>Non applicable</i></p>
<p>12. Vous écoutez une musique d'intensité confortable, venant en face de vous. Il y a également un ventilateur bruyant de côté. Pouvez-vous entendre clairement la musique ?</p>	<p><i>Non, pas du tout</i> <i>Oui, parfaitement</i></p> <p> ----- ----- ----- ----- ----- ----- ----- ----- ----- ----- </p> <p>0 1 2 3 4 5 6 7 8 9 10 <input type="checkbox"/></p> <p style="text-align: right;"><i>Non applicable</i></p>

Questionnaire d'audition spatiale (SHQ) version 1.0

<p>19. Est-il facile de déterminer exactement d'où vient la voix d'un enfant s'il est derrière vous?</p>	<p><i>Non, pas du tout</i> <i>Oui, parfaitement</i></p>  <p>0 1 2 3 4 5 6 7 8 9 10 <input type="checkbox"/></p> <p style="text-align: right;"><i>Non applicable</i></p>
<p>20. Est-il facile de déterminer exactement d'où vient une musique, si elle vient de derrière vous.</p>	<p><i>Non, pas du tout</i> <i>Oui, parfaitement</i></p>  <p>0 1 2 3 4 5 6 7 8 9 10 <input type="checkbox"/></p> <p style="text-align: right;"><i>Non applicable</i></p>
<p>21. Est-il facile de déterminer où se trouve un avion dans le ciel si vous ne le voyez pas?</p>	<p><i>Non, pas du tout</i> <i>Oui, parfaitement</i></p>  <p>0 1 2 3 4 5 6 7 8 9 10 <input type="checkbox"/></p> <p style="text-align: right;"><i>Non applicable</i></p>
<p>22. Vous entendez une voiture au loin mais vous ne la voyez pas. Est-il facile de dire précisément d'où elle vient?</p>	<p><i>Non, pas du tout</i> <i>Oui, parfaitement</i></p>  <p>0 1 2 3 4 5 6 7 8 9 10 <input type="checkbox"/></p> <p style="text-align: right;"><i>Non applicable</i></p>
<p>23. Imaginez-vous au bord d'une route, les yeux fermés. Est-il facile de dire dans quel sens passe la voiture que vous entendez?</p>	<p><i>Non, pas du tout</i> <i>Oui, parfaitement</i></p>  <p>0 1 2 3 4 5 6 7 8 9 10 <input type="checkbox"/></p> <p style="text-align: right;"><i>Non applicable</i></p>
<p>24. Vous êtes dans une pièce d'une maison et vous entendez un grand bruit. Est-il facile de dire à quelle distance s'est produit le bruit?</p>	<p><i>Non, pas du tout</i> <i>Oui, parfaitement</i></p>  <p>0 1 2 3 4 5 6 7 8 9 10 <input type="checkbox"/></p> <p style="text-align: right;"><i>Non applicable</i></p>

Questionnaire d'audition spatiale (SHQ) version 1.0

<p>25. Dans l'ensemble, les questions vous ont-elles semblées faciles ?</p>	<p><i>Non, pas du tout</i> <i>Oui, parfaitement</i></p> <p> ----- ----- ----- ----- ----- ----- ----- ----- ----- ----- </p> <p>0 1 2 3 4 5 6 7 8 9 10 <input type="checkbox"/></p> <p style="text-align: right;"><i>Non applicable</i></p>
<p>26. Les questions reflètent-elles bien les situations rencontrées tous les jours ?</p>	<p><i>Non, pas du tout</i> <i>Oui, parfaitement</i></p> <p> ----- ----- ----- ----- ----- ----- ----- ----- ----- ----- </p> <p>0 1 2 3 4 5 6 7 8 9 10 <input type="checkbox"/></p> <p style="text-align: right;"><i>Non applicable</i></p>

Merci de ne pas oublier de mentionner l'heure de fin de remplissage du questionnaire :