Disentangling preference ratings of concert hall acoustics using subjective sensory profiles

Tapio Lokki,[a] Jukka Pätynen, Antti Kuusinen, and Sakari Tervo
Department of Media Technology, Aalto University School of Science, P.O. Box 15500,
FI-00076 AALTO, Finland

(Received 19 January 2012; revised 31 July 2012; accepted 11 September 2012)

Subjective evaluation of acoustics was studied by recording nine concert halls with a simulated symphony orchestra on a seat 12 m from the orchestra. The recorded music was spatially reproduced for subjective listening tests and individual vocabulary profiling. In addition, the preferences of the assessors and objective parameters were gathered. The results show that concert halls were discriminated using perceptual characteristics, such as Envelopment/Loudness, Reverberance, Bassiness, Proximity, Definition, and Clarity. With these perceptual dimensions the preference ratings can be explained. Seventeen assessors were divided into two groups based on their preferences. The first group preferred concert halls with relatively intimate sound, in which it is quite easy to hear individual instruments and melody lines. In contrast, the second group preferred a louder and more reverberant sound with good envelopment and strong bass. Even though all halls were recorded exactly at the same distance, the preference is best explained with subjective Proximity and with Bassiness, Envelopment, and Loudness to some extent. Neither the preferences nor the subjective ratings could be fully explained by objective parameters (ISO3382-1:2009), although some correlations were found.

[http://dx.doi.org/10.1121/1.4756826]

PACS number(s): 43.55.Gx, 43.55.Hy [LW] Pages: 3148–3161

I. INTRODUCTION

Despite numerous earlier studies, human perception of concert hall acoustics is not fully understood yet. Recently, a sensory evaluation methodology for concert hall acoustics quality assessment was proposed,[1] to better understand the human perception of concert hall acoustics. This methodology uses individual vocabulary profiling[2] (IVP) to extract descriptive characteristics of concert halls and to create sensory profiles of the studied halls. In this paper, the methodology is further developed and applied to nine concert halls to study the mapping of individually elicited attributes, objective parameters, and subjective preferences. This approach allows a direct comparison between the subjective preference, objective parameters, and the sensory profiles of the halls, leading to a better understanding of human perception of concert hall acoustics.

Concert hall acoustics studies often concentrate on finding subjective, objective, or preference ratings of a selection of halls. The subjective evaluation is done *in situ* either by listening to the concerts and filling out questionnaires[3–8] or in laboratory conditions via virtual acoustics. Virtual acoustics techniques are based on convolving anechoic music signals with impulse responses, either captured from real halls[9,10] or simulated via room acoustics modeling.[11–15] Such techniques enable simultaneous comparisons of concert halls, even though the authenticity of the *in situ* listening is lost to some extent.

The objective measures of concert halls are straightforward to calculate using the ISO3382-1:2009 standard.[16] They can be calculated for the simulated or measured impulse responses. A few recent articles[1,17,18] suggest that the current standard objective metrics cannot explain all subjective perceptions. However, standard objective parameters are applied in this paper as there is no evidence that some other measures would perform any better.

Preference mapping[19] refers to a group of multivariate statistical techniques that are used to obtain a deeper understanding of the relationships between a descriptive sensory profile and subjective preferences of test subjects. Although preferences and acceptance of products are actively studied in the context of consumer and food science, there are only a few studies that have assessed the subjective preferences of audio or acoustics. In the domain of concert hall acoustics, preferences have been addressed by Beranek,[8] Schroeder *et al.*,[9] Soulodre and Bradley,[10] and Ando,[20] as well as Kahle.[5] These studies have mainly employed questionnaires and paired-comparisons in performing the preference judgments. In short, the results indicate that the overall acoustical preference is influenced by several factors, such as loudness, reverberance and clarity. There is also evidence that, in general, listeners can be divided into at least two groups according to their preference data: One that prefers reverberant or enveloping sound and another that prefers clear or defined sound. However, these investigations somewhat lack a refined methodology in order to reveal the sensory characteristics best predicting the preference ratings.

This paper presents three contributions to the field of concert hall acoustic studies. First, nine concert halls are measured for comparison with a loudspeaker orchestra, which simulates a symphony orchestra, such that the listening position is the same in all halls. Second, signal processing in

---

[a] Author to whom correspondence should be addressed. Electronic mail: tapio.lokki@aalto.fi

stimuli creation is utilized to render high quality spatial sound samples for listening tests. Third, the data analysis is further developed by including the mapping of individually elicited attributes, objective parameters, and subjective preferences of the nine concert halls studied.

This paper is organized as follows. The procedure to create the stimuli for the listening test and the methodology of the applied listening test are reviewed first. Then the main results of the subjective listening test with an IVP method are shown. In addition, objective and preference results are presented. Finally, all data are analyzed to understand the links between objective, subjective, and preference data. With the unraveled links, the preference ratings can be explained with the subjective characteristics, and it is shown that objective data can neither explain perfectly the subjective, nor preference data.

II. METHODS

The previous study by Lokki *et al.*[1] applied a loudspeaker orchestra as acoustic excitation to measure the halls and a three-dimensional sound capturing and coding algorithm to reproduce it in the laboratory. In addition, they used a listening test methodology that was based on individual attributes of the assessors. In the present study, some details of the processes were changed to raise the quality of samples and to make the listening test less time consuming. In this section the methods are briefly described.

A. Impulse response measurements with a loudspeaker orchestra and music

The studied concert halls were recorded by measuring the spatial impulse responses from all 24 channels (having, in total, 33 loudspeakers) of an enhanced version of the loudspeaker orchestra reported by Pätynen *et al.*[21] The used loudspeakers were Genelec model nos. 1029A, 1032A, and 8030. The layout of the loudspeaker orchestra is shown in Fig. 1. Although the directivities of the loudspeakers differ from the directivity of musical instruments,[22] the mismatch in directivities is not very large with the applied configuration.[23] In each receiver position, spatial impulse responses were captured twice with a six-channel intensity probe (Type 50 VI-1, G.R.A.S., Denmark). The first measurement was performed with a 100 mm spacer, and the second one, with a 25 mm spacer. The use of two spacers enabled the computation of good figure eight microphone response signals at a wide frequency range[24] when six omnidirectional responses are converted to a first order B-format impulse responses. Each loudspeaker on the stage was calibrated in each hall by measuring 87 dBA at 1 m distance when the loudspeaker emitted bandpass filtered (200–1000 Hz) white noise. All microphones were calibrated with the B&K 4231 calibrator (Brüel and Kjær, Nærum, Denmark).

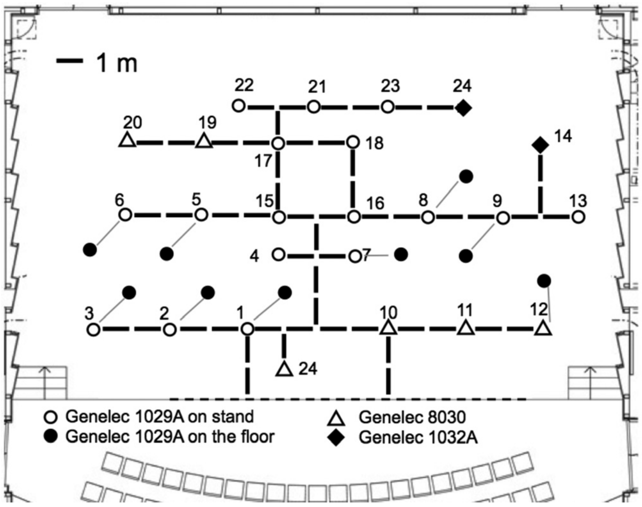


FIG. 1. Layout of the loudspeaker orchestra on the stage of a concert hall.

For spatial sound reproduction in the laboratory, the B-format impulse responses were first processed with the spatial impulse response rendering (SIRR) algorithm.[25,26] It divides a B-format impulse response in the time-frequency domain into individual impulse responses, one for each reproduction channel. In this study, one measured spatial impulse response was distributed to a 14-channel spatial sound reproduction system, consisting of eight loudspeakers at ear level at 45° intervals, four loudspeakers horizontally equispaced at 55° elevation above the ear level, and two loudspeakers 40° below ear level at azimuth angles $-22°$ and 22°. The processing of one measurement is illustrated in Fig. 2. In total, SIRR processing produced 672 impulse responses (24 source channels × 14 reproduction channels × 2 frequency ranges, crossover at 1 kHz) for convolution with the anechoic music.

The musical excerpts[27] convolved with SIRR processed impulse responses were as follows:

(a) *W. A. Mozart* (1756–1791), An aria of Donna Elvira from the opera *Don Giovanni,* Act II, Scene III, bars 1–5, 7 s;
(b) *L. van Beethoven* (1770–1827), Symphony No. 7, movement I, bars 14–16, 7 s; and
(c) *A. Bruckner* (1824–1896), Symphony No. 8, movement II, bars 41–46, 7 s.

The signals of individual instruments were convolved with the SIRR processed responses of the loudspeaker orchestra channels as presented previously.[1] As only one of each string instrument was recorded, the section sounds were done by copying the recordings. Each copy was individually processed with time varying delay, pitch shifting, amplitude modulation, and varying the microphone used in the recording, as string instruments have different timbre when recorded from different directions.[22] When these copies were reproduced from spatially separated loudspeakers, a natural and convincing string section sound was achieved.[28]

B. Concert halls

The studied concert halls are located in southern Finland. They are all used regularly for symphony orchestra concerts, although some of them are relatively small. Figure 3 illustrates the plans of the halls, configuration of the loudspeaker orchestra, and the recording position in each hall. The recording position was always 12 m from the nearest loudspeakers. Thus, the position was on row seven, eight,

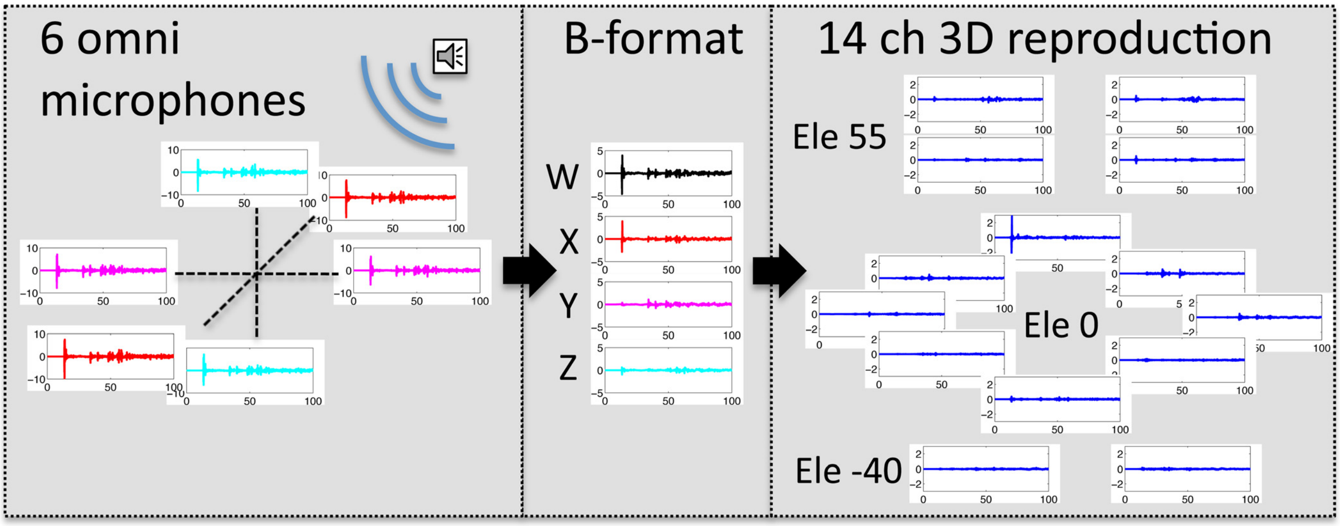
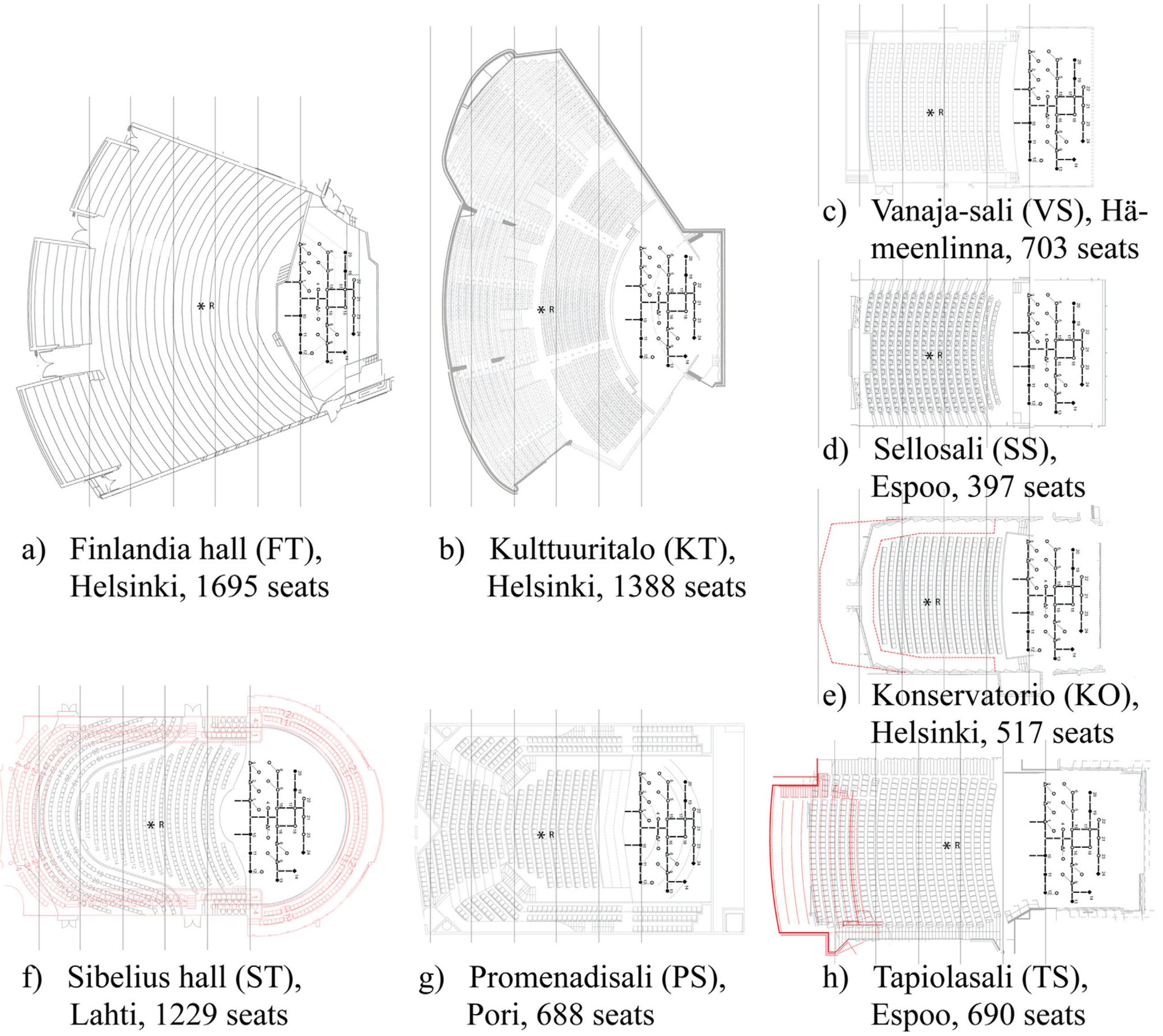


FIG. 2. (Color online) Processing of each measured spatial impulse response.



a) Finlandia hall (FT), Helsinki, 1695 seats

b) Kulttuuritalo (KT), Helsinki, 1388 seats

c) Vanaja-sali (VS), Hämeenlinna, 703 seats

d) Sellosali (SS), Espoo, 397 seats

e) Konservatorio (KO), Helsinki, 517 seats

f) Sibelius hall (ST), Lahti, 1229 seats

g) Promenadisali (PS), Pori, 688 seats

h) Tapiolasali (TS), Espoo, 690 seats

FIG. 3. (Color online) Plans of the studied concert hall in scale (distance between gray lines is 5 m). *R* is the recording position used in the listening test.

or nine, depending on the hall, see Fig. 3. This position is quite close, but as some halls are relatively small it is considered to be a central position in the main audience area. The ninth hall in the listening test was a hybrid hall, which had the direct sound for each source from hall ST [Fig. 3(f)], random artificial early reflections for each source, and the late reverberation, linearly faded in between 50 and 100 ms, from hall KO [Fig. 3(e)]. The 11 artificial early reflections were randomly distributed in time, with an echo density of 150/s. The level followed the 1.8 s early decay time (EDT) curve to be sure that reflections were not too loud compared to the direct sounds. The directions of reflections were semi-random such that first reflections came from frontal directions. Each reflection was filtered with the impulse response of the measurement loudspeaker from the direction defined by the direction of reflection.

C. Implementation of individual vocabulary profiling

The listening test was done with a sensory evaluation method called individual vocabulary profiling (IVP).[1,2] The screening of assessors and the IVP process were completed in three 2 h sessions for each assessor. The first session, designed to determine to the appropriateness of an assessor, worked as an introduction to samples and started the attribute elicitation process. The session began with pure tone audiometry and followed by a triangular AAB forced choice discrimination test. It used 24 pairs of the same samples that were evaluated in the whole process, thus simultaneously familiarizing the assessors with the samples. When completing the discrimination test, the assessors had to write down the discriminating feature of the sounds after every comparison. After the AAB test, this list of perceived differences was used as a starting point for verbal elicitation process in which the assessors listened to samples in free order at their own pace. After half an hour of listening, the assessors were asked to define four attributes with anchors and definitions with which he could order the samples. Affective attributes, such as preference or acceptance, were not allowed.

The second listening session started with listeners ordering the samples based on their own previously elicited attributes. After half an hour of listening, the test supervisor discussed the attributes with the assessor, to become convinced that the assessor felt confident with his attributes. At the end of the second session, the assessor performed the first rating with four of his own attributes and with three musical excerpts, i.e., he rated 12 stimuli sets each consisting of nine samples. The assessors did not know that the samples represented nine different concert halls; they were only ordering samples on a continuous scale with the attributes describing the perceived differences.

The final listening session was the second rating, in which the assessor had to rate the samples, presented in random order, with his own attributes, on a 120-point continuous unstructured line scale. Finally, to complete the whole process, the assessor rated the samples in his preference order with each musical signal. By asking the preference only in the end of the whole listening test process, it was guaranteed that the assessor was familiar with the samples and that the individual vocabulary evaluation process was not disturbed with preference.

Even though the whole process was quite extensive for each assessor, nobody complained about the length of the test and no listener fatigue was noticed. Each listening session for an assessor was on a separate day. During sessions, the assessors could have breaks when needed and some of them used this option for small breaks every now and then.

III. RESULTS

A. Reliability of the assessors and attributes

When performing sensory evaluations, it is mandatory to select assessors with care to ensure the quality of collected data. The suitability of assessors is typically reviewed in terms of their discrimination ability and reliability.[29,30] The assessors do not need to be experts in concert hall acoustics or classical music. It is more important that the assessors can hear differences between samples and can verbalize well what they hear. In our experience, however, people who often go to concerts and actively listen to recordings of classical music are motivated and good candidates. Therefore, potential assessors were openly invited with an article published in a national magazine of classical music. In addition, invitations were sent to student orchestra mailing lists, as well as to students of musicology and music.

Finally, 23 candidates (13 males), each of them with a musical background and between the ages of 19 and 75 years (average age of 35), participated in the listening tests. The screening of the assessors was performed with an audiometry and the AAB discrimination test. In addition, the reliability of assessors was addressed by checking whether assessors could replicate their ratings between the first and second ratings. As ratings with one attribute were done with all signals, the correlation of two matrices (3 signals × 9 halls) can be checked, e.g., with the RV coefficient with the Pearson type III approximation.[31] The *p*-value of the RV coefficient, indicating if the correlation is significant or not, was calculated with the FACTOMINER package.[32] For the whole data, the correlations of all 92 individual attributes are presented in Fig. 4. It can be seen that 60 out of 92 have $p < 0.05$, meaning that they were consistently and reliably repeated.

Table I collects the information of the screening and reliability analysis. All 23 candidates performed all tests, but the data of candidate numbers 3, 4, 6, 9, 14, and 22 show that they did not provide reliable enough data during the whole process. Main reasons for not including those six candidates are as follows. AS3 had a hearing loss (a threshold exceed 15 dB in at least one frequency band), AS4 and AS6 had too many errors (more than 6 out of 24) in the discrimination test, and AS3, AS9, AS14, and AS22 could reproduce none or only one reliable attribute rating. Possible reasons for unreliability are that they have changed their interpretation between the two ratings or these candidates would have required more training. The rest of the candidates, 17 assessors in total (average age of 31, 11 males), had no hearing problems, passed the discrimination test, and could reliably replicate ratings with 2–4 attributes. Therefore, 60 reliable

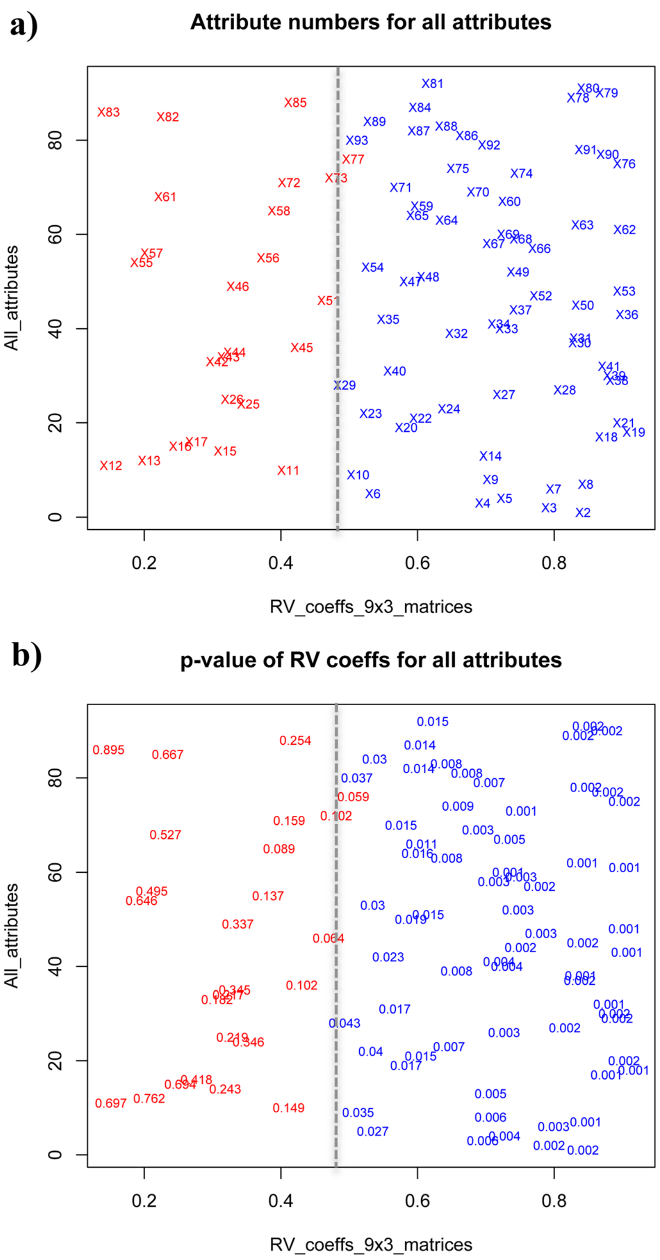


FIG. 4. (Color online) RV coefficients and their p values per attribute between first and second ratings (9×3 matrices).

TABLE I. Assessors[a] and screening results.

AS	Hearing	AAB errors	Number of reliable attributes	Selected
1	passed	1	4	YES
2	passed	0	4	YES
3	not passed	3	1	NO
4	passed	9	1	NO
5	passed	1	4	YES
6	passed	10	3	NO
7	passed	3	3	YES
8	passed	0	4	YES
9	passed	3	0	NO
10	passed	3	4	YES
11	passed	1	4	YES
12	passed	3	3	YES
13	passed	3	3	YES
14	passed	2	1	NO
15	passed	1	4	YES
16	passed	0	4	YES
17	passed	4	2	YES
18	passed	6	2	YES
19	passed	2	3	YES
20	passed	2	4	YES
21	passed	0	4	YES
22	passed	6	1	NO
23	passed	0	4	YES

[a] In total, 60 attributes from 17 selected assessors were included for the final analysis.

attributes, listed in Table II, were included in the final analysis.

B. Analysis based on the elicited attributes

The data of an IVP study can be analyzed with various multivariate methods. Often applied statistical methods are hierarchical clustering, Euclidean distance matrix, multiple factor analysis (MFA), and linear discriminant analysis.[1] Here, MFA[33,34] was used to extract the main principal components of the multidimensional space ordinating the samples. The results are presented in Table III revealing that the main principal component explains half of the variance in the

TABLE II. All 60 elicited attributes with their definitions.[a]

Group	Attribute	Low anchor	High anchor	Definition
Clarity	balance	strange emphasis	natural ensemble and location	how naturally different instrument groups sound in ensemble
	clarity	cut high frequencies	emphasized high frequencies	perception of even cut or emphasized high frequencies
	clear	no descant	too much descant	high pitches are perceived strongly
	clearness	haze	clearness	sounds are clearly separable and in balance
	distinguishable sources	stuffy sound	separating sound	how well individual instruments are distinguishable
	reverberance	thick	clear	how long the sounds reverberate
Definition	articulation	clumpy	defined	how clearly the tones can be distinguished
	clear	smudgy	clear	clearness of articulation
	definition	messy	clear	audibility and balance of different tones
	focus	blurred	focused	how sharp individual instruments can be localized

TABLE II. (*Continued*).

Group	Attribute	Low anchor	High anchor	Definition
Reverberance	amount of reverb	dry	a lot of space	ratio of the direct sound and reflections
	amount of reverb	dry	reverberant	influence of the space
	fullness	no colors	colorful	the tones of music, spatial impression
	reverberance	dry	wet	spaciousness in music / soloist vs orchestra
	reverberance	dry	reverberant	is reverberation dominating or does music sound dry
	reverberance	not much reverb	a lot of reverb	some samples have more reverberation
	reverberant	dry	reverberant	reverberant means music consisting a lot of reverberation
	strength	weak	strong	how loud the sample is
Envelopment	deepness	thin, narrow	full	thickness and size of sound/timbre
Spaciousness	distance	distant	intimate	distance of the sound source
Loudness	envelopment	frontal	enveloping	how sound envelops the listener
	fullness	thin	full	music is warm and seems to fill the space
	fullness	poor	rich	can I hear all instruments/tones well
	loudness	quiet	loud	sometimes music is louder
	loudness	quiet	loud	overall impression of loudness
	openness	filtered	open	how freely the sound is emitted from sources
	presence	absent	comprehensive	spatial impression of music to the audience
	reverberance	anechoic	reverberant	amount of reverb
	reverberation	dry	reverberant	how much reverberation the recording has
	reverberation	anechoic	reverberant	how much sound is reverberated in space
	shape of space	auditorium	church	quality and quantity of timbre and reverberation
	size of hall	tube-like, long, narrow	wide	how big is the hall (and what is the shape) where I am sitting
	size of orchestra	small	defined	how big area the orchestra covers
	width	narrow	wide	sound comes from the side when it is wide
	width	narrow, tube-like	broad, close to conductor	how spacious it feels
	width of stereo image	mono-like	wide stereo	how wide/narrow is the sound image
Bassiness	bassiness	poor bass	rich bass	how well low frequencies are reproduced
Warmth	bassiness	lack of bass	lot of bass	how much there is base line
Softness	bottom	no bottom	a lot of bottom	overridden bottom, is bass clear or muddy/short-handed
	darkness	cut low frequencies	emphasized low frequencies	impression of lack of low frequencies (cut) or emphasized low frequencies
	fleshy	no bass	a lot of bass	amount of bass and depth of space
	fullness	narrow	wide	is sound full (musical) or does something pop out
	juicy	cold	warm	how cold/warm it feels
	low tones	without bass	with bass	how well low tones are heard
	openness	tight	open	has the sound wide range or is it tight
	reverberation	dying	reverberant	sound is reverberated, it stays longer
	richness	rough	rich	rich sound consists of clarity, definition, softness, and roughness, everything in good balance
	sharpness	sharp	round	starts and ends of tones, naturalness of tones at low and high frequencies
	softness	hard	soft	soft timbre (ensemble sound) or is some instrument louder
	softness	row	soft	how individual tones are pop out from music
	warmth	cool	warm	how warm is timbre
Proximity	depth	restricted	deep	wide spectrum, spaciousness, three-dimensional
	distance	distant	close	how far the music seems to come
	distance	far away	near	some samples are near, some far away
	distance	distant	intimate	how far away are the musicians
	distance of source	far	close	in which place in a hall I think I am sitting
	intimacy	distant	present	feeling of naturally close music, or distant source, possibly distorted
Undefined	balanced	unbalanced	balanced	instruments/parts are in balance in music
	penetrating	pungent	soft	is music penetrating unpleasantly
	sharpness	sharp	soft	related to sound quality, wittiness of sound

[a] Grouping is based on AHC with three main principal components found with MFA analysis.

TABLE III. MFA analysis, variances explained by first ten components.

Component	Eigenvalue	Percentage of variance	Cumulative percentage of variance
1	13.40	49.92	49.92
2	2.70	10.07	59.99
3	1.90	7.07	67.06
4	1.21	4.52	71.58
5	1.01	3.76	75.34
6	0.79	2.93	78.27
7	0.71	2.66	80.93
8	0.66	2.47	83.39
9	0.61	2.26	85.65
10	0.53	1.96	87.60

whole data. In addition, the contribution of higher dimensions is rather small, although dimensions 2 and 3 together explain 17.14% of the variance. Dimensions from 4 to 27, explaining 32.94% of the total variance, are not believed to have any meaningful information, as the contribution of individual dimensions is negligible. Therefore, it was decided to group the elicited attributes based on the contribution of the attributes to the first three common principal dimensions. This grouping was performed with agglomerative hierarchical clustering (AHC) based on Euclidean distances, i.e., each data vector starts in its own cluster, and pairs of clusters are merged as one moves up the hierarchy. The clustering is done in conjunction with Wards minimum variance method,[35] i.e., squared Euclidean distance between data vectors.

The result of attribute grouping with AHC is presented in Fig. 5. The attributes are divided into three main groups, which are all further subdivided into smaller groups. The first main branch, consisting of Definition and Clarity attribute groups highlight the differences in clearness, articulation, and definition between the concert halls; see definitions of individual attributes in Table II. The main cluster is also divided into two subgroups, Reverberance and Envelopment/Loudness. The main cluster has the highest number of attributes and shows apparent differences in reverberance, loudness, openness, and width between samples. Finally, the third cluster is further divided into three subgroups containing attributes related to bassiness, richness, distance, and sharpness.

1. Clustering validation with Cronbach's α

To validate the attribute groups, Cronbach's α^{36} was used to investigate to what extent the attributes in one cluster are measuring the same thing. Cronbach's α is the sum of the individual variances of attributes divided by the total variance of the attributes inside a group. Thus, it is a measure of reliability or internal consistency of a multi-item scale. It is useful for evaluating how well different items of a multi-item scale measure the same underlying construct.

Table IV shows Cronbach's α's for the attribute groups. The groups having the highest number of attributes have the highest α, suggesting high inter-item correlation between individual attributes. It is known that Cronbach's α increases when the number of items rises, but correlations as high as those in Table IV suggest high inter-item correlation. If the group consists of only a few items, as the rest of the groups do, higher correlation is needed for the same α-value that is obtained with lower correlation between many items. Therefore, it is interpreted that individual Proximity attributes are highly correlated, Clarity and Definition groups have significant inter-item correlations, but three individual Undefined attributes are clearly not correlating, as suggested by their verbal definitions in Table II.

2. Ordination with multiple factor analysis

The ordination of the data is done with MFA and as mentioned earlier only the first three principal axes are considered meaningful in explaining variance of the data, as shown in Table III. Figures 6(a) and 6(b) show all 60 individual attributes on the factorial space defined by dimensions 1–2 and 1–3. By computing average directions with attributes in each group, defined in Table II, the average perceptual dimensions can be visualized. Figures 6(c) and 6(d) reveal that the variance of the Definition group is best explained by dimensions 1 and 2 in the north-west direction. The variance by the Clarity attribute group is better explained with dimension 3 as shown in Fig. 6(d). The main clusters—Envelopment/Loudness and Reverberance—are mainly explained by

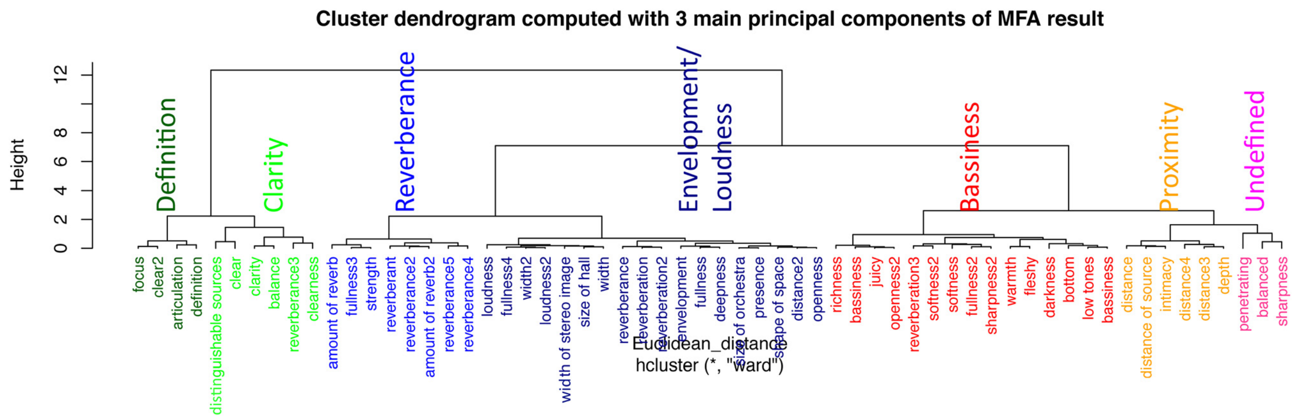


FIG. 5. (Color online) AHC clustering with the attributes contribution to three main principal components.

TABLE IV. Cronbach's α values for attribute groups found in Fig. 5 and Table II.

Attribute group	Number of attributes	Cronbach's α
Clarity	6	0.71
Definition	4	0.81
Reverberance	8	0.94
Envelopment/Loudness	18	0.98
Bassiness	15	0.96
Proximity	6	0.90
Undefined	3	0.22

dimension 1, although Reverberance contributes to the second dimension as well. The Proximity and Bassiness are clearly separated from the largest clusters in dimensions 1 and 2, although Bassiness is contributing to the third dimension as well. Finally, Figs. 6(e) and 6(f) reveal that the studied concert halls have significantly different acoustics as confidence ellipses[37] overlap only in a few cases. The confidence ellipses can be seen as contour lines of a bivariate normal distribution covering 95% of the bootstrapped values. In this case the bootstrap re-sampling is done for positions of all assessors with all three music samples.

C. Preference ratings

The 17 selected assessors were included in the analysis of the preference rating data. There is large variance between the assessors as can be seen in Fig. 7(a), which plots the means between the three music selections. Possible grouping

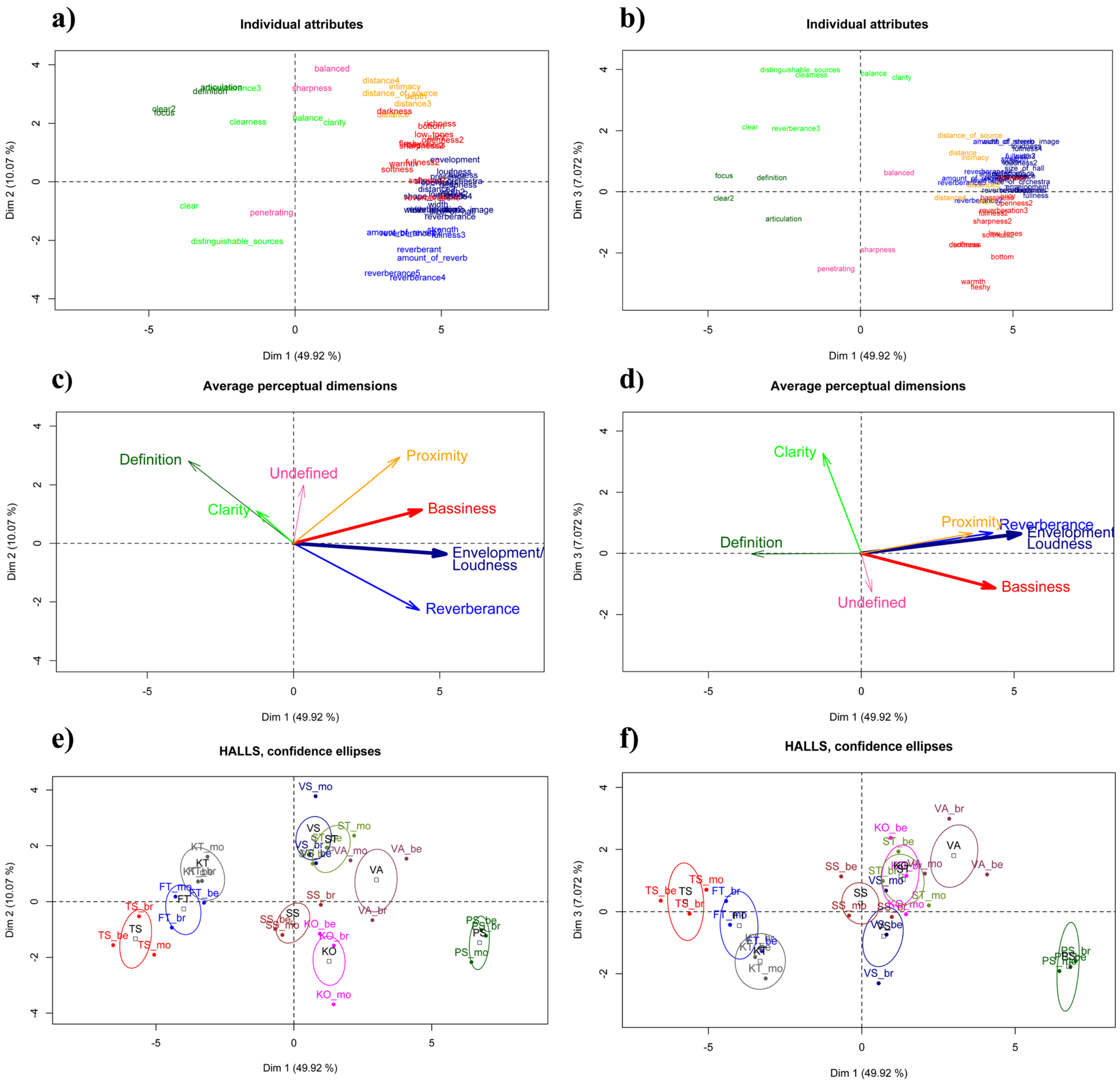


FIG. 6. (Color online) (a), (b) MFA with all 60 attributes. (c), (d) MFA with average vectors of attribute groups. The width of a vector is defined by the number of individual attributes in each group. (e), (f) ordination of concert halls with confidence ellipses.

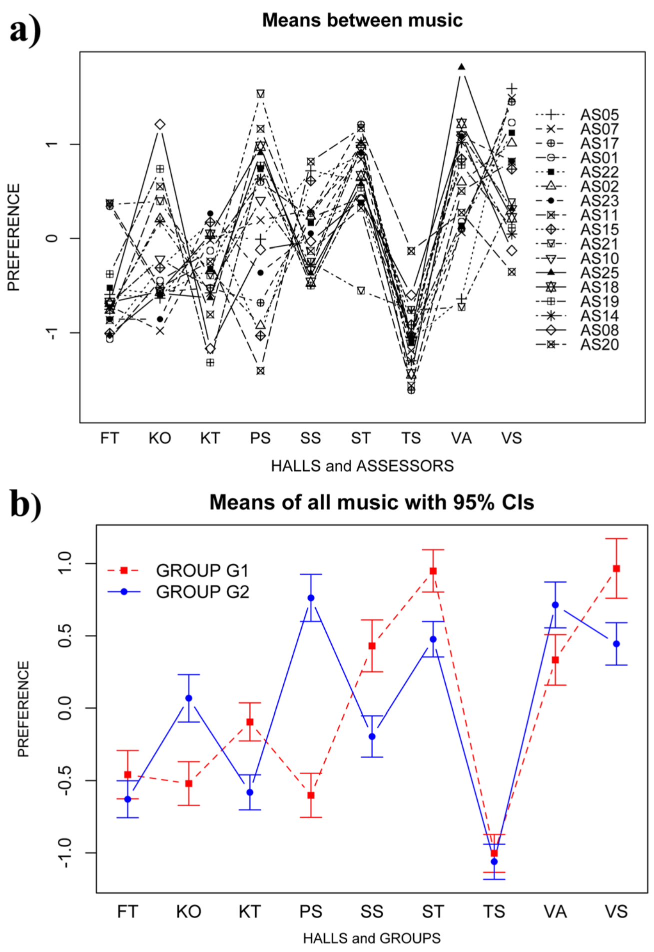


FIG. 7. (Color online) (a) Mean preference ratings of individual assessors show a large variance between individual preferences. (b) Means and 95% confidence intervals of assessor groups found with agglomerative hierarchical clustering.

of the assessors was analyzed with AHC and the analysis revealed that assessors can be grouped into two groups. The groups consist of seven (group G1) and ten (group G2) assessors and means of both groups are plotted in Fig. 7(b).

D. Objective parameters

The objective data, i.e., room acoustic parameters, were analyzed from the impulse responses measured from 24 loudspeaker channels to the receiver position in each hall. Table V shows the means of 24 values for each receiver position, computed according to the guidelines of the ISO3382-1:2009 standard.[16] The standard suggests the objective parameters and their relevant octave bands to describe subjective listener aspects, including strength (G), early decay time (EDT), clarity (C80), early lateral energy fraction (J_{LF}), and late lateral sound level (L_J). In addition, some other octave bands are added to cover a wider frequency range. Note that the measurements were not strictly according to the standard as the sound sources were not omnidirectional at all octave bands, although in practice, the used loudspeakers are omnidirectional up to 1000 Hz.

IV. MAPPING BETWEEN SUBJECTIVE, OBJECTIVE, AND PREFERENCE DATA

Preference-mapping techniques[19] allow the representation and preservation of the individuality of listener responses and allow the identification of listeners that tend to like the same types of sounds or have similar expectations for the sensory characteristics of a stimulus. There are namely two preference mapping methods: internal and external preference mapping. The internal preference mapping relies only on hedonic scores to determine the multidimensional representation of stimuli, whereas external mapping extends this approach by combining the descriptive sensory characteristics and the hedonic data. The term "mapping" is used because the results are graphically communicated and interpreted by a two-dimensional representation of the products in the sensory space.

Here, the preference mapping is done in common factorial space, thus it is considered neither internal nor external mapping. In contrast, the common factorial space is computed with all data to see the ordination of concert halls and to understand the relations between subjective, objective,

TABLE V. Acoustic quantities grouped according to listener aspects (in bold) according to ISO 3382-1 (2009) standard.[a,b]

Subjective listener aspect	Acoustic quantity	Averages of octave bands	Concert halls								
			FT	VS	KT	KO	ST	PS	SS	TS	VA
	G_lows	(dB) 125 and 250	4.17	6.36	5.71	5.38	6.75	9.71	5.40	2.98	3.84
Subjective level of sound	**G_mids**	**(dB) 500 and 1000**	**2.36**	**4.68**	**2.81**	**5.03**	**4.41**	**6.25**	**4.60**	**2.55**	**2.87**
	G_highs	(dB) 2000 and 4000	1.41	1.77	0.74	3.18	2.54	3.62	2.68	2.24	1.73
	EDT_lows	(s) 125 and 250	1.76	1.84	2.25	2.41	2.17	2.59	2.04	2.04	2.55
Perceived reverberance	**EDT_mids**	**(s) 500 and 1000**	**1.95**	**1.81**	**1.94**	**2.02**	**1.94**	**2.60**	**1.49**	**2.09**	**2.44**
	EDT_highs	(s) 2000 and 4000	1.78	1.53	1.59	1.78	1.47	2.00	1.36	1.41	2.02
	C80_lows	(dB) 125 and 250	−1.46	−1.09	−2.95	−3.49	−2.18	−4.57	−3.16	−2.77	−8.62
Perceived clarity of sound	**C80_mids**	**(dB) 500 and 1000**	**−0.70**	**0.18**	**−0.62**	**−2.81**	**0.90**	**−3.55**	**1.17**	**−1.41**	**−4.84**
	C80_highs	(dB) 2000 and 4000	2.24	2.31	2.61	1.05	3.47	0.24	2.20	3.53	−0.77
Apparent Source Width	**(ASW) J_{LF}**	**(%) 125–1000**	**16**	**21**	**14**	**20**	**22**	**27**	**19**	**18**	**31**
Listener Envelopment	**(ENV) L_J**	**(dB) 125–1000[c]**	**−10.3**	**−7.6**	**−8.5**	**−7.4**	**−7.3**	**−3.7**	**−8.5**	**−10.4**	**−8.4**

[a] Reference 16.

[b] Note that G and L_J are only relative values because the sources were not omnidirectional as defined in the standard (Ref. 16).

[c] Energy averaged.

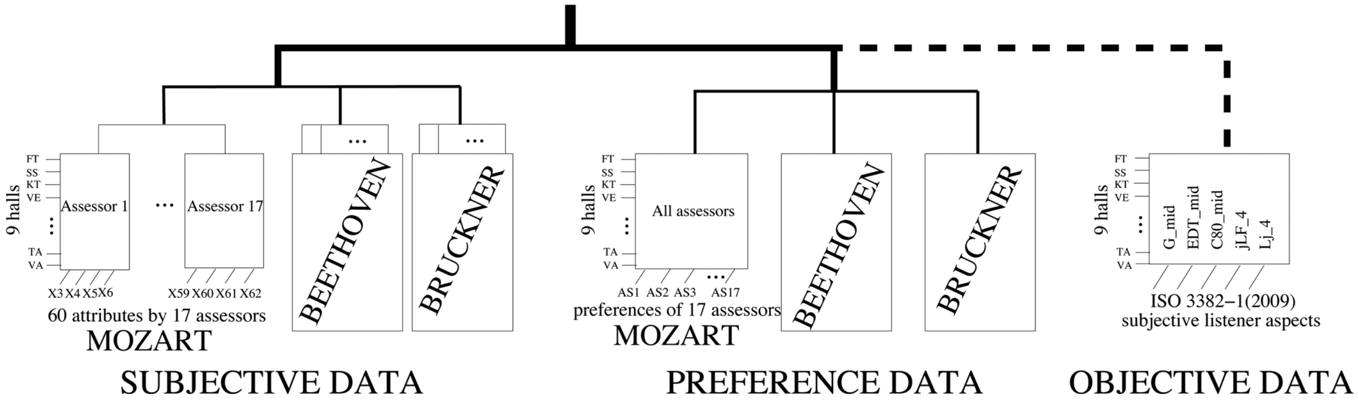


FIG. 8. Organization of the data for the Hierarchical MFA analysis. Different data can be linked in many ways; here subjective and preference data are analyzed first. The second analysis links takes also into account the objective data.

and preference data. Such analysis can be done, e.g., with hierarchical multiple factor analysis (HMFA).[38] The data are organized as shown in Fig. 8 and HMFA applies the MFA first for the subjective data and for the preference data of each musical piece. Finally, results of subjective and preference MFAs are linked with equal weights (33.3%) to the principal component analysis of the objective data to enable the comparison of all data in common factorial space. The objective data are scaled with just noticeable differences[16,39] to maintain the possible large variance in any of the parameters.

The analysis is done first with subjective and preference data. The variances explained by the first four principal components are seen in Table VI. As indicated by low eigenvalues on higher dimensions, only the first two dimensions provide meaningful results. The first visualization reveals the ordination of the concert halls. Figure 9(a) shows the ordination suggested by subjective and preference data. It can be seen that preference data pull data points more apart on the second dimension. However, this plot makes more sense when perceptual dimensions and directions explaining the variance in preference data are visualized in Fig. 10(a). Note that both subjective and preference data are averages of all music and all assessors. First, the orientation of the preference group G1 vector reveals that group G1 prefers concert

TABLE VI. HMFA analysis with subjective, preference, and objective data, variances explained by the first four components.

Component	Eigenvalue	Percentage of variance	Cumulative percentage of variance
Subjective and Preference data			
1	1.86	40.89	40.89
2	0.86	18.96	59.86
3	0.53	11.61	71.46
4	0.35	7.63	79.09
Subjective, Preference, and Objective data			
1	2.48	40.95	40.95
2	1.20	19.73	60.68
3	0.74	12.15	72.84
4	0.49	8.02	80.85

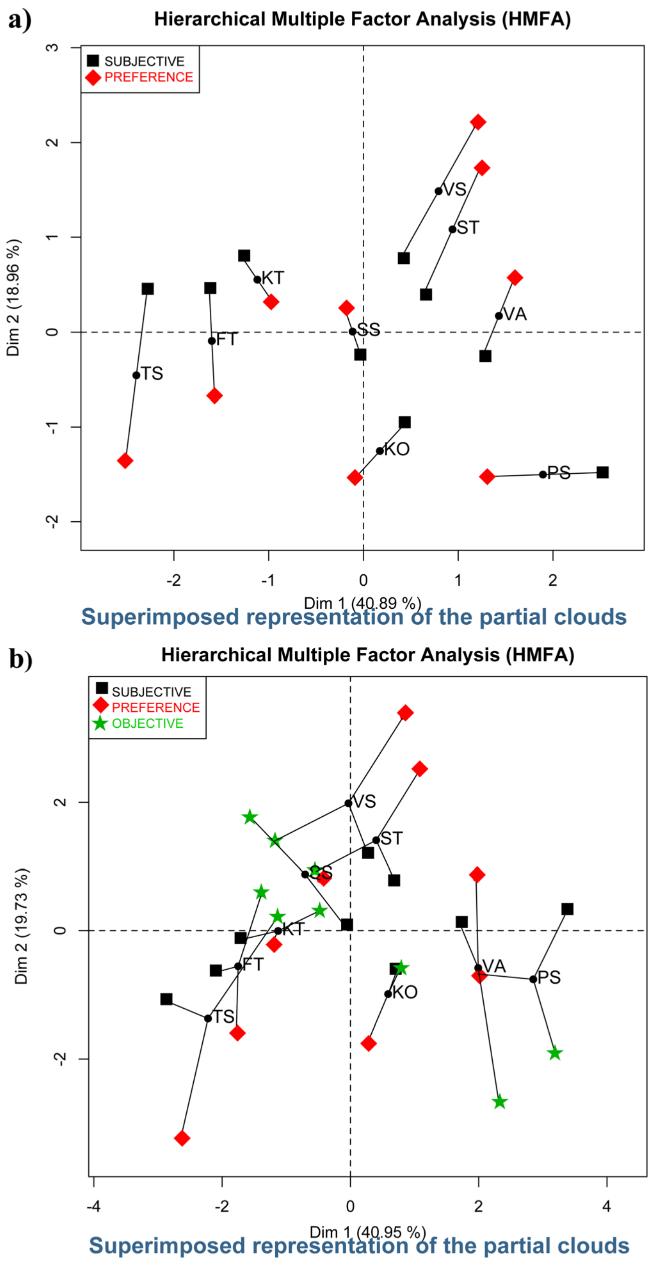


FIG. 9. (Color online) Ordination of concert halls in common factorial spaces. (a) HMFA result when subjective and preference data are analyzed together. (b) HMFA result for all data.

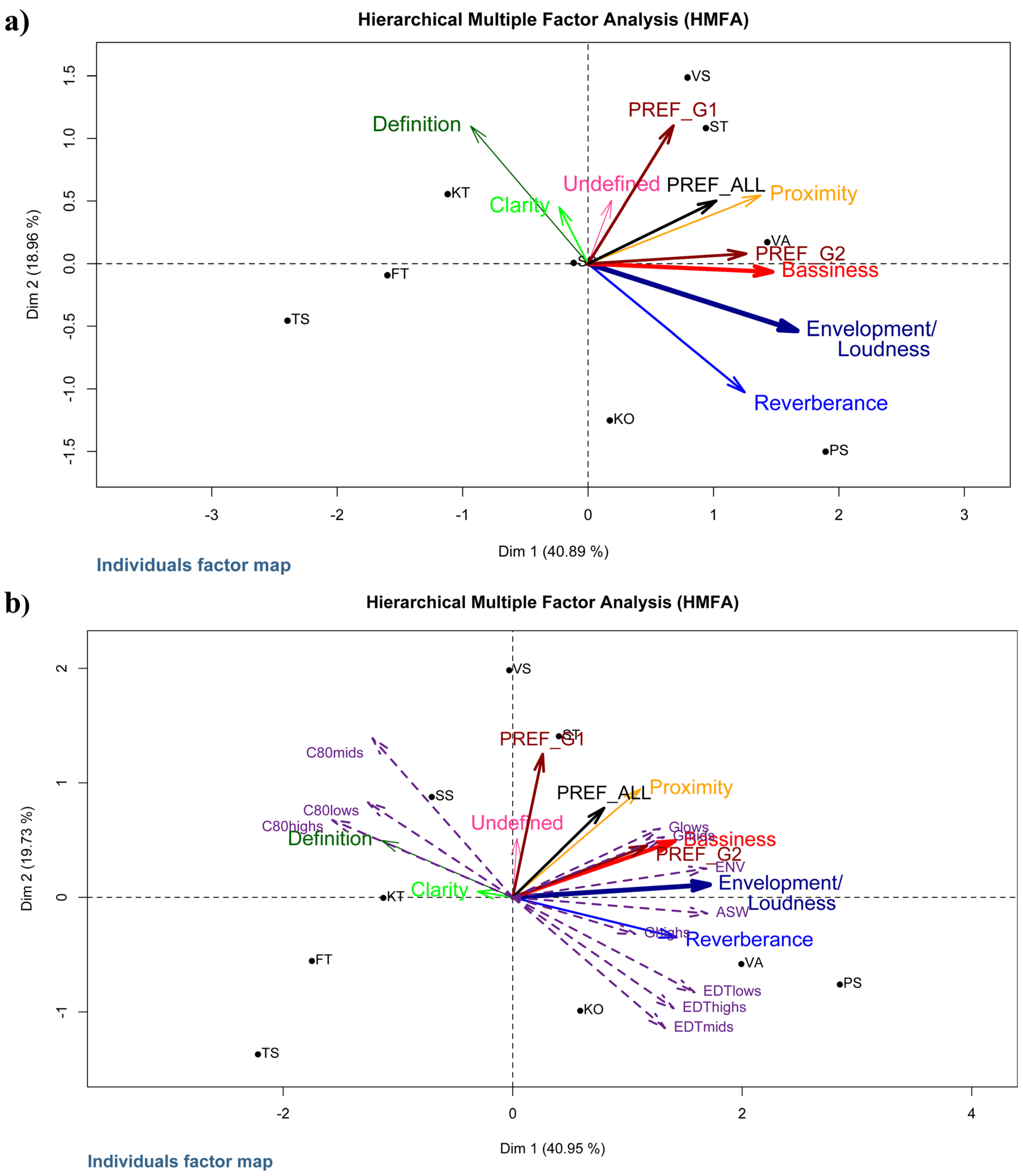


FIG. 10. (Color online) HMFA with average vectors of subjective attribute groups and average of all music. (a) Result with subjective and preference data. (b) Result with subjective, preference, and objective data.

halls VS and ST, i.e., the halls with relatively intimate and proximate sound with good definition. In other words, in these halls it is quite easy to hear individual instruments and melody lines and the Reverberance is moderate. In contrast, group G2 prefers louder and more reverberant sound with good envelopment and strong bass. They do not seem to pay attention to Definition, i.e., the sound could be muddier. Halls PS and VA are the most preferred by the assessors in group G2, as already indicated in Fig. 7(b).

When the objective data presented in Table V is linked to the analysis, it can be seen [Fig. 9(b)] that locations of halls SS and VA change more than other halls. In addition, it can be interpreted that objective data does not match well with the subjective data as in this joint analysis the objective data pull the data points to different directions than subjective data. Figure 9(b) clear shows the mismatch between objective and other data.

The analysis of all data in common factorial spaces [Figs. 9(b) and 10(b)] reveals that the subjective data are mainly explained by the first dimension, which consists of attributes related to Bassiness, Loudness, and Envelopment. In contrast, the preference data have more variance in the second dimension. This means that the preference order of concert halls cannot be explained only with the subjective difference in Loudness and Envelopment. The presented results suggest that preference can be explained better with differences in Definition, Proximity, and Reverberance.

The objective data separate the halls mainly on the Reverberance–Definition axis; see Fig. 10(b). However, objective parameters EDT and C80 at mid-frequencies are not perfectly aligned with the subjective Reverberance and Definition (and Clarity) as suggested by the ISO3382-1:2009.[16] Bassiness and Envelopment/Loudness are well correlated with low and mid-frequency G, L_J, and J_{LF}. An interesting fact is that neither Definition and Reverberance nor EDT and C80 explain the preference at all. In contrast, the preference is best explained with subjective Proximity and with Bassiness, Envelopment, and Loudness to some extent. Further, there is no objective measure that correlates to Proximity and overall average of preference.

A. Sensory profiles for studied concert halls

Based on the grouping of the individual attributes the sensory profiles of the halls can be formed. Such profiles are often visualized with spider plots.[1] Here, Fig. 11 visualizes "unwrapped spider plots" with a novel method to show profiles of all nine halls. In addition, the preference data are shown with the same method, i.e., ordering the halls with the means of the data.

First, on top of Fig. 11(a) the same data as in Fig. 7(b) is seen. Below, the sensory profiles of halls are visualized. The average preference order of the halls is closest to the average of the Proximity attributes, confirming the interpretation of the HMFA results. Three groups of concert halls, namely TS-FT-KT, VS-ST, and KO-VA-PS, also share similar profiles, to some extent. TS-FT-KT halls are the least preferred and they seem to render distant sound with the lack of bass, loudness, and reverberance. Figure 11(b) shows that objective G and L_J predicts the subjective result for these halls. In contrast subjective Reverberance is not well predicted with EDT, e.g., TS has the third longest EDT at mid-frequencies, but the lowest subjective Reverberance. Further, Fig. 10(b) shows that EDT orders the halls in the orthogonal direction than preference. This contradicts strongly with the conclusion by Beranek.[40]

Halls VS and ST were the most preferred by assessors in group G1. These two halls have pretty similar sensory profiles. They render the most intimate sound that contains enough bass and loudness. They have mild reverberance with well-defined sound. With these two halls the objective parameters predict the subjective attributes quite well, although the objective and subjective data locate these halls quite differently; see Fig. 9(b). Interestingly, the change in location is to the direction of subjective Proximity, the direction that none of the objective parameters explains.

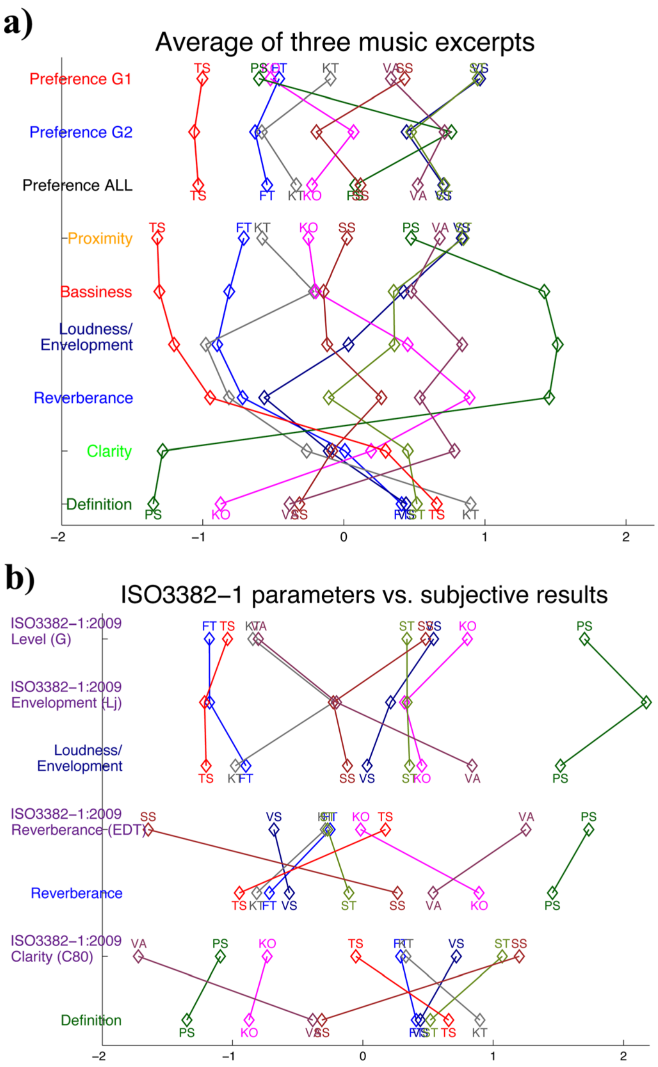


FIG. 11. (Color online) (a) Sensory profiles of the studied concert halls. (b) Subjective listener aspects proposed in the ISO3382-1 (2009) standard compared with subjective results of the IVP process.

Halls KO, VA, and PS have all different profiles. The most preferred halls by group G2 (PS and VA) render close sound with a lot of bass, loudness, envelopment, and reverberance. The definition is very low, but subjective clarity is very diverse within these three halls. The hall VA has very unusual objective parameters because there was no diffuse early energy in the responses due to sharp artificial early reflections resulting in much less early energy than in real measured impulse responses.

V. DISCUSSION

The main overall preference driver in this study was an attribute cluster interpreted as Proximity (related to distance), which correlates highly with the average of all preference ratings. In addition, it is very interesting that the individual differences in preference judgments are manifested in the second perceptual dimension, which is composed on one side by Reverberance attributes and on the other side by Clarity and Definition. In other words, it seems that although high Proximity is something essential for acoustical engagement, the different acoustical "tastes" are manifested by the levels of reverberation, fullness, clarity, and definition. However, it should be kept in mind that common to all preferences is loud enough and enveloping sound as all preference ratings correlates with the first dimension in Fig. 10(a).

The influence of different music as an excitation is not presented here in detail. With all music the results are quite close to each other, but there are also some significant differences. Mozart contains a soprano soloist and the assessors commented that they often concentrated on listening to her. This is probably the main reason why subjective Mozart results are slightly different than the results obtained with Bruckner and Beethoven. In particular, the Proximity attributes for Mozart gave slightly different results than with other music as the Proximity of the singer is very easy to evaluate. Most assessors also preferred halls that render close and intimate human voice. The detailed analysis with different music is left as a future work.

A. Results related to previous preference and subjective studies

Several studies with various techniques have been done in the past. Here, the presented results are compared with some of them.

Hawkes and Douglas[3] found four to six individual factors in their studies involving listening to real symphony orchestras *in situ*. The same factors were found here, such as *reverberance*, *definition*, *brilliance*, and *intimacy*. Soulodre and Bradley[10] found that preference correlated best with *clarity* and *treble,* but also to *loudness*. Sotiropoulou *et al.*[6] found that ordinary concert-goers describe their acoustical experiences with *body* (full-bodied, full, voluminous), *clarity* (clear, distinct), *tonal quality* (of smooth tone, of rich tone), and *proximity* (near, enveloping). These findings are well in line with the results of this study; however, they did not study preference as such.

Here, it was found that assessors can be grouped to two preference groups. Similar grouping has been found also earlier by Schroeder *et al.*,[9] who found similar preference groups related to loud sound and clear sound. Barron[4] divided assessors into groups by intimacy and reverberance. There, results also correlate with the results presented here; one group preferred clear and intimate sound and another group preferred loud, enveloping, and reverberant sound.

The high correlation between overall preference and subjective Proximity was surprising considering that all halls were recorded exactly at the same distance from the loudspeaker orchestra. As none of the standardized objective parameters could explain this, it raises a question of what makes sound close, intimate, and engaging. Recently, it has been suggested that the phase of early reflections affects the perceived bass and engagement.[41] In addition, the sound could be perceived closer if there are lateral early reflections, instead of median plane reflections.[42] The presented results support these ideas as the less intimate sound was perceived in fan-shaped halls (i.e., no lateral early reflections). In addition, the closest sound was perceived in two halls in which the first two early reflections are from flat large surfaces from the side (i.e., the reflections are coherent having the same phase at all frequencies with the direct sound). More investigations are needed to validate these findings.

VI. CONCLUSION

A loudspeaker orchestra was used as an acoustic excitation source, in order to listen to the exact same music in various concert halls. The sound in the concert halls, at exactly the same distance in each hall, was reproduced with spatial impulse responses and convolution, resulting in nine concert hall presentations in which all other variables except the hall were fixed. Seventeen out of 23 potential assessors completed the individual vocabulary profiling process to provide subjective sensory profiles of the concert halls. In addition, they ordered the halls with preference judgments.

The collected subjective, objective, and preference data were analyzed in common factorial space. The results show that the main discriminative attributes between halls are loudness, envelopment, and reverberance. The second large cluster of attributes consists of bassiness and proximity attributes. The third main perceptual dimension has definition and clarity attributes. The preference judgments were divided into two groups of assessors, the first preferring concert halls with loud, enveloping and reverberant sound. The second group preferred concert halls that render intimate and close sound with high definition and clear sound. All assessors dislike the concert halls with weak and distant sound. The best correlation with average preference ratings of all assessors was found to be with subjective proximity. This was quite interesting as the halls were recorded exactly at the same distance. Finally, none of the standardized objective room acoustical parameters could explain the proximity and preference data.

ACKNOWLEDGMENTS

The authors thank Heikki Vertanen for help in implementing the listening test and Dr. Philip Robinson for proofreading and comments. The research leading to these results has received funding from the Academy of Finland (Project Nos. 218238 and 140786) and the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013)/ERC Grant Agreement No. 203636.

[1] T. Lokki, J. Pätynen, A. Kuusinen, H. Vertanen, and S. Tervo, "Concert hall acoustics assessment with individually elicited attributes," J. Acoust. Soc. Am. **130**, 835–849 (2011).

[2] G. Lorho, "Individual vocabulary profiling of spatial enhancement system for stereo headphone reproduction," in *The 119th Audio Engineering Society Convention*, New York (2005), Paper No. 6629.

[3] R. Hawkes and H. Douglas, "Subjective acoustics experience in concert auditoria," Acustica **24**, 235–250 (1971).

[4] M. Barron, "Subjective study of British symphony concert halls," Acustica **66**, 1–14 (1988).

[5] E. Kahle, "Validation d'un modèle objectif de la perception de la qualité acoustique dans un ensemble de salles de concerts et d'opéras (Validation of an objective model for characterizing the acoustic quality of a set of concerts hall and opera houses)," Ph.D. thesis, Université du Maine, Le Mans (1995).

[6] A. Sotiropoulou, R. Hawkes, and D. Fleming, "Concert hall acoustic evaluations by ordinary concert-goers: I, Multi-dimensional description of evaluations," Acustica **81**, 1–9 (1995).

[7] A. Sotiropoulou and D. Fleming, "Concert hall acoustic evaluations by ordinary concertgoers: II, Physical room acoustic criteria subjectively significant," Acustica **81**, 10–19 (1995).

[8] L. Beranek, *Concert and Opera Halls—How They Sound* (Acoustical Society of America, New York, 1996).

[9] M. Schroeder, G. Gottlob, and K. Siebrasse, "Comparative study of European concert halls: Correlation of subjective preference with geometric and acoustics parameters," J. Acoust. Soc. Am. **56**, 1195–1201 (1974).

[10] G. Soulodre and J. Bradley, "Subjective evaluation of new room acoustic measures," J. Acoust. Soc. Am. **98**, 294–301 (1995).

[11] O. Warusfel, C. Lavandier, and J. Jullien, "Perception of coloration and spatial effects in room acoustics," in *Proceedings of the 13th International Congress on Acoustics (ICA'89)*, Belgrade, Yugoslavia (1989), pp. 173–176.

[12] C. Lavandier, "Validation perceptive d'un modèle objectif de caractérisation de la qualité acoustique des salles (Perceptual validation of an objective model for characterizing the acoustical quality of rooms)," Ph.D. thesis, Université du Maine, Le Mans, France (1989).

[13] J. Bradley and G. Soulodre, "Objective measures of listener envelopment," Université du Maine, Le Mans **98**, 2590–2597 (1995).

[14] Y. J. Choi and F. R. Fricke, "A comparison of subjective assessments of recorded music and computer simulated auralizations in two auditoria," Acta Acust. Acust. **92**, 604–611 (2006).

[15] J. Pätynen and T. Lokki, "Evaluation of concert hall auralization with virtual symphony orchestra," J. Build. Acoust. **18**, 349–366 (2011).

[16] ISO 3382-1:2009, *Acoustics—Measurement of Room Acoustic Parameters—Part 1: Performance Spaces* (International Standards Organization, Geneva, Switzerland, 2009).

[17] J. Bradley, "Review of objective room acoustics measures and future needs," Appl. Acoust. **72**, 713–720 (2011).

[18] L. Kirkegaard and T. Gulsrud, "In search of a new paradigm: How do our parameters and measurement techniques constrain approaches to concert hall design?" Acoust. Today **7**, 7–14 (2011).

[19] J. Carroll, "Individual differences and multidimensional scaling," Multidimens. Scal.: Theory Appl. Behav. Sci. **1**, 105–155 (1972).

[20] Y. Ando, "Calculation of subjective preference at each seat in a concert hall," J. Acoust. Soc. Am. **74**, 873–887 (1983).

[21] J. Pätynen, S. Tervo, and T. Lokki, "A loudspeaker orchestra for concert hall studies," Acoust. Bull. **34**, 32–37 (2009).

[22] J. Pätynen and T. Lokki, "Directivities of symphony orchestra instruments," Acta Acust. Acust. **96**, 138–167 (2010).

[23] J. Pätynen, "A virtual symphony orchestra for studies on concert hall acoustics," Ph.D. thesis, Aalto University, Doctoral dissertations 86/2011 (2011), pp. 41–52.

[24] J. Merimaa, "Analysis, synthesis, and perception of spatial sound—binaural localizaion modeling and multichannel loudspeaker reproduction," Ph.D. thesis, Helsinki University of Technology, Laboratory of Acoustics and Audio Signal Processing, Report No. 77 (2006), pp. 16–20.

[25] J. Merimaa and V. Pulkki, "Spatial impulse response rendering I: Analysis and synthesis," J. Audio Eng. Soc. **53**, 1115–1127 (2005).

[26] V. Pulkki and J. Merimaa, "Spatial impulse response rendering II: Reproduction of diffuse sound and listening tests," J. Audio Eng. Soc. **54**, 3–20 (2006).

[27] J. Pätynen, V. Pulkki, and T. Lokki, "Anechoic recording system for symphony orchestra," Acta Acust. Acust. **94**, 856–865 (2008).

[28] J. Pätynen, S. Tervo, and T. Lokki, "Simulation of the violin section sound based on the analysis of orchestra performance," in *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, New Paltz, NY (2011), pp. 173–176.

[29] N. Zacharov and G. Lorho, "What are the requirements of a listening panel for evaluating spatial audio quality?" in *Proceedings of the Spatial Audio and Sensory Evaluation, Techniques Workshop*, University of Surrey, UK (2006), available at http://www3.surrey.ac.uk/soundrec/ias/papers/Zacharov_Lorho.pdf (Last viewed 27 June 2012).

[30] G. Lorho, G. L. Ray, and N. Zacharov, "eGauge—a measure of assessor expertise in audio quality evaluations," in *AES 38th International Conference on Sound Quality Evaluation*, Piteå, Sweden (2010), pp. 186–195.

[31] J. Josse, J. Pagès, and F. Husson, "Testing the significance of the RV coefficient," Comput. Stat. Data Anal. **53**, 82–91 (2008).

[32] S. Lê, J. Josse, and F. Husson, "FactoMineR: An R package for multivariate analysis," J. Stat. Softw. **25**, 1–18 (2008).

[33] B. Escofier and J. Pagès, "Multiple factor analysis," Comput. Stat. Data Anal. **18**, 121–140 (1990).

[34] H. Abdi and D. Valentin, "Multiple factor analysis," in *Encyclopedia of Measurement and Statistics*, edited by N. Salkind, Thousand Oaks, CA (2007), pp. 657–663.

[35] A. Lucas, *amap: Another Multidimensional Analysis Package* (2011), available at http://CRAN.R-project.org/package=amap, R package version 0.8-7 (Last viewed 27 June 2012).

[36] L. J. Cronbach, "Coefficient alpha and the internal structure of tests," Psychometrika **16**, 297–334 (1951).

[37] C. Dehlholm, P. B. Brockhoff, and W. L. P. Bredie, "Confidence ellipses: A variation based on parametric bootstrapping applicable on multiple factor analysis results for rapid graphical evaluation," Food Qual. Pref. **26**, 278–280 (2012).

[38] S. L. Dien and J. Pagès, "Hierarchical multiple factor analysis: application to the comparison of sensory profiles," Food Qual. Pref. **14**, 397–403 (2003).

[39] I. Bork, "Report on the 3rd round robin on room acoustical computer simulation—Part II: Calculations," Acta Acust. Acust. **91**, 753–763 (2005).

[40] L. Beranek, *Concert Halls and Opera Houses: Music, Acoustics, and Architecture*, 2nd ed. (Springer, New York, 2004), Table IV.2, pp. 505–506.

[41] T. Lokki, J. Pätynen, S. Tervo, S. Siltanen, and L. Savioja, "Engaging concert hall acoustics is made up of temporal envelope preserving reflections," J. Acoust. Soc. Am. **129**, EL223–EL228 (2011).

[42] T. Lokki and J. Pätynen, "Lateral reflections are favorable in concert halls due to binaural loudness," J. Acoust. Soc. Am. **130**, EL345–EL351 (2011).