Psychomusicology: Music, Mind, and Brain

Studies on the Perception of Bass in Four Concert Halls

Henna Tahvanainen, Jukka Pätynen, and Tapio Lokki

Online First Publication, May 11, 2015. http://dx.doi.org/10.1037/pmu0000063

CITATION

Tahvanainen, H., Pätynen, J., & Lokki, T. (2015, May 11). Studies on the Perception of Bass in Four Concert Halls. *Psychomusicology: Music, Mind, and Brain*. Advance online publication. http://dx.doi.org/10.1037/pmu0000063

Psychomusicology: Music, Mind, and Brain
© 2015 American Psychological Association
0275-3987/15/$12.00 http://dx.doi.org/10.1037/pmu0000063

Studies on the Perception of Bass in Four Concert Halls

Henna Tahvanainen, Jukka Pätynen, and Tapio Lokki
Aalto University School of Science

It is known that the perception of a warm and full sound in a concert hall requires a rich bass. However, earlier research offers an incomplete understanding on the perception of bass in concert halls, especially regarding how the excess attenuation of low frequencies due to seats affects the perception of musical instruments with low-frequency fundamentals. This article studies the level of perceived bass and its clarity in 4 concert halls via paired comparison listening tests. The results suggest that the perceived level of bass is strongly related to the seat-dip effect, in particular to its main attenuation frequency with both short musical excerpts and individual notes. Moreover, the perceived clarity of bass seems to depend on the musical content and instruments. The results also indicate a complicated relationship between level and clarity of bass, as clarity can be enhanced with an increased level of bass in some cases.

Keywords: concert hall acoustics, seat-dip effect, bass, musical instruments, perceptual studies

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

The importance of bass to the overall sound quality in concert halls was highlighted at the opening of the Philharmonic Hall in New York in 1962 (Beranek, Johnson, Schultz, & Watters, 1964). The hall provided almost no immediate lateral reflections, and bass could barely be heard on the stalls (Barron, 2009). The parenthetical comments brought up particularly the weakness of the cellos and the double basses. Apart from critique, this observation led to the discovery of the seat-dip effect (Schultz & Watters, 1964; Sessler & West, 1964). The seat-dip effect is the term assigned for the attenuation at low frequencies due to sound propagating almost parallel, that is, at near-grazing angles to the surface formed by the tops of the seat backs. The seat-dip effect is characterized by a main attenuation dip between 80 and 300 Hz, and an attenuation bandwidth that can extend up to 1 kHz.

The seat-dip effect is objectively present in all concert halls with seats, but in many halls no noticeable lack of bass is perceived as such. In particular, the shoebox-shaped hall are generally characterized by a rich bass in the stalls (Schultz, 1965). This indicates that the seat-dip attenuation is corrected as the sound energy distribution in the concert hall develops over time. The lack of bass in the direct sound is compensated by reflections that contain the bass and arrive at the listener at nongrazing angles. This phenomenon is referred to as seat-dip recovery or correction.

Henna Tahvanainen, Jukka Pätynen, and Tapio Lokki, Department of Computer Science, Aalto University School of Science.

Special thanks to Nathalie Lurin for the help in running the listening tests and Antti Kuusinen for the test interface and discussions on the statistical models. The research leading to these results has received funding from the Academy of Finland, project no. 257099.

The preliminary results of the first study were reported in the International Symposium of Musical Acoustics in Le Mans, France, July 2014.

Correspondence concerning this article should be addressed to Henna Tahvanainen, Department of Computer Science, Aalto University School of Science, PO Box 15500, 00076 Aalto, Finland. E-mail: henna.tahvanainen@aalto.fi

However, it remains unclear when the bass should arrive at the listener's ears. It was hypothesized that the seat-dip effect could be compensated by increasing the reverberation time at low frequencies (Schultz & Watters, 1964; Barron, 1995). However, Bradley and Soulodre (1997) found that the perceived level of bass is not related to the reverberation time at low frequencies. More recent evidence suggest that the bass should arrive with the early reflections (Beranek, 2011; Davies, Cox, & Lam, 1996; Pätynen, Tervo, & Lokki, 2013; Soulodre & Bradley, 1995). Furthermore, Kahle (1995) suggested that sound energy arriving between 80 and 160 ms augments the perception of low frequencies. Additionally, Bradley and Soulodre (1997) linked the perceived level of bass to both the early sound level and late sound level. The phase of the reflections may also alter the perceived level of bass (Laitinen, Disch, & Pulkki, 2013; Lokki, Pätynen, Tervo, Siltanen, & Savioja, 2011).

It is scarcely understood how the seat-dip effect influences the perception of bass. Orchestral instruments that have their tuning range within the seat-dip attenuation range, such as double bass, tuba, and cello, may lack bass and articulation due to the seat-dip effect. For example, Bradley (1991) estimated that the attenuation of the double bass due to the seat-dip effect can be as much as 6 dB. Barron and Marshall (1981) observed that spatial impression is reduced in the presence of lateral reflections that contain the seat-dip attenuation. Davies et al. (1996) obtained a threshold of audibility for the seat-dip attenuation to be −5.7 dB in the 200 Hz octave band of the early energy over 0 to 40 ms. Although the seat-dip effect is generally considered to hamper the perception of bass, recent psychoacoustic research suggests that if the early reflections retain the bass, the lack of bass in the direct sound may actually enhance the overall perception of bass (Walther, Robinson, & Santala, 2013).

In addition to the seat-dip effect, the perception of bass in concert halls is affected by other factors, such as the residual absorption by the seats, walls, floor, and ceiling (Beranek, 2004, 2011). In general, hard thick surfaces absorb less bass than thin panels. Second, the vibrations of the stage floor have been proposed to augment the sound of the double bass and the cello connected to the floor via pins. So far, the augmented bass has been shown to be audible at the stage, but not among the audience (Abercrombie & Braasch, 2010; Wulfrank, Lyon-Caen, Jurkiewicz, Brulez, & Kahle, 2013). Third, bass may be masked by higher frequencies in the presence of a full orchestra (Nishihara & Hidaka, 2012).

In general, the perceived level, or loudness, may not be enough to describe the perception of bass in concert halls. The low frequencies are often described as "muddy" or "boomy"; these are terms that refer to the quality of bass. Sound quality is often described with clarity, which refers to the degree at which sounds can be distinguished in music (Beranek, 2004) or the ability to hear musical detail (Barron, 2009). According to Reichardt, Alim, & Schmidt (1975), subjective clarity has at least two aspects: how well notes are separated in time and how well different instruments playing simultaneously can be distinguished. Based on this, music from different periods or styles require different degree of clarity. High clarity in a concert hall is often associated with low perceived reverberation (Barron, 2009). However, some interplay exists between the perceived level and clarity in concert halls, as the results obtained by Soulodre and Bradley (1995) suggest that increased loudness can lead to increased clarity.

To summarize, research on the perception of bass in concert halls and its relation to the seat-dip effect shows inconclusive and incomplete results. Most studies rely on the use of standard room acoustical parameters (ISO 3382–1, 2009), which consistently exclude the lowest octave band at 63 Hz, and are thus regarded as insufficient to describe the low frequencies in a concert hall (Kirkegaard & Gulsrud, 2011; Nishihara & Hidaka, 2012). This article studies how musical instruments with low fundamental frequencies are perceived in terms of level and clarity of bass in concert halls that are characterized by different types of seat-dip attenuation. Individual notes and musical excerpts with a few bars were used with the following instruments: double bass, cello, tuba, and trombone. Therefore, the term *clarity* cannot refer to the same term as obtained with music by, for example, Reichardt et al. (1975). Instead, clarity was defined to the assessors with the following three aspects: separability of both notes and instruments, and audibility of the attack. The attack was chosen as an additional definition, as the ability to hear the transient sounds of the musical instruments may help separate the instruments and melody lines. This became apparent in one of the stimuli, where the tuba melody was fairly inaudible as it fused to strings. Additionally, when individual tones were presented as stimuli, listening to the attack was more meaningful than listening to the separability of instruments.

The first listening test with musical excerpts showed that the level of bass is perceived quite consistently across assessors for the three musical excerpts. A higher seat-dip frequency indicated a higher perceived level of bass. However, the perception of clarity yielded less straightforward results, as expected with different musical content. Therefore, a second listening test was run with individual notes to determine whether the perception of bass and clarity is affected by the partials of musical instruments that fall to the seat-dip attenuation range. In addition, based on verbal feedback of the assessors, the individual notes were divided into two frequency ranges that were compared separately. The results of the second test show that the perceived level of bass with all the instruments is linked to the seat-dip frequency, and that the perceived clarity depends on the instrument.

The Studied Concert Halls

This study includes two shoebox and two vineyard halls: Berlin Konzerthaus (BK), Vienna Musikverein (VM), Berlin Philharmonie (BP), and Helsinki Music Centre (MT). Their geometrical and acoustical properties are listed in Table 1. The open seat type refers to a seat with an underpass, so that the seat backrest does not extend all the way to the floor. The closed seats have a fixed seat backrest extending until the floor, or the possible underpass is blocked by a stepwise-raked floor, effectively extending the seat backrest to the floor.

The seat-dip attenuation results from a descructive interference of the direct sound and reflections from between the seats. The effective height of the seat backrest corresponds to a quarter wavelength of the main seat-dip frequency, as the destructively interfering reflection has to travel half a wavelength more than the direct sound (Bradley, 1991; Ishida, 1995). The vineyard halls have a lower seat-dip frequency than the shoebox halls due to the shorter effective seat backrests. In the shoebox halls, the effective seat backrest height corresponds to the actual seat backrest height. The floor in the vineyard halls is stepwise raked and the steps block the seat underpasses, thus extending the effective seat backrest all the way to the floor.

To describe the perception of bass objectively, many parameters have previously been proposed. The most frequently cited parameter, strength G, denotes the ratio of sound energy between the omnidirectional source measured at a seat and at 10 m in the free field over the entire impulse response. Bradley and Soulodre (1997) showed that the strength at the 125 Hz octave band is related to the perceived level of bass. Already earlier, the same authors (Soulodre & Bradley, 1995) correlated the perceived level of bass with the early strength (computed from the first 50 ms of the impulse response) at low frequencies G_{50}(125–500Hz). Previously, Bradley (1991) also linked the early strength at low frequencies G_{40} with the perceived level of bass. Most recently,

Table 1
The Properties of the Measured Unoccupied Concert Halls

Abbreviation	Name	N	Hall type	Floor type	Seat type	RT [s]	C_{80} [dB]	G(low) [dB]	BI [dB]	SDE [Hz]
BK	Berlin Konzerthaus	1600	S	Flat	Open	2.3	−1.9	5.7	3.5	182
VM	Vienna Musikverein	1700	S	Flat	Open	2.8	−4.0	3.6	0.1	177
BP	Berlin Philharmonie	2200	V	Raked	Closed	2.1	0.8	0.0	−1.4	96
MT	Helsinki Music Centre	1700	V	Raked	Closed	2.4	1.1	0.5	−0.4	125

Note. S = shoebox; V = vineyard; Reverberation time (RT) is averaged over 63–2000 Hz octave bands, standard clarity C_{80} over 500–1000 Hz, and strength at low frequencies *G(low)* over 63–125 Hz. Bass index (BI) is defined as $BI = G(63\text{–}125\ \text{Hz}) - G(500\text{–}1000\ \text{Hz})$.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Beranek (2011) proposed a strength-based criterion for the perception of bass in concert halls called the Bass Index: $BI = G(125\text{Hz}) - G(500\text{–}1000\text{Hz})$. Here the BI is extended until the 63 Hz octave band.

Clarity is objectively measured as the ratio of the early-to-late sound, with the most common measure being the clarity index C_{80} proposed by Reichardt et al. (1975). The clarity index does not consider the low frequencies, as it compares the early sound energy (0–80 ms) to the late sound energy (80 ms-∞) at the 500–1000 Hz octave bands.

Figure 1 shows the plans of concert halls with the measurement loudspeaker positions (1–4: Genelec 1029A, 5–6: Genelec 1021A) and the receiver position R, which is located at a distance of 19 m from the line formed by the loudspeakers 1–3. The loudspeakers correspond to the locations of the cello section (1–3), double bass section (4–6), tuba (6), and trombone (6) in a sym-

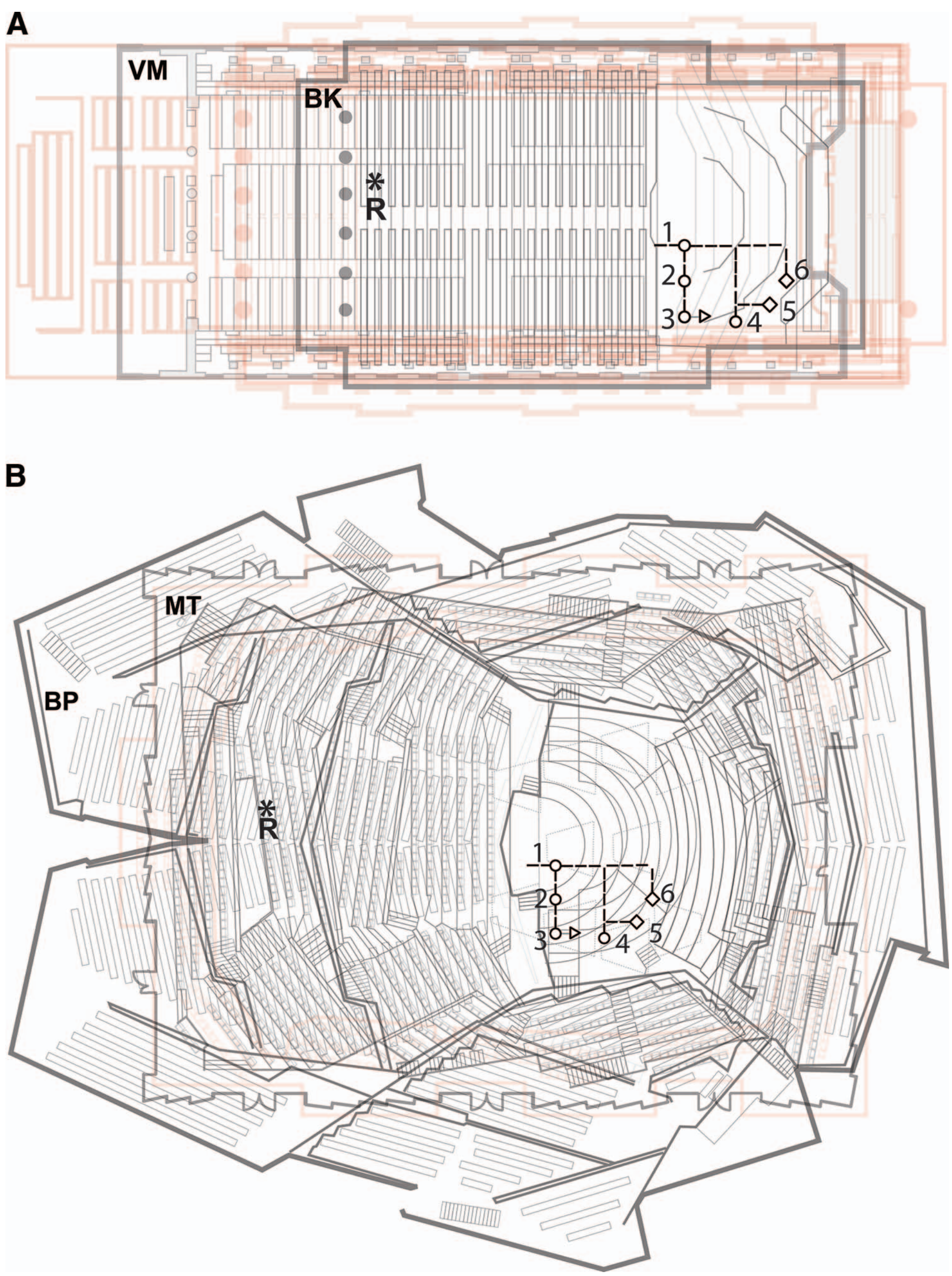

Figure 1. The plans of the studied concert halls superimposed according to hall type, and the loudspeaker orchestra measurement setup with the number of sources (1–6), and the receiver location R indicated. The distance from the receiver R to the line formed by the sources 1–3 is 19 m. Each dashed line on stage represents a distance of 1 m, so that, for example, sources 1 and 2 are 2 m apart. See the online article for the color version of this figure.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

phony orchestra. For channel 3, an auxiliary loudspeaker (Genelec 8020A) was placed on the floor facing upward for better estimation of the directivity of the cellos. The values in Table 1 are computed with the setup in Figure 1 at location R.

Figure 2 shows the cumulative time-frequency development of the impulse response in each concert hall averaged over the loudspeaker sources (1–6) in 10-ms time window increments, starting from 20 ms after the direct sound (see Pätynen et al. (2013) for more details on the method). The 20-ms curve is drawn with a thick line. The second highest curve shows the frequency response at 200 ms, and the highest curve the overall frequency response. The fundamental frequencies of the notes played during the listening test as well as their second and third partials are plotted on the frequency axis for the double bass and the tuba (dashed line), and for the cello and the trombone (solid line).

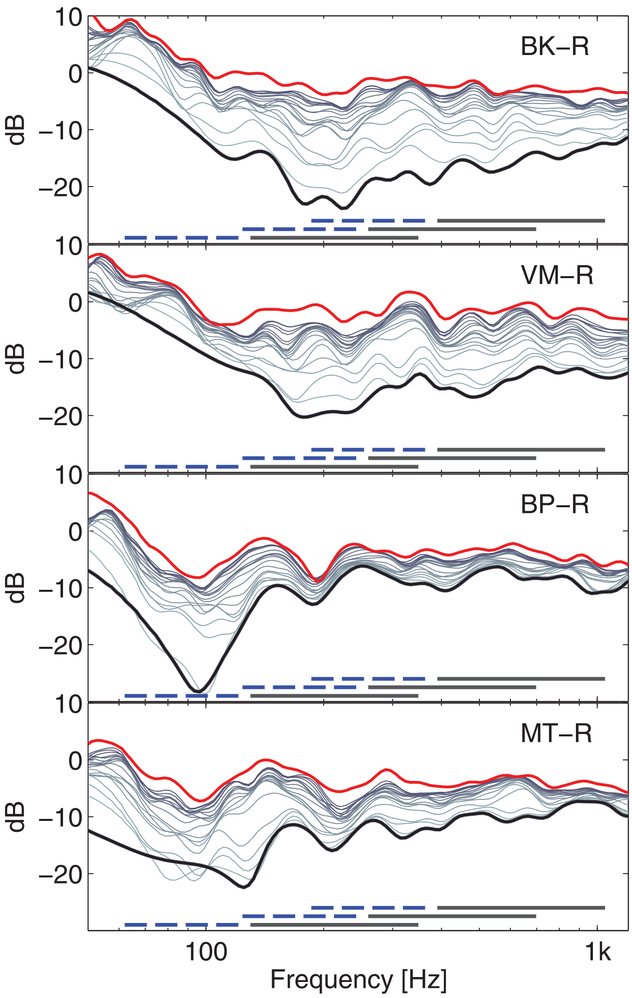

Figure 2. Time-frequency development of the concert halls measured at R with loudspeaker sources (1–6). One-third octave smoothing has been applied. The tuba and double bass range (60–120 Hz) is marked with a dashed line underneath the curves. The cello and trombone range (120–350 Hz) is marked with a solid line. The three levels of the lines indicate the fundamental, second, and third partial, respectively. See the online article for the color version of this figure.

The vineyard halls feature a more narrowband seat-dip attenuation at the 20-ms curve than the shoebox halls. However, MT is an exception to this, as at 30 ms the attenuation extends from 125 Hz down to below 80 Hz. It is mostly likely caused by a canopy reflection arriving out-of-phase with the direct sound at the receiver position. Furthermore, the vineyard halls have a smaller difference between the 20-ms and the overall frequency response curve than the shoebox halls. This means that compared with the vineyard halls, the shoebox halls have relatively more energy arriving at early/late sound field rather than at the direct sound field, when direct sound is considered to extent until 20 ms.

In addition, in the vineyard halls the fundamentals of the double bass and the tuba (62–120 Hz) lie at the main attenuation frequencies while the higher partials are well beyond the seat-dip attenuation. In the shoebox halls, the fundamental frequencies of the double bass and tuba remain below the seat-dip attenuation, while all their other partials, as well as cello, and trombone fundamentals are located within the attenuation band.

Methods

Participants

Nine assessors (two female) ages 21–43 (M = 30, *SD* = 4) participated in the first test, including the authors. Thirteen assessors (two female) ages 21–43 (M = 29, *SD* = 5) participated in the second test, including all the participants of the first test. All assessors had background in acoustics, formal or informal training in some musical instrument, and prior experience in listening tests and attending concerts. None of the assessors reported a hearing impairment.

The authors' responses were obtained under limited experimental conditions. The first author prepared the stimuli, but had no extensive listening experience in the chosen concert halls. The two other authors were not involved in the stimuli selection, and were thus unaware of the halls listened to. As the stimuli were presented in a random order, it further reduced the authors' bias to the results. No significant change in the results of the listening tests was obtained with the omission of the authors' responses, so they were included in the analysis.

Procedure

The level and clarity of bass in the concert halls was evaluated using a paired comparison listening test in which the assessor had to choose whether the sample A or B possesses more of a certain attribute. The listening test had two parts: one for assessing the level of bass and one for clarity. The order of both the hall pairs and the parts was fully randomized. The assessors could listen to the sample pair as many times as necessary before deciding. At the end, they were also asked for verbal feedback.

Two paired comparison tests were run: (a) with musical excerpts as stimuli, where both parts of the listening test consisted of 54 comparison pairs (6 hall pairs, 1 receiver position, 3 musical excerpts, 3 repeats per pair) and (b) with random individual notes, 48 comparison pairs (6 hall pairs, 1 receiver position, 4 sound excerpts, 2 repeats per pair). The first test took on average 35 min (Level of bass—15 min, Clarity—20 min), and the second test 43 min (Level of bass—20 min, Clarity—23 min).

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Stimuli

Musical excerpts. Three excerpts were chosen from anechoic recordings of the musical pieces (II. movement of Bruckner's Eighth Symphony, and I. movement of Beethoven's Seventh Symphony), and only the parts for the cello, double bass, and tuba were included. Their length varied between 6 and 11 s. Figure 3 presents the musical scores of the excerpts M1–M3.

Individual notes. The stimuli of individual notes with bass instruments (double bass section, trombone, tuba) were generated from an audio corpus extracted from the anechoic recordings (Kuusinen, 2014). The each stimulus was 6 s long and included random notes from one instrument at a selected frequency range (60–120 Hz or 120–200 Hz) with a random silence between notes (maximum length of silence between tones 1 s). Because the corpus contained very little material with the tuba above 120 Hz, it was replaced by the trombone above this frequency.

Stimuli Production

The production of the stimuli consisted of three stages: recording the instruments in an anechoic room, measuring the room impulse responses of the concert halls with the loudspeaker orchestra, and convolving these two signals for reproduction via a spatial audio system. Because the low frequencies are at the reproduction limit of the equipment, they were compared at the convolution and the final listening stage.

Production steps. First, single representatives of each musical instrument in a symphony orchestra were recorded individually in an anechoic room with several musical pieces played by professional musicians (Pätynen, Pulkki, & Lokki, 2008). For the current research, only tuba, trombone, cello, and double bass were used. The room was fully anechoic above 125 Hz, but the decay time of the instruments is typically higher than that of the room modes at the low frequencies. Therefore, it can be assumed that the recordings are reliable also below this frequency limit. The string sections were created artificially by copying the recordings with time-varying parameters (Pätynen, Tervo, & Lokki, 2011).

Then, spatial room impulse responses were recorded in the concert halls with a six-probe G.R.A.S. measurement microphone (G.R.A.S. Sound & Vibration A/S, Holte, Denmark) using logarithmic sweeps from a calibrated loudspeaker orchestra on stage as a sound source. The loudspeaker orchestra simulates a symphony orchestra in terms of locations on the stage and the approximate instrument directivities (Pätynen, 2011). The receiver location remained at a constant distance from the stage across the measured halls to allow direct comparison between the halls. It is worth noting that the measurements were conducted in unoccupied halls whereas concerts in reality have an audience. However, the results obtained with unoccupied halls can be considered valid for low frequencies for at least two reasons. First, the audience does not seem to alter the seat-dip attenuation below 800 Hz (Sessler & West, 1964). Second, the low frequency absorption coefficients have no considerable difference between empty and occupied seats (Beranek, 2004).

Finally, the room impulse responses with the directional estimates were used to create a 24-channel convolution reverberation so that the sound was mapped to the closest possible reproduction loudspeaker corresponding to the estimated direc-

Figure 3. The musical scores of the excerpts used in the listening test. M1 (on the top): Bruckner, Symphony no. 8 in C-minor, II movement, bars 25–32, M2: Bruckner, Symphony no. 8 in C-minor, II movement, bars 37–40, and M3: Beethoven, Symphony No. 7 in A-major op. 92, I movement, bars 29–30.

tion (see Pätynen, Tervo, & Lokki (2014) for details). The direction of arrival as a function of time was estimated from the spatial room impulse responses with the recently developed Spatial Decomposition Method (Tervo, Pätynen, & Lokki, 2013). The concert halls were auralized by convolving the anechoic recordings with the 24-channel reverberation. The reproduction system consisted of 20 Genelec 8020B and four Genelec 1029A calibrated loudspeakers in an acoustically treated room that complies with the ITU-R BS.1116 (1997) recommendation for subjective audio evaluation systems. The loudspeakers were positioned on five levels of elevation: at 0° (ear level) [azimuth angles 0°, ±22.5°, ±45°, ±67.5°, ±90°, ±135°, 180°], 30° [azimuth angles 0°, ±45°, ±135°], 45° [azimuth angles ±90°], 90° (on top of the listening position), and at 35° [azimuth angles ±40°, ±150°]. The nominal loudspeaker distance was 1.5 m from the listening position. The background noise level (A-weighted, slow) was measured to be 30 dB. The level of reproduction (A-weighted equivalent) in the listening room varied between 50 and 75 dB depending on the stimulus. The large level difference between stimuli stems from the real orchestral level of the instruments and the acoustics of the concert hall, and therefore it was left unchanged for the listening test.

Reproduction of the low frequencies. As the lowest fundamental note of listening test signals is played by the double bass at B_1 = 61.7 Hz, the reproduction of low frequencies of the spatial sound system was checked because of operating at the lower limit of the audio equipment. According to the manufacturer specifications, the cutoff frequency of the reproduction loudspeakers lies at 66 Hz and below that the frequency response decreases by 24 dB/octave.

The reproduction was compared between the input and output signals of the system. The input signal refers to the convolution between the anechoic recordings of the musical instruments and

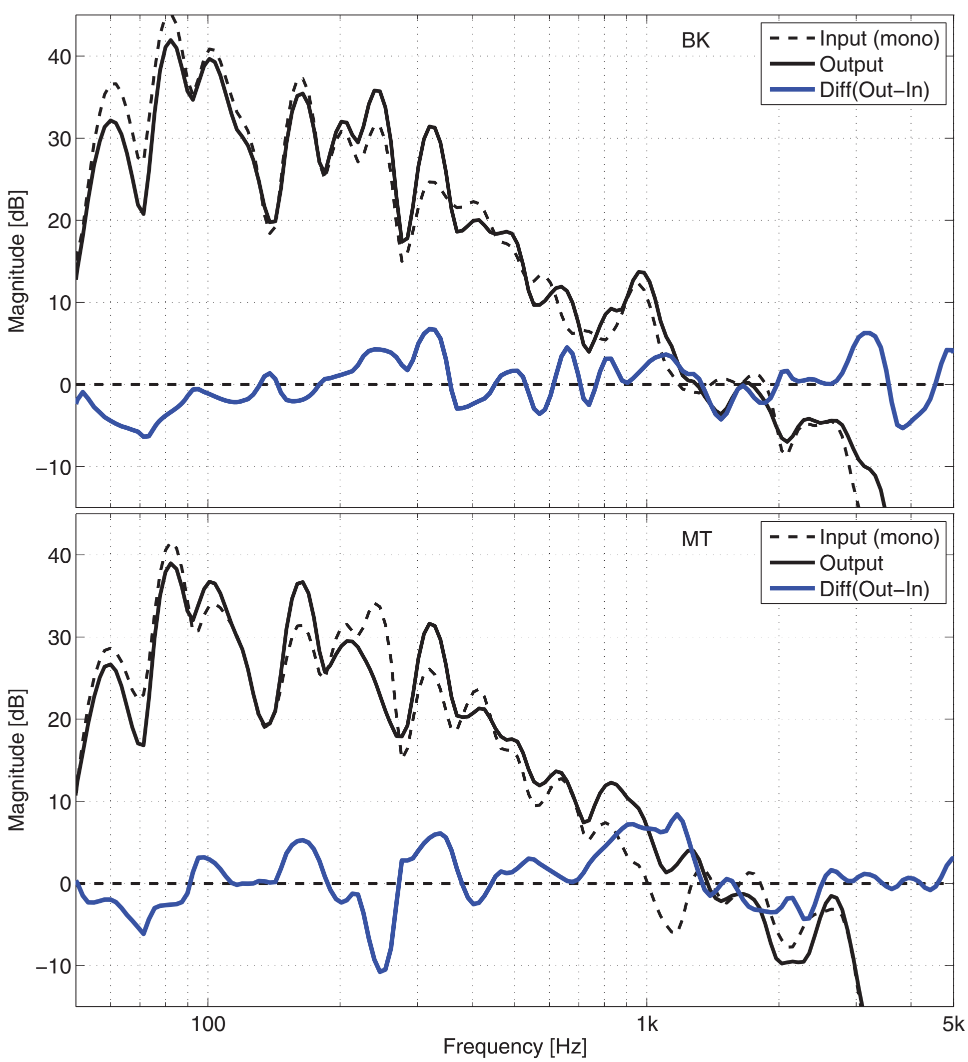

Figure 4. The reproduction of low frequencies in BK and MT. The dashed black curve refers to the input signal before reproduction, the thick black curve to the output signal reproduced by the listening room loudspeakers, and the thick blue curve denotes their difference. See the online article for the color version of this figure.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

the spatial impulse responses of the six loudspeaker orchestra sources summed into a mono signal. The output signal was measured with a G.R.A.S. measurement microphone in the sweet spot of the listening room when the input signal was played. Hence, the direction-dependent spatial hearing was not taken into account.

An example of the frequency responses of these two signals and their difference is plotted in Figure 4 for BK and MT with a musical excerpt (M3) containing double basses between 62 and 124 Hz. Based on the listening tests conducted in this study, the assessors considered these two halls to have the most and the least amount of bass, respectively, so they yield a maximal perceptual difference. It can be seen that the spatial reproduction system is able to reproduce the low frequencies down to 50 Hz very similarly in both halls, and the maximum difference between the input and output signal in each hall occurs at around 72 Hz for the very low frequencies. The other halls and stimuli exhibit similar results, so the ability of the spatial sound system to reproduce low frequencies is sufficient and does not differ greatly between stimuli.

Results

Test With Musical Excerpts

Figure 5 shows the box-and-whiskers plot of the first listening test when the stimuli are grouped by musical excerpt. By visual inspection, it is clear that BK was considered to have the most bass, with the exception of M2, where BK and VM were considered to have a comparable level of bass. In general, the differences in the perceived level of bass between the shoebox halls (BK, VM) and the vineyard halls (BP, MT) are significant (Kruskal–Wallis test $H(1) = 70.35$, $p < .001$), so that the shoebox halls have a higher perceived level of bass than the vineyard halls. This means that the halls with a high

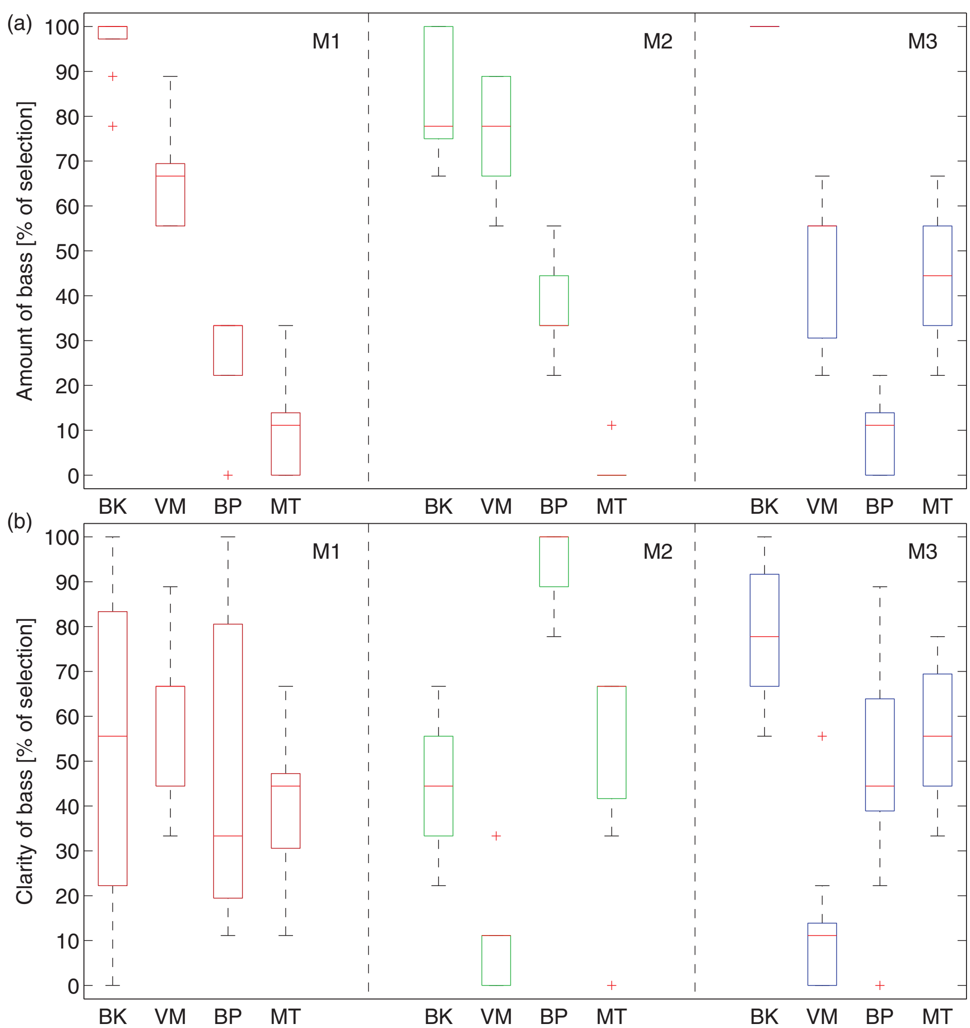

Figure 5. Analysis of the paired comparison test with musical excerpts (M1–M3), (a) for the level of bass and (b) clarity. The *y*-axis indicates the percentage of cases when an assessor has chosen the concert hall against any other concert hall (when the particular concert hall was available for selection). See the online article for the color version of this figure.

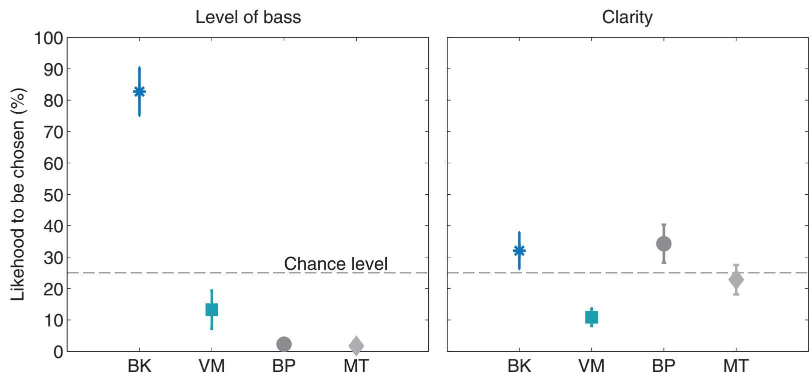

Figure 6. The probablity of the concert halls to be chosen to have the highest level of bass and the highest clarity. See the online article for the color version of this figure.

seat-dip attenuation frequency have a high level of perceived bass. The perceived level of bass and the seat-dip frequency have a relatively high correlation ($r(2) = .80$, $p < .001$).

The results with clarity are not as clear as those with the level of bass. With M1, the hall rankings in clarity are very equal, with VM having the highest mean, although the mean ranks are not significantly different ($H(3) = 3.38$, $p < .34$). M1 includes a tremolo passage with low-register strings, and assessing clarity from such a passage may not be meaningful. With M2, the highest clarity was perceived in BP, followed by MT. However, in this case only the mean ranks of the vineyard halls and the shoebox halls are significantly different ($H(1) = 16.67$, $p < .001$). With M3, BK was perceived the clearest and VM the least clear. However, only the mean rank of VM is significantly different from the others ($H(3) = 20.65$, $p < .001$).

The data from the paired comparison test were analyzed with the Bradley–Terry–Luce (BTL) model, which indicates the alternative that is most probably chosen over a certain attribute (Wickelmaier & Schmid, 2004). The goodness of fit of the BTL model was found sufficient for all the analyzed cases when the likelihoods of the BTL model and a saturated model fitting the data perfectly were compared ($p > .05$ for all cases presented below). Figure 6 shows the results of the BTL model with the 95% confidence intervals (CI) for the case when all musical excerpts are combined. The model shows that BK is very likely to be chosen to have the highest level of bass, and the rest of the halls are below the chance level of selection (25%). Furthermore, BK and BP have almost equal probability of being selected as the clearest, while the probability of VM and MT lies below chance level.

Figure 7 represents BTL models separately for the musical

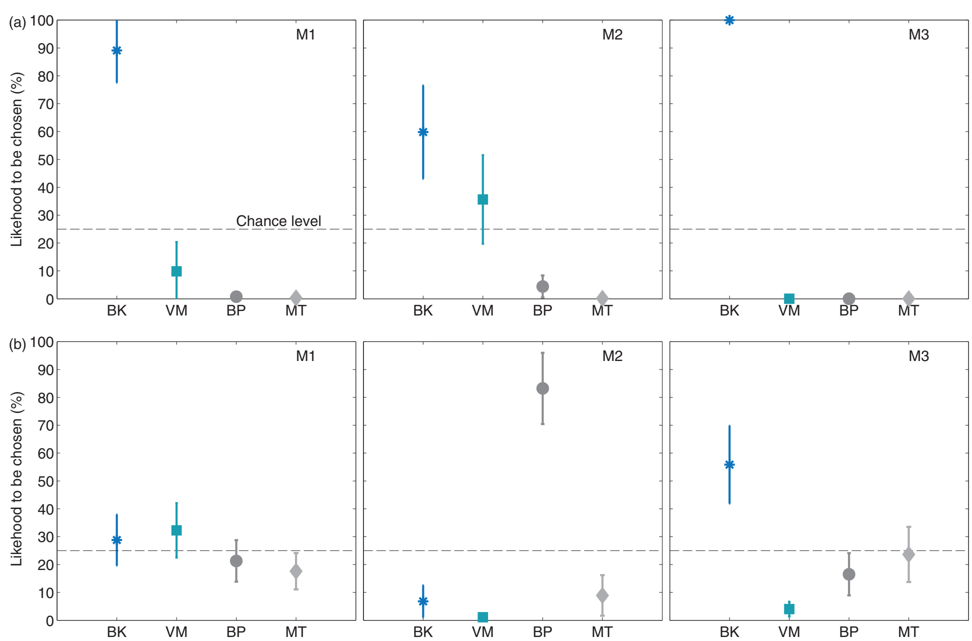

Figure 7. The probability of the concert halls to be chosen to have (a) the highest level of bass, and (b) the highest clarity with different musical excerpts. See the online article for the color version of this figure.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

excerpts M1–M3, and it shows that clarity depends on the musical excerpt. With M1, VM has a higher likelihood than the others, probably because the tremolo parts benefit from reverberation, which results in a good blending. In this case, clarity or the audibility of the attack may not be considered important, and reduced articulation could even be intended in the interpretation. With clear tuba melody line in M2, BP was considered to have the highest clarity. Thus, clarity seems important for both the melody line and the tuba as an instrument. Interestingly, with M2 VM was the least likely to be chosen to have the highest clarity, most probably due to the high level of reverberation in VM. In VM, the melody played by the tuba was poorly audible, as it fused with string instruments. In M3, the string sections were playing an accompaniment, and for this excerpt BK was ranked the highest both in the level of bass and clarity. Both attributes may thus be important for accompaniment with string instruments.

Finally, all assessors found the comparison of clarity more difficult than that of the level of bass, and on average more time was spent to complete the part on clarity.

Test With Individual Notes

Figure 8 shows the box-and-whiskers plot of the perceived level of bass and clarity when the stimuli are grouped by the instruments and their frequency range. BK and VM attain the highest perceived level of bass, but their mean ranks do not differ significantly from one another across the instruments or the frequency ranges. BP and MT attain the highest clarity, except with the double bass at 60–120 Hz, where the mean ranks of none of the halls is significantly different.

Figure 9 shows the corresponding BTL model. For the highest level of bass, BK is most likely to be chosen, except with the trombone at 120–200 Hz, where VM is most likely chosen. Both BK and VM have a markedly higher seat-dip attenuation frequency than BP or MT (see Figure 2). The perceived level of bass and the seat-dip frequency have a relatively high correlation for comparisons with individual notes ($r(2) = .61$, $p < .001$).

BP is most likely chosen to have the highest clarity, except with the double bass at 60–120 Hz, where the likelihood of all the halls is almost equal. In general, it can be seen that assessing clarity with the double bass yields more equal likelihoods between the halls than with the brass instruments. Thus, the perceived clarity seems to depend on the musical instrument, and the BTL models of such grouping are shown in Figure 10. With the double bass, none of the halls has clearly been chosen to have the highest clarity. With the tuba and the trombone, BP has the highest perceived clarity. These brass instruments are characterized by a rich spectrum at high frequencies, and the high partials define the attack of the instrument. Therefore, for comparisons with individual notes, clarity may be more easily estimated with brass instruments than with string instruments.

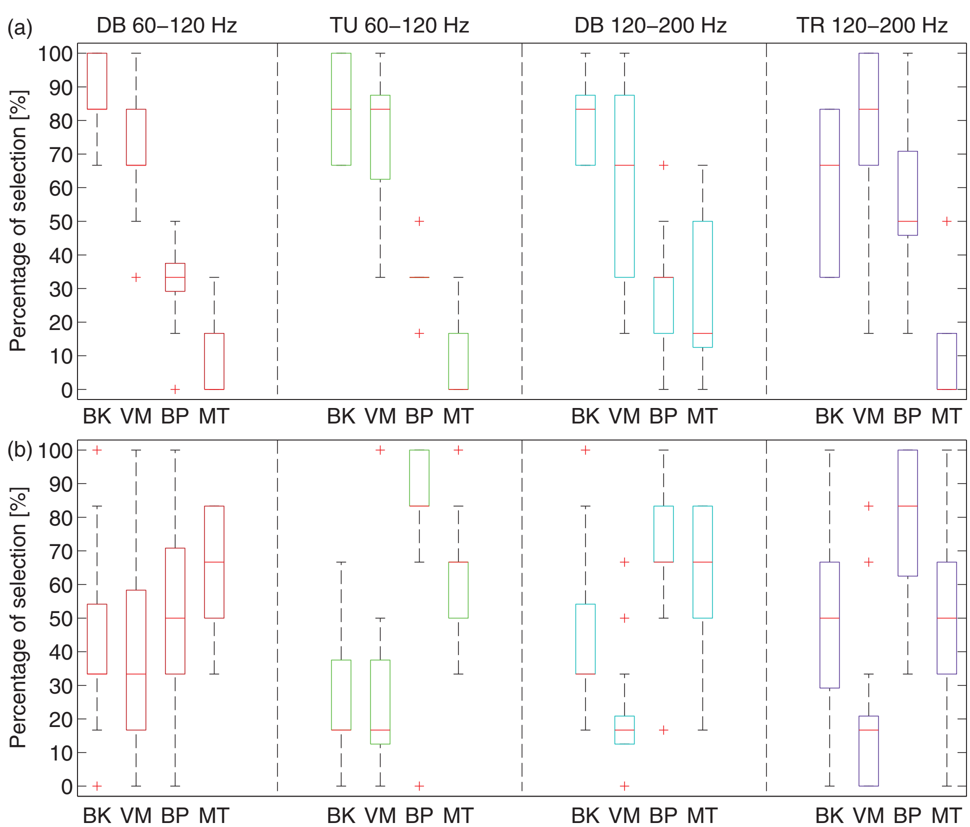

Figure 8. Box-and-whiskers plot of the paired comparison test with individual notes (DB = double bass section; TR = trombone; TU = tuba), (a) for the level of bass and (b) clarity. The *y*-axis indicates the percentage of cases when an assessor has chosen the concert hall against any other concert hall (when the particular concert hall was available for selection). See the online article for the color version of this figure.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

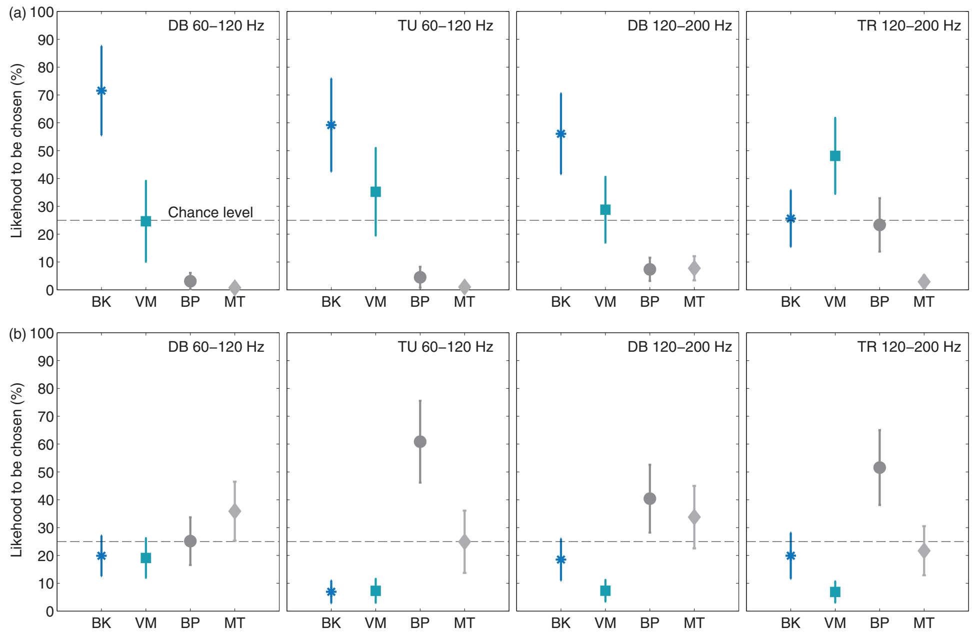

Figure 9. The probability of the concert halls to be chosen to have (a) the highest level of bass and (b) the highest clarity with different musical instruments and frequency ranges. See the online article for the color version of this figure.

Finally, five assessors found the comparison of level of bass more difficult than that of the clarity, whereas six assessors thought assessing clarity was more difficult.

Discussion

The results show clearly that the perceived level of bass is higher in halls with a high seat-dip attenuation frequency than in halls with a low seat-dip attenuation frequency. A high seat-dip frequency also seems to help retain a high objective level of bass. The shoebox halls, which have a high seat-dip frequency due to open seats, have a level of bass (50–200 Hz) 7.5 dB higher than the vineyard halls at 20 ms after the direct sound, and 4.3 dB higher in the overall frequency response. In addition, the shoebox halls have more early lateral reflections than the vineyard halls that most likely compensate for the lack of bass in the direct sound (Pätynen et al., 2013).

Some assessors commented that source width and level of reverberation influenced the perceived level of bass. Bradley (1991) found earlier that the perceived level of bass is affected by the level of reverberation, but not by the reverberation time. The current study supports the findings by Bradley, as VM is not perceived to have a considerably higher level of bass, even though it maintains a longer reverberation time than the other halls. Other factors that may have influence on the perceived level of bass include the seat upholstery and the absorption of low frequencies by the walls and the ceiling (Beranek, 2004). According to the building data of the four concert halls, their wall structures are very

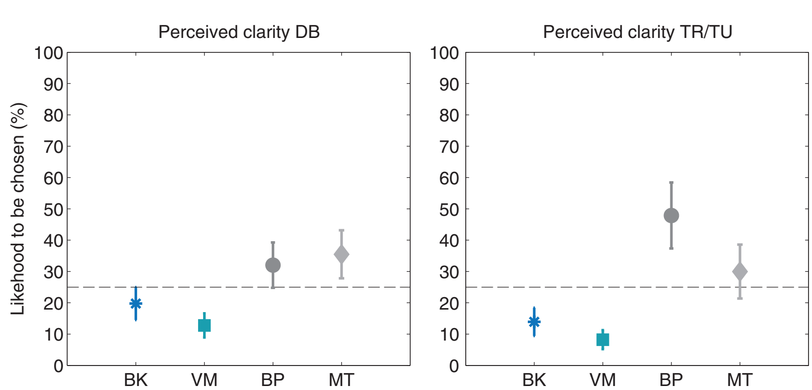

Figure 10. The probability of the concert halls to be chosen to have the highest clarity with the double bass (DB) and the brass instruments (TR/TU). See the online article for the color version of this figure.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Table 2
Strength G(low) and Bass Index (BI) Computed With Different Frequency Bands

Concert hall	G(63 Hz) [dB]	G(125 Hz) [dB]	G(63–125 Hz) [dB]	BI(125 Hz) [dB]	BI(63–125 Hz) [dB]
Berlin Konzerthaus	7.3	4.8	5.7	2.6	3.5
Vienna Musikverein	6.0	3.5	3.6	0.0	0.1
Berlin Philharmonie	0.1	0.6	0.0	−0.8	−1.4
Helsinki Music Centre	0.4	2.3	0.5	1.4	−0.4

Note. The values have been averaged over the sources 1–6 at receiver location R, and the concert halls have been ordered in a decreasing order of perceived level of bass.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

similar and they feature a light seat upholstery. Thus, these factors are considered to account for only a small part of the differences. The stage properties that may affect the perception of bass have not been considered in this study, as the stage thicknesses are not known. The contribution of source width to the perceived level of bass is an interesting phenomenon and requires future research.

The variety of room acoustical parameters aiming to describe the perceived level of bass in concert halls deserves some attention. Table 2 lists the values of some common parameters and their extended, although uncommonly employed, versions. The omission of the 63-Hz frequency band proves to be problematic especially for MT, where the low-frequency attenuation spans around 80–125 Hz. MT and VM become very close in strength when the standard parameter *G*(125 Hz) is considered, although perceptually the level of bass differs quite drastically. BI experiences the same problem, as it is derived from *G*(125 Hz). Even with the inclusion of the 63-Hz band, BI does not yield a large difference between BP, BK, and MT. Therefore, *G*(63–125 Hz) seems to be the best candidate to explain the perceived level of bass over the parameters listed in Table 2, although it yields BP and MT in the wrong order. The seat-dip attenuation could also be captured in measures of early bass level, such as G_{low}^{50ms} proposed by Soulodre and Bradley (1995) or G^{40ms} by Bradley (1991), but further studies are required to determine the effect of arrival time to the perceived level of bass.

As for the quality of bass, the results suggest that the perceived clarity in concert halls differs between single instruments playing individual notes and several instruments playing music together. The results complement the conclusion by Reichardt et al. (1975) that the perceived clarity depends on the type of music played. With clear melody lines clarity seems to be more important, and easier to assess than with tremolo parts. Furthermore, reverberation may adversely influence the perception of clarity.

In addition to that, perception of clarity or attack depends on the musical instrument, as it seems to be more easily judged from the brass instruments than from the string instruments. The results suggest that brass instruments, which are often determined by the mid- and high-frequency partials, do not suffer from major clarity problems in concert halls where the seat-dip effect is at a lower frequency (below 125 Hz) while their clarity may be reduced with a higher seat-dip frequency (above 170 Hz). In the case of string instruments, no clear connection between the seat-dip frequency and clarity can be obtained.

The perceived clarity has been linked with the perceived loudness by Soulodre and Bradley (1995) when a full orchestra in the concert hall was employed, and the results of the current study reinforce this link. In the test with the musical excerpts, eight out of nine assessors (including two authors) reported that they experienced difficulties in comparing clarity between a hall with a blurred bass and a hall with very little bass, and that often the concert hall with more bass appeared to have higher clarity as well. With individual notes, six of 13 assessors (including two authors) reported that the presence of more bass improved the articulation of the instruments. When the cellos and the double basses played an accompaniment (M3), the presence of both bass and clarity was assessed highest in the hall with a high seat-frequency and moderate reverberation (BK). This further suggests that high strength *G* at low frequencies may be favorable to the perception of clarity for string instruments playing accompaniment.

Conclusions

The level of bass and clarity in four concert halls were studied with two paired comparison listening tests using bass instruments (tuba, trombone, and sections of double bass and cello) as excitation signals. The stimuli presented for the assessors consisted of musical excerpts and individual notes. The perceived level of bass was found clearly stronger in concert halls where the main attenuation frequency of the seat-dip effect is quite high. In other words, the lower the main seat-dip frequency, the less bass was perceived. The perception of clarity or attack was found to depend on the content of the musical excerpt (such as melody lines, tremolo, accompaniment) as well as on the musical instruments. The results further indicate a complex relationship between the perceived clarity and level of bass in concert halls.

Future research could be directed toward obtaining a measurable relationship between the level of bass and clarity. Furthermore, the use of the entire orchestra in the auralization would complement the research conducted in this study, as the bass frequencies might be perceived differently when the whole orchesta is playing the music. Finally, the seat-dip effect may slightly vary from seat to seat, and in the first rows of the balcony the effect is generally absent. Thus, the methods in this study could be extended to the comparison of more seats within and between halls.

References

Abercrombie, C., & Braasch, J. (2010). Auralization of audio-tactile stimuli from acoustic and structural measurements. *Journal of the Audio Engineering Society, 68,* 818–827.

Barron, M. (1995). Bass sound in concert auditoria. *Journal of the Acoustical Society of America, 97,* 1088–1098.

Barron, M. (2009). *Auditorium acoustics and architectural design* (2nd ed). New York, NY: Spon Press.
Barron, M., & Marshall, A. H. (1981). Spatial impression due to early lateral reflections in concert halls: The derivation of a physical measure. *Journal of Sound and Vibration, 77,* 211–232.
Beranek, L. L. (2004). *Music, acoustics, and architecture* (2nd ed). New York, NY: Springer-Verlag.
Beranek, L. L. (2011). The sound strength parameter G and its importance in evaluating and planning the acoustics of halls for music. *Journal of the Acoustical Society of America, 129,* 3020–30260.
Beranek, L. L., Johnson, F. R., Schultz, T., & Watters, B. (1964). Acoustics of Philharmonic Hall, New York, during its first season. *Journal of the Acoustical Society of America, 36,* 1247–1262.
Bradley, J. S. (1991). Some further investigations of the seat dip effect. *Journal of the Acoustical Society of America, 90,* 324–333.
Bradley, J. S., & Soulodre, G. A. (1997). Factors influencing the perception of bass. *Journal of the Acoustical Society of America, 101,* 3135.
Davies, W. J., Cox, T. J., & Lam, Y. W. (1996). Subjective perception of seat dip attenuation. *Acta Acustica United with Acustica, 82,* 784–792.
Ishida, K. (1995). Investigation of the fundamental mechanism of the seat-dip effect – Using measurements on a parallel barrier scale model. *Journal of the Acoustical Society of Japan (E), 16,* 105–114.
ISO 3382–1. (2009). *Acoustics – Measurement of Room Acoustic Parameters. I: Performance spaces*. Geneva, Switzerland: International Standards Organisation.
ITU-R BS. 1116. (1997). *Methods for the subjective assessment of small impairments in audio systems including multichannel sound systems*. Geneva, Switzerland: International Telecommunications Union.
Kahle, E. (1995). *Validation d'un modéle 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]*. (Unpublished doctoral dissertation). Laboratoire d'Acoustique de l'Université du Maine, Le Mans, France.
Kirkegaard, L., & Gulsrud, T. (2011). In search of a new paradigm: How do our parameters and measurement techniques constrain approaches to concert hall design? *Acoustics Today,* 7–14.
Kuusinen, A. (2014). An anechoic audio corpus for room acoustics and related studies. In *Proceedings of EAA Joint Symposium on Auralization and Ambisonics*, Berlin, Germany.
Laitinen, M.-V., Disch, S., & Pulkki, V. (2013). Sensitivity of human hearing to changes in phase spectrum. *Journal of the Audio Engineering Society, 61,* 860–876.
Lokki, T., Pätynen, J., Tervo, S., Siltanen, S., & Savioja, L. (2011). Engaging concert hall acoustics is made up of temporal envelope preserving reflections. *Journal of the Acoustical Society of America Express Letters, 129,* EL223–EL228.
Nishihara, N., & Hidaka, T. (2012). Loudness perception of low tones undergoing partial masking by higher tones in orchestral music in concert halls. *Journal of the Acoustical Society of America, 132,* 799–803.
Pätynen, J. (2011). *A virtual loudspeaker orchestra for studies on concert hall acoustics*. PhD dissertation, Aalto University School of Science, Espoo, Finland.
Pätynen, J., Pulkki, V., & Lokki, T. (2008). Anechoic recording system for symphony orchestra. *Acta Acustica United with Acustica, 94,* 856–865.
Pätynen, J., Tervo, S., & Lokki, T. (2011). Simulation of the violin section sound based on the analysis of orchestra performance. In *Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA 2011)* (pp. 173–176). New Paltz, NY.
Pätynen, J., Tervo, S., & Lokki, T. (2013). Analysis of concert hall acoustics via visualisations of time-frequency and spatiotemporal responses. *Journal of the Acoustical Society of America, 133,* 842–857.
Pätynen, J., Tervo, S., & Lokki, T. (2014). Amplitude panning decreases spectral brightness with concert hall auralizations. In *Proceedings of the 55th International Coference of the Audio Engineering Society*, Helsinki, Finland.
Reichardt, W., Alim, O. A., & Schmidt, W. (1975). Definition und messgrundlage eines objektiven masses zur ermittlung der grenze zwischen brauchbarer und unbrauchbarer durchsichtigkeit beim musikdarbietung. *Acustica, 32,* 126–137.
Schultz, T. (1965). Acoustics of the concert hall. *IEEE Spectrum Magazine, 2,* 56–67.
Schultz, T., & Watters, B. (1964). Propagation of sound across audience seating. *Journal of the Acoustical Society of America, 36,* 885–896.
Sessler, G., & West, J. (1964). Sound transmission over theatre seats. *Journal of the Acoustical Society of America, 36,* 1725–1732.
Soulodre, G. A., & Bradley, J. S. (1995). Subjective evaluation of new room acoustic measures. *Journal of the Acoustical Society of America, 98,* 294–301.
Tervo, S., Pätynen, J., & Lokki, T. (2013). Spatial decomposition method for room impulse responses. *Journal of Audio Engineering Society, 61,* 1–13.
Walther, A., Robinson, P., & Santala, O. (2013). Effect of spectral overlap on the echo threshold for single reflection conditions. *Journal of the Acoustical Society of America, 134,* EL158–EL164.
Wickelmaier, F., & Schmid, C. (2004). A Matlab function to estimate choice model parameters from paired-comparison data. *Behaviour Research Methods, Instruments, and Computers, 36,* 29–40.
Wulfrank, T., Lyon-Caen, I., Jurkiewicz, Y., Brulez, J., & Kahle, E. (2013). Recent experiments with vibration of stage and audience floors in concert halls. In *Proceedings of Meetings on Acoustics, 19,* 015029. http://dx.doi.org/10.1121/1.4800261

Received June 21, 2014
Revision received November 26, 2014
Accepted January 26, 2015 ■

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.