Limitations of spatial perception in room auralizations

Michael Kohnen, Josep Llorca-Bofí, Michael Vorländer
Institute of Technical Acoustics, 52056 Aachen, Germany, Email: mko@akustik.rwth-aachen.de

Abstract
In a collaboration between architecture and engineering a test is conducted to see whether scale drawings made by architects purely listening to a virtual acoustic scene can be fitted to the simulation model by compensating an inherent, individual bias. A listening test regarding source distance perception and volume perception was conducted and mismatches found correlated to deviations of the drawings from the simulation model. Even though the findings of the source distance perception test are according to those found in literature, results indicate no correlation between mismatches for source distances, distances to walls of the room or to the perceived volume.

Introduction
Modern 3-D audio auralization techniques are evolving with their need due to 3-D visual capturing, synthetization and reproduction approaches. On the one hand, real-time rendering of room acoustics is a costly but crucial part to excite a sensation of acoustic plausibility or authenticity. On the other hand, our perception of complex acoustic auralizations is usually superposed by multi-modal stimuli, especially visual cues. Yet, research on acoustic sound presentation typically focuses on a pure acoustic stimulus presentation. This study investigates whether people are inherently biased in their perception when listening to a scene without visual feedback. This work builds up on the idea presented in [1]. Students of architectural faculty made scale drawings while listening to an auralized scene. The task was to draw their imagination of the scene, including the room and present sources. The ability of architectural students to express their sensation through true to scale drawings opens up a new way of psychoacoustic measurement, especially as they deliver the tested parameters without knowing the actual purpose of the test. The approach here is to first let the naive participant scale draw a complex acoustic scene. Later on they were tested on source distance, room volume and room geometry (i.e. length of the room) perception. The correlation of occurring mismatches found between drawings and simulation model to the mismatches found in the listening test are the key values of this investigation.

Methods
Auralization and playback
All room acoustics auralizations in this paper were done with RAVEN [2] using an HRTF dataset of the artificial head made at the Institute of Technical Acoustics measured with a three degree resolution in azimuth and elevation. RAVEN is geometric acoustics simulation software combining the image source method for early reflection with an ray-tracing approach for the late reverberant field. To compensate the headphone transfer-function (HpTF) of each participant the individual HpTFs were measured at the blocked ear canal using Sennheiser KE3 capsules and silicon rubber domes. The measurements were done in a separate room before entering the hemi-anechoic chamber. The equalization was calculated according to Masiero et al. [3] and applied for all parts of the listening test.

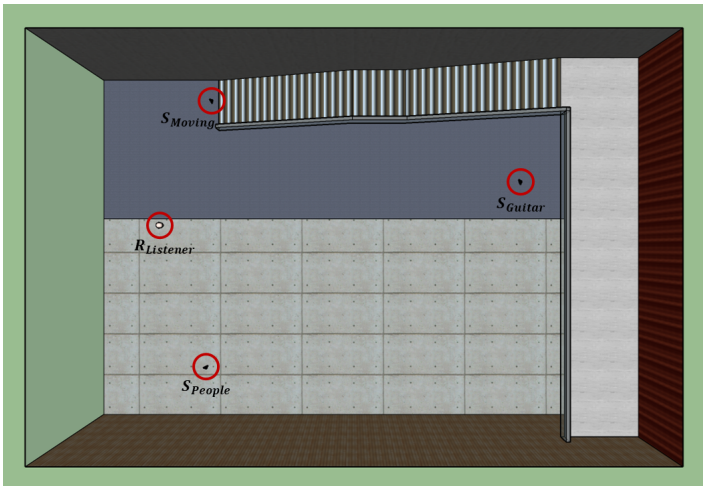
Drawings


Figure 1: Model of the room under test for the drawing part. The wall materials vary and the three main sources are indicate by 'S' while the receiver is denoted as 'R'. The ceiling was an absorptive acoustic ceiling material. The moving sources is mainly auralized as foot steps along the stairs and gallery.

Before drawing the scene a short introduction video was played to introduce the students to the basic concepts of room acoustics (especially room impulse responses) and their perceptual aspects. After the video presentation the actual drawing part started by playing back the auralization. The participants were told to take about 15 minutes for this part. The acoustic scene as shown in fig. 1 consisted of two fixed sources (a guitar player and a group of people at a table talking and typical sounds of dishes being used) and four different moving sounds with distinct walking sounds along the stairs and the gallery, coughing and greeting. The four sources are audible only for a few seconds each. The guitar player was located at distance of 10.5m, the group of people at 4.3m. From the listeners point of view the left side floor was attached with carpet material, the right side with concreteform. Left side wall was chosen to be concrete, ride side to wood panels, frontal wall to carpet, rear wall to glass and the ceiling to be an acoustic ceiling. All material data was taken from the RAVEN database [2]. The duration of the scene was 1:49 minutes. The participant were free to move to any point in time during playback and repeat the scene as often as desired. The task was to scale draw the scene from three different perspectives: plan (top down), section (median plane of the listener) and military perspective (3D side view). For the section and plan view the listener was printed on the sheets and participants were asked to include source positions (and the receiver position for the military perspective). This part will from now on be referred to as 'drawing' part in contrast to the 'listening test' part which will be explained in the following sections. The drawings as

such can be found in the architectural evaluation [4] of this work.

Distance perception
To evaluate the general performance of the participants distance perception a listening test was conducted after the drawing part. Distance perception in anechoic environment relies on loudness cues, binaural effects for near-field sources and probably a change in frequency due to air absorption for sources farther away then 15m [5]. All these cues need a prior knowledge (or at least familiarity) with the presented source signal to estimate its sound power (in terms of loudness) and frequency spectrum, otherwise the distance estimate can only be relative (e.g. closer, farther). For other environments an absolute estimate of the distance might be found when taking the ratio of direct to reverberant energy into account including a room specific coloration of the reverberation. Yet, these spaces have to be made familiar by listening to sound in the room or at least by getting a grasp of the acoustic visually. For the distance perception test an excerpt of a violin playing in a anechoic chamber was taken. The violin was taken to get a stimulus that is referencing a real-life experience situated in an plausible setting. The violin was auralized at different distances (2m 3m 4m 5m 7m 9m 12m) in the room which is shown in fig. 1 including room acoustics but assuming far-field conditions for the HRTF dataset used. The incidence direction deviated 25 degree from the frontal direction to the right to allow enough space between wall and the source at twelve meter distance and to avoid a frontal direction with very little binaural cues. The stimuli consisted of a 2m reference (same input signal) followed by the auralized target distance. The reference sound is intended to be a cue for the sources loudness and its frequency spectrum as well as the rooms general coloration of reverberation and direct to reverberant ratio. A slider with a resolution of 0.5m was used to state the perceived distance. The participants were able to repeat the stimulus as often as desired. Each stimulus was tested three times. The stimulus order was randomized.

Space perception
The perception of room acoustics is a complex matter and is difficult to measure especially when aiming at a specific (quantitative) compensation for the mismatch of perception and simulation model. The approach here is to access the volume of the room as it can be easily calculated from the drawings made and seen to be a key feature of acoustically room perception as it directly influences the reverberation time. To test different room volumes of a simple room different parameters can be varied: width, length or height. Furthermore, the relative position between source and receiver is important as well as their position relative to the room. Both will mainly influence the early reflections. As the number of independent variables influences the length (and complexity) of the listening test the independent variables were narrowed down to the length (in listeners view and in source direction) of the room and a fixed receiver source combination. Acoustic parameters of the room under test can be found in 1. Width and height were six meter each. The receiver was located 0.7m from the rear wall, 2.8m from the left and 3.2m from the right wall to avoid symmetric effects of early reflections. The source was located at 9m distance and eight degrees to the left to avoid effects of room symmetries. The distance was chosen to be big compared to the critical distance of direct to reverberant energy ratio as it was intended to test the effects of the room rather than the source position and characteristics. The input signal was the same as chosen in distance perception test. Side walls were applied with acoustic parameters of wood panels, floor with cork and ceiling with acoustic ceiling parameters. The rear wall was simulated with absorptive curtain material to avoid flutter echoes (especially for the shorter rooms). The frontal wall was modeled as concrete wall to support the perception of the room length. A visualization of the two test rooms with the shortest and longest room length can be found in 2. Again, participants were able to repeat the stimuli as often as

Table 1: (Acoustical) Parameters of the tested rooms

Length	Volume	RT(mid)	EDT
10m	360m^3	1.29s	1.33s
15m	540m^3	1.51s	1.53s
20m	720m^3	1.71s	1.65s
25m	900m^3	1.76s	1.80s
30m	1080m^3	1.97s	1.90s
40m	1440m^3	2.01s	2.11s
50m	1800m^3	2.12s	2.30s

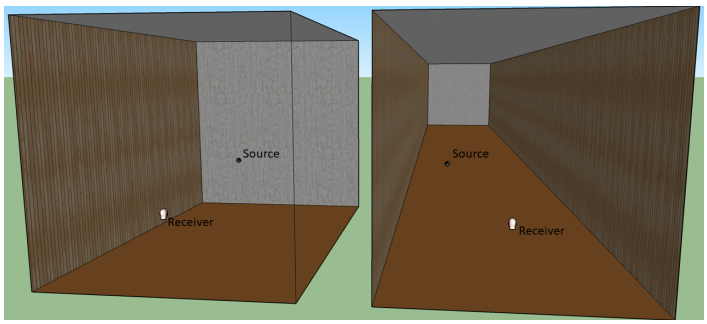


Figure 2: Models of the tested rooms with the minimum length of 10 meters and 50 meters. The wall in frontal direction is simulated as a concrete wall allowing for a distinct reflection to estimate its distance.

Figure 3: Setup of the listening test. The test was conducted in the hemi-anechoic chamber, providing a larger space than a hearing booth and low environmental noise. Additionally, spotlights at the desk should further fade out the environment and its geometries.

desired. Perceived room length was stated as free numerical input and participants asked to state their perceived room length in source direction.

Setup

The listening test was set up in the hemi-anechoic chamber at the Institute of Technical Acoustics to limit the psychological room divergence effect between auralized room and listening room [6] as the chamber is far more spacier than a hearing booth. To further excite a feeling of a wide space the desk were illuminated using spotlights, trying to fade out the room geometries in the dark environment. Six people were measured at the same time using six laptops. The HpTF measurements were done beforehand and loaded to the laptops. A picture of the setup can be found in fig. 3.

Results

41 particpants (23 females and 19 males) between 19 and 35 years old (mean: 23.4) took part in the drawing part and listening test. The overall duration was about one hour. In 13 out of 42 drawings front back confusion is visible for the guitar source. One drawing was discarded due to a lack of boundary walls and missing sources.

Distance perception

Fig. 4 shows the result of the distance perception test. The general trend (as e.g. in [5]) of overestimating the source distance for close distances and underestimating them for larger distances can be seen. The point where simulated and perceived distance intersect usually is found to be a lot lower than shown in the results here and is most likely due to the reference at two meters which introduces an offset especially for closer sources.

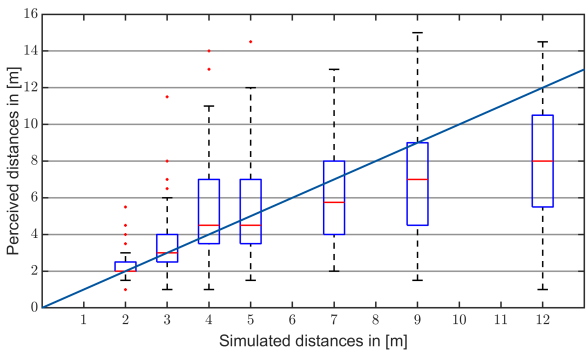


Figure 4: Setup of the listening test. The test was conducted in the hemi-anechoic chamber, providing a larger space than a hearing booth. Additionally, spotlights at the desk should further fade out the environment.

Space perception

Different room lengths were tested in the listening test and the stated answers of the perceived room length can be found in fig. 5. Noticeably is the saturation effect for a perception of 15m for all rooms longer than 15m. Clearly, the attribute rated is not longer related to the wall distance anymore. Tab. 1 shows the increasing reverberation time and the perceptually more relevant early decay time. Both are increasing with room size and therefore seem not to be a good predictor in this case also clarity values as suggested in [7] are not fitting here. As the source distance and especially the distance to the early reflections from the concrete wall behind the source is way higher than the critical distance of direct to reverberant energy (which calculates to be less than 2m for any of the room geometries) the diffuse reverberant part may be dominant and cover the early reflections. This includes the assumption that a change in reverberation time might not be directly linked to one dimensional room geometry.

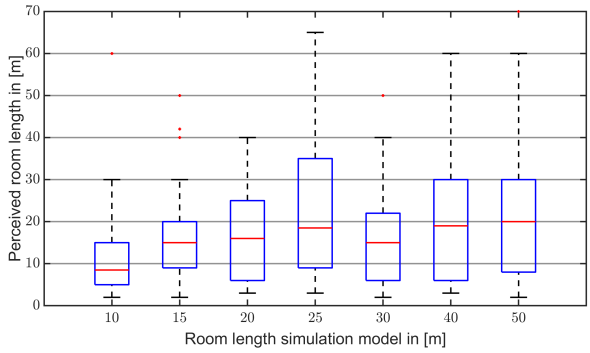


Figure 5: Listening test results for the perception of the same room and source receiver combination, but different room lengths. Height and width of the simulated model are six meters so that the room length times 36 results the volume.

Correlation

To compare the drawings with results of the listening test fig. 6 shows the perceived distances in the listening test part as well as the drawn source and wall distances in the drawing part exemplary for one source and room dimension. As described in the methods section two sources were fixed in their position and constantly playing. These two sources were used as references to compare the results of the distance perception test with the drawings. The top figure shows the results for the group of people talking (and making dining sounds with their dishes). In blue the individual mean of the perceived distances in the listening test for a source at four meter distance is plotted. Note that the participants on the x-axis are sorted by this distance. Red dots indicate the drawn source distance of the group of people. A simple Pearson correlation reveals a correlation of 0.03 stating no correlation at all which can be also taken from the figure as the red dots do no to react on the increasing blue ones. The frontal guitar player results the same findings and a correlation coefficient of 0.08 when compared to the perceived distances of a nine meter distant source. It should be noted that any shifting of the curves will not influence the correlation, therefore, correlating for e.g. the differences between auralized and perceived distances will not change the correlation. The bottom of fig. 6 illustrates the correlation between the perceived room length in the listening test (blue) and the drawn distances to the frontal wall (green) as well as the total room length (red) which takes also into consideration the distance to the drawn rear wall. The latter two do not correlate to the first one with absolute correlation coefficients smaller than 0.02. The same observation is made for the right side wall (which is also in the direction of a source) and the room width. Correlation coefficient between listening test and room width is -0.28 and for the wall distance from the receiver -0.02. The evaluation of wall distances and parts of the room geometry is an attempt to narrow down the independent variables of the investigation and of course complexity and duration of the listening test. Yet, room perception as such is more related to the volume as to the presented scales. Therefore, fig. 7 plots the correlation between the perceived volumes in

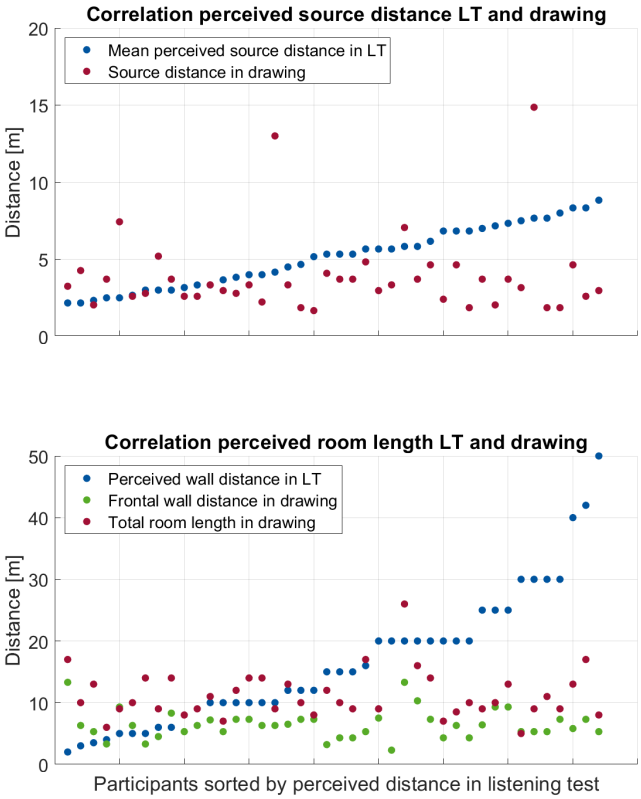


Figure 6: Correlation between drawings and listening test results. Top picture shows the sound source g labeled group of people talking and using dishes which is located at the right side of the listener at a distance of 4.3m. The followed listening test tested a source distance of 4m. Blue dots show the perceived distance of a violin in the listening test, red the drawn distance of the perceived group of people in the drawing part. Bottom: In blue is plotted the perceived room length in the listening test for a room length of 15m, in red the drawn distance from receiver to frontal wall and in green the total drawn room length.

the listening test (calculated from the perceived room length times the width and height of the model) in blue and the drawn volume (red) and floor area (green) in the drawing part. Again, absolute correlation coefficients of lower than 0.25 indicate no correlation.

Discussion and outlook

The listening test indicates that a calibration based on inherent biases of participants is not possible, at least not with the results from this test. Post-calibration of the drawings would lead to non-matching and inconsistent volumes and source distances distributed as arbitrary to each other as they were before. The distance perception parts suffers from the bias of the reference source and the idea of getting familiar with the source and environment is missing in the drawing part, as there is no fix reference available. For the space perception part it might be a better measure to let the participants draw a shoe box room to access the perceived volume. This way all dimensions can be increased evenly and together with a more homogeneous choice of materials may lead to more linearly perceived room acoustics.

Acknowledgement

The authors like to thank the participants of the listening test. For the implementation and evaluation of the listening test the ITA-Toolbox [8] was used. This research was

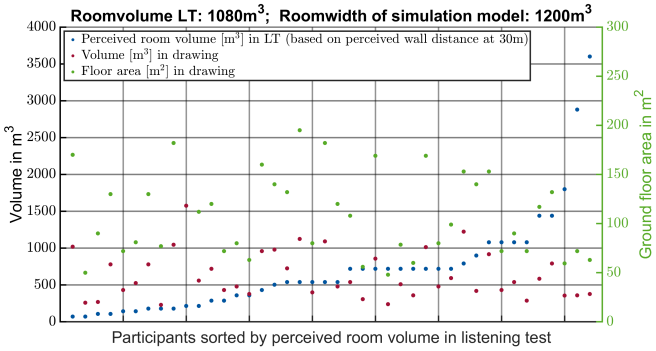


Figure 7: Results of the listening test sorted by perceived room volume in blue (calculated from the perceived wall distances times the real width and height of the simulation model). Red denotes the drawn volume and green the drawn floor area.

financed by the Head Genuit Foundation under the project ID P-16/4-W.

References

[1] Josep Llorca-Bofí, Vicent Llorca-Bofí, and Ernest Redondo Domínguez. Representation of the soundscape in the architectural design process. *EGA Revista de Expression Grafica Arquitectonica*, 24(37):192–203, 2019.

[2] Sönke Pelzer, Lukas Aspöck, Dirk Schröder, and Michael Vorländer. Integrating real-time room acoustics simulation into a CAD modeling software to enhance the architectural design process. *Building Acoustics*, 4(2):113–138, apr 2014.

[3] Bruno Masiero and Janina Fels. Perceptually robust headphone equalization for binaural reproduction. In *130th AES Convention, Convention Paper 8388*, London, UK, 2011.

[4] Josep Llorca-Bofí, Michael Kohnen, and Michael Vorländer. Disentangling spatial sound perception with drawings. In *Fortschritte der Akustik - DAGA 2020*, Hannover, 2020.

[5] Pavel Zahorik, Douglas S Brungart, and Adelbert W Bronkhorst. Auditory distance perception in humans: A summary of past and present research. *Acta Acust. united Ac.*, 91(3):409–420, 2005.

[6] Stephan Werner, Florian Klein, Thomas Mayenfels, and Karlheinz Brandenburg. A Summary on Acoustic Room Divergence and its Effect on Externalization of Auditory Events. *2016 Eighth International Conference on Quality of Multimedia Experience (QoMEX)*, pages 1–6, 2016.

[7] Claudiu B. Pop and Densil Cabrera. Auditory room size perception for real rooms. *Annual Conference of the Australian Acoustical Society 2005, Acoustics 2005: Acoustics in a Changing Environment*, (Experiment II):91–97, 2005.

[8] Marco Berzborn, Ramona Bomhardt, Johannes Klein, Jan-Gerrit Richter, and Michael Vorländer. The ITA-Toolbox: An open source MATLAB toolbox for acoustic measurements and signal processing. In *Fortschritte der Akustik – DAGA 2017*, pages 222–225, Kiel, Germany, 2017.