

**Katharina Groß-Vogt, Marian Weger,
Matthias Frank, and Robert Höldrich**

Institute of Electronic Music and Acoustics
University of Music and Performing Arts Graz
Inffeldgasse 10/3, 8010 Graz, Austria
{vogt, weger, frank, hoeldrich}@iem.at

Peripheral Sonification by Means of Virtual Room Acoustics

Abstract: Peripheral interaction is a new approach to conveying information at the periphery of human attention in which sound is so far largely underrepresented. We report on two experiments that explore the concept of sonifying information by adding virtual reverberation to real-world room acoustics. First, to establish proof of concept, we used the consumption of electricity in a kitchen to control its reverberation in real time. The results of a second, in-home experiment showed that at least three levels of information can be conveyed to the listeners with this technique without disturbing a main task being performed simultaneously. This number may be increased for sonifications that are less critical.

Scenario. While I am at home, a few sounds, random as well as intentional, appear in the background of my attention: traffic and the twittering of birds through an open window, a radio program, the rattling of dishes while I cook. All is part of the soundscape with specific room acoustics with which I am implicitly familiar. Suddenly, I become aware of a change in the room's reverberation: The room acoustics are virtual and are actually a sonification of the household's electrical power consumption. The boiler started heating!

Our hearing can be exploited at the periphery of our attention to convey information about anything that is relevant to us. We can realize this in an unobtrusive way by manipulating sounds that exist in the soundscape. This idea is explored and discussed in this article.

Background

Following Weiser and Brown's (1996) well-known call for calm computing, new modes of interaction are needed in times of ubiquitous computing and ever-growing amounts of data. Using the recent taxonomy of the interaction-attention continuum introduced by Bakker and Niemantsverdriet (2016), peripheral displays (Matthews, Hsieh, and Mankoff 2009) link the gap between systems of focused

interaction (e.g., as those in most of our usual computing devices and in many sonifications), on the one hand, and systems of implicit interaction (e.g., automatic lighting or doors, or an alarm clock that wakes one up at the most appropriate moment in the sleep cycle), on the other. Many other related concepts might be mentioned here: ambient information systems (Pousman and Stasko 2006), ambient intelligence and persuasive technology (Verbeek 2009), auditory augmentation (Bovermann, Tünnermann, and Hermann 2012), embedded sonification (Barrass and Barrass 2013), and ambient sonification systems (Ferguson 2013). Ideally, a new display type should be developed that facilitates the alternation between the required level of attention, from peripheral to focused interaction and back again. Ambient systems have been shown to provide such an alternative. One example is the Student Feedback Orb (Hazlewood, Stolterman, and Connelly 2011), which uses ambient light. In a case study, an ambient orb was used to show lecturers the results of their course evaluation surveys in their offices. The orb alternated between red and green depending on the weekly results of student evaluations.

Light is often used in prototypes of peripheral interaction, whereas audio is generally avoided, possibly due to acoustic and psychoacoustic issues (e.g., interplay with the soundscape or room response, overall noise level, or interference with other signals). These problems may be partly overcome by existing methods of digital signal processing and expert tuning, however. State-of-the-art technology in 3-D audio includes, among other techniques, virtual room acoustics, interactive audio scenes

Computer Music Journal, 44:1, pp. 71–88, Spring 2020
doi:10.1162/COMJ_a.00553
© 2021 Massachusetts Institute of Technology.

(e.g., orientation-tracking headphones with binaural rendering), 3-D loudspeaker arrays, or adaptive digital signal processing systems (e.g., Pulkki and Karjalainen 2015; Zotter and Frank 2019). Although active noise cancellation is standard, similar algorithms may augment certain audio features or even perceptual entities in the auditory scene.

Current research on auditory ambient displays shows us how to cope with the strengths and weaknesses of the audio modality. The need for audio to be unobtrusive for nonusers may be met by using personalized sounds. Butz and Jung (2005), for example, experimented with a prototype of encoded Muzak for a mall or museum. By adding specific instruments or motifs to the arrangement, employees can be summoned without annoying the customers. In an interesting experiment with ringtones, John Brown (2016) attempted to show that using personalized sounds can “encalm” ringtone interaction. The evaluation of this EEG-based experiment produced no significant results, however. The examples of Butz and Jung and of Brown display only binary information (call / no call). Richer data streams have been sonified using ambient systems, but these were proofs of concept with little or no evaluation. These sonifications have usually focused on everyday interactions to exploit hearing as an additional information channel—for example to induce more-economical driving attitudes (Hammerschmidt, Tünnermann, and Hermann 2014), when the sound produced by knocking on a door is used to reveal whether or not someone is inside (Tünnermann, Hammerschmidt, and Hermann 2013), or when opening a wardrobe door is used to trigger an auditory weather report (Ferguson 2013).

Auditory augmentation provides an ingenious method for displaying richer information in an unobtrusive way. Bovermann, Tünnermann, and Hermann (2012) introduced auditory augmentation with a focus on structure-borne sounds. In their *WetterReim* project, for example, the sounds of keys during typing change depending on weather parameters. We have generalized the concept of auditory augmentation. In a recent study, summarized in the *PilotKitchen* Experiment section, we used artificial reverberation in a kitchen to convey its consump-

tion of electrical power in real time (for further details, cf. Groß-Vogt et al. 2018). In this case, we found that only extraordinary changes come to the foreground of attention—for example, when the small kitchen reverberates as if it were a church. In our broader definition, auditory augmentation makes use of either object-specific sounds caused by interaction, or adds environmental sounds that blend into the sound scene.

Time tagging, which is a similar concept in ubiquitous music (“ubimus”) research, uses local acoustic cues for aesthetic decision making. Time tagging is based on local resources, provides synchronous feedback, and demands unobtrusive audio-processing techniques (Keller et al. 2010; Farias et al. 2015; Keller 2018; Keller, Schiavoni, and Lazzarini 2019).

Hypothesis and Related Research

We hypothesized that auditory augmentation can be used to convey user-relevant information with embedded sounds that are easily learned and have little distraction potential. In this way we can exploit the auditory modality at the periphery of attention. We conducted two experiments to explore our hypothesis. In the *PilotKitchen* experiment, we set up an adaptive system for room acoustics in the kitchen of the Institute of Electronic Music and Acoustics at the University of Graz that was controlled by the electrical power consumption of kitchen appliances. In the *RadioReverb* experiment, we explored how many levels of information can be conveyed using this method.

Parthy, Jin, and van Schaik (2004) used reverberation for ambient data communication. Their study found that reverberant decay time must increase by approximately 60 percent or decrease by approximately 30 percent to be clearly perceived. We followed a different argument for tuning our reverberation levels, discussed below. The *Weakly Intrusive Ambient Soundscape* is an older concept by Kilander and Lönnqvist (2002). Kilander and Lönnqvist conveyed individual notifications in a ubiquitous service environment, with a sound associated to each user. Playback volume and

Figure 1. The kitchen at the Institute of Electronic Music and Acoustics (IEM), where the PilotKitchen experiment was conducted, is a small room of 16 square meters that was used regularly by

15 staff members plus students and visitors. Microphone, computer, and loudspeakers were mounted unobtrusively on the ceiling and on top of the cupboards (a). The measurement of the

electrical power consumption was conducted over simple plugs that sent the data wirelessly (b). (Photo: Till Bovermann.)



(a)



(b)

reverberation were used to convey three levels of intensity for the notification. In a more recent experiment, Lockton et al. (2014) used sonification to communicate in-home electricity use. First, Lockton and coworkers experimented with abstract sounds; when they found that these were not sufficiently unobtrusive, however, they turned to birdsong. Our experiments pursued a similar design scenario.

For evaluating peripheral displays, Matthews, Hsieh, and Mankoff (2009) developed design dimensions and criteria based on the framework of activity theory. As evaluation criteria they listed awareness, the detection of the system's breakdowns, distraction, and appeal. To these standard factors they added learnability, which is necessary for automated operation (and consequently, for peripheral display). For peripheral sonification, we needed to add effectiveness of information transfer. This factor was explored in detail in the experiment described in the RadioReverb Experiment section.

PilotKitchen Experiment

This section gives an overview of the PilotKitchen experiment that we reported in a previous paper (Groß-Vogt et al. 2018). The PilotKitchen was the prototype of a system for auditory augmentation that conveyed information on the electric power consumption of the kitchen of the Insti-

tute of Electronic Music and Acoustics (IEM; <https://iem.kug.ac.at>). The purpose of the study was to raise awareness of power consumption. Data from the kitchen appliances were collected in real time. Pictures of the kitchen and the wall plugs measuring the electric power consumption are shown in Figure 1. A demo video with binaural audio recording of the PilotKitchen can be found online at https://dx.doi.org/10.1162/comj_a_00553.

System Design of the PilotKitchen

Data on the electric power consumption of five kitchen appliances were collected from wall plugs using the Fibaro intelligent-home system (www.fibaro.com/en/the-fibaro-system/wall-plug) that transmitted the data over Z-Wave (Yassein, Mardini, and Khalil 2016). The data from a dishwasher, a coffee maker, a water kettle, a microwave oven, and a refrigerator showed two patterns—one stemming from the technical cycle of each appliance, and the other from the interaction of the kitchen users. We implemented a self-adapting system that conveyed direct feedback from the actual electrical power consumption and also related this information to the typical weekly user pattern. The algorithm of data preprocessing consisted of an initialization and three iterative steps: (1) smoothed real-time electrical power consumption (EPC), (2) baseline week EPC, and (3) the difference relation.

In the initialization we gathered three weeks of total EPC data and averaged the data to obtain one typical week. This was our initial baseline week EPC.

For the smoothed real-time EPC we collected real-time EPC data every second. Because of the rapid fluctuations in the data caused by the nature of the electric devices and the measurement, it was necessary to smooth the raw data. This was done with a leaky integrator over the last 15 minutes that retained a fraction a of the past output and added $1 - a$ of the actual input. The smoothing constant a expressed the relation between the filter time constant $\tau = 900$ sec (15 minutes) and the sampling interval $T = 1$ sec,

$$a = e^{-\frac{T}{\tau}} = 0.99.$$

With this smoothing constant we collected a new EPC value every second.

For the baseline week EPC, the smoothed real-time value obtained above was used to update the baseline week of typical EPC. We applied another leaky integrator with $a = 1/3$. Thus, the EPC of the present moment was added with a weight of two-thirds to the updated baseline, while the average over all previous weeks contributed with a weight of one third.

For the difference relation, we compared the real-time value to the value of the baseline week that corresponded to the present moment. For instance, if it was Tuesday at 13:05:10, we subtracted the corresponding value of the baseline week from the present value. The result could be positive, when the actual consumption was higher than usual, or negative, when it was lower. Outliers were clipped, and the resulting number was mapped to a range of $[-1, +1]$. This number was further used in real time to control the reverberation level.

The sonification design was based on virtual acoustics by adding reverberation to the room. In our setup, we installed sound absorbers in the kitchen to lower its real reverberation. Then we recorded the environmental sound with a microphone, applied filters and a reverberation algorithm, and played the signal back over loudspeakers in real time.

The Preset 0 of the virtual room acoustics was then tuned to approximately restore the original reverberation. The difference relation controlled the level of reverberation within three sound presets: 0 corresponded to a “typical” kitchen use, that is, a plausible but virtually added kitchen reverberation; whereas +1, that is, a particularly high consumption, led to reverberation resembling that of a church. Atypically low consumption, that is, a value of -1 , turned the virtual reverberation off. The reverberation levels were interpolated between these extremes.

Evaluation and Discussion of the PilotKitchen

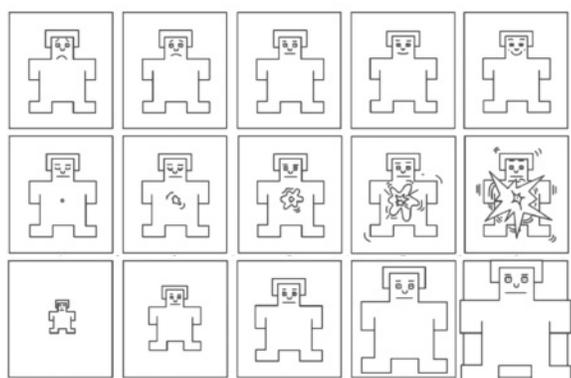
The evaluation of the PilotKitchen experiment aimed at assessing users’ perception and evaluation of the system. In the first round of the evaluation, the system was installed for ten days. Kitchen users were notified about the experiment but given no additional information about it. After the experiment was finished, the users were asked to answer open questions in a questionnaire. Their comments included assessments of their affective reactions to our prototype system and included subjective interpretations of how to control the system (for details, see Groß-Vogt et al. 2018).

For the second round of evaluation, the participants were introduced to the audio system but not to the underlying algorithm—that is, we did not reveal the relationship between the EPC and the reverberation level. This was due to our aim to assess the user’s ability to perceive the peripheral sonification but not to estimate the actual EPC level. (Later we did take up user estimation of reverberation levels, this is discussed in the RadioReverb Experiment section.) Furthermore, the levels of control and valence found in the first experiment were explored more systematically, and complemented by adding the dimension of arousal. Here we followed Stevens, Murphy, and Smith (2016), who used self-assessment manikins (SAMs) to study affective reactions to soundscapes. Our evaluation used five-scale SAMs for the dimensions of control, valence, and arousal in a diary study because these

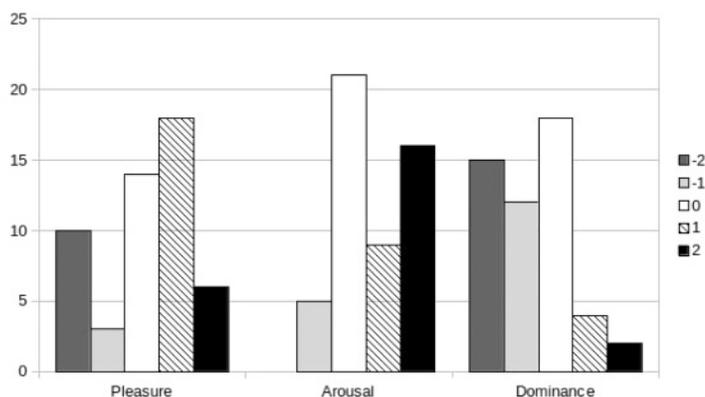
Figure 2. Five-scale self-assessment manikin (SAM) (a). The SAM represents the dimensions valence (top row), arousal (middle), and control (bottom), thus measuring

the affective reaction to our prototype system. (Pictograms taken from Soares et al. 2013.) For analysis the five states of the above diagram were given the numbers -2 ,

0 , 1 , and 2 , where 0 marks the neutral SAM in the middle. The results for the SAM rating showed a wide range of affective responses in the diary entries (b).



(a)



(b)

provide a quick and intuitive visual rating scheme (see Figure 2a). The experiment was conducted over a period of two weeks.

The resulting data from 60 diary entries by 14 participants did not show a correlation between the perception of reverberation and the difference relation driving the sonification. This might be due to statistical deficiency, to interpersonal differences in perceiving the virtual acoustics, or to insufficient testing of the tuning of the reverberation levels. Qualitative results of the SAM ratings revealed a heterogeneous attitude towards the system (see Figure 2b).

The results of the PilotKitchen experiment showed that the chosen evaluation techniques were not sufficient. In general (following Hazlewood, Stolterman, and Connelly 2011, p. 877), the challenge for prototypes of peripheral interaction is to explore

[h]ow to provide in-depth evaluations on something that is defined as blending with the surrounding world, and meant to be (in some respects) ignored.

Hazlewood and colleagues argued strongly in favor of long-term and in situ measurements. They reported an interesting case study (“clouds and lights”), in which three ambient displays were installed in a publicly shared atrium of a large building in an effort to persuade passers-by to use the stairs instead of the elevators. Interestingly,

quantitative data showed that staircase use increased significantly. When interviewed, however, users denied that they had changed their behavior. This example shows that at-the-periphery-of-attention qualitative evaluation alone is not sufficient. On the other hand, if indirect quantitative measurement is not possible, qualitative data are needed to resolve issues of test and evaluation design. This was also discussed by Hazlewood and coworkers, who reported that indirect measurement in the above-mentioned Student Feedback Orb study failed for reasons that did not become evident until the post hoc interviews.

For our PilotKitchen experiment, we concluded that a simplified experimental design is needed to conduct a thorough evaluation of the method. The RadioReverb experiment was designed to meet this criterion.

RadioReverb Experiment

The RadioReverb experiment, designed and carried out in April 2020 in Austria, was originally planned as a listening experiment in our sonic interaction design lab at the IEM. Due to the COVID-19 pandemic, we redesigned the experiment as an in-home test using an app on participants’ smartphones and headphones.

On the basis of the results of the PilotKitchen experiment, we posed the following research question:

When the environmental sounds of a room are augmented by changes in the room acoustics, how much information can be conveyed at the periphery of attention (without disturbing the main task the user wants to perform)?

Design of the RadioReverb Experiment

To investigate our research question, we set up the scenario of a user who was casually listening to a half-hour radio feature at home as the main task. The reverberation of the room changed randomly, and the participants of the experimental group (EG) were asked to estimate the changes they perceived on a simple GUI on their smartphones. Participants in a control group (CG) listened only to the radio feature at constant reverberation. Then, both groups were requested to answer questions on the content of the feature. This ensured that the EG participants completed the task by listening mainly to the radio and only peripherally to the reverberation, as opposed to altering the main task by tracking changes in the reverberation. The radio signal was played back via a virtual loudspeaker, and the room response was carefully simulated to correspond to that of a standard living room and rendered binaurally.

Audio Material and Participants

To meet the prerequisite of providing an information source that should be relevant to the participants (Hazlewood, Stolterman, and Connelly 2011), we chose two episodes from a current radio feature “Ö1-Radiokolleg,” produced by ORF (the Austrian Broadcasting Corporation). These episodes on the topic of food plants (fig and millet) were chosen with the expectation that they would be of general interest. Episodes are around 15 minutes each and were played back consecutively as one file. Approximately half of the 15-minute episode contained speech in a dry studio; the other half consisted of interviews recorded in a variety of settings. This material would not be ideal for a strict listening test, but it supported our in-home, environmental test design.

We recruited 33 participants; although some did not finish the experiment owing to technical issues. Ultimately, 12 participants (six men, six women) were taken into account in the CG, with the remaining 17 (13 men, 4 women) making up the EG. All members of the EG were staff or students of our institute, with the exception of two musicians. All were advanced listeners and were assumed to have normal hearing. The participants of the CG were recruited from among the acquaintances of the authors, and, with the exception of two sound engineers, had no special audio or music background. The experiment was conducted in German, the first language of all participants.

Test Procedure

The experiment was conducted on each user’s smartphone using headphones (not earbuds) with the MobMuPlat app, described below. The users were asked to first install the environment and load their personalized preset, then to find a comfortable sitting place in a normal room with little distraction. The app guided the participant through the experiment and provided a slider on the GUI as the experimental interface. The experiment consisted of four phases:

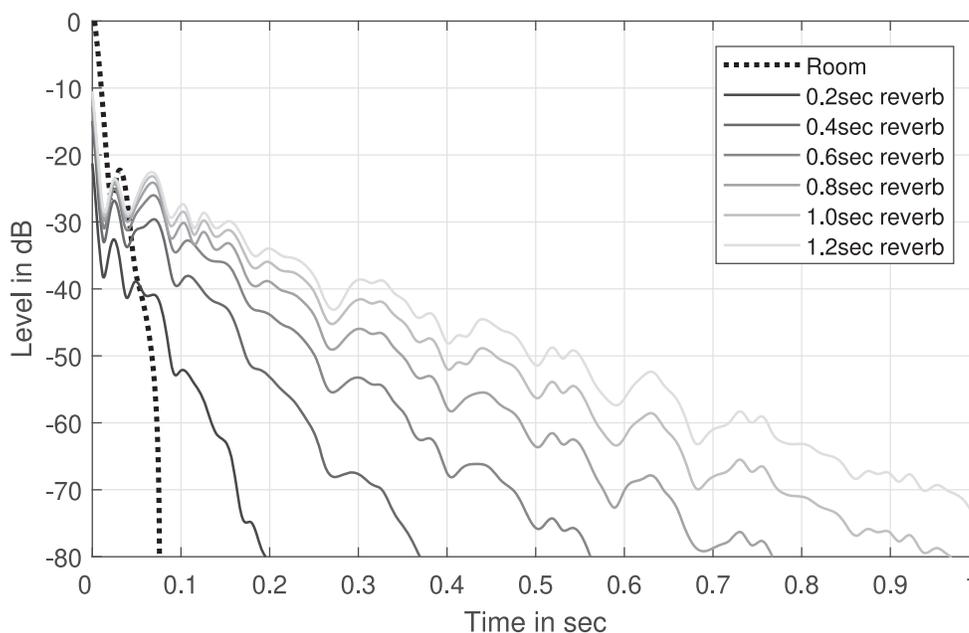
1. Training I,
2. Training II,
3. Experiment, and
4. Post hoc survey.

In Training I, a 46-sec excerpt from a radio feature was played. During playback, the reverberation was iterated through twelve discrete levels, starting from minimum, up to the maximum, and then back to minimum reverberation. The slider showed the correct setting simultaneously. This procedure could be repeated on demand.

In Training II, a 3.5-sec speech sample was played repeatedly at different levels of reverberation. For each trial, the participants were asked to estimate the corresponding slider position. The correct answer was revealed through an additional slider in the background as soon as the sample ended. After a pause of 1.5 sec, the next trial was presented. The level of reverberation was randomized by presenting

Figure 3. Peak envelope (10-msec window) of impulse responses for dry room reflections from the image-source model and

the additional diffuse FDN reverberation with different reverberation times.



consecutive rows of the twelve levels in random order. The procedure was repeated for at least two such rows (24 random jumps) and could be continued as long as the participant wanted. After completion of Training II, the experiment proper was carried out, in which the radio feature was played for half an hour. The participant was instructed to estimate each change on the slider as rapidly as possible. For the comfort of the participants, the interface could be released between changes because the slider retained the last value.

Finally, in the post hoc survey, the participants were asked to return the gathered data and fill out an online survey about the content of the broadcast and about the experiment.

Reverberation Levels

We chose the reverberation levels based on perceptual considerations, taking into account just-noticeable differences (JNDs) from the literature and the physically plausible range as given by the size of the simulated room ($4 \times 5 \times 3$ meters). Previous work by members of our group (Weger, Hermann, and

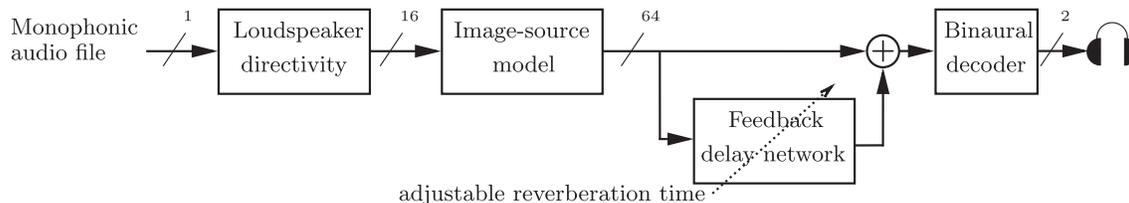
Höldrigh 2018) has shown that limiting the auditory augmentation to a plausible range leads to a calm sonification that fits naturally and seamlessly into the everyday acoustic environment. According to Connell and Keane (2006, p. 95):

A highly plausible scenario is one that fits prior knowledge well: with many different sources of corroboration, without complexity of explanation, and with minimal conjecture.

In our case, the lower threshold was tuned to a dry reverberation of a standard living room with discrete reflections only, while the upper threshold simulated conditions in a room of the same dimensions that was tiled as a bathroom, resulting in 1.2 sec of diffuse reverberation. We argue that this is a plausible range because it can be achieved physically for the given room dimensions.

Within this range, we achieved seven steps of reverberation (room only, room + 0.2-sec diffuse reverberation, room + 0.4-sec diffuse reverberation, . . . room + 1.2-sec diffuse reverberation) that clearly differed from each other (see Figure 3). The exact number of perceptually distinguishable steps

Figure 4. Block diagram for rendering virtual room acoustics. Numbers indicate how many audio channels are rendered at each step.



in terms of JNDs could not be calculated in a straightforward way because the different settings differed not only with respect to the reverberation time but also in the direct-to-reverberant energy ratio. Taking only the reverberation time and the corresponding JNDs for short reverberation times from Niaounakis and Davies (2002) would result in 27 distinguishable steps between dry room and the maximum reverberant setting. Considering the direct-to-reverberant energy ratio, however, only four JNDs were spanned within the chosen range of settings (Larsen et al. 2008). For center time (Cox, Davies, and Lam 1993), as well as clarity (Bradley, Reich, and Norcross 1999), eleven JNDs would fit into that range. Informal listening revealed that the seven steps were clearly distinguishable, and that we as human beings are able to perceive a range of around ten JNDs. Note that the range of seven items also has a psychological rationale. In a seminal article, George A. Miller (1956) postulated—based on experimental data of different modalities—that our ability to process information is “the magical number seven, plus or minus two.” The choice of seven levels was thus deemed sufficient.

During the whole radio episode, all possible transitions between the seven levels (i.e., at least 42 transitions) were presented in random order for each participant, with each setting lasting for at least 7 sec.

Sound Rendering for Virtual Room Acoustics

The sound was rendered with the Reaper digital audio workstation (www.reaper.fm) using Ambisonics technology (Zotter and Frank 2019) and the IEM Plug-in suite (<https://plugins.iem.at>). The individual processing steps are shown in Figure 4.

Monophonic audio was fed into a DirectivityShaper plug-in that created 16 channels to rep-

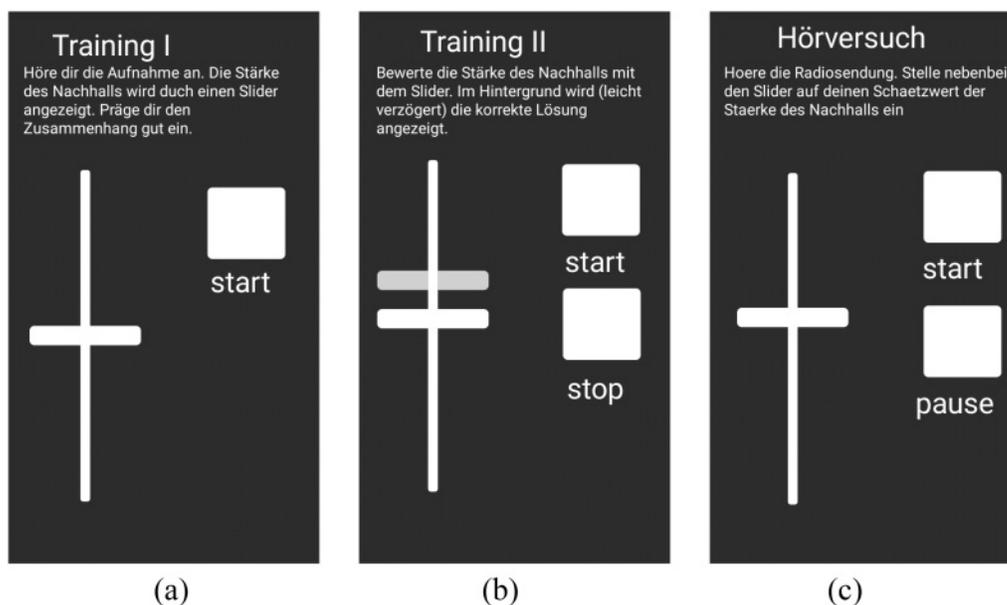
resent typical frequency-dependent directivity of a loudspeaker in third-order spherical harmonics. The orientation of the directivity pattern was rotated by 180° to face the listener using the SceneRotator plug-in. The rotated directivity pattern was sent into the RoomEncoder, which used an image-source model of a 5 × 4 × 3 m shoebox-shaped room with 236 reflections to create basic, dry living room acoustics. The virtual loudspeaker was placed at (1.7, 0.3, -0.5) m and the listener at (-0.3, 0.3, -0.5) m relative to the center of the room, resulting in a typical listening distance of 2 m. The walls of the shoebox room had a reflection coefficient of $\Gamma_{0..8\text{kHz}} = -2$ dB for lower frequencies and $\Gamma_{8\text{kHz}..20\text{kHz}} = -7$ dB for higher frequencies. To simulate the effect of a carpet on the floor, the floor reflections were attenuated by 2 dB. The output of the RoomEncoder was 64 channels so that seventh-order Ambisonics could make use of the highest possible spatial resolution. The output of a 64 × 64 feedback delay network (FDN; cf. Stautner and Puckette 1982), was added to these 64 channels according to the different levels of reverberation using the FDNReverb plug-in. The parameters of the plug-in were chosen to take into account the average free path of the room and a 0.1-sec fade-in to increase diffusivity (Blochberger, Zotter, and Frank 2019) for a smooth blending into the early reflections of the image-source model. A frequency-dependent reverberation time with about half the reverberation time at higher frequencies was achieved by high-shelf attenuation at 8 kHz. The resulting envelopes of the dry room and the additional diffuse reverberation levels were depicted in Figure 3. The BinauralDecoder created the headphone signals from the resulting 64-channel Ambisonics stream with state-of-the-art decoding technology (Schörkhuber, Zaunschirm, and Höldrich 2018; Zaunschirm, Schörkhuber, and Höldrich 2018). The detailed parameter settings of

Figure 5. Screenshots of the experiment GUI during Training Phase I (a), Training Phase II (b), and the actual Listening Experiment (Hörversuch) (c). Onscreen instructions

for the three phases were as follows. Training I: Listen to the recording. The slider shows the strength of the reverberation. Commit this relationship to

memory. Training II: Assess the strength of the reverberation on the slider. After a slight delay, the correct position will be shown in the background.

Listening Experiment: Listen to the broadcast. While listening, move the slider to match your estimate of reverberation strength.



each plug-in can be found as supplementary material at https://dx.doi.org/10.1162/comj_a_00553.

Collecting In-Home Data via MobMuPlat

We used the free MobMuPlat app (Iglesia 2016, see also www.danieliglesia.com/mobmuplat), which allowed the participants to conduct the experiment themselves at home on their smartphones or tablets. MobMuPlat is a “mobile music platform” for Android and iOS that provides a GUI for the audio engine in Pure Data (Pd). The experiment procedure was coded in Pd and controlled via the GUI. The GUI is shown in Figure 5.

The participants returned their recorded data in the form of a text file containing the slider values sampled at a rate of 100 Hz. Of the 17 participants, 3 accidentally performed the experiment with the wrong audio sampling rate: Participant 16 listened to the 44.1-kHz version with a 48-kHz sampling rate, while for Participants 42 and 44 it was the other way around. The resulting change in playback speed (−8% or +9%) was not considered relevant for our experiment, as JNDs for diffusivity and

reverberation time are defined relatively, not as absolute values. For unknown reasons, the data from two others (Participants 6 and 40) were 1 sec too long. All collected data were therefore resampled to the same length of 163,662 samples (about 27 minutes 17 seconds).

Evaluation in the Post Hoc Online Survey

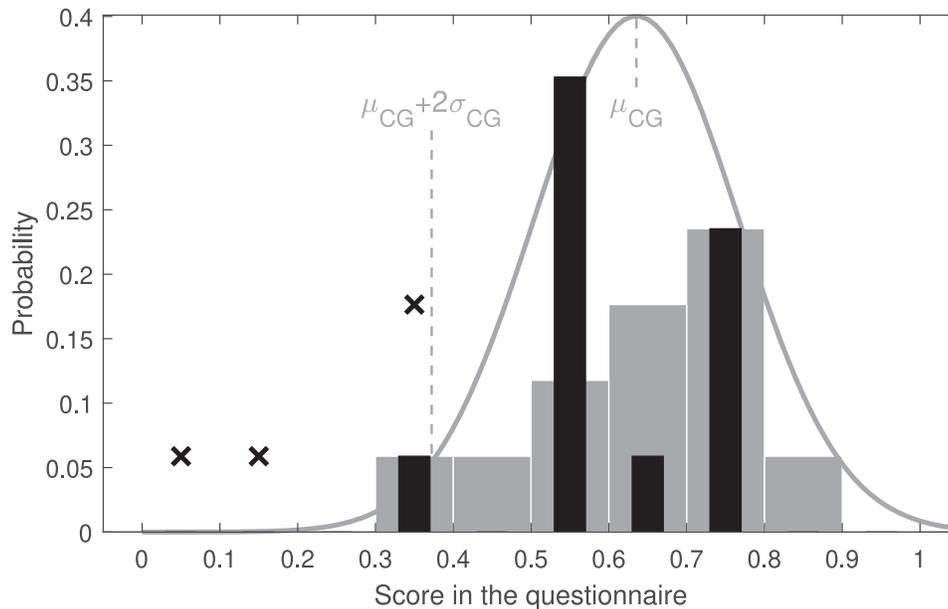
After the experiment, the participants were asked to fill out an online survey. This was anonymized for the CG participants but not for the EG participants, as the data needed to be linked to the quantitative data collected via the MobMuPlat app. The questionnaire contained a multiple-choice test on the content of the radio feature, and the following questions on the circumstances of the experiment and on the subjective experience of the sound, whereby the CG participants were asked the first three questions only:

1. Multiple-choice, multiple-response statements on Episode 1 (fig).
2. Multiple-choice, multiple-response statements on Episode 2 (millet).

Figure 6. Distribution of the scores of the multiple-choice questionnaire for the experimental group (EG) and control group (CG). Data for CG (gray) are shown as a histogram and

the corresponding fitted normal distribution with mean μ and standard deviation σ . Data for EG (black) are also shown as a histogram. Five outliers (crosses) indicate scores below $\mu - 2\sigma$ from CG

participants. As these five participants seemed to be least focused on the content, they were categorized as sound-focused (SF).



3. Description of the room in which the experiment was conducted, and of possible distractions or technical issues during the experiment (formulated as an open question).
4. "How well do you think you were able to estimate the reverberation level?" (Likert item, 1-7).
5. "How much were you challenged by simultaneously listening to the radio and estimating the reverberation level?" (Likert item, 1-7).
6. "How realistically did the virtual radio and its reverberation blend into your listening environment?" (Likert item, 1-7).

by the participants on a given scale. In our case, a slider was used with a given beginning and end point but no further segmentation. Many effects are known for magnitude estimation, regardless of the presented modality (see Petzschner, Glasauer, and Stephan 2015). When we analyzed our data we took note of behavioral effects, such as the regression effect, where the reproduced range is smaller than the physical one, as well as sequential effects. We expected to find a hysteresis in the data, that is, a bias in estimates towards the recent history of the stimuli experienced.

The participants answered sets of detailed questions on the content of the radio feature, in two parts corresponding to the two episodes. Each set contained 14 statements on figs and millets, respectively, all of which were factually correct statements. The participants had to indicate whether or not the statement had actually been part of the radio feature. To validate the difficulty of this task, the CG participants were asked the same questions as the EG participants.

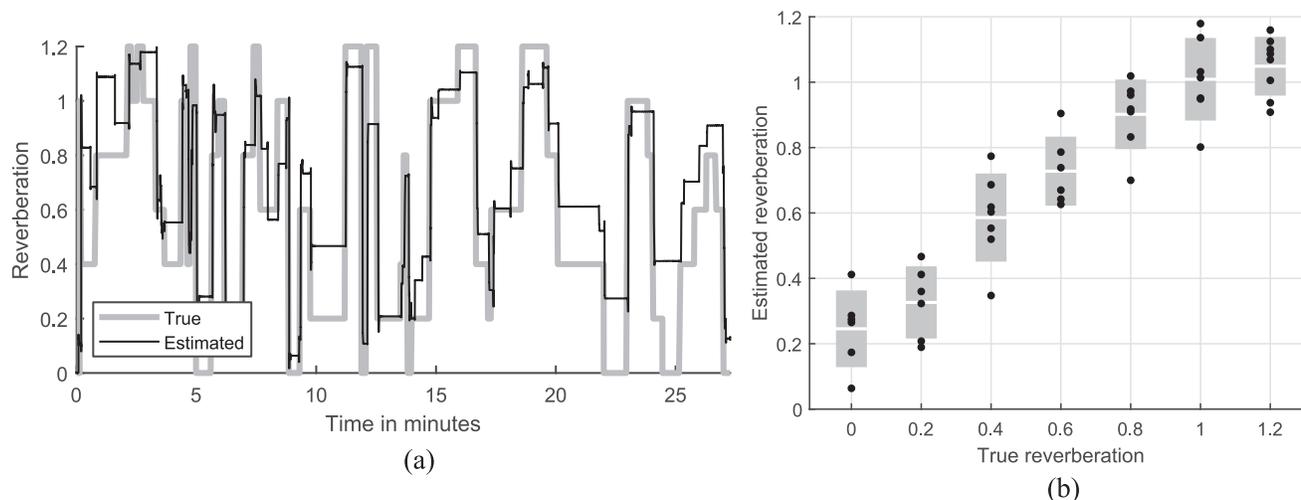
Our test design was a magnitude-estimation experiment: A certain stimulus was to be estimated

Results of the RadioReverb Experiment

First the data from the multiple-choice questionnaire were analyzed. Figure 6 shows the distribution of the EG data in comparison with that of the CG data. When we excluded the data from the five worst-performing participants, the results of the questionnaire from the CG were not significantly different from those of the remaining participants of the EG ($p = 0.33$). When we compared the results of the CG with those of the EG, we found that some of the EG participants were much less focused on

Figure 7. Experiment data for Participant 16. Time series of true reverberation with reference in bold gray and estimated reverberation in black (a) and estimated versus real value of reverberation (b).

In the latter graph, the corresponding values, their mean values (horizontal white bars), and standard deviations (gray rectangles) are shown for each level.



listening to the radio than the other EG participants. We call this group the sound-focused listeners (five participants with IDs 4, 6, 11, 23, and 42), whereas the other EG participants behaved as content-focused listeners in a more “peripheral” way, as we had intended in the experiment. The EG participants were therefore subdivided into sound-focused (SF) and content-focused (CF) listeners.

The data obtained from the MobMuPlat app were time series of estimated reverberation levels (response) in relation to the true reverberation levels (reference) for each participant (see one example in Figure 7a). To simplify comparisons, both are given in quantities of reverberation time T_{60} , as was also done for the added FDN reverberation. As a first step in the analysis, we estimated the delay of the response with respect to the reference via the maximum of their cross-correlation. This average delay was between 0.54 sec and 2.22 sec for all participants (mean 1.48 sec, standard deviation 0.55 sec). A moving-average instantaneous delay over time was computed in the same way by using a sliding rectangular window of four minutes in length. This instantaneous delay ranged from 0 to 3.88 sec for all participants, with an average delay over time of 0.99 sec.

For further analysis, the time series were divided into segments of constant true reverberation

(plateaus are shown in Figure 7a in the gray reference curve). The first segment was excluded from further analysis. In addition, the first 2.22 sec (i.e., the maximum average delay), as well as the last second of each segment were removed; for the remainder, the average over the response (i.e., the estimated reverberation) was calculated. After collecting average responses according to the corresponding reference level, we obtained six to nine estimates per participant and reverberation level, as shown in Figure 7b.

For each participant and level, the distribution of estimated reverberation was tested for normality by the Lilliefors test. The null-hypothesis of normal data was rejected in only 4 percent of the cases (and for a maximum of one level per participant). Therefore, we generally assumed a normal distribution for further statistical analysis. For each participant, we performed pairwise one-tailed Welch’s t -tests on all possible pairs of reference levels, with a 5 percent threshold for significance. For all participants, a difference of three reverberation levels (e.g., fifth versus second level) always resulted in estimates that differed significantly from each other. Below that level, the estimates of some participants were not significantly different from each other. When pooled over participants, all pairwise comparisons were significant; that is, the estimated reverberation

of each level was always significantly higher than that of all levels below and lower than the levels above, respectively.

When analyzing the estimated versus the true values, several different approaches can be used to define the number of levels that participants were able to identify correctly. As can be seen in the example of Figure 7b, it was quite obvious that the participants did not reach our seven reverberation levels. Due to the small number of data points per level and the fact that participants each have their own nonlinear mapping, we decided to use the statistical measure of effect size to analyze the data. Note that this choice follows the assumption that our levels are equally spread and that the perceptual distance between adjacent levels is similar over the whole range of reverberation. The data indicate a linear behavior; Figure 7b, for example, shows a rather linear function of average estimates over reverberation levels. The same behavior can be seen in the analysis of the average values when the data are pooled over participants.

The effect size measures the magnitude of an effect, in our case the effect observed in the answers when comparing a lower with the next higher level of reverberation, as measured with Cohen's d . For each of six steps between two adjacent levels A and B , Cohen's $d_{A,B}$ was calculated as the difference between the mean values μ_A and μ_B divided by the pooled standard deviation s , as was done for two independent samples of unequal size and variance:

$$d_{A,B} = \frac{\mu_A - \mu_B}{s}.$$

We summed up all six steps and compared this sum to different threshold values d_t as given in Table 1. Values of d_t of 1, 1.2, and 2 are usually interpreted to signify a "large," a "very large," and a "huge" effect, respectively. When we added the one level that can always be heard as a baseline, the resulting N in

$$N = \frac{\sum_1^6(d_{A,B})}{d_t} + 1$$

yielded the number of levels that our participants were able to discriminate with the given threshold effect size d_t . Any value of effect size can also be

Table 1. Numbers of Discriminable Levels

d_t	1.19	1.81	2.33	2.77
P_S	0.8	0.9	0.95	0.975
N_{all}	6.1	4.3	3.6	3.2
N_{CF}	5.8	4.1	3.4	3.0
N_{SF}	6.8	4.8	4.0	3.5

For different levels of superiority P_S , corresponding to the thresholds d_t , the table shows the results for the number of levels discriminable by all listeners, those who were content focused (CF), and those who were sound focused (SF). There was a minimum of 3.0 discernible levels and a maximum of 6.8, taking into account a lower probability of correct estimates.

expressed in terms of "probability of superiority" P_S , that is, the probability that a participant decides for the correct answer. This probability is actually the same as the area under the receiver operating characteristic curve.

The results of this analysis are given in Figure 8, averaged over all participants, as well as averaged, separately, over the groups of SF and CF listeners. In Figure 8a, N was plotted against a continuous P_S , whereas for Figure 8b, four selected P_S (corresponding to four values of d_t) were chosen. When choosing a d_t of 2.77 (corresponding to a P_S of 0.975), CF listeners achieved an average of 3.0 discriminable levels (standard deviation 0.4), whereas SF listeners achieved an average of 3.5 levels (standard deviation 0.4). The number of perceived levels was significantly higher for SF listeners than for CF listeners: $t(9.44) = 2.161$, $p = 0.029$.

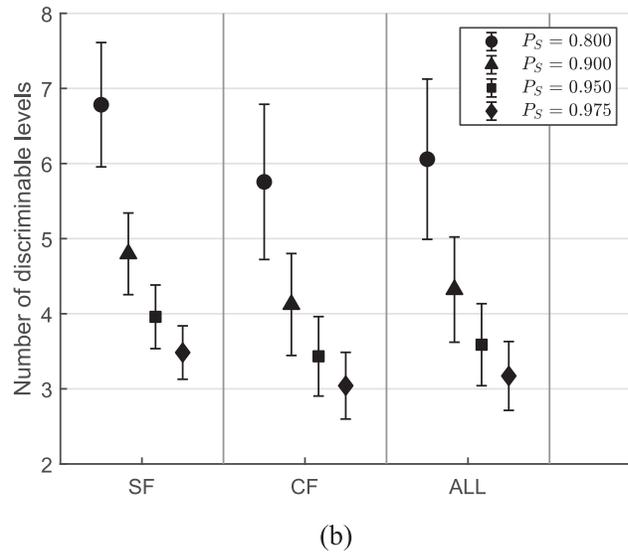
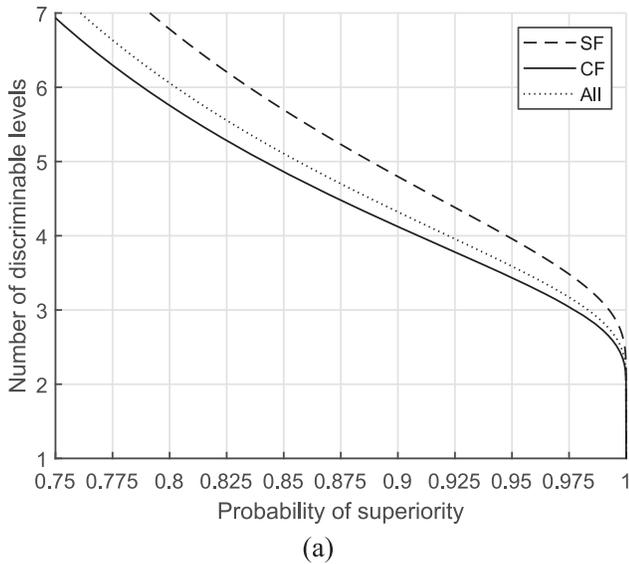
For each reverberation level, the individual estimates can be grouped into upward jumps (the prior reference level was lower than the current one) and downward jumps (the prior reference level was higher than the current level). In this way, the estimated reverberation can be plotted against the true reverberation for upward and downward movements separately. Due to the unavoidably small number of data points, a per-subject analysis would not be meaningful. Figure 9 shows the resulting curve, pooled over all participants. The plot shows two effects that can be expected when dealing with magnitude estimation (cf. Petzschner, Glasauer, and Stephan 2015). First, the regression effect is the

Figure 8. Number of discriminable levels for different thresholds d_t (Table 1) and their corresponding probability of superiority (P_S). Sound-focused (SF)

listeners reach slightly higher values than content-focused (CF) listeners: the whole range of levels over P_S (a); results for four thresholds we recommend for further use

in sonifications (b). The error bars in the latter graph show the standard deviation over participants' individual results. For a huge effect size, about three levels can

be discriminated with a probability of $P_S = 0.975$. CF listeners achieved more than five levels with $P_S = 0.8$.



tendency that estimates are systematically biased towards the center of the distribution (i.e., the objective range of 0 to 1.2 is mapped to approximately 0.2 to 1.1). Second, the judgments depend on the recent history of stimuli, known as sequential effect. Estimates after a large previous stimulus tend to be larger, while estimates after a small previous stimulus tend to be smaller, leading to a perceptual hysteresis curve that is clearly shown in Figure 9. Note that for Levels 1 and 7 only one direction exists. For Levels 2 to 5, upward jumps were estimated to be significantly higher than downward jumps ($t \geq 2.40, p \leq 0.011$), based on pairwise one-tailed Welch's t -tests. For Level 6, however, the upward jump was significantly lower than the downward jump: $t(29.24) = -1.93, p = 0.032$. This may be explained by the extreme reverberant tail of Level 7 that was masking the small transition to Level 6, leading to the result that participants simply did not notice this change.

The instantaneous delay that was discussed above was further analyzed in a similar way as the estimated reverberation, by computing the average instantaneous delay per time segment and participant. When these data were pooled over all participants, however, no significant effects could be observed: There was no significant difference

in delay, neither between different reverberation levels, nor between jump distances, nor between jump directions.

Finally, Figure 10 shows the numerical results of the questionnaires. Participants in EG indicated that their listening experience seemed more virtual compared to the CG participants. Mapping of the original Likert scale [1,7] to a range of [-1,1] produced mean values for CG -0.1 (standard deviation 0.9) and for EG -0.6 (standard deviation 0.4). The EG participants did not feel excessively challenged by the need to listen to the radio while estimating the reverberation (mean lies in the center, at 0.1 with standard deviation 0.3); however, they found it more challenging than only listening to the radio (as in CG). Furthermore, they were rather confident of their estimate (mean 0.4 with standard deviation 0.3).

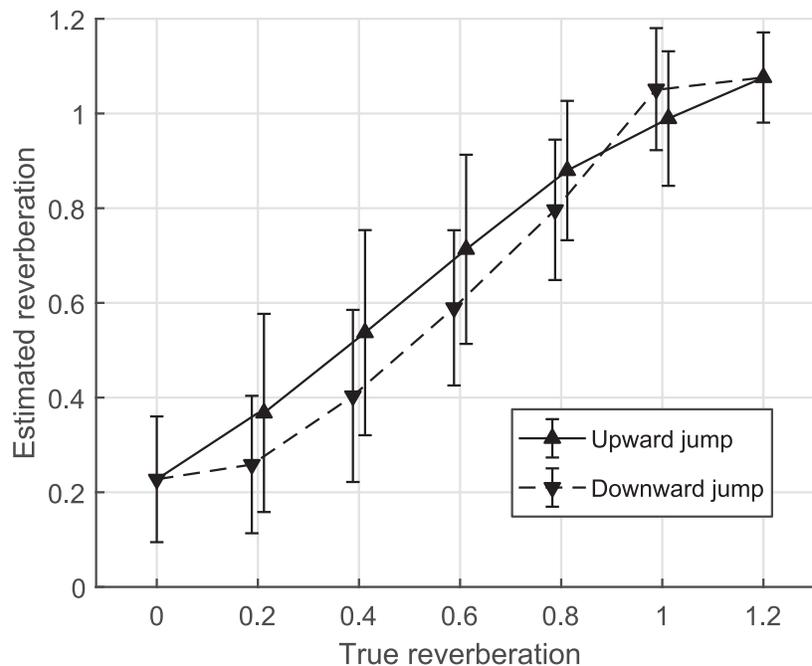
Discussion of the RadioReverb Experiment

As we expected, we found that listeners cannot discern as many levels of reverberation in a magnitude-estimation experiment as in a JND experiment with pairwise comparisons. The results of the RadioReverb experiment showed that we may reliably

Figure 9. Estimated versus real reverberation, pooled over all participants, split into upward and downward level changes. For instance, with an

upward jump the number sums all changes of reverberation level coming from any lower level. Error bars indicate the standard deviation. The results

show a clear bias of listeners depending on the sequence of the presented change.



convey three levels (about 1.5 bits) of information by virtual reverberation in real time at the periphery of attention, while the users are focused on another task. This number corresponds to probability of superiority of 0.975, that is, 97.5 percent correct answers. We may increase the number of bits when the probability of superiority is lowered, for instance to 0.8, resulting in 5.8 levels or about 2.5 bits. The choice of the necessary probability of superiority should be based on the criticality of the conveyed information. Another measure for adapting a peripheral sonification based on reverberation is to take into account the over- and underestimation of upward jumps and downward jumps, respectively, that is, the sequential bias found in magnitude estimation, shown in Figure 9. Our data show estimates that exaggerate the difference between consecutive reverberation settings. Changes from high to low reverberation values (downward jumps) were underestimated as compared to jumps from low to high (upward jumps). The latter showed a more linear behavior.

Our analysis of the temporal correlation between reference reverberation levels and response

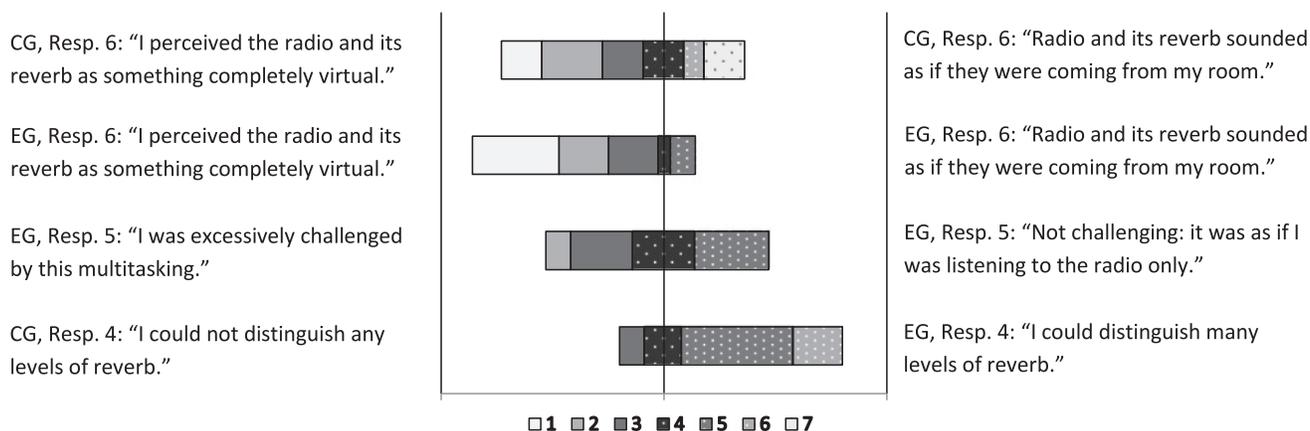
showed that this sonification works well in real time. The average instantaneous delay of 1 sec includes the estimation of the participant and the setting of the slider; thus we may assume a real-time-capable display within physiological reaction times.

As was evident both in the qualitative responses in the questionnaire and in the time-series response data, we found that some listeners concentrated more on content and others more on sound (CF versus SF listeners in Figure 6). We preferred to take only the CF listeners into account for the results of this experiment, as they fulfilled our criterion for peripheral listening to the reverberation. The RadioReverb experiment was designed as an ecologically valid experiment; however, true peripheral listening without any experimental task—in our case, setting a slider—might lead to different results. Nevertheless, the fact that we were able to distinguish CF listeners from SF listeners on the basis of both the questionnaire and the data from the experiment gives evidence that our goal of designing a peripheral listening experiment was to some extent achieved.

Figure 10. Results of questions on the participant's impressions. The normalized number of participants giving the same answer is shown as length of the segment 1–7: In their responses to

Question 6 (see the Evaluation in the Post Hoc Online Survey section), both CG and EG participants indicated how they experienced the “virtuality” of the sound. In Question 5, the EG

participants assessed how challenging they found the multitasking. In Question 4, the EG participants indicated how confident they were of their ability to distinguish levels of reverberation.



Participants did not feel excessively challenged by the task in the experiment, and they felt confident of their estimates. Even if no questions addressed these subjective reactions explicitly, the responses indicate that the sound was accepted and was not too disturbing. These issues should be addressed explicitly in a long-term, in situ, follow-up experiment.

The EG participants experienced the radio and its reverberation as being very unnatural. The responses of the CG participants were less negative; this is perhaps to be expected as they were exposed to a constant reverberation. From this, we may conclude that in Question 6 in the above description of the post hoc online survey, the participants were assessing the reverberation more than the radio. Nevertheless, these responses indicate that the situation was experienced as being less natural than we had expected, and this reduces the ecological validity of the experiment. From personal observation and informal comments of some participants, we may say that the radio blended well into the customary soundscape. The abstract formulation of the question might explain why the questionnaire responses were so negative in this regard.

The binaural test design of the experiment was due to the fact that the experiment was conducted during the COVID-19 pandemic. Initially, the experiment was planned and already set up in a studio using Ambisonics technology and loudspeakers. Comparing the listening experience

of the lab environment with the binaural room simulation, we may say that these were comparable and therefore would probably have produced similar results. Although the lab version would have been more controlled in its test procedure, the binaural design was more ecologically valid due to the at-home setting.

Conclusions

We conducted two experiments to explore peripheral sonification by altering a room's reverberation with virtual acoustics. First, in the PilotKitchen experiment we installed a prototype in the real surrounding of our institute's kitchen. Real-world sounds were recorded by a microphone and played back over loudspeakers to add virtual reverberation, depending on the actual electrical power consumption of the kitchen. The evaluation of this experiment was carried out in the form of a diary study and led to ambiguous results, revealing a heterogeneous attitude towards the system. As a proof of concept, this experiment permitted us to draw conclusions concerning the implementation of such a system. The second experiment, RadioReverb, was conducted with the total of 29 participants in their homes using their smartphones to audition binaural renderings of a virtual radio and room response. The experiment task was to first listen to the radio feature and second to track changes of

reverberation and estimate them on a GUI slider. As discussed above, we may state that within a plausible range of virtual room response, three levels of reverberation can be distinguished reliably. More levels can be distinguished if the criticality of the sonified information is lower and a larger number of incorrect estimates can be tolerated. The tuning of the sonification has been described in detail in the supplementary material to this article and can be used for creating peripheral sonifications with the tools we used or with other tools.

Overall, we conclude that the method works well as a peripheral auditory display. The sound design is rather unobtrusive, whereas adding sounds to an environment that is already noisy, especially in an at-home display, can often be difficult. Augmenting existing sounds produces better results. Furthermore, learnability, which may be low in abstract sonifications, can be assumed to be higher, since tracking room characteristics such as reverberation are an evolutionary feature of human hearing. Although our experiments did not make use of musical sounds, we believe that our results could be extended to experiments that use reverberation as an informational layer in background music or music performances.

Our outlook for future research includes a plan to install a system of peripheral sonification long-term and in situ in a semipublic or at-home environment to test all evaluation criteria, such as those given by Matthews, Hsieh, and Mankoff (2009). Especially the criteria appeal and acceptance, which could not be measured in our studies, are cornerstones for the successful application of any new technology and should be explored in future studies. Furthermore, on a conceptual level, our finding that the EG participants could be divided in SF and CF listeners, both from qualitative and quantitative data, might be linked to typologies of listening, for example, starting with Schafer (1993) and ubiquitous music participation (cf. Keller, Schiavoni, and Lazzarini 2019).

Acknowledgments

We would like to thank our test participants, mainly our colleagues and students, who were first exposed

to our virtual reverberation in the institute's kitchen, and then installed the experiment environment on their private smartphones. Finally, many thanks to Henry Fullenwider and to the editors for their thorough proofreading of the manuscript.

References

- Bakker, S., and K. Niemantsverdriet. 2016. "The Interaction-Attention Continuum: Considering Various Levels of Human Attention in Interaction Design." *International Journal of Design* 10(2):1-14.
- Barrass, S., and T. Barrass. 2013. "Embedding Sonifications in Things." In *Proceedings of the International Conference on Auditory Display*, pp. 149-152.
- Blochberger, M., F. Zotter, and M. Frank. 2019. "Sweet Area Size for the Envelopment of a Recursive and a Non-Recursive Diffuseness Rendering Approach." In *International Conference on Spatial Audio*, pp. 151-157.
- Bovermann, T., R. Tünnermann, and T. Hermann. 2012. "Auditory Augmentation." In *Innovative Applications of Ambient Intelligence: Advances in Smart Systems*. Hershey, Pennsylvania: IGI, pp. 98-112.
- Bradley, J. S., R. Reich, and S. Norcross. 1999. "A Just Noticeable Difference in C50 for Speech." *Applied Acoustics* 58(2):99-108.
- Brown, J. N. A. 2016. "'Unseen, Yet Crescive': The Unrecognized History of Peripheral Interaction." In S. Bakker, D. Hausen, and T. Selker, eds. *Peripheral Interaction: Challenges and Opportunities for HCI in the Periphery of Attention*. Berlin: Springer, pp. 31-38.
- Butz, A., and R. Jung. 2005. "Seamless User Notification in Ambient Soundscapes." In *Proceedings of the International Conference on Intelligent User Interfaces*, pp. 320-322.
- Connell, L., and M. T. Keane. 2006. "A Model of Plausibility." *Cognitive Science* 30(1):95-120.
- Cox, T. J., W. J. Davies, and Y. W. Lam. 1993. "The Sensitivity of Listeners to Early Sound Field Changes in Auditoria." *Acta Acustica united with Acustica* 79(1):27-41.
- Farias, F. M., et al. 2015. "Bringing Aesthetic Interaction into Creativity-Centered Design: The Second Generation of mixDroid Prototypes." *Journal of Cases on Information Technology* 17(4):53-72.
- Ferguson, S. 2013. "Sonifying Every Day: Activating Everyday Interactions for Ambient Sonification Systems." In *Proceedings of the International Conference on Auditory Display*, pp. 77-84.

- Groß-Vogt, K., et al. 2018. "Augmentation of an Institute's Kitchen: An Ambient Auditory Display of Electric Power Consumption." In *Proceedings of the International Conference on Auditory Display*, pp. 105–112.
- Hammerschmidt, J., R. Tünnermann, and T. Hermann. 2014. "EcoSonic: Towards an Auditory Display Supporting a Fuel-Efficient Driving Style." In *Proceedings of the Conference on Sonification of Health and Environmental Data*, pp. 979–982.
- Hazlewood, W. R., E. Stolterman, and K. Connelly. 2011. "Issues in Evaluating Ambient Displays in the Wild: Two Case Studies." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 877–886.
- Iglesia, D. 2016. "The Mobility is the Message: The Development and Uses of MobMuPlat." In *Proceedings of the Pure Data Conference*, pp. 56–61.
- Keller, D. 2018. "Challenges for a Second Decade of Ubimus Research: Knowledge Transfer in Ubimus Activities." *Revista Música Hodie* 18(1):148–165.
- Keller, D., F. Schiavoni, and V. Lazzarini. 2019. "Ubiquitous Music: Perspectives and Challenges (Editorial)." *Journal of New Music Research* 48(4):309–315.
- Keller, D., et al. 2010. "Anchoring in Ubiquitous Musical Activities." In *Proceedings of the International Computer Music Conference*, pp. 319–326.
- Kilander, F., and P. Lönnqvist. 2002. "A Whisper in the Woods: An Ambient Soundscape for Peripheral Awareness of Remote Processes." In *Proceedings of the International Conference on Auditory Display*. Available online at hdl.handle.net/1853/51336. Accessed October 2020.
- Larsen, E., et al. 2008. "On the Minimum Audible Difference in Direct-to-Reverberant Energy Ratio." *Journal of the Acoustical Society of America* 124(1):450–461.
- Lockton, D., et al. 2014. "Bird-Watching: Exploring Sonification of Home Electricity Use with Birdsong." In *Conference on Sonification of Health and Environmental Data*. Available online at www.york.ac.uk/media/c2d2/media/sonihedconference/Lockton_etal_SoniHED_2014.pdf. Accessed October 2020.
- Matthews, T., G. Hsieh, and J. Mankoff. 2009. "Evaluating Peripheral Displays." In *Awareness Systems*. Berlin: Springer, pp. 447–472.
- Miller, G. A. 1956. "The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information." *Psychological Review* 63(2):81–97.
- Niaounakis, T. I., and W. J. Davies. 2002. "Perception of Reverberation Time in Small Listening Rooms." *Journal of the Audio Engineering Society* 50(5):343–350.
- Parthy, A., C. Jin, and A. van Schaik. 2004. "Reverberation for Ambient Data Communication." *Proceedings of the International Conference on Auditory Display*. Available online at hdl.handle.net/1853/50906. Accessed October 2020.
- Petzschner, F. H., S. Glasauer, and K. E. Stephan. 2015. "A Bayesian Perspective on Magnitude Estimation." *Trends in Cognitive Sciences* 19(5):285–293.
- Pousman, Z., and J. Stasko. 2006. "A Taxonomy of Ambient Information Systems: Four Patterns of Design." In *Proceedings of the Working Conference on Advanced Visual Interfaces*, pp. 67–74.
- Pulkki, V., and M. Karjalainen. 2015. *Communication Acoustics: An Introduction to Speech, Audio and Psychoacoustics*. Hoboken, New Jersey: Wiley.
- Schafer, R. M. 1993. *The Soundscape: Our Sonic Environment and the Tuning of the World*. New York: Simon and Schuster.
- Schörkhuber, C., M. Zaunschirm, and R. Höldrich. 2018. "Binaural Rendering of Ambisonic Signals via Magnitude Least Squares." In *Fortschritte der Akustik: Deutsche Jahrestagung für Akustik*, pp. 339–342.
- Soares, A. P., et al. 2013. "Affective Auditory Stimuli: Adaptation of the International Affective Digitized Sounds (IADS-2) for European Portuguese." *Behavior Research Methods* 45(4):1168–1181.
- Stautner, J., and M. Puckette. 1982. "Designing Multi-Channel Reverberators." *Computer Music Journal* 6(1):52–65.
- Stevens, F., D. Murphy, and S. Smith. 2016. "Emotion and Soundscape Preference Rating: Using Semantic Differential Pairs and the Self-Assessment Manikin." In *Proceedings of the Sound and Music Computing Conference*, pp. 455–462.
- Tünnermann, R., J. Hammerschmidt, and T. Hermann. 2013. "Blended Sonification: Sonification for Casual Information Interaction." In *Proceedings of the International Conference on Auditory Display*, pp. 119–126.
- Verbeek, P.-P. 2009. "Ambient Intelligence and Persuasive Technology: The Blurring Boundaries between Human and Technology." *Nanoethics* 3(3):231–242.
- Weger, M., T. Hermann, and R. Höldrich. 2018. "Plausible Auditory Augmentation of Physical Interaction." In *Proceedings of the International Conference on Auditory Display*, pp. 97–104.

-
- Weiser, M., and J. S. Brown. 1996. "Designing Calm Technology." *PowerGrid Journal* 1(1):75–85.
- Yassein, M. B., W. Mardini, and A. Khalil. 2016. "Smart Homes Automation Using Z-Wave Protocol." In *Proceedings of the International Conference on Engineering and MIS*, pp. 93–98.
- Zaunschirm, M., C. Schörkhuber, and R. Höldrich. 2018. "Binaural Rendering of Ambisonic Signals by Head-Related Impulse Response Time Alignment and a Diffuseness Constraint." *Journal of the Acoustical Society of America* 143(6):3616–3627.
- Zotter, F., and M. Frank. 2019. *Ambisonics: A Practical 3-D Audio Theory for Recording, Studio Production, Sound Reinforcement, and Virtual Reality*. Berlin: Springer.