

The Virtual Reality Lab: Realization and Application of Virtual Sound Environments

Volker Hohmann,^{1,2,3} Richard Paluch,^{3,4} Melanie Krueger,^{2,3} Markus Meis,^{3,5} and Giso Grimm^{1,2,3}

To assess perception with and performance of modern and future hearing devices with advanced adaptive signal processing capabilities, novel evaluation methods are required that go beyond already established methods. These novel methods will simulate to a certain extent the complexity and variability of acoustic conditions and acoustic communication styles in real life. This article discusses the current state and the perspectives of virtual reality technology use in the lab for designing complex audiovisual communication environments for hearing assessment and hearing device design and evaluation. In an effort to increase the ecological validity of lab experiments, that is, to increase the degree to which lab data reflect real-life hearing-related function, and to support the development of improved hearing-related procedures and interventions, this virtual reality lab marks a transition from conventional (audio-only) lab experiments to the field. The first part of the article introduces and discusses the notion of the *communication loop* as a theoretical basis for understanding the factors that are relevant for acoustic communication in real life. From this, requirements are derived that allow an assessment of the extent to which a virtual reality lab reflects these factors, and which may be used as a proxy for ecological validity. The most important factor of real-life communication identified is a closed communication loop among the actively behaving participants. The second part of the article gives an overview of the current developments towards a virtual reality lab at Oldenburg University that aims at interactive and reproducible testing of subjects with and without hearing devices in challenging communication conditions. The extent to which the virtual reality lab in its current state meets the requirements defined in the first part is discussed, along with its limitations and potential further developments. Finally, data are presented from a qualitative study that compared subject behavior and performance in two audiovisual environments presented in the virtual reality lab—a street and a cafeteria—with the corresponding field environments. The results show similarities and differences in subject behavior and performance between the lab and the field, indicating that the virtual reality lab in its current state marks a step towards more ecological validity in lab-based hearing and hearing device research, but requires further development towards higher levels of ecological validity.

Key words: Audiology, Audiovisual environments, Ecological validity, Hearing acoustics, Hearing aids, Hearing loss, Virtual reality.

Abbreviations: ADM, Adaptive directional microphone; dB SPL, Sound pressure level in decibel; EEG, Electroencephalography; EMA, Ecological momentary assessment; EOG, Electrooculography; EXPU, Experienced hearing aid users; FTU, First-time hearing aid users; HA, Hearing aid; MD, Median value; NH, Normal-hearing listeners; TASCAR, Toolbox for Acoustic Scene Creation and Rendering

(Ear & Hearing 2020;41:31S–38S)

¹Department of Medical Physics and Acoustics, University of Oldenburg, Oldenburg, Germany; ²HörTech gGmbH, Oldenburg, Germany; ³Cluster of Excellence “Hearing4all,” Oldenburg, Germany; ⁴Department of Social Sciences, University of Oldenburg, Oldenburg, Germany; and ⁵Hörzentrum Oldenburg GmbH, Oldenburg, Germany.

Copyright © 2020 The Authors. Ear & Hearing is published on behalf of the American Auditory Society, by Wolters Kluwer Health, Inc. This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

INTRODUCTION

Current hearing devices are technologically advanced and support acoustic communication in adverse conditions characterized by noise and reverberation. Noise reduction and speech enhancement algorithms show significant benefit in terms of speech reception in the lab (e.g., Völker et al. 2015). However, they do not allow a full rehabilitation of hearing impairment in the individual patient. Listeners with hearing loss still experience great difficulty with listening in noise, especially in complex listening conditions and with cochlear implants (e.g., Lesica 2018). One reason may be the discrepancy between established lab-based evaluation methods, which mainly employ static listening conditions, limited acoustic complexity, and unidirectional communication, and the dynamically evolving interactive acoustic communication in real life. In fact, even for standard directional microphones, which should work robustly in all acoustic environments and give a small but significant benefit in signal to noise ratio, a discrepancy between lab results and real-life results has been reported (Cord et al. 2004; Bentler 2005; Wu et al. 2019). The discrepancy is expected to be even higher for more complex algorithms, for example, for novel visually guided beamformers (Grimm et al. 2018; Jennings & Kidd 2018), which require dynamic evaluation conditions, or algorithms based on machine learning (e.g., Pandey & Wang 2019), which require real-life conditions to learn and adapt their functioning. To overcome these limitations of established lab-based procedures, two general approaches are currently being pursued. One approach is to use virtual reality technology to increase the level of realism in the lab (e.g., Oreinos & Buchholz 2014; Grimm et al. 2016; Pausch et al. 2018; Ahrens et al. 2019). This virtual reality lab has the advantage that environments and experiments are scalable and reproducible, which is relevant for the design and evaluation of novel devices and algorithms as well as for characterizing similarities and dissimilarities in behavior across different subject groups. The disadvantage is that subjects in the virtual reality lab may not feel fully involved in the artificial scene, and thus findings may not be indicative of real-life subject behavior and performance. The extent to which the virtual reality lab in its current developmental stage is ecologically valid, that is, reflects real-life hearing-related function, is discussed in this article. The second approach is to gather meaningful data from the field, for example, by ecological momentary assessment (EMA; Shiffman et al. 2008). This approach has the advantage that hearing interventions can be tested directly in the real-life environment of the subject. The disadvantage is that it is difficult to control, to reproduce, and to scale, in terms of acoustic and cognitive complexity.

The ecological validity of the two approaches is currently being discussed in the research community (e.g., Campos & Launer, 2020, this issue, pp. 99S-106S; Holube et al., 2020,

this issue, pp. 79S-90S; Smeds et al., 2020, this issue, pp. 20S-30S). Although we work on the virtual reality lab and report on it here, we do not consider the virtual reality lab to be the only way to achieve meaningful data. Rather, we assume that experiments in the virtual reality lab may inform field studies, and vice versa, in gathering complementary data in pursuit of purposes A (Understanding) and B (Development) for striving for higher levels of ecological validity in hearing research; see Keidser et al. (2020, this issue, pp. 5S-19S). Purpose A refers to the need for better understanding the role of hearing in everyday situations, while purpose B refers to the need for better evaluation protocols for supporting the development of new procedures and interventions. To theoretically underpin these efforts, we outline the communication loop model in the next section, which in our opinion may serve the purpose of identifying factors of ecological validity and corresponding metrics.

REQUIREMENTS FOR REFLECTING REAL-LIFE COMMUNICATION IN THE LAB: THE COMMUNICATION LOOP

Human communication in real life cannot be described based on a simple signal sender and receiver model of information transmission. In fact, communication evolves in an interactive loop, which links the sound field (representing all sounds in the environment, including conversation partners), the hearing device (if present) and the actively behaving subject and conversation partners (including their cognitive and perceptual processes) in a dynamic and complex way. Figure 1 shows a sketch of this *communication loop*, the notion of which is established as a basis for reasoning in cognitive sciences and neurosciences (e.g., Wilms et al. 2010), and has also been used to model speech communication (e.g., Moulin-Frier et al. 2012). Until now, it has not been applied to assess ecological validity in lab-based hearing device evaluation and hearing assessment.

Given the vast variability of the different building blocks of the loop in real-life environments, it appears questionable whether the simple loop model sketched in Figure 1 can enable us to derive relevant experimental factors in hearing research that can be used as a proxy for ecological validity. Nevertheless, we argue that the simple loop model is sufficient to at least identify several major factors that influence acoustic communication and that are common to many relevant real-life scenarios. The major assumption underlying our approach is that the degree of ecological validity of an experiment, be it in the lab or in the field, depends on the level to which the subject feels embedded (or involved) in the communication loop in a way that matches the tested communication scenario. Furthermore, we assume that controllable experimental factors and corresponding quantitative metrics can be identified, which determine the level of involvement, and thus can be used as a proxy of ecological validity.

One example is that the distraction of the subject caused by the interaction with a smartphone during EMA binds cognitive resources and takes the subject partially out of the communication loop. This in turn will limit the ecological validity of EMA procedures in the field, according to the model proposed here. To quantify the level of involvement, and thus ecological validity, however, experimental factors and corresponding metrics have to be identified that determine the level of involvement. Finding these factors and metrics from EMA experiments in

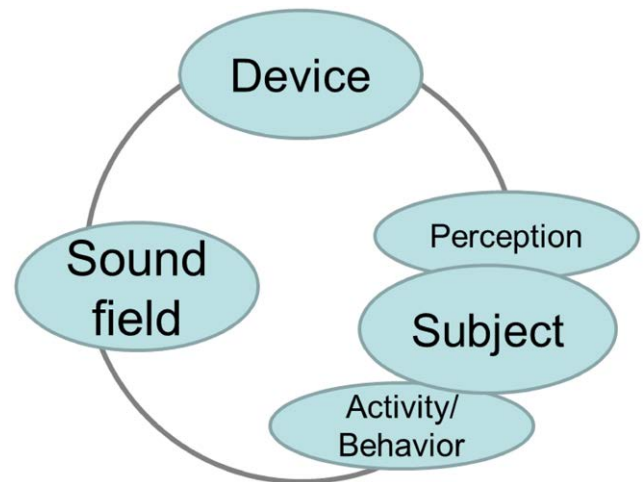


Fig. 1. Communication loop representing the interaction between sound field, device (if present), and actively behaving subject.

the field may prove difficult. In the virtual reality lab, however, conditions are scalable and reproducible, which may help in designing experiments that find relations between experimental factors, suitable metrics, and the level of involvement in a communication loop. As participants in lab studies may feel somewhat insecure about the situation, and thus behave partially unnaturally, field data are necessary to complement lab data. Knowledge of the relations between experimental factors and corresponding metrics from lab studies may inform experiments in the field so they can reduce the level of distraction in EMA procedures by designing and testing appropriate interaction techniques that show less reduction in involvement. This combined approach may help to exploit the respective strengths of tests in the lab and in the field and to gain complementary and significant data.

The proposed approach can be illustrated by the example of head movements that typically occur in acoustic communication conditions. Hendrikse et al. (2020) found in the virtual reality lab that individual head-movement tracks recorded in turn-taking conversations (Hendrikse et al. 2019) differentially affected the benefit in signal to noise ratio provided by a standard adaptive directional microphone and that head movement significantly reduced signal to noise ratio benefit of the adaptive directional microphone on average. This demonstrates the relevance of realistic head movement for ensuring the ecological validity of hearing aid outcome measures (benefit in signal to noise ratio, in this case). It is therefore important to optimize the experimental paradigms in the lab (and also in the field) in a way that the head-movement behavior is as realistic as possible to increase the level of ecological validity of hearing aid outcome measures. To be able to control the realism of head movement, the experimental factors influencing the movement need to be identified. In that regard, Hendrikse et al. (2018) found that subjects moved their heads and eyes much less in turn-taking conversations if no visual stimulation was provided. Furthermore, Hendrikse et al. (2018) showed that subjects changed their movement behavior dependent on the head and gaze behavior of animated characters used as simulated conversation partners. This shows that measures of head and eye movements may be used as a metric to assess experimental paradigms and their level of ecological validity*.

Whereas head and eye movements have been investigated in some detail in recent years in the context of ecological validity to assess the influence of visual and acoustic factors (see Grimm et al., 2020, this issue, pp. 48S–55S), other relevant experimental factors have not been investigated in such detail, and the corresponding quantitative metrics are yet largely unknown. Table 1 shows a (noncomplete) list of experimental factors that may reduce the level of involvement in the communication loop (i.e., reduce ecological validity) and the (potential) corresponding metrics that may be used as a proxy to quantify the level of involvement.

Whereas the quality of acoustic and visual stimulation (items 1 and 2) can be controlled parametrically using current virtual reality technology, and, as mentioned earlier, several studies relating the factors of acoustic and visual stimulation with the corresponding metrics have been published (e.g., Hendrikse et al. 2018, 2019, 2020), items 3 and 4 are largely under-investigated, to the knowledge of the authors. On the basis of the notion of the relevance of bodily acts in communication (Mead 1956, p. 98), we may assume that head movements, as well as other gestures, such as leaning forward or backward, or arm and hand movements, are important observables/metrics for estimating the level of involvement, and thus the level of ecological validity, also for items 3 and 4. Furthermore, Wilms et al. (2010) reported, citing Frith (2007, p. 175), that “gaze is also known to ‘connect’ human beings in everyday life situations by means of a ‘communication loop’ in which interactors impact reciprocally on each other’s behavior.” Gaze can therefore be assumed to be another observable/metric for involvement in the communication loop. In fact, recent studies pave the way toward more interactive closed-loop paradigms. Wilms et al. (2010) emphasized the need for new experimental paradigms that “allow studying social encounters in a truly interactive manner by establishing ‘online’ reciprocity in social interaction” and developed an animated virtual character “whose behavior becomes ‘responsive’ to being looked at allowing the participant to engage in ‘online’

interaction with this virtual other in real-time.” According to Straub’s human–robot interaction study (Straub 2016), subjects were more communicatively involved when the counterpart greeted, used ice breakers, responded to questions, or engaged in small talk. These approaches from cognitive and social sciences may support the development of new interactive closed-loop approaches for hearing (aid) research. Another recent study by Hadley et al. (2019) investigated speech, movement, and gaze behaviors during dyadic conversation in noise, where the interlocutors were both subjects, and their behavior was a dependent variable. Extension of these approaches to using virtual responsive animated characters, as suggested by Wilms et al. (2010) and Straub (2016), may lead to scalable and reproducible closed-loop paradigms with high levels of involvement, and thus ecological validity. For this, models of behavior need to be developed that can be used to make the animations sufficiently realistic to promote realistic subject behavior. Studies of self-motion and interaction in communication conditions, as discussed by Carlile and Keidser (2020, this issue, pp. 56S–67S) and Grimm et al. (2020, this issue, pp. 48S–55S), may help in developing such models.

TECHNOLOGY OF THE VIRTUAL REALITY LAB

To implement the communication loop in the virtual reality lab, interactive low-delay multimedia rendering techniques with integrated sensors for closing the loop are required. This section gives a brief summary of the current state of development of the virtual reality lab at Oldenburg University. For a more comprehensive overview of the technology, see Llorach et al. (2018), and for a detailed description of the setup, see Hendrikse et al. (2018, 2019).

Whereas game-engine-based real-time technology for interactively rendering visual content on screens or head-mounted displays is readily available, for example, the open-source

TABLE 1. Factors that may reduce the level of involvement in the communication loop (i.e., reduce ecological validity), examples of these factors, and corresponding metrics that may be used as a proxy to quantify the level of involvement (with references, if available)

	Factor	Example	Metric	References
1	Visual stimuli differ from those experienced in the real-life condition under test	Low video reproduction quality; use of animated characters with low quality of appearance	Similarity of head and eye movement behavior to real-life behavior; Questionnaires/subjective scaling	Hendrikse et al. (2018)
2	Acoustic stimuli differ from those experienced in the real-life condition under test	Low sound field reproduction quality; static (nonmoving) sound sources	Quantitative measures of the function of a (simulated) hearing device in comparison to its functioning in the real-life condition under test Questionnaires/subjective scaling Head- and eye-movement behavior	Grimm et al. (2016), Hendrikse et al. (2020) Hendrikse et al. (2018) To be investigated
3	Unrealistic behavior of conversation partners in the (simulated) experiment	Animated characters do not move heads and eyes or do not show facial expressions	Head- and eye-movement behavior	To be investigated
4	Lack of interactivity of the experimental paradigm	The paradigm does not promote a closed-loop communication similar to the real-life condition under test	Subjective scaling of the involvement; gaze behavior or facial expression	To be investigated

*Note that head and eye movements were used both as a metric and as experimental factors by Hendrikse et al. (2018): head and eye movements of communication partners in the (simulated) experiment were used as an experimental factor that was found to influence the metric, that is, the head and eye movement of the subjects in the experiment.

†The sound events and levels experienced by each subject during the walks were not controlled. This means that the sound levels in the field and in the lab were the same on average, but not for each subject individually. Individual dosimetry would have been required to make individual adjustments.

‡Because HA amplification was the same in both conditions and no noise reduction was employed, this finding suggests that the background noise level may have been somewhat higher in the field than in the lab and may have partially masked the sounds from the bicycles.

software tool blender (Roosendaal & Wartmann 2003; <https://www.blender.org/>), real-time, low-delay, high-quality interactive audio rendering software had to be developed for the purpose of implementing the virtual reality lab for hearing research. The open-source TASCAR software toolbox (Grimm et al. 2019) was developed for this purpose. It allows the researcher to design dynamic acoustic scenes and render them using different techniques, such as 3D higher-order Ambisonics (Daniel, Reference Note 1; Heller & Benjamin 2014), to headphones or loudspeaker arrays. The simulation method focuses on a time-domain simulation of the direct path and a geometric image source model, which simulates air absorption and the Doppler effect. To establish the feasibility of the approach, the interaction between reproduction method and technical and perceptual hearing aid performance measures was investigated using computer simulations for regular circular loudspeaker arrays with 4 to 72 channels (Grimm et al. 2015). The results confirm that the physical sound field accuracy can be rendered sufficiently well even for 4-microphone binaural beamforming to work properly, if adequate rendering techniques are applied and the number of loudspeakers is high enough (depending on the targeted frequency range).

On the basis of the developed audiovisual technology, a set of five different audiovisual environments was designed: a cafeteria, a lecture hall, a train station, a street with car traffic, and a living room (Hendrikse et al. 2019; see <https://www.youtube.com/channel/UCAXZPzxbOJM9CM0IBfgvoNg> for a video demonstration). Environments were selected to represent relevant conditions for daily life, according to Wolters et al. (2016). Following a gamification approach, animated characters can be placed in the scene together with other sound sources, for example, trains or cars, and story lines can be implemented to embed the subject in the scene. Depending on the experiment, speech recognition tasks, detection tasks, divided attention tasks, etc. can be implemented using the TASCAR/blender combination. The animated characters show a coarse lip movement driven by the input speech signal in low-delay real time (see Llorach et al. 2016 for technical details). Furthermore, animated characters automatically turn their head towards the currently active speaker in turn-taking conversations (see Hendrikse et al. 2018, 2019), further increasing the perceived naturalness of the character's behavior.

For closing the loop, several sensors have been integrated, in particular, optical motion tracking, gaze tracking (optical and electrooculography), electroencephalography (cEEGrid; Bleichner & Debener 2017), and video monitoring.

Figure 2 shows a picture of the virtual reality lab with a subject sitting in its center. Because of its interactive low-delay rendering of sound and video, as well as the in-the-loop sensor technology, the lab in its current state enables in principle the investigation of different factors affecting ecological validity and the corresponding metrics outlined in the previous section. The next section presents an experiment that investigated several of these factors and metrics in a comparative field versus lab study.

COMPARATIVE STUDY OF SUBJECT BEHAVIOR AND PERFORMANCE IN THE VIRTUAL REALITY LAB AND IN REAL LIFE

This section presents an overview of earlier studies comparing subject behavior in the field and in the audiovisual virtual



Fig. 2. Virtual reality lab at Oldenburg University (photo: Hörtech gGmbH). Projectors (yellow) project a 300-degree field-of-view image on a cylindrical screen. Sound is rendered over a spherical multi-loudspeaker array with 29 full-range loudspeakers (blue) and 4 subwoofers. Sensors include EOG sensors for eye gaze, EEG sensors, and head tracking (red). Furthermore, the subject can be video recorded during an experiment (green). EEG indicates electroencephalography; EOG, electrooculography.

reality lab (Paluch et al. 2017, 2018a, b, 2019; Paluch 2019) and discusses the findings in relation to some of the factors of ecological validity outlined earlier. Two environments relevant for daily life (Wolters et al. 2016) were investigated: road traffic and cafeteria. In a mixed-methods approach, qualitative evaluation tools were used for the field and quantitative evaluation tools were used for the virtual reality lab: behavior analyses using field notes for developing hypotheses in the field, and video annotation for testing these hypotheses in the virtual reality lab. Furthermore, questionnaire data from surveys addressing loudness perception, annoyance, and localization were collected in the field and in the virtual reality lab and compared with each other. Participants included age-matched people with normal and impaired hearing. Seven experienced HA users (EXPU), seven first-time HA users (FTU), and seven age-matched normal-hearing (NH) subjects participated. Two different presentation modes were tested in the FTU group, unaided and aided in omnidirectional microphone mode.

Figure 3 shows the virtual environments used in the experiment. They simulate the environments tested in the field and were presented in the virtual reality lab described above with the subject standing (street) or sitting (cafeteria) in the center of the loudspeaker ring. In this experiment, the field of view was restricted to 120°, as only one projector was used. In the street scene, a busy environment was simulated with several pedestrians, bicycles, cars, trucks, buses, and an ambulance passing by, while the subject was moved slowly on the sidewalk (virtual walk). The subject was asked to stop at a virtual bus stop and was asked closed-ended (categorical) questions by a female animated character. The experimenter noted the subject's answers. In the second part of the street scene, the subject was moved to a quiet environment with typical background sounds such as birds singing and cyclists and cars passing by from time to time. The loudness of the street scene varied with changes in the activity of the surrounding sound sources. The animated character also appeared in this quiet street environment at a bus stop to ask questions about the subject's experience in the virtual environment. The virtual street walk took about 22 min.



Fig. 3. Virtual reality lab animations: Noisy street (upper panel) and cafeteria situation (lower panel, with a subject attending the scene). Photos: CC BY-SA 3.0 Giso Grimm.

In the simulated cafeteria situation, the subject sat at a simulated table together with four animated characters (two female, two male, at azimuthal directions of 45° , 15° , -15° , and -45° , facing towards the subject; Hendrikse et al. 2018). Tables, chairs, lamps, etc., were shown, and cafeteria noise sounded in the background. The animated characters began to talk to each other, showing coarse synchronized lip movement (Llorach et al. 2016), with turn-taking initiated by head movements of the animated characters to the character that started talking (Hendrikse et al. 2018, 2019). The subject listened to two different casual conversations between the four animated characters, each lasting for approximately 90 sec. Afterwards, one of the four animated characters asked the subject a set of closed-ended questions to assess the experience of the animated situations. The cafeteria scene took about 5 min.

The sound levels from the real street (noisy street $L_{Aeq,60sec} = 69$ dB SPL A-weighted; quiet street $L_{Aeq,60sec} = 65$ dB SPL A-weighted) and the real cafeteria situation ($L_{Aeq,60sec} = 72$ dB SPL A-weighted) were measured, and levels in the lab were set accordingly¹.

Via closed-ended questionnaires, the subjects rated psychoacoustic characteristics. Three items were used in the street: (1) loudness of sound sources, (2) annoyance caused by sound sources, and (3) localization of sound sources. In the cafeteria, three items were used: (1) loudness of sound sources, (2) annoyance caused by background noise, and (3) localization of background noise. In the noisy and quiet streets, the loudness of trucks/buses, cars, bicycles, and speech were evaluated separately, as were the annoyance and localization of trucks/buses, cars, and bicycles. In the cafeteria, test subjects could evaluate the loudness of speech and background noise. For evaluation,

the questionnaire had a Likert-type scale ranging from 1 to 5. Loudness was rated from 1 = very soft, 2 = soft, 3 = moderate, 4 = loud, 5 = extremely loud; annoyance from 1 = not annoyed, 2 = slightly annoyed, 3 = moderate, 4 = severely annoyed, 5 = extremely annoyed (ISO/TS 15666:2003); and localization by 1 = very good, 2 = rather good, 3 = moderate, 4 = rather poor, to 5 = very poor.

To evaluate the level of ecological validity achieved in the laboratory, it is interesting to look at the comparison of laboratory and field results. The results showed a statistically significant difference in perceived loudness ($p < 0.05$, Wilcoxon). In the laboratory, the trucks/buses and cars were perceived as louder than in the field by the subjects with NH (median value MD = 3.0 in the field and MD = 4.0 in the lab). The aided FTU and the EXPU subjects stated that the bicycles were louder in the laboratory than in everyday life (aided FTU: MD = 2.0, EXPU: MD = 1.0 in the field; both MD = 3.0 in the lab). This means that the loudness of the bicycles was rated as too quiet in the field and as moderate in the laboratory; several subjects stated they could not hear any bicycles with their HA in the field setting². Similar results can be seen for the annoyance ratings of the NH subjects and the FTU (both unaided and aided). In the lab, these subjects were more annoyed by the trucks/buses, cars, and bicycles than in the field ($p < 0.05$, Wilcoxon, see Tab. 3). The NH and the unaided FTU subjects reported the same annoyance ratings: they were not annoyed at all at the truck/buses and cars in the field (MD = 1.0), whereas in the laboratory they were moderately annoyed (MD = 3.0). Additionally, the bikes were not annoying in the field at all (MD = 1.0), but in the laboratory, they were found to be slightly annoying (MD = 2.0). A similar result was reported by the aided FTU subjects. In the field, they were somewhat annoyed by trucks/buses and cars (MD = 2.0) because of their HA, but even more in the laboratory (MD = 4.0). Since they did not hear the bikes in everyday life, they were not annoyed by them (MD = 1.0), but in the laboratory they were (MD = 3.0); this is consistent with the loudness evaluation. There was little significant difference in localization perception between NH and unaided FTU subjects and between aided FTU and EXPU subjects. The FTU (unaided) subjects found it harder to localize the sound sources in the field and in the laboratory than the subjects with NH. For the aided FTU subjects, it was easier to locate the sound sources than for the EXPU with the exception of bicycles in the field. Although it is debatable why the aided FTU evaluated their localization of sound sources better than EXPU, it has to be considered that these are minor differences and that they are not statistically significant ($p > 0.05$, Wilcoxon).

In the field, a notable expressive orientation reaction was found in the aided, but not in the unaided FTU. This finding was reflected to a smaller extent in the lab by an (not significant) increase in small torso and large head movements in the aided compared to the unaided FTU group. In the field, the subjects showed a noticeable reaction to the environment, whereas in the lab, their behavior was less pronounced (e.g., similar to watching a television screen). In an everyday environment, communication takes place by exchanging information, whereas passive listening was predominant in the lab experiment. This may explain the difference in behavior between field and lab. Other behaviors may appear when interference occurs (e.g., wearing hearing aids for the first time and strongly turning to sound sources). One could assume that, when the aided FTU get used

to their hearing aids and become EXPUs, they would also show less pronounced movements. Similarly, the situation in the lab equalizes the subjects' behavior (they are more passive than active) because it represents an insecure situation for all subject groups involved (Nassehi 2019, p. 91). However, if the lab becomes more ecologically valid, the behavior of the subjects may also become more like that in the field.

In summary, data from the virtual reality lab reflect real-life hearing-related function to a certain extent, but relevant differences between field and lab remain. One hypothesis is that subjects in the field felt more involved in the listening situation than in the lab, which may have led to the observed differences in movement behavior. Using more interactive procedures may increase involvement, which may reduce the differences in behavior between lab and field found in the studies by Paluch et al. (2017, 2018a, b, 2019) reported here. This remains to be tested in future studies. The studies also showed the difficulties encountered when comparing field and lab data. The use of field notes in the field and video annotation in the lab to assess similar classes of behavior is a promising technique as it is nonintrusive to the subject and allows deriving testable hypotheses. The comparative data from the studies may be used to derive metrics that can be used both in the lab and the field to assess movement behavior and relate them to performance in hearing-related tasks. Specifically, questionnaires covering the motion-related factors identified here may be derived that can be used both in the field and in the lab to assess hearing-related function.

GENERAL DISCUSSION AND CONCLUSIONS

The available data and experiences suggest that the virtual reality lab is generally applicable in hearing research and has the potential to assess and predict the ability of subjects with and without hearing devices to accomplish real-life hearing-related tasks. In that sense, the virtual reality lab in its current state may already provide higher ecological validity than established audio-only methods. In particular, Hendrikse et al. (2018) found that visual stimulation influenced head- and eye-movement behavior in normal-hearing subjects. Movement behavior was found to be similar for visual stimulation with animated characters and with videos of real conversations, showing that animated characters may be used to promote realistic movement behavior. Audio-only presentation led to a significantly different behavior. In that case, subjects virtually stopped moving their heads. Because head movement reduces the benefit in signal to noise ratio provided by spatial filtering (Hendrikse et al. 2020), realistic head movement may be a crucial factor in ecologically more valid hearing aid evaluation in the lab.

Current limitations include the limited resolution of the visual systems and the lack of precise automatic lip-syncing procedures, as well as the facial and body gestures of the animated characters (see Hendrikse et al. 2019). Whereas video resolution can be assumed to improve automatically with the further development of video technology, improving lip-syncing and facial expressions appears to be more difficult. This is because lip movement and facial expressions need to be generated in real-time to enable interactivity, and this limits rendering quality. Deep-learning-based techniques, which are currently widely used to recognize facial expression (e.g., Zeng et al. 2018) may lead to improved solutions.

Another factor is the differences in loudness perception between lab and field; this requires further consideration, as shown in earlier studies (e.g., Smeds et al. 2006). If sounds presented in the lab continue to be louder than in the field even with more interactive measurement paradigms and with audiovisual stimulation, guidelines need to be derived for reducing the sound level in the lab compared to that in the field.

Furthermore, current experimental paradigms do not fully meet the requirements for putting the subject in a communication loop, as they still primarily follow the concept of unidirectional communication, with a signal sender and a signal receiver. Therefore, future work will be towards improving the visual technology and developing “subject-in-the-loop” paradigms to further increase the ecological validity of the virtual reality lab. One approach to closing the loop may be to use acted conversations (e.g., Beechey et al. 2019, 2020) in combination with virtual interactive animated characters in the virtual reality lab. This could be done via embodied conversational agents (Lorach et al. 2019), which are animated characters in the virtual reality that are controlled by a virtual conversational agent (e.g., Lorach & Blat 2017) or an actor (interactive puppeteering, e.g., Husinsky & Bruckner 2018). Another potential way to close the loop is to combine virtual-reality glasses with an omnidirectional treadmill, which may increase the interactivity and involvement of the subject while they perform hearing-related tasks. Lau et al. (2016), for example, showed that a combination of a hearing-related and a mobility-related task implemented in virtual reality with a treadmill may lead to more effective approaches to assessing hearing-related function.

Another emerging and potentially useful technology is augmented reality, for example, Miller et al. (2019). Augmented reality glasses with audio and video augmentation may be used to design tests in the field that are more similar to experiment designs in the virtual reality lab and use less distractive response techniques. This way, information learned from the virtual reality lab on how to design interactive tests may be transferred to the field, and vice versa, leading to a more integrated approach between hearing assessment in the lab and in the field.

ACKNOWLEDGMENTS

The authors thank Matthias Latzel (Sonova AG) for providing the hearing aid devices.

This study was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—Project-ID 352015383—SFB 1330 and under Germany's Excellence Strategy—EXC 2177/1—Project ID 390895286, the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 675324 (ENRICH), as well as financially supported by the Evangelisches Studienwerk Villigst.

Parts of this study have been presented at the meeting of the German Audiological Society 2017 and 2018 (Paluch et al. 2017, 2018b), the junior meeting of the Working Group Philosophical Anthropology of the Section Cultural Sociology (German Sociological Association) and Helmholtz Plessner Society 2018 (Paluch 2019) and the ISAAR symposium 2017 (Paluch et al. 2018a).

The article was written by V.H. and R.P. and critically revised for important intellectual content by G.G.; R.P., M.M., G.G., and V.H. contributed to the design of the study presented in the section “Comparative study of subject behavior and performance in the virtual reality lab and in real life,” which was conducted and analyzed by R.P. and M.K. G.G. designed the virtual reality lab (section “Technology of the virtual reality lab”) and contributed to

the concepts presented in the section “Requirements for reflecting real-life communication in the lab: The communication loop.”

The authors have no conflicts of interest to disclose.

Address for correspondence: Volker Hohmann, Department of Medical Physics and Acoustics and Cluster of Excellence Hearing4all, University of Oldenburg, D-26111 Oldenburg, Germany. E-mail: volker.hohmann@uol.de

Received May 29, 2020; accepted July 15, 2020.

REFERENCES

- Ahrens, A., Lund, K. D., Marschall, M., Dau, T. (2019). Sound source localization with varying amount of visual information in virtual reality. *PLoS One*, *14*, e0214603.
- Beechey, T., Buchholz, J. M., Keidser, G. (2019). Eliciting naturalistic conversations: A method for assessing communication ability, subjective experience, and the impacts of noise and hearing impairment. *J Speech Lang Hear Res*, *62*, 470–484.
- Beechey, T., Buchholz, J. M., Keidser, G. (2020). Hearing aid amplification reduces communication effort of people with hearing impairment and their conversation partners. *J Speech Lang Hear Res*, *63*, 1299–1311.
- Bentler, R. A. (2005). Effectiveness of directional microphones and noise reduction schemes in hearing aids: A systematic review of the evidence. *J Am Acad Audiol*, *16*, 473–484.
- Bleichner, M. G., & Debener, S. (2017). Concealed, unobtrusive ear-centered EEG acquisition: cEEGrids for transparent EEG. *Front Hum Neurosci*, *11*, 163.
- Campos, J. L., & Launer, S. (2020). From healthy hearing to healthy living: A holistic approach. *Ear Hear*, *41*(Suppl 1), 99S–106S.
- Carlile, S., & Keidser, G. (2020). Conversational interaction is the brain in action: Implications for the evaluation of hearing and hearing interventions. *Ear Hear*, *41*(Suppl 1), 56S–67S.
- Cord M., Surr R., Walden B., Dyrland O. (2004) Relationship between laboratory measures of directional advantage and everyday success with directional microphone hearing aids. *J Am Acad Audiol*, *15*, 353–364.
- Frith, C.D. (2007). *Making Up the Mind*. Blackwell Publishing.
- Grimm, G., Hendrikse, M., Hohmann, V. (2020). Survey of self motion in the context of hearing and hearing device research. *Ear Hear*, *41*(Suppl 1), 48S–55S.
- Grimm, G., Luberadzka, J., Hohmann, V. (2019). A toolbox for rendering virtual acoustic environments in the context of audiology. *Acta Acustica United Acust*, *105*, 566–578.
- Grimm, G., Kayser, H., Hendrikse, M., Hohmann, V. (2018). A gaze-based attention model for spatially-aware hearing aids. *Speech Communication*; *13th ITG-Symposium*, Oldenburg, Germany (pp. 1–5). VDE.
- Grimm, G., Kollmeier, B., Hohmann, V. (2016). Spatial acoustic scenarios in multichannel loudspeaker systems for hearing aid evaluation. *J Am Acad Audiol*, *27*, 557–566.
- Grimm, G., Ewert, S., & Hohmann, V. (2015). Evaluation of spatial audio reproduction schemes for application in hearing aid research. *Acta Acustica United with Acustica*, *101*, 842–854.
- Hadley, L. V., Brimijoin, W. O., Whitmer, W. M. (2019). Speech, movement, and gaze behaviours during dyadic conversation in noise. *Sci Rep*, *9*, 10451.
- Heller, A., & Benjamin, E. (2014). The ambisonic decoder toolbox: Extensions for partial-coverage loudspeaker arrays. *Proceedings of the Linux Audio Conference*, Karlsruhe, Germany (pp. 1–9). <http://lac.linuxaudio.org/2014/papers/17.pdf>.
- Hendrikse, M. M., Llorach, G., Grimm, G., Hohmann, V. (2018). Influence of visual cues on head and eye movements during listening tasks in multi-talker audiovisual environments with animated characters. *Speech Commun*, *101*, 70–84.
- Hendrikse, M. M. E., Llorach, G., Hohmann, V., Grimm, G. (2019). Movement and gaze behavior in virtual audiovisual listening environments resembling everyday life. *Trends Hear*, *23*, 2331216519872362.
- Hendrikse, M. M. E., Grimm, G., Hohmann, V. (2020). Evaluation of the influence of head movement on hearing aid algorithm performance using acoustic simulations. *Trends Hear*, *24*, 2331216520916682.
- Holube, I., von Gablenz, P., Bitzer, J. (2020). Ecological momentary assessment (EMA) in audiology: Current state, challenges, and future directions. *Ear Hear*, *41*(Suppl 1), 79S–90S.
- Jennings, T. R., & Kidd, G., Jr. (2018). A visually guided beamformer to aid listening in complex acoustic environments. *Proceedings of Meetings on Acoustics 175th ASA*, Minneapolis, MN, USA (Vol. 33, No. 1, p. 050005). ASA.
- Husinsky, M., & Bruckner, F. (2018). Virtual stage: Interactive puppeteering in mixed reality. *2018 IEEE 1st Workshop on Animation in Virtual and Augmented Environments (ANIVAE)*, Reutlingen, Germany (pp. 1–7). IEEE.
- Keidser, G., Naylor, G., Brungart, D., Caduff, A., Campos, J., Carlile, S., Carpenter, M., Grimm, G., Hohmann, V., Holube, I., Launer, S., Lunner, T., Mehra, R., Rapport, F., Slaney, M., Smeds K. (2020). The quest for ecological validity in hearing science: What it is, why it matters, and how to advance it. *Ear Hear*, *41*(Suppl 1), 5S–19S.
- Lau, S. T., Pichora-Fuller, M. K., Li, K. Z., Singh, G., Campos, J. L. (2016). Effects of hearing loss on dual-task performance in an audiovisual virtual reality simulation of listening while walking. *J Am Acad Audiol*, *27*, 567–587.
- Lesica, N. A. (2018). Why do hearing aids fail to restore normal auditory perception? *Trends Neurosci*, *41*, 174–185.
- Llorach, G. & Blat, J. (2017). Say hi to Eliza. *An embodied conversational agent on the web. Proceedings of the 17th International Conference on Intelligent Virtual Agents (IVA)*, Stockholm, Sweden. Springer.
- Llorach, G., Agenjo, J., Blat, J., Sayago, S. (2019). Web-based embodied conversational agents and older people. In S. Sayago (Ed.), *Perspectives on Human-Computer Interaction Research with Older People* (pp. 119–135). Springer.
- Llorach, G., Evans, A., Blat, J., Grimm, G., Hohmann, V. (2016). Web-based live speech-driven lip-sync. *2016 8th International Conference on Games and Virtual Worlds for Serious Applications (VS-GAMES)*, Barcelona, Spain (pp. 1–4). IEEE.
- Llorach, G., Grimm, G., Hendrikse, M. M., Hohmann, V. (2018). Towards realistic immersive audiovisual simulations for hearing research: Capture, virtual scenes and reproduction. *Proceedings of the 2018 Workshop on Audio-Visual Scene Understanding for Immersive Multimedia*, New York, NY, USA (pp. 33–40). ACM.
- Mead, G. H. (1956). *The Social Psychology of George Herbert Mead*. The University of Chicago Press.
- Miller, M. R., Jun, H., Herrera, F., Yu Villa, J., Welch, G., Bailenson, J. N. (2019). Social interaction in augmented reality. *PLoS One*, *14*, e0216290.
- Moulin-Frier, C., Laurent, R., Bessière, P., Schwartz, J. L., Diard, J. (2012). Adverse conditions improve distinguishability of auditory, motor, and perceptuo-motor theories of speech perception: An exploratory Bayesian modelling study. *Lang Cogn Process*, *27*, 1240–1263.
- Nassehi, A. (2019). *Muster: Theorie der digitalen Gesellschaft*. C.H. Beck.
- Oreinos, C., & Buchholz, J. (2014). Validation of realistic acoustic environments for listening tests using directional hearing aids. *2014 14th International Workshop on Acoustic Signal Enhancement (IWAENC)*, Juan-les-Pins, France (pp. 188–192). IEEE.
- Paluch, R. (2019). Die technisch vermittelte Umweltbeziehung des leiblichen Selbstes in virtuellen Welten. In J. F. Burow, L.-J. Daniels, A.-L. Kaiser, C. Klinkhamer, J. Kulbatzki, Y. Schütte, & A. Henkel (Eds.). *Mensch und Welt im Zeichen der Digitalisierung. Perspektiven der Philosophischen Anthropologie Plessners* (pp. 145–164). Nomos.
- Paluch, R., Krueger, M., Hendrikse, M. M. E., Grimm, G., Hohmann, V., Meis, M. (2018a). Ethnographic research: The interrelation of spatial awareness, everyday life, laboratory environments, and effects of hearing aids. *Proc ISAAR*, *6*, 39–46. <https://proceedings.isaar.eu/index.php/isaarproc/article/view/2017-06>.
- Paluch, R., Krueger, M., Meis, M. (2018b). The technization of self-care in hearing aid research. *21. Jahrestagung der Deutschen Gesellschaft für Audiologie*. <https://www.dga-ev.com/fileadmin/dga2018/site/data/final/0056.pdf>.
- Paluch, R., Krueger, M., Grimm, G., & Meis, M. (2017). Moving from the field to the lab: Towards ecological validity of audio-visual simulations in the laboratory to meet individual behavior patterns and preferences. *20. Jahrestagung der Deutschen Gesellschaft für Audiologie*. CD-ROM.
- Paluch, R., Krueger, M., Hendrikse, M. M. E., Grimm, G., Hohmann, V., Meis, M. (2019). Towards plausibility of audiovisual simulations in the laboratory: Methods and first results from subjects with normal hearing or with hearing impairment. *Z Audiol*, *58*, 6–15.
- Pandey, A., & Wang, D. (2019). A new framework for CNN-based speech enhancement in the time domain. *IEEE/ACM Trans Audio Speech Lang Process*, *27*, 1179–1188.
- Pausch, F., Aspöck, L., Vorländer, M., Fels, J. (2018). An extended binaural real-time auralization system with an interface to research hearing aids for experiments on subjects with hearing loss. *Trends Hear*, *22*, 2331216518800871.

- Roosendaal, T., & Wartmann, C. (2003). *The Official Blender Game Kit: Interactive 3D for Artists*. No Starch Press.
- Shiffman, S., Stone, A. A., Hufford, M. R. (2008). Ecological momentary assessment. *Annu Rev Clin Psychol*, 4, 1–32.
- Smeds, K., Keidser, G., Zakis, J., Dillon, H., Leijon, A., Grant, F., Convery, E., Brew, C. (2006). Preferred overall loudness. II: Listening through hearing aids in field and laboratory tests. *Int J Audiol*, 45, 12–25.
- Smeds, K., Gotowiec, S., Wolters, F., Herrlin, P., Larsson, J., Dahlquist, M. (2020). Selecting scenarios for hearing-related laboratory testing. *Ear Hear*, 41(Suppl 1), 20S–30S.
- Straub I. (2016). 'It looks like a human!' The interrelation of social presence, interaction and agency ascription: A case study about the effects of an android robot on social agency ascription. *AI Soc*, 31, 553–571.
- Völker, C., Warzybok, A., Ernst, S. M. (2015). Comparing binaural pre-processing strategies III: Speech intelligibility of normal-hearing and hearing-impaired listeners. *Trends Hear*, 19, 2331216515618609.
- Wilms, M., Schilbach, L., Pfeiffer, U., Bente, G., Fink, G. R., Voigeley, K. (2010). It's in your eyes—using gaze-contingent stimuli to create truly interactive paradigms for social cognitive and affective neuroscience. *Soc Cogn Affect Neurosci*, 5, 98–107.
- Wolters, F., Smeds, K., Schmidt, E., Christensen, E. K., Norup, C. (2016). Common sound scenarios: A context-driven categorization of everyday sound environments for application in hearing-device research. *J Am Acad Audiol*, 27, 527–540.
- Wu, Y. H., Stangl, E., Chipara, O., Hasan, S. S., DeVries, S., Oleson, J. (2019). Efficacy and effectiveness of advanced hearing aid directional and noise reduction technologies for older adults with mild to moderate hearing loss. *Ear Hear*, 40, 805–822.
- Zeng, N., Zhang, H., Song, B., Liu, W., Li, Y., Dobaie, A. M. (2018). Facial expression recognition via learning deep sparse autoencoders. *Neurocomputing*, 273, 643–649.

REFERENCE NOTE

- Daniel, J. (2001). Représentation de champs acoustiques, application à la transmission et à la reproduction de scènes sonores complexes dans un contexte multimedia. PhD Thesis. Université Pierre et Marie Curie (Paris VI), Paris, France, Vol. 1. <http://pcfarina.eng.unipr.it/Public/phd-thesis/jd-these-original-version.pdf>.