DOI: 10.61782/fa.2023.0259

TOWARDS INVESTIGATING LISTENING COMPREHENSION IN VIRTUAL REALITY

Carolin Breuer[1*] **Larissa Leist**[2] **Stephan Fremerey**[3]
Alexander Raake[3] **Maria Klatte**[2] **Janina Fels**[1]
[1] Institute for Hearing Technology and Acoustics, RWTH Aachen University, Germany
[2] Cognitive and Developmental Psychology, RPTU Kaiserslautern-Landau, Germany
[3] Audiovisual Technology Group, Technische Universität Ilmenau, Germany

ABSTRACT

The investigation of listening comprehension in auditory and visually complex classroom settings is a promising method to evaluate children's cognitive performance in a realistic setting. Many studies were able to show that children are more susceptible to noise than adults. However, it has recently been suggested that established monaural listening situations could overestimate the influence of noise on children's task performance. Therefore, new, close-to real-life scenarios need to be introduced to investigate cognitive performance in everyday situations rather than artificial laboratory settings. This study aimed at extending a validated paper-and-pencil test towards a virtual reality setting. To get first insights, into different interaction methods, a pilot study with adult participants was conducted. In contrast to other recent studies, the virtual environment had little influence on this listening comprehension paradigm, since comparable results were obtained in the paper-and-pencil test and in the virtual reality variants for all user interfaces. Thus, the presented paradigm has proven to be robust and can be used to further investigate the usage of virtual reality to evaluate children's cognitive performance.

Keywords: *listening comprehension, binaural technology, virtual reality, interaction system, classroom*

**Corresponding author*: *carolin.breuer@akustik.rwth-aachen.de.*

1. INTRODUCTION

Everyday, our communication via speech signals is impaired by noisy environments. For adults, such scenarios include offices and other crowded spaces. For children, the noise exposure is especially high in schools. Considering that children are more susceptible to noise than adults [1] and that German educational facilities seem to need acoustic improvement [2], the detrimental effects of noise on children's listening comprehension need to be investigated in more detail. Previous work regarding the speech perception and listening comprehension was conducted by Klatte et al. [3] and Leist at al. [4] who showed that children are more impaired by background noise than adults. Using a word identification task, Leist et al. [4] also found an effect of the reproduction method (monaural vs. binaural), indicating that children and adults were better at identifying a word in a binaurally presented noise than in a monaurally presented noise at the same level. Since participants could make use of binaural cues and thus separate the noise and target stimuli, this could indicate that monaural investigations oversimplify the acoustic scenes and thus might overestimate the effects of background noise as opposed to binaural scenes [4]. Both studies raise the importance of more complex and realistic investigations of cognitive processes. Thus, the suitability of using virtual reality (VR) for this kind of listening comprehension task should be evaluated. However, there are persistent debates on the safety of virtual reality for children [5,6]. These concerns stem from the content and high level of immersion in children [7], but also from equipment that is designed for adults and is thus too big for children. Therefore, while this study is motivated with the target group of children in mind, a first pilot was conducted with adults to protect this vulnerable group.

With the trend towards using VR for research within the last years, multiple studies suggested that VR can be used for cognitive research [8–11]. Nevertheless, its applicability and reproducibility need to be thoroughly investigated for each paradigm separately. The transfer of established paradigms to VR environments needs to be thoroughly validated, since immersive virtual environments might influence the cognitive performance [8, 11]. Although the respective cognitive processes are not yet fully understood, there have been indications for, e.g., an enhanced attention [8]. However, an enhanced cognitive load in VR may not always be reflected directly in the task performance. Here, Redlinger et al. found increased cognitive load in VR only in EEG measurements but did not see performance differences in the task scores [10]. Further, the introduced gamification in the virtual world can lead to more engagement [12]. This could increase the participants' immersion in the scene and enhance the overall plausibility [7]. This all indicates that performance measures obtained in classical paradigms might differ in VR. However, it needs to be noted that the listed studies are visually dominated.

The question remains, how natural or plausible a VR experiment needs to be in order to mimic the behaviour in the real world. To get first insights into the suitability of VR to investigate listening comprehension in close-to real-life classroom scenarios, an established paradigm was transferred into VR and the applicability of different interaction systems was examined.

2. METHOD

2.1 Participants

To get first insights in the interaction in virtual reality, 18 adults participated in the experiment. Due to insufficient German language skills (below C1), three participants had to be excluded for the evaluation, resulting in a final group of 15 participants (age range= 20–35 years, M = 25.7 years, SD = 3.7 years, 7 female). All participants had normal hearing within 20 dB[HL] according to a pure tone audiometry and had (corrected to) normal sharpness and color vision. Informed consent was given before the experiment.

2.2 Audiovisual Reproduction

Since the focus of this experiment was the exploration of the interaction in virtual reality, a simple acoustic setup was chosen. Using the Virtual Acoustics auralization framework [13], a binaural reproduction with a generic head-related transfer function via headphones was realized. The oral instructions were played from in front of the participant and no additional noise sources were used. Thus, the auditory conditions were kept constant through the experiment for all participants. During the experiment, the participants were equipped with an HTC Vive Pro Eye head-mounted display (HMD) and the respective controllers. They were seated in the virtual classroom illustrated in Figure 1, which was created based on a real elementary classroom in Ilmenau, Germany [14]. The instructions and trials were displayed on the blackboard in front of the participants. The seating position in the virtual classroom was chosen so that all items from one trial were in the participants' field of view when looking straight ahead. Nevertheless, the participants were encouraged to look around the virtual classroom.

Figure 1. Picture of the virtual classroom model [14].

2.3 Listening Comprehension Task

The listening comprehension task used by Klatte et al. [3] and Leist et al. [4], was previously conducted as a paper-and-pencil task. In this original task, the participants were shown a list of items, which were printed on response sheets and lying in front of them. The participants were instructed orally to mark specific items. An exemplary trial is given in Figure 2. In this trial, the participants were instructed to "Draw a cross under the book that lies next to the chair". Therefore, the participants needed to identify the correct book and mark it in the specified way. Other options for marking were, e.g., circles or lines. Correctly solving the task included finding the wanted item and marking it in the specified way. To adapt the difficulty of the task, the number of items displayed and marked per trial could be changed. Therefore, in some trials up to 22 items were displayed at the same time, while in others a minimum of nine items was visible. To create a similar level of difficulty within all trials, the number of items to be marked varied between two and seven. Previous studies by Klatte et al. [3] and Leist et al. [4] showed that the task performance was comparable between the trials and blocks.

The experiment was performed in three blocks, each of which consisted of eight trials. This resulted in a total of 24 trials per experiment. By changing the number of objects in each trial, also maximum score varied. To achieve comparability, the style of the trials was similar over the blocks and thus the total score for each block was kept constant.

Each trial was initiated by a school bell sound, followed by the oral instructions and a fixed response interval of 18 seconds to complete the task. The participants were allowed to start answering during the instructions. As a first step towards a more realistic VR setting, a simple transfer of the items into a VR classroom was chosen and different input methods were compared.

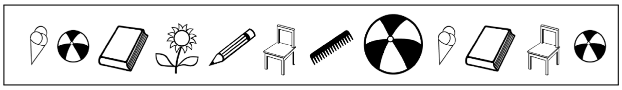

Figure 2. Examplary trial for the listening comprehension task. Corresponding instructions in German: "Male ein Kreuz unter das Buch, das neben einem Stuhl liegt." ("Draw a cross under the book that lies next to a chair.") [3, 4]

2.4 Input Methods

In the validated listening comprehension paradigm, the answers were given using a pencil on a paper answer sheet. This allowed for a very easy to use and yet complex input in terms of choosing a location and style to mark an item, e.g., "draw a cross under the book". Since the direct transfer of a paper-and-pencil based paradigm into a VR environment is difficult, the acceptability and impact of different input methods was explored. The methods needed to be intuitive to use and offer flexible interactions, since a variable number of items needed to be marked in each trial using different symbols and locations, e.g., in one trial a cross had to be put on an item, in another trial a line had to be drawn under an item. For the evaluation of the tasks, the correct item, symbol and its placement were considered.

Referring to the original input method, which allowed for a high level of flexibility, a drawing method was implemented. Using the HTC Vive controllers, the participants were able to draw directly on the virtual blackboard, where the items for each trial were displayed, see Figure 3. Using a laser beam emitted from the end of the virtual controller model, the participants were able to aim at the respective position on the items and start the drawing by holding the trigger button with their index finger. The input in this method was expected to be intuitive, but also difficult due to a lack of accuracy. A major drawback of this method was that the responses still had to be evaluated manually.

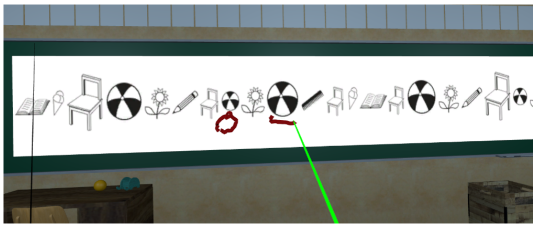

Figure 3. Example for the painting input.

Considering the interaction with established software and games, the selection of buttons is a very common method in everyday life. Thus, the input using buttons can be expected to be intuitive. Thus, aiming was required to select the correct button. A positive aspect of this input method is that pressing a button can directly be evaluated using algorithms, which is a considerable benefit compared to a paper-based input. Due to the complexity of not only choosing an item but also the type of marker, the input system was divided into two stages. First, the respective marker had to be selected. In the virtual environment, the participants could aim at the buttons with a laser pointer and select the respective button using the trigger input of the HTC Vive controllers. In the example in Figure 4, the single line was selected as marker. Secondly, the participant had to select the position to be marked, e.g., in Figure 4 the position under the comb was selected.

The order of the input methods (paper-and-pencil vs. button vs. painting) was balanced over all participants. Before each input method, a set of training tasks were implemented to familiarize the participants with the new approach.

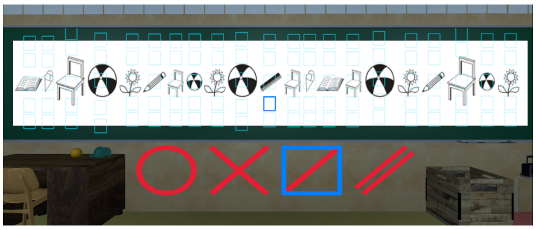

Figure 4. Example for the button input.

2.5 Questionnaire

For more detailed insights into the participants' assessment of the different input methods, a questionnaire was answered after completing the whole experiment (see Table 1). Questions included the preference of each input method (Q1 - Q3), as well as the intuitiveness (Q4 & Q5) and time to complete the task (Q6 - Q8). Further questions evaluated the acceptance of the virtual environment (Q9 - Q12), e.g., whether the participants felt like they were sitting in an actual classroom and if they felt uncomfortable or dizzy during the experiment. All questions were rated on a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree).

3. RESULTS

Since the presented study is a pilot study to examine different input methods for future investigations and considering the small sample size, the results need to be considered as preliminary.

To put the presented results into perspective, the mean performance scores were compared to those obtained in a previous study using the same paradigm by Leist et al. [4]. The mean performance scores in the Leist study ($M = 85\%$, $SD = 10\%$) did not differ significantly from the current study ($M = 80\%$, $SD = 9\%$), which is supported by a t-test, $t(49) = 1.825$, $p = .074$, $d = .561$, thus, the results of both studies seem to be comparable.

As illustrated in Figure 5, the performance appears to be similar for all input methods ($M_{button} = 77.13\%$, $SD_{button} = 2.4\%$ vs. $M_{painting} = 79.08\%$, $SD_{painting} = 2.5\%$ vs. $M_{paper} = 79.46\%$, $SD_{paper} = 2.4\%$). This was also confirmed statistically using an rm-ANOVA, $F(2, 28) = .230$, $p = .796$, $\eta_p^2 = .016$.

Regarding the questionnaire, all results regarding the preference are rather inconclusive due to high standard deviations. There might be a slight tendency towards preferring the paper-and-pencil method over the virtual reality sessions ($M_{Q1} = 3.53, SD_{Q1} = 1.36$). No distinct trend for preference was found regarding the VR input methods. However, the participants' responses were consistent in questions 2 and 3, which were formulated in a contradictory manner to get an idea of the response reliability ($M_{Q2} = 2.73, SD_{Q2} = 1.00$, $M_{Q3} = 3.13, SD_{Q3} = 1.26$). Still, both VR input methods were rated as intuitive ($M_{Q4} = 3.93, SD_{Q4} = 0.68$, $M_{Q5} = 3.93, SD_{Q5} =$

Table 1. Questionnaire answered after the experiment. Questions as well as mean and standard deviation are reported.

Question		M	SD
Q1	I preferred the paper-and-pencil version over the virtual reality sessions.	3.53	1.36
Q2	The input via buttons was more fun than the painting method in VR.	2.73	1.00
Q3	I preferred the drawing in VR over the buttons.	3.13	1.26
Q4	The input via buttons in virtual reality was intuitive.	3.93	0.68
Q5	The input via painting in virtual reality was intuitive.	3.93	1.00
Q6	I had enough time to complete the tasks in the button version.	3.27	1.00
Q7	I had enough time to complete the tasks in the drawing method.	4.20	1.05
Q8	I had enough time to complete the tasks in the paper-and-pencil version.	4.33	1.07
Q9	I felt uncomfortable in the virtual environment.	1.80	1.05
Q10	I felt dizzy during or after leaving the virtual environment.	1.53	0.96
Q11	I felt like I was sitting in a real classroom.	2.93	0.77
Q12	I forgot the real world around me when I was in the virtual reality.	3.80	0.65

1.00). Further, the participants indicated insufficient time to complete the task using the button method ($M_{Q6} = 3.27, SD_{Q6} = 1.00$) but less timing problems using the painting input in VR ($M_{Q7} = 4.20, SD_{Q7} = 1.05$) and the paper-and-pencil version ($M_{Q8} = 4.33, SD_{Q8} = 1.07$).

In terms of the virtual reality environment, the participants felt rather comfortable ($M_{Q9} = 1.80, SD_{Q9} = 1.05$) and did not show prominent signs of dizziness ($M_{Q10} = 1.53, SD_{Q10} = 0.96$). The participants reported that while the virtual classroom did not seem very realistic ($M_{Q11} = 2.93, SD_{Q11} = 0.77$), they could focus on the virtual environment while wearing the HMD ($M_{Q12} = 3.80, SD_{Q12} = 0.65$).

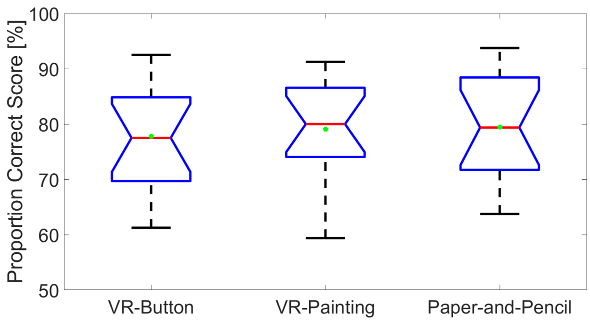

Figure 5. Performance scores for the three input methods: Button-VR, Painting-VR and Paper-and-Pencil. Mean values are represented by green asterisks, median values are shown by a red line. Lower and upper limits of the box represent the 25th and 75th percentile, respectively. The whiskery indicate the minimum and maximum values.

4. DISCUSSION

First of all, the present study was able to reproduce the results previously obtained by Leist et al. [4], which indicates an overall stability of the paper-and-pencil listening comprehension paradigm. Thus, the presented paradigm can be used for future studies investigating the listening comprehension in a virtual classroom environment, bringing established methods closer to real-life scenarios.

Although this pilot study found no significant differences between the proposed input methods, the overall acceptance of the interaction systems was inconclusive. No clear preference was found between the input methods. Despite the intuitiveness of all presented options, both drawing-based methods seemed to be preferred. This could be due to a lack of time to complete the tasks using the button input. This time was kept constant for all trials and all input methods. However, future studies should consider introducing a varied time interval which could also allow for the investigation of reaction times.

Further, the accuracy of the painting method is in direct relation to the distance to the canvas that is painted on, i.e., the accuracy rises the closer to the canvas. Figure 3 clearly shows the difficulty of drawing the marks which is visible through shaky lines. Similarly, the handling of the button method required aiming at and selecting specific buttons on the canvas. The difficulty of this task also increases the further away the participants stood from the blackboard. However, the minimal distance to the blackboard was limited by the partipants' field of view. To minimize the time to find the relevant item, all of the items needed to be visible when looking directly to the front. This also reduced fast head movements which can induce motion sickness. None of the participants showed signs of motion sickness, which was supported by the questionnaire. Still, a more advanced questionnaire, such as the Simulator Sickness Questionnaire [15] or the Slater-Usoh-Steed Presence Questionnaire [16], should be implemented for further studies to get more detailed insights into the participants' well-being and immersion into the virtual world.

The difficulty of the interaction systems was partly given by the complexity of the presented task, which requires a variety of markers and marked locations, e.g., "put a cross under the book" or "mark the chair with a circle". Combined with the possibility of giving partly correct answers, e.g., marking the wrong item but using the correct marker style, this intricacy calls for an elaborate evaluation scheme which cannot easily be implemented. Thus, although a partly automatic evaluation was implemented for the button version, the final evaluation had to be completed manually for all versions. Therefore, the paradigm and its evaluation strategy need to be revised, while keeping in mind the options and limitations offered by an immersive virtual environment. To exploit the benefits of virtual reality and create a close-to-real-life scenario, the paradigm needs to be extended further. By introducing direct interaction with the virtual world and allowing the participants to navigate through the classroom, the level of immersion could be increased. Also, adding realistic classroom noise referring to the previous study by Leist et al. [4] needs to be a next step in creating an interactive and ecologically valid paradigm to evaluate the listening comprehension in a virtual classroom.

5. CONCLUSION

Since there were no statistical differences between the previous paper-and-pencil study by Leist at al. [4] and the newly introduced input methods, it can be concluded that the presented listening comprehension paradigm can be used for further investigations in virtual reality. Based on the finding by Klatte et al. [3] and Leist et al. [4] regarding the detrimental effects of background noise, a room acoustic simulation as well as a realistic background noise should be included in further investigations.

6. ACKNOWLEDGMENTS

The authors would like to thank Lalita Angsuphanich for assisting in the paradigm design and data collection, as well as Karin Loh for sharing her thoughts and giving advice on designing the paradigm.

The research described in the paper was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under the project ID 444697733 with the title "Evaluating cognitive performance in classroom scenarios using audiovisual virtual reality – ECoClass-VR". This project is part of the priority program "AUDICTIVE – SPP2236: Auditory Cognition in Interactive Virtual Environments".

7. REFERENCES

[1] M. Klatte, K. Bergström, and T. Lachmann, "Does noise affect learning? a short review on noise effects on cognitive performance in children," *Frontiers in psychology*, vol. 4, p. 578, 2013.

[2] K. Loh, M. Yadav, K. Persson Waye, M. Klatte, and J. Fels, "Toward child-appropriate acoustic measurement methods in primary schools and daycare centers," *Frontiers in Built Environment*, vol. 8, 2022.

[3] M. Klatte, T. Lachmann, M. Meis, *et al.*, "Effects of noise and reverberation on speech perception and listening comprehension of children and adults in a classroom-like setting," *Noise and Health*, vol. 12, no. 49, p. 270, 2010.

[4] L. Leist, C. Breuer, M. Yadav, S. Fremerey, J. Fels, A. Raake, T. Lachmann, S. J. Schlittmeier, and M. Klatte, "Differential effects of task-irrelevant monaural and binaural classroom scenarios on children and adults; speech perception, listening comprehension, and visual verbal short-term memory," *International Journal of Environmental Research and Public Health*, vol. 19, no. 23, 2022.

[5] P. Kaimara, A. Oikonomou, and I. Deliyannis, "Could virtual reality applications pose real risks to children and adolescents? a systematic review of ethical issues and concerns," *Virtual Reality*, 08 2021.

[6] R. Zender, J. Buchner, S. Caterina, D. Wiesche, K. Kelly, and L. Tueshaus, "Virtual reality für schüler:innen: Ein beipackzettel für die durchführung immersiver lernszenarien im schulischen kontext," *MedienPädagogik: Zeitschrift für Theorie Und Praxis Der Medienbildung*, no. 47, pp. 26 – 52, 2022.

[7] G. Makransky, T. S. Terkildsen, and R. E. Mayer, "Adding immersive virtual reality to a science lab simulation causes more presence but less learning," *Learning and Instruction*, vol. 60, pp. 225–236, 2019.

[8] C. Breuer, K. Loh, L. Leist, S. Fremerey, A. Raake, M. Klatte, and J. Fels, "Examining the auditory selective attention switch in a child-suited virtual reality classroom environment," *International Journal of Environmental Research and Public Health*, vol. 19, no. 24, 2022.

[9] R. N. Meghanathan, P. Ruediger-Flore, F. Hekele, J. Spilski, A. Ebert, and T. Lachmann, "Spatial sound in a 3d virtual environment: All bark and no bite?," *Big Data and Cognitive Computing*, vol. 5, no. 4, 2021.

[10] E. Redlinger, B. Glas, and Y. Rong, *Enhanced Cognitive Training Using Virtual Reality: Examining a Memory Task Modified for Use in Virtual Environments*, p. 1–8. New York, NY, USA: Association for Computing Machinery, 2021.

[11] G. Li, J. A. Anguera, S. V. Javed, M. A. Khan, G. Wang, and A. Gazzaley, "Enhanced Attention Using Head-mounted Virtual Reality," *Journal of Cognitive Neuroscience*, vol. 32, no. 8, pp. 1438–1454, 2020.

[12] J. Lumsden, E. A. Edwards, N. S. Lawrence, D. Coyle, and M. R. Munafò, "Gamification of cognitive assessment and cognitive training: A systematic review of applications and efficacy," *JMIR Serious Games*, vol. 4, p. e11, Jul 2016.

[13] Institute for Hearing Technology and Acoustics, RWTH Aachen University, "Virtual Acoustics – A real-time auralization framework for scientific research." http://www.virtualacoustics.org/. Accessed on 2022-03-16.

[14] S. Fremerey, C. Reimers, L. Leist, J. Spilski, M. Klatte, J. Fels, and A. Raake, "Generation of audiovisual immersive virtual environments to evaluate cognitive performance in classroom type scenarios," *Tagungsband - DAGA 2021: 47. Jahrestagung für Akustik : 15.-18. August 2021, Wien und Online : Tagungsband der DAGA 2021*, pp. 1336–1339, 2021.

[15] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal, "Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness," *The International Journal of Aviation Psychology*, vol. 3, no. 3, pp. 203–220, 1993.

[16] M. Usoh, E. Catena, S. Arman, and M. Slater, "Using Presence Questionnaires in Reality," *Presence: Teleoperators and Virtual Environments*, vol. 9, pp. 497–503, 10 2000.