2022 IEEE on Conference Virtual Reality and 3D User Interfaces (VR) | 978-1-6654-9617-9/22/$31.00 ©2022 IEEE | DOI: 10.1109/VR51125.2022.00042

Empirical Evaluation of Calibration and Long-term Carryover Effects of Reverberation on Egocentric Auditory Depth Perception in VR

Wan-Yi Lin*
College of Computer Science
National Yang Ming Chiao Tung Univ.

Ying-Chu Wang†
Dept. of Computer Science
National Yang Ming Chiao Tung Univ.

Dai-Rong Wu‡
College of Computer Science
National Yang Ming Chiao Tung Univ.

Rohith Venkatakrishnan§
School of Computing
Clemson University

Roshan Venkatakrishnan¶
School of Computing
Clemson University

Elham Ebrahimi‖
Dept. of Computer Science
UNC Wilmington

Christopher Pagano**
Department of Psychology
Clemson University

Sabarish V. Babu††
School of Computing
Clemson University

Wen-Chieh Lin‡‡
Dept of Computer Science
National Yang Ming Chiao Tung University

Abstract

Distance compression, which refers to the underestimation of egocentric distance to objects, is a common problem in immersive virtual environments. Besides visually compensating the compressed distance, several studies have shown that auditory information can be an alternative solution for this problem. In particular, reverberation time (RT) has been proven to be an effective method to compensate distance compression. To further explore the feasibility of applying audio information to improve distance perception, we investigate whether users' egocentric distance perception can be calibrated, and whether the calibrated effect can be carried over and even sustain for a longer duration. We conducted a study to understand the perceptual learning and carryover effects by using RT as stimuli for users to perceive distance in IVEs. The results show that the carryover effect exists after calibration, which indicates people can learn to perceive distances by attuning reverberation time, and the accuracy even remains a constant level after 6 months. Our findings could potentially be utilized to improve the distance perception in VR systems as the calibration of auditory distance perception in VR could sustain for several months. This could eventually avoid the burden of frequent training regimens.

Keywords: Depth Perception, Auditory Reverberation, Calibration, Perceptual Learning

Index Terms: Computing methodologies [Computer Graphics]: Graphics systems and interfaces—Perception Human-centered computing [Human computer interaction (HCI)]: Interaction paradigms—Virtual reality

1 Introduction

Distance perception in virtual reality (VR) is crucial when users need to react accurately to some spatial information in applications like driving to an assigned location, or when distinguishing targets in visualizations [32, 36]. Distance perception is also the pre-requisite for the accurate perception of higher level information such as size, shape, and speed perception [4, 29, 35, 43]. However, there is a well-known distance perception problem in VR, *distance compression*. It is a phenomenon wherein users of VR simulations tend to estimate the distance between themselves and objects around them as closer than the modeled location. Renner et al. analyzed studies on distance compression and found that the perceived egocentric distances (distance from the observer to an object) in VR was about 74 percent of the actual distance to the object [40]. Although head mounted displays (HMDs) have iteratively improved over the years, allowing for more accurate estimations of depth, the distance compression issue continues to persist and is yet to be completely resolved [8, 9, 26, 34].

When perceiving distances in the real world, people may integrate information from multiple perceptual channels towards estimating overall distance. Traditionally, VR experiences tended to involve users primarily relying on visual perceptual information towards the estimation of depth more than information from other sensory channels. This was because these other modalities of sensory information were seldom provided. However, with modern immersive virtual environments (IVEs) increasingly featuring multisensory experiences including visual, auditory, tactile and even olfactory perceptual information, it stands to reason that distance estimation in contemporary VR depends on more than just visual information. Along these lines, research has shown that in VR, users' estimations of depth can be enhanced by leveraging auditory information in a room sized acoustic environment [44]. Researchers have thus investigated several auditory information and components associated with distance perception including those of audio intensity/level, reverberation, spectral cues, binaural cues, and dynamic cues [27].

Reverberation is the accumulation of sound-waves in a space and is created by sound reflecting off the surfaces in the environment [25]. Auditory reverberation has recently been shown to enhance users' spatial perception [16, 31]. Altmann et al. have shown that increasing the reverberation time could improve the distance estimation of farther sources [3]. With respect to auditory cues, Bailey and Fazenda demonstrated that stimuli with late reverberation resulted in better distance estimation than those without late reverberation in VR [5]. Furthermore, Huang et al. have shown that the mere presence of auditory reverberation in addition to visual information compensated for depth compression and enhanced distance perception in VR [21]. Overall, research seems to indicate that the reverberation component of auditory information can be exploited to alter how users perceive depth in IVEs.

In IVEs, users' perceptual judgments can be improved after calibration or attunement [7, 37]. It has been shown that compression of visually perceived depth in medium field distances in IVEs can be improved by closed-loop feedback of travel and locomotion in a calibration session. For instance, prior research in IVEs have demonstrated the existence of carryover effects when providing visual or haptic feedback in a calibration session [10, 13]. It was

*e-mail: waynei.pcs09g@nctu.edu.tw
†email: joannaahj@gmail.com
‡e-mail: azusa871227.cs10@nycu.edu.tw
§email: rohithv@g.clemson.edu
¶email: rvenkat@g.clemson.edu
‖email: eebr248@gmail.com
**email: cpagano@clemson.edu
††email: sbabu@clemson.edu
‡‡e-mail: wclin@cs.nctu.edu.tw

978-1-6654-9617-9/22/$31.00 ©2022 IEEE
DOI 10.1109/VR51125.2022.00042

found that these carryover effects enhanced depth perception in the post-calibration phase with users' estimates being more accurate than those provided in the pre-calibration phase [13]. While there have been such works that have investigated calibration, perceptual learning and carryover effects in the context of distance estimation, these have largely been limited to visual information. There is hence a lack of research that has investigated if users can attune to auditory information provided in a calibration phase let alone looking into the longitudinal effects (over time) of calibration. There hence exists an avenue of research worth investigating that looks at users' ability to calibrate to auditory information and if/how these carryover effects are affected over time.

In this paper, we investigate the perceptual learning and carryover effects of calibration to auditory reverberation associated with the perceived depth of a target in VR. A motivation of this study is to explore the feasibility of applying the carryover effects to improve distance perception in IVEs. We thus aim to discover if users can calibrate their distance perception with reverberation stimuli in VR, and whether this can be exploited to make users potentially perceive distances to objects more accurately and closer to that of the modeled distance. This could work well especially in acoustic environments like caves and sports stadiums to name a few. Moreover, if the calibration carryover effect of auditory reverberation is shown to persist over a longer duration, then it attests to the strength of this stimuli in producing long term persistence of enhanced depth perception in VR, and thereby minimizes the need for users to be recalibrated in VR. This would greatly benefit many VR applications involving repeated usage some of which include training simulations, educational application, games, etc. [33].

Specifically, the research questions that this study aims at answering are (1) whether participants can calibrate to auditory reverberation towards enhancing egocentric distance estimation in IVEs; (2) whether the effects of calibration carry over periods of 1 month and 6 months. To answer these research questions, we rendered a virtual environment where the users were given an opportunity to learn how to use reverberation time (RT) in a calibration phase, as information to differentiate between different target distances in an accurately scaled virtual environment. Reverberation time (RT) is one of the reverberation parameters that can be linearly adjusted to manipulate the users' distance perception based on auditory information [22]. Our experiment consisted of a pre-test phase, a calibration phase and a post-test phase. In both the pre-test and post-test phases, participants provided depth judgments to the perceived distance of a sound source location without feedback (open-loop depth judgment task) in VR. In the calibration phase, participants received feedback and were allowed to adjust their depth judgments to the actual distance of the sound source location with in VR. At the end of the calibration phase, participants were immediately transitioned to the post-test phase, which was identical to the pre-test phase. Participants were then recalled after one month and six months to test their depth perception in an attempt to examine the persistence of calibration effects to auditory reverberation information in VR.

In summary, the contributions of our empirical evaluation are as follows: (1) a novel study in examining the perceptual learning and carryover effects of auditory reverberation information on egocentric distance perception in VR; (2) we evaluate to what extent reverberation time could be effective information for distance perception and the carryover effects potentially persisting over a longer duration (one to six months).

2 Related Work

Virtual reality is an excellent testbed for experimentation, training, and education because different real-world characteristics can be controlled and manipulated in virtual environments. However, limitations in virtual technology present some challenges to VR developers. One of these challenges is distance misperception, which is a well-studied field where a considerable amount of work has been focused on reducing the disparities between real and virtual environments [15]. Previous studies showed that many factors contribute to distance misperception in VE, such as carrying a heavy backpack [38], weight and forces from the head-mounted display (HMD) [47], action capabilities of the body [19, 30], the field of view [24], the visual fidelity of the VE [42]. However, the exact cause of distance underestimation is still inconclusive.

In most cases, the virtual rendering contains enough visual, audio, and haptic information that the user needs to complete a task successfully. However, the user has not been given an opportunity to learn how to attune to (detect) and use that information [45]. Some previous work suggested that user interaction with the environment helps with perceptual learning and allows the user to attune and calibrate to the new environment, consequently improving the distance estimation [2, 48]. Attunement is an important type of differentiation where the human learns to detect different sources of perceptual information [11]. Along with the attunement, the user also becomes calibrated during perceptual learning. Via calibration, the users come to properly scale their use of the perceptual information they attune to [1, 7]. Richardson and Waller [41] provided error correction feedback as perceptual information to users. They found that users calibrated to the information and, consequently, distance estimation improved in a blind walking task. They showed that the effect of distance calibration persisted for at least one week after the experiment. Mohler et al. [33] investigated the perceptual feedback in VE. They provided three different forms of feedback. They found that perceptual feedback increases the accuracy of the distance estimation, not its type, and in some cases, the calibration effect was carried over to the real world.

Moreover, previous research showed that VE users calibrated their physical reach even when different perturbations of perceptual feedback were presented. Ebrahimi et al. [13] investigated the effect of visual and proprioceptive information on near-field distance estimation. They showed that users were able to calibrate their physical reach in the presence of visual and proprioceptive feedback even with different levels of perturbation in visual information. This kind of calibration effect has been shown to carry over to either real-world [33, 51] or to another virtual scene when feedback is absent [13]. Different studies also suggest that calibration to perceptual information occurs relatively quickly when closed-loop interaction with the environment was available [1, 12]. In terms of distance estimation, the effect of closed-loop interactions and perceptual learning within an environment has been well studied. Researchers spent a considerable amount of work studying the visuomotor calibration through closed-loop interactions with the environment [28, 29]. Still, the auditory component has received substantially less scientific attention than the visual counterpart in virtual environments.

A few studies investigated the effect of audio, visual, and audio-visual distance perception in VE [18, 39]. Rebillant et al. [39] compared distance estimation in audio only, visual only, and audio-visual conditions. They found similar distance compression in the audio- and visual-only conditions. In a similar study, Finnegan et al., [18] investigated audio-visual feedback on reducing distance compression. They showed that incongruent audio-visual information could produce more accurate distance estimation. Other studies looked at different characteristics of sound to better understand it and explore factors that could reduce distance compression [21, 50]. Huang et al. [21] looked at the reverberation time in distance perception in VE. Their results indicate that users associate longer reverberation time with farther distances and shorter reverberation time with closer distances. They also found that the presence of reverberation time is more effective in near-field distance estimation. However, increasing reverberation time could introduce a mismatch between visual and auditory information that could influence distance estimation.

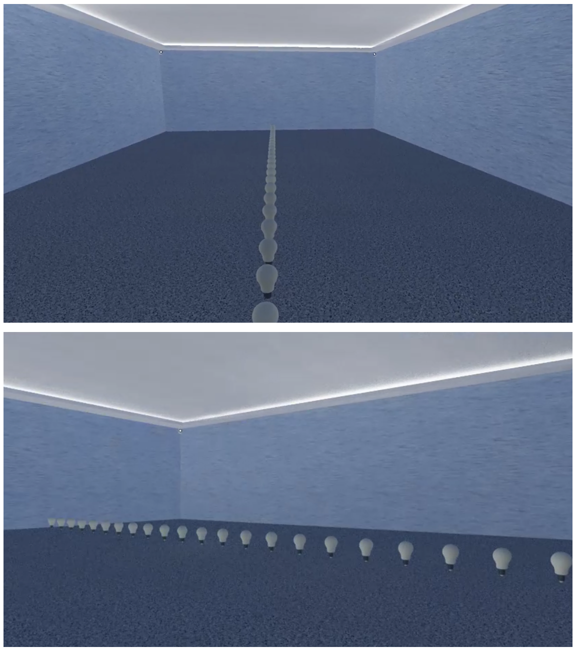


Figure 1: Virtual room in our experiment

Table 1: The audio stimuli used in the pretest, calibration, and three post-test phases. RT is the reverberation time of an audio stimulus whose source location is in front of the participants with *actual distance*.

Calibration Phase		Pretest / Post-tests	
Actual Distance	RT	Actual Distance	RT
3 meters	0.6 seconds	2 meters	0.4 seconds
5 meters	1.0 second	4 meters	0.8 seconds
7 meters	1.4 seconds	6 meters	1.2 seconds
9 meters	1.8 seconds	8 meters	1.6 seconds
11 meters	2.2 seconds	10 meters	2.0 seconds

3 System Description

Virtual Scene We built a virtual room (Fig. 1) using Unreal game engine (Version 4.23.1). The size of the room was set as 8.72m × 3.5m × 15m. The light sources and texture were set appropriately in order to let participants perceive the room and objects without difficulty. We also avoided using any textures with periodic patterns to prevent the participants from exploiting the patterns as additional visual information for distance estimation. There were 27 light bulbs lined on the floor in front of the participants. The nearest light-bulb was 0.5 meters away from the participant, while the farthest bulb stood 14 meters away. The interval between the light bulbs was set at 0.5 meter. The size of each light bulb is 0.17m × 0.17m × 0.29m.

Audio Stimuli The audio stimulus used in this study was a four-second horn anechoic recording [23]. We added reverberations with different RTs to the anechoic recording by using WAVES IR1 Convolution Reverb framework. In this study, we used the same anechoic clip with ten different reverberation time (from 0.2 seconds to 2.2 seconds with an increment of 0.2 seconds), and other audio properties (e.g., intensity) are invariant. To avoid that the participants simply memorize the stimulus for each actual distance, we separate these audio stimuli into two groups such that the actual distances (and corresponding RT) were different in the calibration and pretest/post-test phase (see Table 1). According to the reverberation time calculation equation in [6], the ideal reverberation time of the virtual room in our experiments should be 0.7 seconds.The RTs of our stimuli are rather exaggerated in an effort to study if participants can distinguish sounds more easily thus also amplifying the power of the experiment. This design decision was made because research into perceptual attunement to auditory reverberation information in egocentric depth estimation remains relatively unexplored. Furthermore, we were first interested in determining the feasibility of calibration to auditory information and the length of time to which these effects persist both immediately post-calibration, and in the long term in VR, thus making us use a rather large range of RTs in the experiment.

Equipment. In this study, we used the HTC Vive Pro HMD with its built-in noise-canceling earphones. The participants performed the selection tasks with an HTC Vive controller. The simulation program was run on a PC, running a Windows 10 operating system, with a 3.1GHz Intel i5 processor, 16GB of RAM, and an NVIDIA GeForce RTX 2060 graphics card. All the experiments were conducted in a controlled laboratory setting in an attempt to avoid any interference from environmental sounds.

4 Experiment Design

4.1 Research Questions

Aiming for understanding the calibration and long-term carryover effects of audio stimuli on egocentric depth perception in an IVE, we investigated the following research questions:

RQ1: Can people utilize perceptual learning (or calibration) to auditory reverberation to enhance egocentric distance estimation in IVEs? Although there have been several studies of perceptual learning on different sensory in IVEs [12, 13], perceptual learning on auditory distance perception remains an unexplored area in VR research. Therefore, the first research question we would like to investigate is whether distance perception could be calibrated based on auditory stimuli. In particular, due to the recent success of utilizing the reverberation time on compensating distance compression [21], we want to first verify whether the reverberation time could be further exploited as perceptual information for calibrating users' distance perception. Specifically, we designed a within-subject experiment including pretest, calibration, and post-test phases to observe participants' distance estimation performance over the calibration process.

RQ2: Can the calibration be carried over for a long period such as 1 month or 6 months? A followup question after RQ1 is then if users' auditory distance perception could be calibrated, can the learning effect of calibration be carried over for a long period of time? The study of RQ2 not only provides deeper understanding of perceptual learning of auditory distance estimation but also establishes a ground for practical applications. For instance, the users of a VR system can perform a calibration procedure once and without the need to re-calibrate for a long period. This could potentially benefit many VR applications that requires accurate distance perception. Therefore, to answer RQ2, we conducted additional post-test phases several months after participants completed their calibration experiment.

RQ3: How does people's performance on distance perception vary with different phases? In addition to RQ1 and RQ2, we want to further quantitatively evaluate the participants' performance and its variations at different phases. The hypothesis would be that the participants could estimate egocentric distance to the sound more accurately and quickly through calibration, but the calibration effect may gradually diminish with span over several months. It would be interesting to quantitatively measure and analyze the decay of calibration effect over time.

4.2 Participants

We first conducted an a priori power analysis to determine the appropriate sample size for our experiment with 5 within subjects sessions. Using G* Power, for an effect size of 0.25 [14], alpha error probability of 0.05, Power of 0.95, number of measurements per session of 25, and correlation among repeated measures of 0.5, we found that we needed a total of 12 participants. However, we ended up recruiting a total of eighteen participants (13 females) for this study. The study protocol was approved by the Research Ethics Center for Human Subject Protection of our institution. Twelve of them were college students aged from twenty-one to twenty-nine; the others were non-students aged around thirty. The inclusion criteria for all participants were the ability to hear from both ears, 20/20 visual acuity or corrected 20/20 vision using contact lenses, and participants' ability to fuse stereoscopic images. All participants had game experience and seven of them had some VR experiences (less than 5 hours) before this experiment. All participants were provided with informed consent and were free to opt out of the study at any time.

4.3 Tasks

In the pretest and post-test phases, the participants performed distance estimation without any feedback. In the calibration phase, the participants performed partially close-loop distance estimation. For each phase, a participant needed to perform 25 trials of the assigned task (5 distances × 5 times = 25, see Table 1). As we adopted the within-subject experimental design, the five reverberation stimuli in each phase were played in random order to prevent the learning effect. We describe the two distance estimation tasks below.

Open-loop distance estimation. A participant pressed the grip button of the controller to start a trial and an reverberated horn sound was played repeatedly. While the sound was playing, the participant estimated the egocentric distance to the sound location by pressing the forward button on the controller to light up a series of light-bulbs one by one from near to far until the one that they considered to be the sound location. The participant could adjust the estimated distance by pressing the forward (backward) button to turn on (off) the light bulbs. The participant pressed the trigger button to confirm their answer and finished the trail. To avoid the interference of cognition on distance perception, the participants were instructed to start estimating the distance as soon as hearing the sound and complete the task quickly.

Partially close-loop distance estimation. Similar to the open-loop distance estimation, the participants initiated a trial by pressing the grip button and used the forward/backward button to estimate the distance to the sound location while a reverberated horn sound was playing. The difference is that once the participants confirmed their answer using the trigger button, they immediately received feedback: two red arrows pointed to the light-bulb located at the target distance. Participants needed to readjust their estimation by lighting up a light-bulb back and forth using the forward and backward button until they reached the one indicated by the red arrows. The indicator arrows would turn green immediately once the participants lighted up the correct light-bulb (Fig. 2). The participants then can press the grip button for next trial.

4.4 Procedure

Upon arrival, all participants completed a standard informed consent form and a pretest survey which includes demographic information and game experience. For sanitary concern, participants were asked testing their hearing ability with personal earphones on Widex's hearing evaluation website [46]. Following that, their interpupillary distance (IPD) was measured using a digital pupil distance meter. The measured IPD was used to set the graphical inter-ocular distance for the experiment simulation, and the HMD was adjusted accordingly for each participant. Participants were instructed to stand on a fixed point for accurate positioning and their eye heights were used as parameters to ensure that their eye height in VR corresponded to their real world counterparts [17]. Experimenters then described the experimental instructions and tasks. Before starting the formal trials, the participants were allowed to practice for three trials to familiarize themselves with the VR scene and controllers.

There were three phases for participants to undergo: pretest, calibration and post-test. All phases were conducted in the same virtual scene: a series of light-bulbs placed on the ground, situated in front of the participant from near to far with 0.5 m apart. The size of the simulated room is 8.72m x 3.5m x 15m.

Each participant began with a baseline pretest phase of distance estimates without feedback. After a thirty-minute break, participants started the calibration phase in which the participant perceived distances with feedback. An immediate post-test phase was brought to evaluate the participant's distance perception after calibration.

After completing all three phases, participants proceeded to fill out the questionnaires for mental and physical demand evaluation and immersiveness evaluation in VR. Each participant was engaged in a short interview with the experimenter to discuss their experience in this study. Upon completion of the interview, participants were debriefed and financially compensated for taking part in the study and were free to leave. It took a participant up to 1 hour to complete the whole procedure.

The study involved participants being invited to attend a post-test session after 1 month and after 6 months, in order to examine to what extent calibration to auditory reverberation time carried over to sound source distance perception, which some of the participants were able to attend. There were 6 participants who did the follow up trials at the one month mark, and 5 participants that did the follow up trials after 6 months. These participants' HMDs and VR positioning were adjusted accordingly, and three practice trials were provided to each participant for recalling the experience of controller manipulation and the VR environment for their depth judgments. The participants performed the open-loop distance estimation task and then were interviewed briefly by the experimenter.

4.5 Measures

We measured the following items in the experiment.

Perceived distance (PD) is a participant's perceived egocentric distance to the location of sound source. That is, PD is a participant's result in a distance estimation task.

Judgement time (JT) is the duration between the time point that a participant starts to estimate the egocentric distance and the time point that the participant press the trigger button to confirm the perceived distance. Specifically, $JT = FT - IT$, where FT denoting *Full Time* is a timestamp at which a participant pressed the trigger button to confirm his/her estimated distance. IT denoting *Initial Time* is a timestamp at which the participant presses the forward (or backward) button at the first time.

Pretest questionnaire acquires participants' demographic information including gender, age, education level, game experience (game playing duration, game genres), and VR experience.

NASA Task Load Index (NASA-TLX) was used for evaluating each participant's perceived task load [20]. In this study, we rated our participants' mental and physical demands with this index.

Igroup Presence Questionnaire (IPQ) was used to evaluate how effective our experimental VR scene is [49].

Interviews. All participants were briefly interviewed after finishing the tasks. The follows are some examples of the questions: *"What is your strategy to perceive distances during the experiment?"*; *"Is this simulation against your daily experience in the real world?"*.

Besides, the eleven participants who took the second post-test session a few months later were asked some questions based on the comparison with their previous experiment experience. Example questions are: *"What is your strategy to perceive distances this time? Please rate your perception accuracy based on your performance*

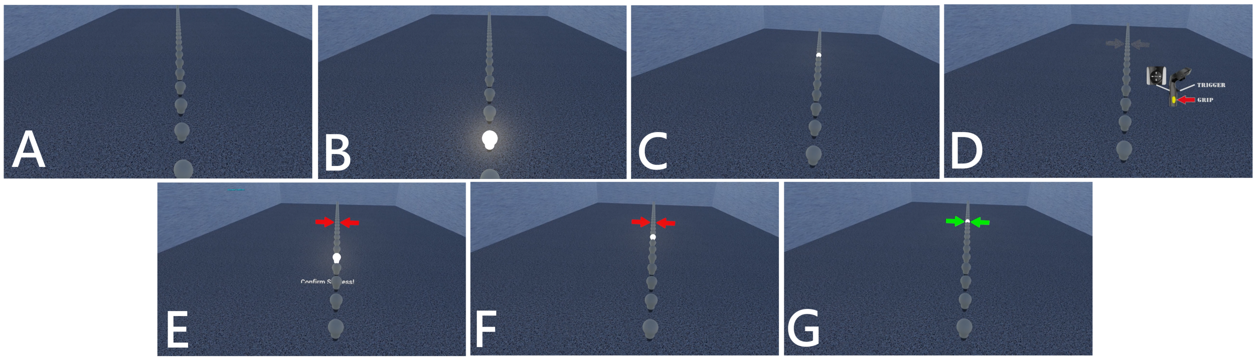


Figure 2: **Open-loop** task is as follows: A, B, C, D. **Partially closed-loop** is as follows: A, B, E, F, G, D. A: Audio stimuli onset. B: Estimate and enter the perceived distance by pressing the forward/backward button. C: Press the trigger button to confirm the answer. D: Press the grip button for next trial. E: Press the trigger to confirm the answer and receive immediate feedback (red arrows) F: Adjust the perceived distance to match the actual one (indicated by red arrows) by pressing the forward/backward button. G: Reach the actual distance and the indicator arrows turn green. **Please see also the supplemental video for demonstration of the experimental tasks.**

among these three phases: pretest, first post-test and second post-test."; *"How will you expect your performance if re-calibrating your perception before the second post-testing?"*.

5 Results

5.1 Multiple Regression Results

In order to conduct a multiple regression analysis and evaluate how phase of measurement of the depth perception (pre-test/calibration/post-test/post-test-at-1month/post-test-at-6months), actual distance, and phase-by-actual distance interaction affect the perceived distance to the targets, a multiple regression was calculated to predict perceived distance based on the independent variables. As is often done ahead of a multiple regression analysis, an analysis of standardized residuals was carried out on the data to identify any outliers, using which data that were beyond + or – 3.0 of the standardized residuals were removed. The final standardized residual minimum was -1.50 to +1.50. Tests to examine if the data met the assumptions of collinearity indicated that multicollinearity was not a concern (Phase, Tolerance = 1.0, VIF = 1.0; Actual Distance, Tolerance = 1.0, VIF = 1.0; Phase-by-Actual Distance, Tolerance = 1.0, VIF = 1.0). The data met the assumptions of independent errors (Durbin-Watson value = 1.47). The histogram of standardized residuals indicated that the data contained approximately normally distributed errors, as did the P-P plot of standardized residuals, which showed that the data were close to a linear regression profile. The scatter-plot of standardized residuals showed that the data met the assumptions of homogeneity of variance and linearity, as well as the data met the assumptions of non-zero variance.

A multiple regression was first conducted to examine if phase and actual distance predicted perceived distance. A significant regression equation was found F(2, 1508) = 1165.44, $p < 0.001$, with an R^2= 0.61. Participants' *Perceived Distance = - 0.12 + 0.77* $\times$ *Actual Distance + 0.37* $\times$ *Phase*; where *Actual Distance* was measured in meters, phase 1 was pretest, 2 was calibration, 3 was post-test, 4 was post-test after 1 month, and 5 was post-test after 6 months. *Perceived Distance* increased by 0.77 meters for every meter of increase in actual distance, perceived distance increased by 0.37 meters for a difference in one phase to the next (i.e., pretest to calibration or calibration to post-test). Both *Actual Distance* ($p < 0.001$) and *Phase* ($p < 0.001$) were significant predictors of *Perceived Distance*.

In order to evaluate the significant interaction effects, the continuous independent variable of actual distance was mean centered to eliminate any multicollinearity effects, and the interaction terms (mean centered) *Actual Distance* $\times$ *Phase* was added to the model in a hierarchical multiple regression. The regression model with the interaction variables was found to be significant, F(3, 1508) = 818.90, $p < 0.001$, with an R^2= 0.623 (with the change in R^2 of 0.013). When including the interaction term, participant's *Perceived Distance = 1.27 + 0.54* $\times$ *Actual Distance + 0.39* $\times$ *Phase – 0.094* $\times$ *(Aual Distance* $\times$ *Phase)*; where *Actual Distance* was measured in meters, phase 1 was pre-test, 2 was calibration, 3 was post-test, 4 was post-test after 1 month, and 5 was post-test after 6 months. *Perceived Distance* increased by 0.54 meters for every meter of increase in actual distance, perceived distance increased by 0.39 meters for a difference in one *Phase* to the next (i.e., pre-test to calibration or calibration to post-test), and perceived distance decreased by 0.094 for a unit of *Actual Distance* by *Phase* interaction. All three predictors, *Actual Distance* ($p < 0.001$), and *Actual Distance* by *Phase* interaction term ($p < 0.001$) were significant predictors of *Perceived Distance*.

By phase, the linear regression equation for the pre-test phase (R^2=0.30) is *Perceived Distance = 1.41 + 0.49* $\times$ *Actual Distance*, the linear regression equation for the initial judgments in the calibration phase is (R^2=0.64) is *Perceived Distance = 0.81 + 0.79* $\times$ *Actual Distance*, the linear regression equation for the post-test phase is (R^2=0.81) is *Perceived Distance = 0.49 + 0.93* $\times$ *Actual Distance*, the linear regression equation for the post-test phase after 1 month is (R^2=0.70) is *Perceived Distance = 0.69 + 0.85* $\times$ *Actual Distance*, and the linear regression equation for the post-test phase after 6 month is (R^2=0.80) *Perceived Distance = 0.70 + 0.80* $\times$ *Actual Distance* (see Fig. 3).

5.2 Depth Perception Accuracy

We computed the accuracy of the participants via the following formula, as used in other relevant studies such as [12].

$$\%Accuracy = \frac{Perceived\ Distance - Actual\ Distance}{Actual\ Distance} \times 100 \quad (1)$$

This equation created a signed error of % accuracy, where negative values represent % underestimation of the actual target distance and positive values represent % overestimation of the actual target distance. Mean % accuracy scores were calculated for each participant for the different phases of the experiment (pre-test, calibration, post-test, post-test after 1 month and post-test after 6 months), and by sets of trial (1^{st} set = trials 1 through 5, 2^{nd} set = trials 6 through

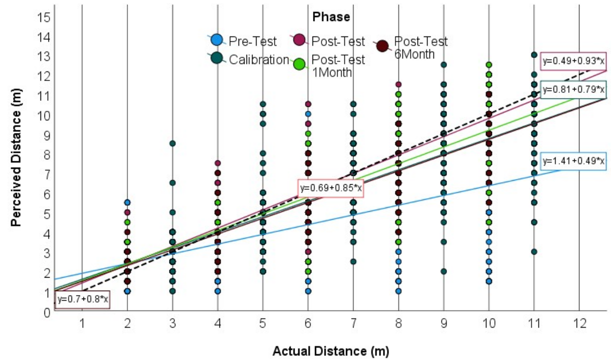


Figure 3: Linear regression models showing the pre-test, calibration, post-test, post-test after 1 month, and post-test after 6 month data on the perceived distance of participants to egocentric depth perception of sound source distance using auditory reverberation in VR. Coefficient of determination: Pre-test R^2=0.30, Calibration R^2=0.64, Post-Test R^2=0.81, Post-Test 1Month R^2=0.70, Post-Test 6Month R^2=0.80.

10 and so on). We aimed to evaluate if the % accuracy differed by the phases of the experiment and over time in each phase when analyzed by sets of trials.

Prior to conducting the parametric repeated measures ANOVA analysis on the accuracy scores, we carefully verified that the underlying assumptions of the test were met. Namely, the data in the samples were normally distributed and error variance between accuracy scores in the different phases were equivalent. We insured that Box's test of equality of covariance matrix was not significant. Levene's test was conducted to verify homogeneity of variance, and Mauchly's test of sphericity was conducted to ensure that error variance between accuracy scores in the different phases of the experiment was equivalent. Pairwise post-hoc tests were conducted using Bonferroni adjusted alpha method.

The mean % accuracy scores were subjected to a 5 (phase) x 5 (trial set) Repeated Measures ANOVA analysis. Phase was a within subjects variable consisting of pre-test, calibration, post-test, post-test after 1 month and post-test after 6 months. Trial set was a variable that examined the effect of experiment duration via sets of trials on % accuracy scores. For trial sets, accuracy scores across 25 trials were averaged into sets of 5 trials each – set 1 (trials 1 through 5), 2 (trials 6 through 10), 3 (trials 11 through 15) etc. The ANOVA analysis revealed a significant main effect of phase F(4, 450) = 15.94, p <0.001, p.η^2 = 0.24. Post-hoc pairwise comparisons using Bonferroni method revealed that participants' mean % accuracy scores were significantly lower in the pre-test phase (M=−18.67%, SD=28.35) as compared to their initial judgment in the calibration phase (M=−7.13%, SD=11.55) p = 0.002, post-test phase (M=4.88%, SD=16.15) $p < 0.001$, post-test after 1 month (M=2.03%, SD=18.96) $p < 0.001$, and post-test after 6 months (M=−3.0%, SD=16.46) p = 0.007. Mean % accuracy in the post-test phase was the highest, and it was also significantly higher than pre-test as well as the initial judgments in the calibration phase $p < 0.001$ (See Fig. 4). No other main or interaction effects were found.

5.3 Judgment Time

We conducted an analysis on the participants' judgment time in seconds, which was computed as the time from the presentation of the stimuli to the participants' judgment via a button press on the controller, as a measure of the participants reaction time to the auditory stimuli of target location. The judgment data was carefully

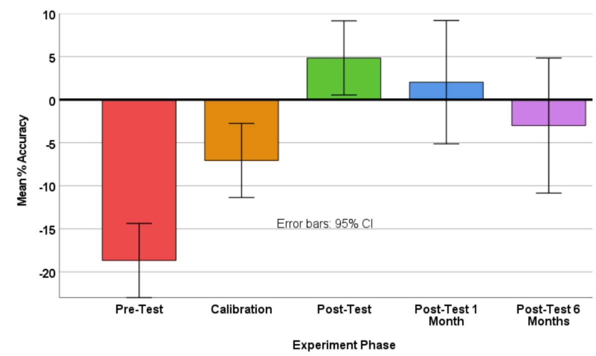


Figure 4: Mean % accuracy scores between different phases of the experiment.

verified to insure that the underlying assumption of a parametric analysis was met. The data was normally distributed and variance was homogeneous among groups of judgement scores between the experiment phases. The mean judgment time between experiment phases were subjected to a one-way repeated measures ANOVA analysis. Post-hoc comparisons were conducted using the Bonferroni method.

A one-way repeated measures ANOVA analysis on the judgment time was significant, F(4, 90) = 5.73, $p < 0.001$, p.η^2 = 0.21. Post-hoc comparisons using the Bonferroni method revealed that the mean judgment time in the post-test phase (M=5.62s, SD=2.94) was significantly lower than the pre-test phase (M=6.65s, SD=3.94) $p < 0.001$. Post-hoc comparisons using the Bonferroni method also revealed that the mean judgment time in the post-test phase after 6 months phase (M=5.6s, SD=3.37) was also significantly lower than the pre-test phase (M=6.65s, SD=3.94) p = 0.020 (See Fig. 5).

5.4 NASA Task Load Index

The NASA-TLX measures participants' self-ratings of a task on six separate factors contributing to their perceived workloads. Ratings are on a scale from 0 to 100 and are multiples of 10 (e.g., 10, 20, 70). Participants rate the relative importance of each factor in comparison with the other five factors in a pair-wise comparison procedure (total number of comparisons = 15). Participants' weighted ratings of these six factors are as follows: Performance(0.26) = Effort(0.26) > Frustration(0.2) = Mental(0.2) > Temporal(0). The highest rated factor is Performance (M=63, SD=22.53) as shown in Fig. 6. Based on qualitative assessment analysis, we can see the Performance workload overpowers other workloads. This might be related to some participants' feedback that after tasks in the calibration phase, the participants were more confident and willing to estimate/perceive distance because of confirming their strategy of distance estimation. So they were more concentrated on the task such that the performance workload overpowered other factors.

5.5 Igroup Presence Questionnaire

IPQ measures the sense of presence experienced in VR. It rates a user's awareness which is relied on interaction between sensory stimulation, environmental factors and internal tendencies to become involved in a VR world. The questionnaire includes fourteen questions and can be divided into three dimensions: spatial presence, involvement, and reality. The analysis of the IPQ resulted in the following statistics: Spatial presence (M=4.6, SD= 0.47); involvement (M=3.72, SD=0.74); experienced realism (M=3.76, SD=0.88); overall presence (M=4.12, SD=0.76). This indicated that the participants has good sense of presence in the experiment, particularly the spatial presence.

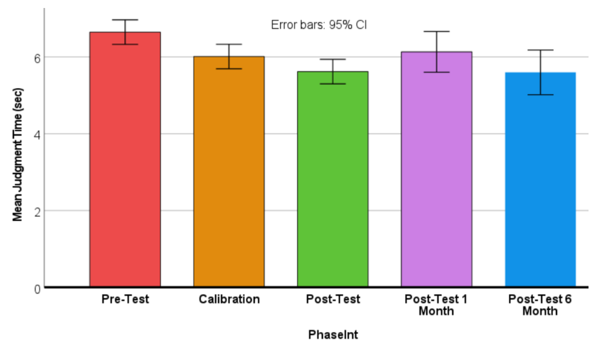
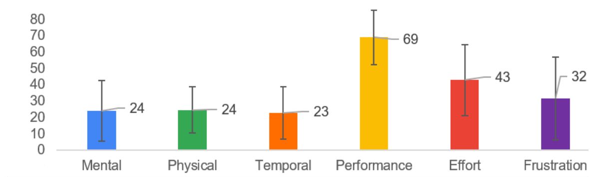


Figure 5: Mean judgment time in seconds between different phases of the experiment.



	Mental	Physical	Temporal	Performance	Effort	Frustration
Mean	24	24	23	69	43	32
SD	18.515	14.234	16.017	16.764	21.640	25.263

Figure 6: NASA-TLX Workload evaluation. The error bars show the SD.

5.6 Interviews Results

In the overall interviews, 16 of the 18 participants described their strategies to perceive distance were reverberated sound stimuli, they used some phrases to explain their ideas, such as *"a continuous sound reflection"*, or *"the variation level of intensity after the first sound came to their ears"*, (which meets the definition of reverberation, the energy declined after the first direct sound). 13 of the 18 participants agreed that the simulation experience in this study corresponded to their daily life experience that a longer reverberated sound might refer to a longer distance in a larger space.

The participants who took part in the second post-test session were further interviewed with questions related to the calibration effect. All of these participants ($N = 11$) considered their performance in the pretest the worst and agreed that the calibration phase improves their accuracy on auditory egocentric distance estimation in this study. Furthermore, ten of the eleven participants rated their performance in the three phases from good to poor as follows: the first post-test, the second post-test (one month or six months later), then the pretest. These participants all mentioned that they felt more confident with their distance perception in the first post-test. One participant considered the second post-test performance is better than the first one, because he thought himself to be more involved in this VR environment.

6 Discussion

The statistical analysis conducted using regression models revealed that participants were able to calibrate their perception of depth to the reverberation time component of auditory stimuli, producing more accurate estimates of egocentric distances after enduring the calibration phase. This finding can be inferred from Fig. 3 which depicts the predicted perceived distance as the interaction between the *Actual Distance* and *Phase* in the linear regression model. The ideal prediction is represented by the dotted line which characterizes the veridical judgement, the slope of which equals one. In other words, the veridical judgement line represents the prediction where participants' perceived distance is equal to the actual distance of the presented stimuli. Thus a regression line that is closer to the veridical judgement line (dotted line) implies that the perceived egocentric distances are more accurate. Along these lines, it is evident that the slope of the pretest phase, in which participants estimate the distances without any feedback, is much lower than the other phases (see Fig. 3). Furthermore, the regression models also show that upon experiencing a calibration phase, participants' performances improved with increases in the slope and decreases in the intercepts of the post test phases as compared to the pretest phase. Additionally, the accuracy scores of the post-test phases are higher than the pre-test phase where participants largely underestimated distances. This trend can be observed in Fig. 4 which illustrates the mean percentage accuracy scores for the different phases. Overall, these results are indicative of the success in employing a calibration phase towards improving participants' egocentric distance estimates in immersive virtual reality experiences. These findings seem to be in line with other works that have shown that users can successfully calibrate their perception of depth based on visual feedback provided during a calibration phase [1, 7, 12]. It hence appears that users can successfully calibrate their perception of depth based on the reverberation time component of an auditory signal in addition to visual information which has been investigated in previous research.

The regression profiles and the analyses conducted make a case for the persistence of calibration effects after prolonged periods of time. This can be seen in Fig. 4, which depicts the mean percentage signed accuracy for the different phases over the trials. Perfect prediction would involve zero percentage accuracy scores with negative signed accuracy scores representing underestimation and positive representing overestimation respectively. As can be seen from the figure, participants largely underestimated distances in the pre-test phase while their post-test estimates reflected significantly higher accuracy scores even at the one-month and six-month marks. Furthermore, we can infer that the effect of calibration can carry over for an extended period of time, with the post-tests at both the one-month and six-month marks being closer to the veridical judgement line (see Fig. 3). On the flip side however, it must be noted that the performance tends to degrade over time, with the accuracy at the six-month mark being lower than the accuracy at the one-month mark (Fig. 4). After six months, participants' started to underestimate distances again, but this was still more accurate than the pre-test phase, suggesting a persistence of calibration even up to about six months after. More investigations are needed to thoroughly examine the extent of time after which the effects of calibration start to wear off.

From the analysis conducted on the mean judgement times taken for each phase of the study, we found that users took less time after encountering the calibration phase. When compared to the pretest phase, the post-tests' mean judgement times were lesser during the immediate post-test, and during the post-tests at both the one month, and six month marks. This seems to be suggestive of increased confidence in making depth judgements after enduring a calibration phase, with some participants explicitly commenting about the same.

Our work serves to demonstrate that egocentric depth estimation can be calibrated using the reverberation time component of an auditory sound signal. Users in our study were able to differentiate various lengths of reverberation times in VR, enhancing egocentric distance estimation by utilizing perceptual learning (or calibration) to auditory reverberation. This learning effect seems to persist for an extended period of time, our study suggesting at least six months. Based on this finding, it might be hence be useful for VR system designers to employ auditory reverberation calibration as a means to improve users' egocentric depth estimates. Furthermore, our findings with respect to the persistence of calibration partially relieves us of the need to periodically train users in simulations that require a highly accurate perception of depth including those like training, medical surgery, etc. It remains to be seen as to when exactly the effects of calibration no longer persist and wear off completely.

More investigations may be needed to determine this time period. In contrast to the work conducted by the authors of [21] who leveraged relative depth judgements to address distance compression in VR, we used absolute depth judgements in determining the ability of users to calibrate to reverberation information in audio signals. From the results obtained in our study, there seems to be potential for the use of calibration to auditory reverberations towards addressing distance compression effects, a problem that continues to plague immersive virtual reality applications.

6.1 Limitations and Future work

For ease of perceptual learning, we used rather long RTs that are exaggerated. This could cause unrealistic spatial perception to the participants as they felt the reverberation sounds like generated in a much larger room than the VR scene they experienced. We could use a more realistic RT in the future and an interesting question then is whether the same calibration effect could still be achieved when correct RT's are used. Besides, if a minimal detectable difference (MDD) of RT can be measured, the calibration procedure should choose the increment of actual distance according to MDD. This could avoid the unrealistic spatial perception due to exaggerated RTs.

Our preliminary study demonstrated that auditory egocentric depth estimation could be made more accurate via attunement and calibration to reverberation, and this effect can be carried over for at least 6 months. The findings are encouraging and could shed light on perceptual learning of distance perception. In particular, we only focused on the auditory perceptual channel in the study. It would be interesting to explore whether and how other sensory information (e.g., visual, haptic) can be used to calibrate users' distance estimation and how the calibration effect would vary longitudinally over time.

7 Conclusion

In this study, we empirically evaluated the feasibility of auditory calibration or auditory perceptual learning with reverberation stimuli to enhance sound source depth perception in IVEs. We find that participants typically underestimate the sound source distance in VR, in a manner similar to distance estimation using visual perception. We investigated whether users' sound source perception can be calibrated via auditory reverberation and whether the calibrated effect can be carried over and persist for a longer duration 1 month later or 6 months later.

Past research have shown the users' depth perception in VR can be improved by reducing distance compression with audio, visual, or audio-visual cues. We sought to examine if Reverberation Time (RT) can be calibrated as a stimulus in perceptual learning. For this purpose, we conducted an experiment requiring participants to perceive distance in different phases that we can observe the calibration effect with varying RTs. We found that auditory reverberation time to distance perception can be calibrated and the effects of calibration seems to persist over time. Our work suggests that users of VR simulations can attune or calibrate to auditory reverberation time information in auditory target stimuli to estimate the distance of these targets in VR. This has implications to users' interactions in multimodal VR simulations for games and education. In such simulations, the use of a calibration phase allows users to accurately perceive the distance of the sound source stimuli, and after the initial experience, users may still perceive accurate depth in such simulations overtime without requiring further recalibration.

In future work, we strive to examine how auditory reverberation in combination with other types of stimuli, such as haptic and visual stimuli, in a multi-modal perception scenario can potentially enhance depth perception in VR. We also aim to examine how congruent and divergent visuo-auditory reverberation stimuli affects sound source depth perception in VR experiences.

Acknowledgments

This material is based upon work partially supported by the Taiwan Ministry of Science and Technology under Grant 109-2221-E-009-123-MY3 and the US National Science Foundation under Grant 2007435. We also gratefully acknowledge the participants of our empirical evaluations.

References

[1] B. M. Altenhoff, P. E. Napieralski, L. O. Long, J. W. Bertrand, C. C. Pagano, S. V. Babu, and T. A. Davis. Effects of calibration to visual and haptic feedback on near-field depth perception in an immersive virtual environment. In *Proceedings of the ACM symposium on applied perception*, pp. 71–78, 2012.

[2] B. M. Altenhoff, C. C. Pagano, I. Kil, and T. C. Burg. Learning to perceive haptic distance-to-break in the presence of friction. *Journal of Experimental Psychology: Human Perception and Performance*, 43(2):231, 2017.

[3] C. F. Altmann, K. Ono, A. Callan, M. Matsuhashi, T. Mima, and H. Fukuyama. Environmental reverberation affects processing of sound intensity in right temporal cortex. *European Journal of Neuroscience*, 38(8):3210–3220, 2013.

[4] P. W. Anderson and P. Zahorik. Auditory/visual distance estimation: accuracy and variability. *Frontiers in Psychology*, 5:1097:1–1097:11, Oct. 2014.

[5] W. Bailey and B. M. Fazenda. The effect of reverberation and audio spatialization on egocentric distance estimation of objects in stereoscopic virtual reality. *The Journal of the Acoustical Society of America*, 141(5):3510–3510, 2017.

[6] L. L. Beranek and T. Mellow. Acoustics: sound fields and transducers. In *Academic Press*, p. 473, 2012.

[7] G. P. Bingham and C. C. Pagano. The necessity of a perception-action approach to definite distance perception: monocular distance perception to guide reaching. *Journal of Experimental Psychology: Human Perception and Performance*, 24:45–68, 1998.

[8] L. Buck, R. Paris, and B. Bodenheimer. Distance compression in the htc vive pro: A quick revisitation of resolution. *Frontiers in Virtual Reality*, p. 157.

[9] L. E. Buck, M. K. Young, and B. Bodenheimer. A comparison of distance estimation in HMD-based virtual environments with different HMD-based conditions. *ACM Trans. Appl. Percept.*, 15(3):21:1–21:15, July 2018.

[10] A. G. de Siqueira, R. Venkatakrishnan, R. Venkatakrishnan, A. Bharqava, K. Lucaites, H. Solini, M. Nasiri, A. Robb, C. Pagano, B. Ullmer, et al. Empirically evaluating the effects of perceptual information channels on the size perception of tangibles in near-field virtual reality. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*, pp. 1–10. IEEE, 2021.

[11] M. Dicks, J. van der Kamp, R. Withagen, and J. Koedijker. Can we hasten expertise by video simulations? Considerations from an ecological psychology perspective. *International Journal of Sport Psychology*, 45(6), 2015.

[12] E. Ebrahimi, B. Altenhoff, L. Hartman, J. A. Jones, S. V. Babu, C. C. Pagano, and T. A. Davis. Effects of visual and proprioceptive information in visuo-motor calibration during a closed-loop physical reach task in immersive virtual environments. In *Proceedings of the ACM Symposium on Applied Perception*, pp. 103–110, 2014.

[13] E. Ebrahimi, B. M. Altenhoff, C. C. Pagano, and S. V. Babu. Carryover effects of calibration to visual and proprioceptive information on near field distance judgments in 3D user interaction. In *2015 IEEE Symposium on 3D User Interfaces (3DUI)*, pp. 97–104. IEEE, 2015.

[14] E. Ebrahimi, L. S. Hartman, A. Robb, C. C. Pagano, and S. V. Babu. Investigating the effects of anthropomorphic fidelity of self-avatars on near field depth perception in immersive virtual environments. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 1–8. IEEE, 2018.

[15] F. El Jamiy and R. Marsh. Distance estimation in virtual reality and augmented reality: A survey. In *2019 IEEE International Conference on Electro Information Technology (EIT)*, pp. 063–068. IEEE, 2019.

[16] P. Etchemendy, E. Abregú, E. Calcagno, M. Eguía, N. Vechiatti, F. Iasi, and R. Vergara. Auditory environmental context affects visual distance perception. *Scientific Reports*, 7, 2017.
[17] I. T. Feldstein, F. M. Kölsch, and R. Konrad. Egocentric distance perception: A comparative study investigating differences between real and virtual environments. *Perception*, 49(9):940–967, 2020.
[18] D. J. Finnegan, E. O'Neill, and M. J. Proulx. Compensating for distance compression in audiovisual virtual environments using incongruence. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pp. 200–212, 2016.
[19] M. Gonzalez-Franco, P. Abtahi, and A. Steed. Individual differences in embodied distance estimation in virtual reality. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 941–943. IEEE, 2019.
[20] S. G. Hart and L. E. Staveland. Development of nasa-tlx (task load index): Results of empirical and theoretical research. In P. A. Hancock and N. Meshkati, eds., *Human Mental Workload*, vol. 52 of *Advances in Psychology*, pp. 139–183. North-Holland, 1988. doi: 10.1016/S0166-4115(08)62386-9
[21] Y.-H. Huang, R. Venkatakrishnan, R. Venkatakrishnan, S. V. Babu, and W.-C. Lin. Using audio reverberation to compensate distance compression in virtual reality. In *ACM Symposium on Applied Perception 2021*, pp. 1–10, 2021.
[22] F. Iazzetta, F. Kon, M. G. de Queiroz, F. Soares, C. da Silva, and M. de Avelar Gomes. AcMus: Computational tools for measurement, analysis and simulation of room acoustics. In *Proceedings of the European Acoustics Symposium*, 2004.
[23] P. J, P. V, and L. T. Anechoic recording system for symphony orchestra. vol. 94, pp. 856–865. S. Hirzel Verlag, 2008.
[24] J. A. Jones, D. M. Krum, and M. T. Bolas. Vertical field-of-view extension and walking characteristics in head-worn virtual environments. *ACM Transactions on Applied Perception (TAP)*, 14(2):1–17, 2016.
[25] N. Kaplanis, S. Bech, S. H. Jensen, and T. van Waterschoot. Perception of reverberation in small rooms: A literature study. In *Audio Engineering Society Conference: 55th International Conference: Spatial Audio*, 2014.
[26] J. W. Kelly, L. A. Cherep, and Z. D. Siegel. Perceived space in the HTC vive. *ACM Transactions on Applied Perception (TAP)*, 15(1):1–16, 2017.
[27] A. J. Kolarik, B. C. Moore, P. Zahorik, S. Cirstea, and S. Pardhan. Auditory distance perception in humans: a review of cues, development, neuronal bases, and effects of sensory loss. *Attention, Perception, & Psychophysics*, 78(2):373–395, 2016.
[28] B. R. Kunz, S. H. Creem-Regehr, and W. B. Thompson. Does perceptual-motor calibration generalize across two different forms of locomotion? investigations of walking and wheelchairs. *PloS one*, 8(2):e54446, 2013.
[29] M. Leyrer, S. A. Linkenauger, H. H. Bülthoff, U. Kloos, and B. Mohler. The influence of eye height and avatars on egocentric distance estimates in immersive virtual environments. In *Proceedings of the ACM SIGGRAPH symposium on applied perception in graphics and visualization*, pp. 67–74, 2011.
[30] S. A. Linkenauger, J. K. Witt, J. K. Stefanucci, J. Z. Bakdash, and D. R. Proffitt. The effects of handedness and reachability on perceived distance. *Journal of Experimental Psychology: Human Perception and Performance*, 35(6):1649, 2009.
[31] D. H. Mershon and L. E. King. Intensity and reverberation as factors in the auditory perception of egocentric distance. *Perception & Psychophysics*, 18(6):409–415, Nov. 1975.
[32] B. J. Mohler, H. H. Bülthoff, W. B. Thompson, and S. H. Creem-Regehr. A full-body avatar improves egocentric distance judgments in an immersive virtual environment. In *Proceedings of the 5th Symposium on Applied Perception in Graphics and Visualization*, p. 194. Association for Computing Machinery, New York, NY, USA, 2008.
[33] B. J. Mohler, S. H. Creem-Regehr, and W. B. Thompson. The influence of feedback on egocentric distance judgments in real and virtual environments. In *Proceedings of the 3rd symposium on Applied perception in graphics and visualization*, pp. 9–14, 2006.
[34] A. Peer and K. Ponto. Evaluating perceived distance measures in room-scale spaces using consumer-grade head mounted displays. In *2017 IEEE symposium on 3d user interfaces (3dui)*, pp. 83–86. IEEE, 2017.
[35] E. Peillard, T. Thebaud, J.-M. Normand, F. Argelaguet, G. Moreau, and A. Lécuyer. Virtual objects look farther on the sides: The anisotropy of distance perception in virtual reality. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 227–236. IEEE, 2019.
[36] L. Phillips, B. Ries, M. Kaeding, and V. Interrante. Avatar self-embodiment enhances distance perception accuracy in non-photorealistic immersive virtual environments. In *2010 IEEE virtual reality conference (VR)*, pp. 115–1148. IEEE, 2010.
[37] K. Ponto, M. Gleicher, R. G. Radwin, and H. J. Shin. Perceptual calibration for immersive display environments. *IEEE Transactions on Visualization and Computer Graphics*, 19(4):691–700, 2013.
[38] D. R. Proffitt, J. Stefanucci, T. Banton, and W. Epstein. The role of effort in perceiving distance. *Psychological Science*, 14(2):106–112, 2003.
[39] M. Rébillat, X. Boutillon, É. Corteel, and B. F. Katz. Audio, visual, and audio-visual egocentric distance perception by moving subjects in virtual environments. *ACM Transactions on Applied Perception (TAP)*, 9(4):1–17, 2012.
[40] R. S. Renner, B. M. Velichkovsky, and J. R. Helmert. The perception of egocentric distances in virtual environments - a review. *ACM Comput. Surv.*, 46(2):23:1–23:40, Dec. 2013.
[41] A. R. Richardson and D. Waller. The effect of feedback training on distance estimation in virtual environments. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, 19(8):1089–1108, 2005.
[42] E. J. Rinalducci. Characteristics of visual fidelity in the virtual environment. *Presence: Teleoperators & Virtual Environments*, 5(3):330–345, 1996.
[43] J. P. Rolland, W. Gibson, and D. Ariely. Towards quantifying depth and size perception in virtual environments. *Presence: Teleoperators & Virtual Environments*, 4(1):24–49, 1995.
[44] A. Turner, J. Berry, and N. Holliman. Can the perception of depth in stereoscopic images be influenced by 3D sound? *Proc.SPIE*, 7863:1–10, 2011.
[45] M. Wexler and J. J. Van Boxtel. Depth perception by the active observer. *Trends in cognitive sciences*, 9(9):431–438, 2005.
[46] Widex. Widex hearing test, 2021.
[47] P. Willemsen, M. B. Colton, S. H. Creem-Regehr, and W. B. Thompson. The effects of head-mounted display mechanical properties and field of view on distance judgments in virtual environments. *ACM Transactions on Applied Perception (TAP)*, 6(2):1–14, 2009.
[48] R. Withagen and C. F. Michaels. The role of feedback information for calibration and attunement in perceiving length by dynamic touch. *Journal of Experimental Psychology: Human Perception and Performance*, 31(6):1379, 2005.
[49] B. G. Witmer and M. J. Singer. Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence: Teleoperators and Virtual Environments*, 7(3):225–240, 06 1998. doi: 10.1162/105474698565686
[50] P. Zahorik, D. S. Brungart, and A. W. Bronkhorst. Auditory distance perception in humans: A summary of past and present research. *ACTA Acustica united with Acustica*, 91(3):409–420, 2005.
[51] J. Zhang, X. Yang, Z. Jin, and L. Li. Distance estimation in virtual reality is affected by both the virtual and the real-world environments. *i-Perception*, 12(3):20416695211023956, 2021.