Improving Effective Throughput Performance using Auditory Feedback in Virtual Reality

Anil Ufuk Batmaz
Concordia University
Montreal, Quebec, Canada

Kangyou Yu
Xi'an Jiaotong-Liverpool University
Suzhou, China

Hai-Ning Liang
Xi'an Jiaotong-Liverpool University
Suzhou, China

Wolfgang Stuerzlinger
Simon Fraser University
Vancouver, British Columbia, Canada

ABSTRACT

During the complex process of motor skill acquisition, novices might focus on different criteria, such as speed or accuracy, in their training. Previous research on virtual reality (VR) has shown that effective throughput could also be used as an assessment criterion. Effective throughput combines speed, accuracy, and precision into one measure, and can be influenced by auditory feedback. This paper investigates through a user study how to improve participants' effective throughput performance using auditory feedback. In the study, we mapped the speed and accuracy of the participants to the pitch of the auditory error feedback in an ISO 9241-411 multi-directional pointing task and evaluated participants' performance. The results showed it is possible to regulate the time or accuracy performance of the participants and thus the effective throughput. Based on the findings of our work, we also identify that effective throughput is an appropriate assessment criterion for VR systems. We hope that our results can be used for VR applications.

ACM Reference Format:
Anil Ufuk Batmaz, Kangyou Yu, Hai-Ning Liang, and Wolfgang Stuerzlinger. 2022. Improving Effective Throughput Performance using Auditory Feedback in Virtual Reality. In *Symposium on Spatial User Interaction (SUI '22), December 1–2, 2022, Online, CA, USA.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3565970.3567702

1 INTRODUCTION

Virtual Reality (VR) systems using Head-Mounted Displays (HMDs) enable a user to immerse themselves into a virtual environment (VE), where the virtual content has been set up in advance by a designer [40]. Such VR systems also permit users to interact with virtual objects in the VEs. Since immersive VEs afford easily adjustable scenes, including changing the illumination, the wind direction, or audio feedback, they are nowadays frequently used as training systems. Moreover, VR systems typically include input devices designed to support data collection around user interactions and thus support user performance assessment. Work on other types of training systems has relied on the use of external measurement tools, such as external inertia measurement units, cameras used to record the position of the users' hands, and image processing algorithms to detect the trajectory of hand motions. In contrast, to collect user data in a VR-based ball play training simulator, it is sufficient to use the sensors embedded in VR controllers to collect the required data. Moreover, VR systems can display a realistic training scenario that can be repeated in safe environments, such as in the trainee's home, as often as the training requires, without exposing trainees to potentially risky conditions, such as hazardous weather.

VR training systems have previously been studied in different fields, including surgery [15, 39, 48], hockey [51], skiing [46], and maintenance [22, 61]. Most of these fields require strong eye-hand coordination. Today, eye-hand coordination training systems and the associated trainee performance assessment are frequently used in such fields to improve the trainees' reaction time, which is correlated to attentional and cognitive functions. Some of these eye-hand coordination systems, such as Nike SPARQ [45] or Batak [52], are used to increase the trainees' motor performance by decreasing their reaction time while maintaining reasonable accuracy. Actions requiring fast and accurate hand movements, such as hitting a ball in volleyball, reaching the puck in hockey, putting the tool-tip of a surgical instrument into the correct location, or reaching a valve in firefighter training, are additional areas where eye-hand coordination training has been shown to be useful to increase performance. Such training can be also used for medical purposes, such as for stroke patients [58].

Recent work investigated eye-hand coordination training systems with VR HMDs to analyze human performance through Fitts' law [21]. Results showed that VR eye-hand coordination training systems have great potential as an alternative to real-world training approaches and motor performance in such systems conforms to Fitts' law [8]. In one such VR-based eye hand coordination training system [10], the authors changed the pitch of the auditory feedback when the participants made an error. Their results revealed that it is possible to affect participants' performance in terms of time, error rate, and/or throughput by changing the frequency of the sound feedback. However, they did not investigate how auditory feedback can be used to help users focus on a specific assessment criterion, such as accuracy.

Here, we use the term error as in previous Fitts' law studies, i.e., when the participants miss the target like selecting a location outside of the target area [21]. Such errors affect users' precision

---

SUI '22, December 1–2, 2022, Online, CA, USA

ACM ISBN 978-1-4503-9948-7/22/12…$15.00
https://doi.org/10.1145/3565970.3567702

and overall performance. In this paper, we focus not only on time or precision but also on increasing participants' effective throughput with auditory feedback.

Previous work (e.g., [4, 7, 10]) has shown that receiving auditory feedback with different pitches can affect user performance in eye-hand coordination training and Fitts' tasks in VR HMDs. For example, when users made an error in a Fitts' task, Batmaz and Stuerzlinger [10] demonstrated the effect of different frequencies with computer-generated sounds as auditory feedback on user performance. Most of this previous work has focused on the effect of discrete sound for auditory feedback but the use of continuous sound auditory feedback in VR training systems has received little attention. Moreover, there is an inherent trade-off between speed and accuracy in a Fitts' task. Although previous research has shown that training efficiency is higher when novices' focus is on precision [4, 5], in most of this previous work participants often were not instructed if they should focus on their speed or accuracy in a task. In addition, they may have used different execution strategies to complete a task, depending on how participants interpreted the instructions given to them, which feedback they received, or how their performance was measured [5].

This work extends previous research on the effect that pitch has on auditory feedback [7, 10] through a user study to investigate two different auditory feedback methods, discrete and continuous, and understand their impact on the user motor performance in an eye-hand coordination training task. Our study mapped participants' speed and accuracy to the pitch of the auditory error feedback. The results showed that participants could be encouraged to focus on different task execution strategies, such as a focus on accuracy or speed, through auditory feedback and that it is possible to alter user performance by mapping sound frequency to the desired assessment criterion.

In short, we present the following contributions:

- The design of appropriate mappings of continuous auditory feedback and their experimental evaluation through a study with the ISO pointing task to identify how such feedback can help users to focus on different task execution strategies.
- Suggestions for the design of continuous auditory feedback to help trainees focus on different task execution strategies in VR training systems.

2 RELATED WORK

In this section, we discuss previous work on Fitts' law, auditory feedback, and training strategies.

2.1 Fitts' law

Fitts' law is a mathematical model for rapid, aimed human movements [21] in Human-Computer Interaction research. MacKenzie's formulation of Fitts' law is shown in Equation 1 [35]:

$$MovementTime\,(MT) = a + b * \log_2\left(\frac{A}{W} + 1\right) = a + b * ID \quad (1)$$

In Equation 1, A and W are the target distance and target size, respectively. The *log* term in the equation represents the task difficulty, or the *index of difficulty*, *ID*. The coefficients a and b are empirically derived via linear regression. We also use the ISO 9241-411:2012 [27] effective throughput as assessment criterion:

$$Throughput = \left(\frac{ID_e}{MovementTime}\right) \quad (2)$$

In Equation 2, movement time is the task execution time for each trial and ID_e is the effective index of difficulty, with accounts for the effect of the combination of user accuracy and precision in ISO pointing tasks [27]:

$$ID_e = \log_2\left(\frac{A_e}{W_e} + 1\right) \quad (3)$$

According to ISO 9241-411, ID_e represents the "measure of user **precision** achieved in accomplishing a task" [27]. In Equation 3, the effective target distance A_e is the real distance traversed to execute the task, and W_e is the effective target width, calculated as $W_e = 4.133 \times SD_x$, where SD_x is the standard deviation of the distance between the target center and the selection coordinates projected onto the task axis (called x), i.e., it is a uni-variate formulation that does not account for deviations in the direction orthogonal to the main task direction. SD_x represents the **accuracy** of the task execution [27, 36, 37].

In this work, we use the ISO 9241:411 multidirectional pointing task with accuracy, precision, and throughput measurements to assess the performance of the participants. Previous work has shown that effective throughput can be used to analyze the motor performance of the users [59].

2.2 Effect of auditory feedback on motor performance

Sound and pitch play an important role in physical actions and can change people's behaviour [20, 34]. For instance, Ley-Flores et al. [34] demonstrated that auditory feedback can alter human movements. Still, previous studies in body movement research showed that changing the pitch does not increase the performance in terms of moving the body to the desired position, but instead reduces the ability to judge a limb position accurately.

Similarly, user performance can potentially be affected by any change in perceptual information [55]. Specifically, in VR systems, users perceive the VE via visual, haptic, and auditory feedback. Thus, auditory feedback is a critical component that can affect users' experience and their performance. Previous work has shown that auditory feedback could affect users' reaction time [47] and that it reduces target acquisition times [44, 49, 50]. Also, Dhruv et al. [28] explored the accessibility of auditory feedback in VR.

As in previous work [23], we define discrete auditory feedback as a short sound used as feedback to make the user aware of the state of the task, e.g., a short 'bing' when the user misses a target. Continuous auditory feedback plays sounds for longer periods as a stimulus, to make the user aware of continuous changes in a relevant variable, such as proximity to an ideal location [23].

Some studies have also used auditory *error* feedback as a signal to remind participants they made errors and need to improve their performance [30, 31, 54]. Konttinen et al. [31] mapped user movements to the pitch in a shooting task, and their results showed that users made fewer errors when receiving higher frequency auditory feedback. Similarly, other work mapped user deviations in rowing-type movements to auditory pitch, and their results showed that users could follow a desired trajectory [17, 24, 29, 38], has focused on mapping errors in movement trajectories to auditory feedback to improve user motor performance. Still, the effect of auditory error feedback on 3D pointing tasks (especially in VEs) has not been explored in depth.

2.3 Auditory Feedback in Fitts' Law Studies

In previous Fitts' task studies [2, 14, 64], sound has been studied for both uni-modal and multi-modal feedback. Akamatsu et al. [1] compared three types of feedback in a pointing task, including haptic, visual, and auditory (i.e., a 2kHz sound). Compared with the positive effect of haptic feedback alone, they found that the combination of haptic, visual, and auditory feedback did not improve user performance further. On the other hand, Brent et al. [26] found that users performed faster with confirmatory auditory feedback using a 1kHz sine wave. Moreover, Sterkenburg et al. [56, 57] found that user throughput performance increased with a combination of visual and continuous auditory feedback when the frequency of the sine wave was increased (i.e., the pitch) as users hit closer to targets. Previous work on 3D UIs also showed that auditory error feedback could increase user motor performance when used in appropriate ways [11, 38].

Several 2D Fitts' law studies provided (positive) auditory feedback when users completed a successful selection through clicking the button of a 2D mouse or tapping the tablet, e.g., [1]. In 3D Fitts' law studies, typically auditory (error) feedback is provided when users make an error in pointing to a target, e.g., [50, 60]. One of the key elements to improve a users' motor performance is to correlate the variation in sound to the desired change in user behaviour [55]. In our work, when users miss a target, i.e., they did not point successfully to the target, we thus provide an augmented form of auditory error feedback to remind them that they should improve their performance.

However, errors can be costly, and thus it is advisable to pay more attention to the error rate of a task. This raises the issue that giving too much *positive* auditory feedback when users perform successfully might even decrease the awareness of negative events, such as missing the target. Thus, we decided to focus on providing *auditory error feedback* as a signal to remind users when they made an error in the pointing tasks.

2.4 Training Strategies

How to efficiently train novices in a system is still an open research question. Because users can follow different learning strategies, the results of previous work that focused on different learning strategies suggest that it is better not to use just time alone as an assessment criterion [5, 6]. These results also suggest that to improve motor skill acquisition in training systems, novice trainees need to pay more attention to the precision of their task execution at the beginning of their training [5, 6]. Later, they can then also pay attention to their speed [4]. Based on this insight, trainers could easily observe and monitor novices' motor performance in such VR training systems and simulators to provide appropriate feedback to users and facilitate their learning process. Moreover, recent studies showed that trainee performance and training efficiency can be improved through an active feedback training method [18, 19, 53].

3 USER STUDY

Based on previous work on auditory error feedback [10], a potential approach to alter user performance is changing the pitch of the sound. However, this previous work only used discrete audio error feedback. Yet, in VR training systems and simulators, it is also possible to use continuous auditory feedback and to map the auditory feedback to various (desired) assessment criteria, regardless if users make an error or not. We explore this idea here.

3.1 Motivation & Hypotheses

In contrast to the above-mentioned previous work [10], we explore here continuous sound feedback, mapped to different task execution strategies. Previous work [7, 10] has also identified that users try to avoid hearing higher frequencies. As such, this behaviour can be used to help participants focus on different task execution strategies.

In this study, we thus investigated the following two hypotheses:

H1 *It is possible to help trainees to focus on different assessment criteria with continuous auditory feedback:* Previous work identified that user performance can be altered by changing the pitch of the auditory error feedback when a trainee misses the target, but the auditory error feedback was only played once, i.e., it was discrete [10]. Until the next selection, participants then do not receive any other stimulus. Yet, we hypothesize that playing a *continuous* sound could help participants to focus better on a particular assessment criterion.

H2 *The participants' motor performance can be controlled when auditory feedback is mapped to time or precision:* As participants tend to avoid higher frequency auditory feedback, mapping continuous feedback to a task execution strategy, e.g., precision or speed, could help participants focus on a specific task execution strategy. Previous work has shown that users prefer the C4 frequency, which is known as the middle C (262Hz), as the pitch for error feedback [10]. Given this, assigning the C4 frequency to a desired time or precision "target" value might help participants to reach that criterion.

3.2 Pilot Study

While most VR applications, including training systems, use electronically generated sounds, one of our conditions used a human voice as sound feedback, which contains different frequencies. Thus, before evaluating the use of different frequencies for continuous feedback to increase the motor performance of the trainee, we decided first to compare several existing methods in a pilot study.

Also, to be able to map distances in user errors within an ISO 9241-411 task to appropriate feedback, we decided to conduct a short user study with 16 participants. The aim of this pilot study was to acquire the 3D position of the selection points in VR to discover a good mapping from the error distance to the pitch of sound feedback. For this purpose, we applied four different auditory error feedback conditions in this pilot study, including no sound, two different frequencies of computer-generated sounds, and one verbal error feedback sound. We used the no sound condition as the control condition for this pilot study, where we did not provide subjects with any auditory error feedback. For the two computer-generated sound conditions, we played a middle C (C4 - 262Hz) or the highest C on a piano keyboard (C8 - 4186Hz) for 0.25 seconds when the participant made an error in the task. We selected these two sounds based on their use in previous work [4, 7]. For the verbal sound condition, we played a human voice that said "miss" when the participant made an error. The voice was recorded by a native American speaker. Our motivation behind choosing C4 and C8 sounds was to test the optimal and highest frequency that we were considering for the main study. We also included the voice feedback to investigate the impact of the voice as a feedback. Since ray casting and virtual hand selection techniques are widely used in mid-air interaction, we were also interested in any potential interaction between these selection techniques and auditory error feedback [33]. To arrive at the same IDs as used in the previous work we also used the same target sizes and distances used in [10] to increase the comparability of our results.

Thus we provided two different selection techniques for participants to select targets, the virtual hand/controller (see Figure 1(a)) and ray casting (see Figure 1(b)). In each of the two conditions (i.e., virtual hand/controller or ray casting), all 11 targets were placed at a visual depth of 0.4 m or 1.5 m in front of the participants, respectively. These visual depths were chosen based on the results of previous work [9], to ensure comparability in terms of conditions. We used three different target sizes (3_{TS} = 1.5, 2.5, and 3.5 cm) and three different target distances (3_{TD} = 12.5, 25, and 27.5 cm) to vary the index of difficulty (ID).

In our analysis of the pilot study results for auditory error feedback, we were not able to observe an effect in terms of time, error rate, or throughput. We only found that participants were more accurate and precise with the C4 auditory error feedback compared to the C8 feedback. Our results on auditory error feedback thus match previous work on auditory error feedback [10]. The results show that we can use these sounds to control the motor performance of the participants for this study. We also learned that voice did not elicit different results compared to other forms of feedback, and we thus decided not to consider voice feedback for the main study.

3.3 Subjects

18 participants (all right-handed; 9 males and 9 females) with an average age of 20.88 ±0.83 were recruited for this study. We recruited predominantly people without VR experience (only three had previous experience with VR HMDs). Seventeen participants reported that they play 0-5 hours of computer games, and one played 5-10 hours weekly. Ten participants reported that they use computers 0-2 hours, three 2-4 hours, three 4-6 hours, and two 6-8 hours daily. The inter-pupillary distance of the headset was adjusted for each participant before the experiment.

3.4 Procedure

We conducted the main user study on-site in an indoor laboratory at a university campus. Before the experiment started, the participants filled out a demographic questionnaire which collected their demographic information Then an experimenter described the procedure of the experiment in detail to participants and helped them to put on the HMD for the experiment. In the VE, participants sat in the middle of an empty room with pictorial depth cues, as shown in Figure 1. During the experiment, participants were then asked to position a blue cursor attached to the virtual representation of the VR controller to point at targets by using their dominant hand. They were asked to select each target by pressing the space bar on a keyboard with their non-dominant hand, which was designed to avoid the adverse consequences of the "Heisenberg" effect [16], i.e., when the user applies a mechanical force on the button to select a target, that force moves the physical controller and thus the cursor, causing additional errors.

We used the ISO 9241-411:2012 multidirectional pointing task [27] for our experiment. In the VE, there were 11 gray sphere targets placed in a circle in front of the participants. Participants used the blue cursor, which was attached to the VR controller, to select each target. There were two different **selection techniques** for participants to select targets, the virtual hand/controller (see Figure 1(a)) and ray casting (see Figure 1(b)). In each of the two conditions, all 11 targets were placed at a visual depth of 0.4 m (for virtual and) or 1.5 m (for ray casting) in front of the participants, respectively. These visual depths were chosen based on an experimental design used in previous research [9], which matches the IDs of the task across the two input conditions, i.e., ensures higher comparability of results across the input conditions. In the virtual hand/controller condition, we placed the 1 cm blue sphere cursor 3 cm above the VR controller to eliminate diplopia. In the ray casting condition, we placed the blue cursor at the intersection of the ray and the 2D target plane, which was always 1.5 m away from the user. To vary the index of difficulty (ID), we used three different target sizes and three different target distances.

During each round of trials, participants were required to select all 11 targets in a sequence. The first target was randomly chosen by the software, and the color of the target would change to orange. Then participants selected the next target diagonally across the circle, progressing either in a clockwise or counter-clockwise direction. Participants selected orange-colored targets by pressing the space bar when the cursor was inside the target. To show when the cursor was inside the target, we highlighted the target in blue as visual feedback. If participants pressed the space bar when the cursor was inside the target, we changed the color of the target to green as positive feedback to show a successful selection. Otherwise, if participants pressed the space bar when the cursor was outside the target, the color would change to red, which means participants made an error.

Different from the pilot study, we mapped participants' speed and precision to auditory feedback and played three different forms of continuous auditory feedback. We asked participants to perform the task as fast and as accurately as possible.

We applied three different **auditory error feedback** conditions: no mapping, precision mapping, and speed mapping. For the control condition (i.e., the no mapping condition), we did not provide any auditory error feedback. For the other conditions, we mapped the speed and precision of the participants to the pitch of the auditory feedback. In the speed-based auditory feedback condition, we played a sound with increasing pitch, from C1 (33Hz) to C8 (4186Hz) in each trial. We played no sound for 0.5s, then increased the sound's pitch to C4 at 1.1s (which was the average task duration in the pilot), and increased to C8 at 2.6s until the end of each trial. In other words, participants heard a lower pitch in fast trials and a higher pitch when they performed more slowly (Figure 2(a)).

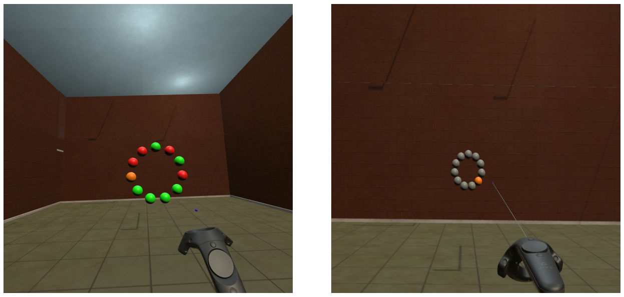

(a) (b)

Figure 1: The user study environment and selection techniques: (a) virtual hand/controller and (b) ray casting

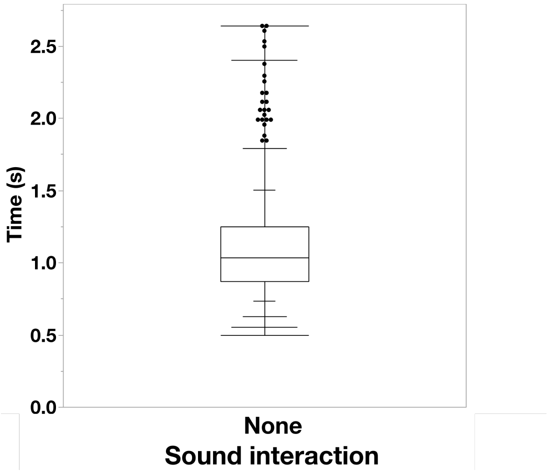
Based on the results of the pilot study, we mapped the precision to each target size in the precision-based auditory feedback condition as follows. For each selection, we calculated the distance between the selection point and the target center. If the target size was 1.5 cm, there was no auditory feedback if the distance was less than 0.1 cm, we linearly mapped 0.1 - 0.4 cm to the C1-C8 frequency range, and played a C8 frequency beyond 0.4 cm. For the 2.5 cm target size, there was no sound if the distance was less than 0.2 cm, a C1 to C8 linearly-mapped auditory feedback between 0.2 - 0.6 cm, and a C8 frequency beyond 0.6 cm. For the 3.5 cm target size, there was no sound if the distance was less than 0.2 cm, for distances between 0.2 - 0.7 cm we used C1 to C8 linearly-mapped auditory feedback, and a C8 frequency beyond 0.7 cm. In other words, participants heard a lower pitch when they performed precisely and a higher pitch when they performed less well (Figure 2(b).

We verified the effectiveness of the sound mappings for time and precision based on a quick evaluation with three participants. In this evaluation, we played auditory feedback starting from 0 seconds and 0 cm distance. However, we observed that the participants were then focusing too much on the feedback. Thus, encouraging participants to only go faster caused them to be (much) less precise with the time mapped to frequency. Similarly, with a precision-mapped frequency focusing on precision right from the start made participants (much) slower. Thus, we decided not to play the sound feedback all the time. Based on that insight, we then verified with another 3 participants that, with the adjusted mappings presented above, participants were able to focus more on the task, not on the feedback - as they had (some) time to plan their action before they got feedback. We ran the main study after this verification.

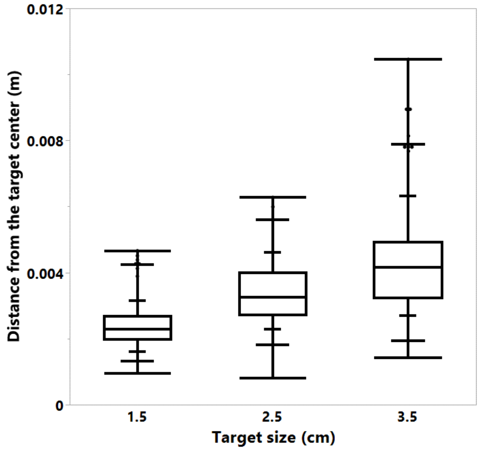
To avoid biasing participants towards specific task execution strategies the participants in this study were *not* given instructions in terms of which strategy to use.

Participants could hear the sounds via the built-in headphones of the HTC VIVE Pro during the task duration. We set the computer volume to 40%.

At the end of the study, participants were asked to fill out a questionnaire about their perception and preferences of the three different auditory feedback conditions. Overall, completing the study took approximately 13 minutes for each participant.



(a)



(b)

Figure 2: The data used for speed and precision mapping. We used the data from the pilot study's no auditory feedback condition to determine the speed (a) and precision mapping (b) in the main study. In speed mapping (a), we used the minimum, mean, and maximum values to map the frequency in execution time. For precision mapping (b), we used the distance from the target center variable's minimum, mean, and maximum values to map the frequency. The boxes show the mean and 25% and 75% quartiles. The top whiskers show the 90%, 95%, and 97.5% quartiles, as well as the maximum (and symmetrical for the bottom whiskers). Individual datapoints outside the boxes are highlighted.

3.5 Experimental Design

This study used a within-subjects design with two factors: (1) **auditory feedback** (**A-Feedback**: 3_{AF} = Speed-based sound feedback, Precision-based sound feedback, and No sound feedback) and (2) **selection techniques** (`S-Technique`: 2_{ST} = ray casting and virtual hand/controller), comprising a $3_{AF} \times 2_{ST}$ design. For assessing the outcome, we measured task execution time (seconds), error rate (%), effective throughput (bits/s), the effective index of difficulty ID_e, which is precision according to ISO 9241-411, and SD_x, which is the standard deviation of selection coordinates along the task axis and which represents the **accuracy** of the task [36]. We used three target sizes (3_{TS} = 1.5, 2.5, and 3.5 cm) and three target distances (3_{TD} = 12.5, 25, and 37.5 cm) to vary the index of difficulty *ID*, which created 9 unique IDs between 2.19 and 4.7 (see Table 1). Each participant performed $3_{AF} \times 2_{ST} \times 9_{ID} \times 11$ repetitions = 594 trials. In order to eliminate potential learning effects, the order of **selection technique** and **auditory feedback** was counterbalanced with a Latin Square (see Table 2) for each participant.

Table 1: Randomized task variables for each condition (TD = Target Distance, TS = Target Size).

TD/cm	12.5	12.5	12.5	25	25	25	37.5	37.5	37.5
TS/cm	1.5	2.5	3.5	1.5	2.5	3.5	1.5	2.5	3.5

Table 2: $3_{AF} \times 2_{ST}$ conditions counterbalanced with a Latin Square design (AF = Auditory Feedback, None = No auditory feedback, SB = Speed-based auditory error feedback, PB = Precision-based auditory error feedback, ST = Selection Technique, VH = Virtual Hand/Controller, and RC = Ray casting).

AF	None	None	SB	SB	PB	PB
ST	VH	RC	VH	RC	VH	RC

3.6 Results

We used Repeated Measures (RM) ANOVA in SPSS 24 to analyze the data. The Skewness (S) and Kurtosis (K) of the data distribution were used to perform normality tests of the data. We considered data to be normally distributed when S and K values were within ±1.5. Otherwise, we used ART [62] before ANOVA. We used the Bonferroni method for post-hoc analyses and applied Huynh-Feldt correction when the ε was less than 0.75. Results are shown in Figure 3 as means and standard error of means. In the results section, we first present the main factor results in Table 3.

Since the main focus of our work is on auditory feedback, we do not detail results for the interaction methods here for *brevity*, except if there were notable effects.

Time results: The time dependent variable was normally distributed (S = 0.31, K = -0.36). According to the results in Table 3 and Figure 3 (a), and compared to no auditory feedback, participants were faster when the sound feedback was mapped to speed and slower when the sound feedback was mapped to precision.

Table 3: One-Way RM ANOVA results of the study.

	Sound Interaction	Selection Technique	ID
Time	F(2, 34) = 210.91 $p<0.001, \eta^2 = 0.925$	F(1, 17) = 7.755 $p<0.05, \eta^2 = 0.313$	F(8, 136) = 266.386 $p<0.001, \eta^2 = 0.940$
Error rate	F(2, 34) = 97.14 $p<0.001, \eta^2 = 0.821$	F(1, 17) = 49.64 $p<0.001, \eta^2 = 0.745$	F(8, 136) = 35.26 $p<0.001, \eta^2 = 0.675$
Throughput	F(2, 34) = 195.76 $p<0.001, \eta^2 = 0.92$	F(1, 17) = 125.34 $p<0.001, \eta^2 = 0.881$	F(8, 136) = 13.467 $p<0.001, \eta^2 = 0.442$
ID_e	F(2, 34) = 90.07 $p<0.001, \eta^2 = 0.841$	F(1, 17) = 406.339 $p<0.001, \eta^2 = 0.96$	F(8, 136) = 313.602 $p<0.001, \eta^2 = 0.949$
SD_x	F(2, 34) = 84.096 $p<0.001, \eta^2 = 0.832$	F(1, 17) = 354.506 $p<0.001, \eta^2 = 0.77$	F(8, 136) = 56.809 $p<0.001, \eta^2 = 0.77$

Error rate results: The error rate dependent variable was not normally distributed (S = 1.86, K = 3.59); thus, we analyzed the data with ART. According to the results in Table 3 and Figure 3(b), participants made fewer errors when auditory feedback was mapped to the precision performance of the participant.

Throughput results: Throughput was normally distributed (S = 0.46, K = -0.24). According to the results, shown in Figure 3(c), participants exhibited higher throughput when the auditory feedback was based on the speed of the user.

Effective Index of Difficulty ID_e results: The effective index of difficulty dependent variable was normally distributed (S = 0.13, K = -0.308). The results showed that participants' precision increased when the auditory feedback was mapped to precision.

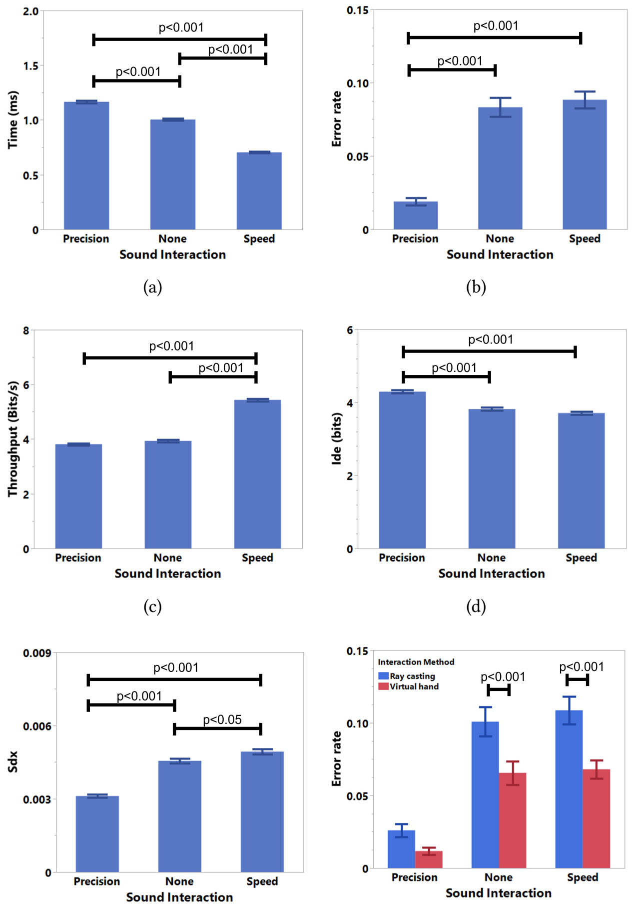
Standard Deviation SD_x results: The standard deviation dependent variable was normally distributed (S = 1.02, K = 1.3). The results in Table 3 and Figure 3(e) illustrate that the accuracy of participants significantly increased when the precision of participants was mapped to the auditory feedback and decreased when their speed was mapped to the auditory feedback.

Two-way Interactions: We found a significant interaction for error rate between selection technique and sound interaction (F(2,34) = 6.223, $p < 0.05$, $\eta^2 = 0.268$), as shown in Figure 3(f). According to the results, participants' error rate increased with ray casting when the speed was mapped to the auditory feedback and when there was no mapping.

3.6.1 *Detailed Analysis.*

According to the detailed results, all individuals were able to understand and focus on the target assessment criterion. When we mapped the sound feedback to the speed of the participants, all the participants got faster. This is shown in Figure 4(a) for each participant. The average execution time of the participants was 0.229 seconds less with speed-mapped sound feedback (No auditory feedback 1.04 s ±0.01, speed-mapped auditory feedback 0.705 ±0.008).

Similarly, when we mapped the feedback to the precision of the participants, we observed an increase in precision for each participant. On average, participants were 0.131 cm closer to the target compared to the no auditory feedback condition (No auditory feedback 0.402 cm ±0.011, precision-mapped auditory feedback 0.276 cm ±0.012). In further analysis, we also observed that participants were getting closer to the center of the targets for each target size, as shown in Figure 4(b). For 1.5 cm, participants were 0.05 cm closer to the target center, for 2.5 cm, 0.09 cm, and for 3.5 cm, 0.11 cm,



(a) (b)

(c) (d)

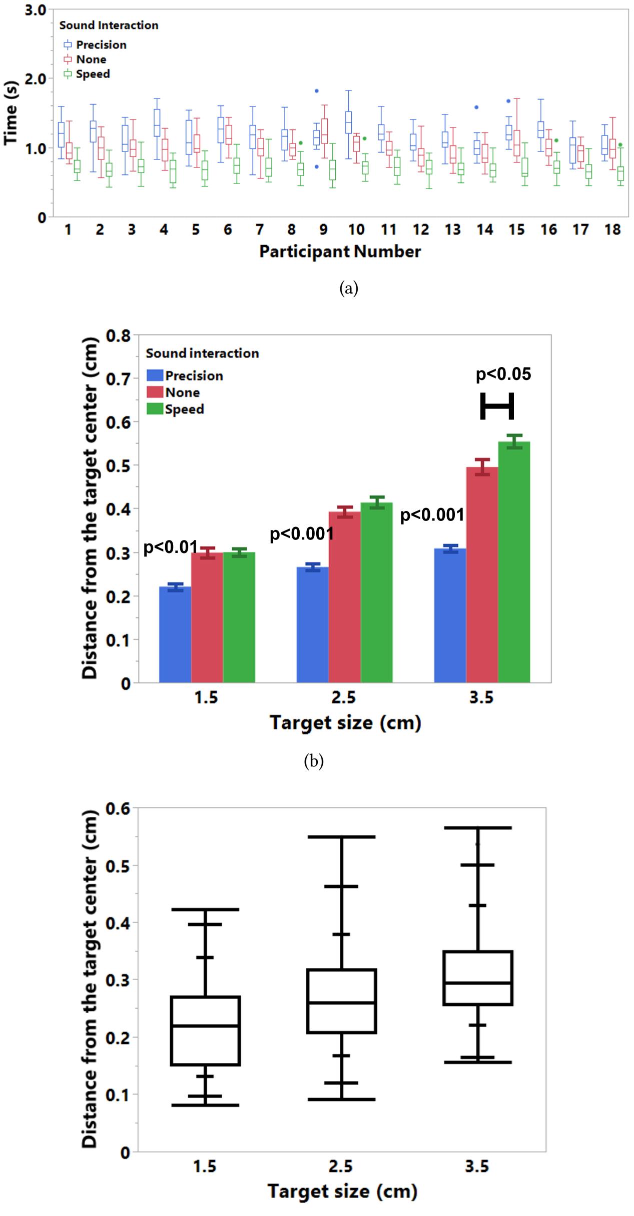
(e) (f)

Figure 3: Study results for (a) time, (b) error rate, (c) throughput, (d) effective index of difficulty, (e) standard deviation and (f) two-way interactions.

compared to the no auditory feedback condition. With the precision-mapped sound feedback, the average distance from the center of the target was 0.22 cm ±0.007 for 1.5 cm, 0.265 cm ±0.008 for 2.5 cm, and 0.308 cm ±0.007 for 3.5 cm. This is also shown in Figure 4.

3.6.2 Subjective Results. After the experiment, we asked participants to fill out a questionnaire about their preferences for the three different auditory feedback conditions. Ten participants preferred the auditory feedback mapped to precision, and eight participants preferred the auditory feedback mapped to their speed. Interestingly, four participants commented that the higher frequency in the auditory feedback mapped to the precision was "*annoying*", i.e., which might be a result of them hearing the precision feedback for longer periods. In contrast, one participant commented "*conditions which have sounds are better than no sound conditions*".

Participants did reported no notable physical (average = 1.16, STD = 0.38) nor mental (average = 1.27, STD = 0.46) fatigue (1 = I feel extremely rested, 7 = I feel extremely fatigued).



(a)

(b)

(c)

Figure 4: Detailed analysis results for auditory feedback mapping. a) Each participant focused on the desired execution strategy. b) Detailed analysis of target size vs auditory feedback. c) Quantile plot of the precision-mapped auditory feedback for each target size.

3.6.3 Fitts' Law Analysis. Using Fitts' law in Equation 1, we found that task execution time can be modeled as MT = 0.3 + 0.19 * ID, R^2 = 0.98 for Study 2. Linear regressions were MT = 0.42 + 0.16 * ID, R^2 = 0.98 for the control condition, MT = 0.41 + 0.22 * ID, R^2 = 0.98 for the precision condition, and MT = 0.07 + 0.18 * ID, R^2 = 0.96 for the speed condition, as shown in Figure 5.

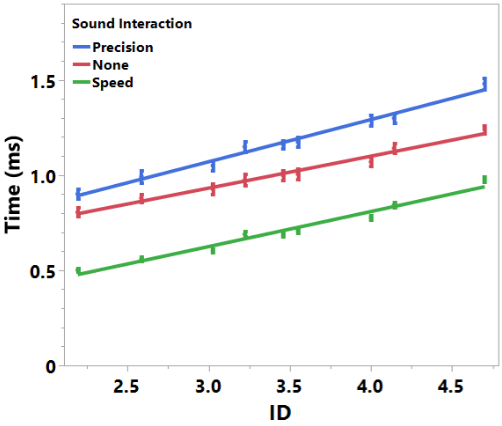


Figure 5: Fitts' law results for auditory feedback mapped to task execution strategy in the study.

4 DISCUSSION

In this study, we aimed to help trainees to focus on a particular task execution strategy and (with this) also to change the trainee's performance to match a pre-determined speed or precision, in a task that uses throughput as an assessment criterion.

4.1 Outcomes

According to the study results, when we played a sound feedback mapped to the user's speed, participants' task execution time decreased. Similarly, when we played an auditory feedback mapped to the user's precision, participants' precision also increased. Thus, the results of our study support **H1**, *that it is possible to help trainees to focus on different assessment criteria with continuous auditory feedback* in an ISO 9241-411 task. The results here also support the findings of previous work on auditory error feedback, which decreased the error rate of the participants [10], and extends the findings to both speed and precision.

While one could argue that 18 participants is a low number, all statistically significant results in this work exhibit a large effect size, i.e., $\eta^2 > 0.14$. As effect size is independent of sample size, this indicates that our results are likely robust and thus replicable.

In our detailed analysis, we observed that participants were able to reach a task execution time of (on average) 0.6 seconds. When individuals heard the change in frequency, they started to get faster with the auditory feedback. We had set the C4 frequency to 1.1 seconds for the speed-mapped sound feedback. Our results revealed that participants did not even wait for the pitch to reach this comfortable level (C4), but selected targets even faster. We believe that participants might have actively avoided hearing higher frequencies, which aligns with results from previous work [7, 10] and that this was a motivating factor for them to get faster. However, when we look at the detailed results for precision, we observe that participants' average selection points were well inside the target size. This means that, by using an appropriate stimulus, we were able to help participants to select the targets within a range that matched what our auditory feedback encouraged them to do. This supports our hypothesis **H2**, *The participants' motor performance can be controlled when auditory feedback is mapped to time or precision.* We believe that participants' performance in terms of speed could fall into a similar range for speed-mapped auditory feedback, if auditory feedback was also mapped to each target size.

Even though the precision-mapped feedback results reveal that it is possible to map the assessment criterion into a specific range, there are still open questions. For instance, we cannot guarantee that the participants were selecting targets as precisely as possible, i.e., that our results might be just a side-effect of our specific mapping. Although we believe that our findings identify the potential for directing user motor performance through appropriate auditory feedback, future work is needed to substantiate this theory. Furthermore, the results for the speed-mapped feedback are less clear, yet we believe that a more carefully chosen mapping might help us control user speed more accurately.

Overall, the results of this study showed that it is possible to use auditory feedback to encourage trainee performance to match a desired assessment criterion (in terms of time and precision). Moreover, with appropriate auditory feedback, we were able to "steer" the participants' average motor performance results into specific ranges. We caution that this approach clearly has limits, e.g., one cannot go beyond the biomechanical limits of the human body. Still, we believe that our results open new avenues for motor skills training systems. Also, our work provides better insights into the role of auditory feedback in ISO 9241-411 tasks.

Previous ISO 9241-411 studies, e.g., [60], suggested using auditory feedback to improve user performance, but did not evaluate this idea. Previous work in VR training systems and simulators followed up on this observation and changed human motor performance [7, 10]. Building on this work, our work extends the literature and suggests a method to actively *control* users' task execution through auditory feedback based on a specific assessment criterion. We believe our results are thus a step towards a deeper understanding of the psychophysics of interacting with a VR training system and simulator. We also believe that our results can be used in the design of VR training applications and simulators as well as mid-air pointing studies.

4.2 Recommendations

Based on our findings, we suggest the following:

Choose an appropriate auditory feedback mapping for VR training applications that use speed, precision, accuracy, or throughput as an assessment criterion. As can be seen from the results of our study, different mappings can elicit different performance results. When we combine our findings with those from previous work on auditory error feedback [10], we can identify that it is possible to help individuals to focus on a particular assessment criterion, for instance, when a trainee needs to focus on their speed, by mapping the auditory feedback accordingly. A similar approach is also valid for precision and error rate.

Test the auditory feedback before starting the training and identify trade-offs. Even though auditory feedback can help trainees focus on different assessment criteria, trainees can (and will) still choose their own trade-offs, which can impact their motor skills. One such example in the work presented here is that our results did not exhibit differences in terms of precision between speed-mapped auditory feedback and no auditory error feedback, but participants' accuracy decreased when we mapped sound pitch to the speed of the participants. Similarly, since the precision and error rate are correlated in an ISO 9241-411 task, we did not observe a significant difference between the no auditory feedback and speed-mapped auditory feedback conditions. Such changes might seem trivial, but motor learning is a complex process which ideally results in permanent behavioral and neurological changes [32]. Since there is no optimal suggested learning method for motor skills, we recommend monitoring each performance criterion and considering carefully their relation to other assessment criteria.

Previous work on VR simulators and training systems for motor performance training, e.g., [8, 12, 13, 43], did not investigate auditory feedback. In VR training systems, the trainees have to focus on the task while they focus on processing the environmental cues provided by the virtual environment, which increases their cognitive workload. For example, they have to focus on the path of the ball coming towards them. Thus, distractions in the virtual environment, any other stimuli from the system, and disruptions during the task execution may all affect the trainees' focus and performance. Since previous work showed that sound feedback has an impact on motor performance [10] and the fact that sound feedback is (in general) underutilized in VR training systems), we decided to extend this previous work. Here, we analyzed user performance with different kinds of auditory feedback within an ISO 9241-411:2012 task [27]. Our results can thus also be used in 3D user interface studies that use throughput to analyze user performance with different input devices.

4.3 Limitations

We chose participants from the local university community, not from a specific trainee group. Thus, the novice results presented here may differ from professional trainees who use a VR training system or simulator on a daily basis. Our choice allowed us to understand the impact of auditory feedback on a more general population in a VR training system and helped us to avoid any potential confounds related to the motor skill development of professionals. To identify the most efficient training strategies, there is thus still a need to study the impact of manipulating the motor performance of users with different experience levels in VR further.

We acknowledge that the previous literature on skill transfer from VR systems to the real-world is somewhat inconclusive [25, 40, 63]. We see this as further evidence that we first need to understand the fundamental effects of VR-based training systems before designing and applying them for specific tasks.

Also, we only investigated the effect on user motor performance by varying the pitch of the auditory error feedback within a limited range, with C8 (4186Hz) as the highest frequency. We selected this value based on previous work [7, 10]. However, this frequency range could be changed based on each individual's hearing capabilities. Moreover, various training systems could use different tools, such as horns or whistles as stimuli for an erroneous action, which could be different from the voice feedback we investigated in our pilot study mentioned here. There are also other dimensions of auditory feedback to consider, such as loudness, quality, and timing. To further understand the impact of auditory feedback on user performance. These other dimensions should be also analyzed and investigated in terms of their usefulness to facilitate motor learning and skill transfer [54].

In this paper, we used an ISO 9241:411 multidirectional selection task to evaluate auditory feedback. Although the ISO task is largely accepted by the HCI community as an accurate way of assessing motor performance, the task is still repetitive and monotonous. This method reduces the cognitive load of the participants so that they can focus on the task itself and also minimizes biases across different task difficulties. However, as a task gets more complex, a user might have to attend to different stimuli simultaneously, which might affect their motor performance. Thus, we expect to observe higher impact when the task complexity increases. We thus suggest studying the impact of auditory feedback with different tasks, including manipulation of and travel within the virtual environment, to further extend our findings. Furthermore, the results presented here may vary for other interaction modalities [41, 42], so our findings also need to be verified for other techniques.

Moreover, we used a linear mapping for the sound frequency of the auditory feedback. There are other possible mappings, such as logarithmic or exponential mappings. Since the higher frequencies tend to annoy users and alter their performance, we expect that a change in duration to reach higher frequencies could have a different impact on user performance. Further, in the precision-mapped auditory feedback, we changed the auditory feedback for each target size but kept the sound the same for each target size in the time-mapped method. One could thus expect different motor performance results when the time is also mapped to each target size. However, such speculation must be tested in future studies.

5 CONCLUSION & FUTURE WORK

In this paper, we explored different forms of auditory feedback that are useful in a VR training system for changing or manipulating trainees' performance in terms of different assessment criteria. Through our mappings that affect the pitch of the auditory feedback, participants were able to focus on different assessment criteria. We showed that it is possible to decrease the execution time or increase the precision of the participants by changing the auditory mapping. Furthermore, it is also possible to help participants' performance to reach a pre-determined range by using auditory feedback. Based on these results, we suggest that designers, practitioners, and developers could use the pitch of the auditory feedback to encourage trainees to focus on a specific assessment criterion. We also recommend careful calibration of the auditory pitch mapping, since the outcomes elicited by the mapping may vary.

In the future, we are planning to extend our work to other features of auditory feedback, such as length, tone, or timbre. We also plan to apply the results of our findings to VR and AR simulators and training systems that are commercially available on the market. We also want to extend our results to other 3D interaction modalities, such as manipulation and navigation. Further, we plan to conduct learning studies to understand the long-term impact of auditory feedback. We also want to apply our results to rehabilitation and medical research [3].

REFERENCES

[1] Motoyuki Akamatsu, I Scott MacKenzie, and Thierry Hasbroucq. 1995. A comparison of tactile, auditory, and visual feedback in a pointing task using a mouse-type device. *Ergonomics* 38, 4 (1995), 816–827.

[2] Oscar Ariza, Gerd Bruder, Nicholas Katzakis, and Frank Steinicke. 2018. Analysis of proximity-based multimodal feedback for 3d selection in immersive virtual environments. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 327–334.

[3] Breton M Asken, Jason P Mihalik, Julianne D Schmidt, Ashley C Littleton, Kevin M Guskiewicz, and Joseph B Hopfinger. 2016. Visual performance measures and functional implications in healthy participants: A sports concussion perspective. *Athletic Training & Sports Health Care* 8, 4 (2016), 145–153.

[4] Anil Ufuk Batmaz. 2018. *Speed, precision and grip force analysis of human manual operations with and without direct visual input.* Ph. D. Dissertation. Strasbourg.

[5] Anil Ufuk Batmaz, Michel de Mathelin, and Birgitta Dresp-Langley. 2016. Getting nowhere fast: trade-off between speed and precision in training to execute image-guided hand-tool movements. *BMC psychology* 4, 1 (2016), 1–19.

[6] Anil U Batmaz, Michel de Mathelin, and Birgitta Dresp-Langley. 2017. Seeing virtual while acting real: Visual display and strategy effects on the time and precision of eye-hand coordination. *PloS one* 12, 8 (2017), e0183789.

[7] Anil Ufuk Batmaz, Michel de Mathelin, and Birgitta Dresp-Langley. 2019. Different sound pitch effects on motor performance of individuals in head-mounted virtual reality. In *Perception*, Vol. 48. SAGE Publications, UK, 21–21.

[8] A. U. Batmaz, A. K. Mutasim, M. Malekmakan, E. Sadr, and W. Stuerzlinger. 2020. Touch the Wall: Comparison of Virtual and Augmented Reality with Conventional 2D Screen Eye-hand Coordination Training Systems. In *Conference on Virtual Reality and 3D User Interfaces (VR '20)*. IEEE, 184–193. https://doi.org/10.1109/VR46266.2020.1581205539914

[9] Anil Ufuk Batmaz, Aunnoy K Mutasim, and Wolfgang Stuerzlinger. 2020. Precision vs. power grip: A comparison of pen grip styles for selection in virtual reality. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. IEEE, 23–28.

[10] Anil Ufuk Batmaz and Wolfgang Stuerzlinger. 2021. Effects of Different Auditory Feedback Frequencies in Virtual Reality 3D Pointing Tasks. In *2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. 189–194. https://doi.org/10.1109/VRW52623.2021.00042

[11] Anil Ufuk Batmaz and Wolfgang Stuerzlinger. 2021. Effects of different auditory feedback frequencies in virtual reality 3D pointing tasks. In *2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. IEEE, 189–194.

[12] Anil Ufuk Batmaz, Xintian Sun, Dogu Taskiran, and Wolfgang Stuerzlinger. 2019. Hitting the Wall: Mid-Air Interaction for Eye-Hand Coordination. In *25th ACM Symposium on Virtual Reality Software and Technology* (Parramatta, NSW, Australia) *(VRST '19)*. ACM, New York, NY, USA, Article 30, 5 pages. https://doi.org/10.1145/3359996.3364249

[13] A. U. Batmaz, X. Sun, D. Taskiran, and W. Stuerzlinger. 2020. Eye-Hand Coordination Training for Sports with Mid-air VR. In *26th Symposium on Virtual Reality Software and Technology (VRST '20)*. ACM, Article 19, 10 pages. https://doi.org/10.1145/3385956.3418971

[14] Joshua D Bell and Kristen L Macuga. 2019. Goal-directed aiming under restricted viewing conditions with confirmatory sensory feedback. *Human movement science* 67 (2019), 102515.

[15] Antonio Bernardo. 2017. Virtual reality and simulation in neurosurgical training. *World neurosurgery* 106 (2017), 1015–1029.

[16] Doug Bowman, Chadwick Wingrave, Joshua Campbell, and Vinh Ly. 2001. Using pinch gloves (tm) for both natural and abstract interaction techniques in virtual environments. (2001), 1–5.

[17] Eric Olivier Boyer, Bénédicte Maria Babayan, Frédéric Bevilacqua, Markus Noisternig, Olivier Warusfel, Agnes Roby-Brami, Sylvain Hanneton, and Isabelle Viaud-Delmon. 2013. From ear to hand: the role of the auditory-motor loop in pointing to an auditory source. *Frontiers in computational neuroscience* 7 (2013), 26.

[18] Kenneth J Ciuffreda and Bin Wang. 2004. Vision training and sports. In *Biomedical engineering principles in sports*. Springer, 407–433.

[19] Alexandra Covaci, Cristian-Cezar Postelnicu, Alina Ninett Panfir, and Doru Talaba. 2012. A virtual reality simulator for basketball free-throw skills development. In *Doctoral Conference on Computing, Electrical and Industrial Systems*. Springer, 105–112.

[20] Zohar Eitan. 2017. Cross-modal experience of musical pitch as space and motion: Current research and future challenges. *Body, sound and space in music and beyond: multimodal explorations* (2017), 49–68.

[21] Paul M Fitts. 1954. The information capacity of the human motor system in controlling the amplitude of movement. *Journal of experimental psychology* 47, 6 (1954), 381.

[22] Stephanie G Fussell and Dothang Truong. 2020. Preliminary results of a study investigating aviation student's intentions to use virtual reality for flight training. *International Journal of Aviation, Aeronautics, and Aerospace* 7, 3 (2020), 2.

[23] BoYu Gao, HyungSeok Kim, Hasup Lee, Jooyoung Lee, and Jee-In Kim. 2018. Effects of continuous auditory feedback on drawing trajectory-based finger gestures. *IEEE Transactions on Human-Machine Systems* 48, 6 (2018), 658–669.

[24] Andrew Godbout and Jeffrey Edwin Boyd. 2012. Rhythmic Sonic Feedback for Speed Skating by Real-Time Movement Synchronization. *International Journal of Computer Science in Sport (International Association of Computer Science in Sport)* 11, 3 (2012), 1–15.

[25] Rob Gray. 2002. Behavior of college baseball players in a virtual batting task. *Journal of Experimental Psychology: Human Perception and Performance* 28, 5 (2002), 1131.

[26] Brent C Hatfield, William R Wyatt, and John B Shea. 2010. Effects of auditory feedback on movement time in a Fitts task. *Journal of motor behavior* 42, 5 (2010), 289–293.

[27] ISO 9241-411:2012. 2012. Ergonomics of human-system interaction – Part 411: Evaluation methods for the design of physical input devices. ISO. https://www.iso.org/standard/54106.html

[28] Dhruv Jain, Sasa Junuzovic, Eyal Ofek, Mike Sinclair, John R. Porter, Chris Yoon, Swetha Machanavajhala, and Meredith Ringel Morris. 2021. Towards Sound Accessibility in Virtual Reality. In *Proceedings of the 2021 International Conference on Multimodal Interaction*. 80–91.

[29] Grega Jakus, Kristina Stojmenova, Sašo Tomažič, and Jaka Sodnik. 2017. A system for efficient motor learning using multimodal augmented feedback. *Multimedia Tools and Applications* 76, 20 (2017), 20409–20421.

[30] Max Kleiman-Weiner and Jonathan Berger. 2006. The sound of one arm swinging: a model for multidimensional auditory display of physical motion. Georgia Institute of Technology, 1–3.

[31] Niilo Konttinen, Kaisu Mononen, Jukka Viitasalo, and Toni Mets. 2004. The effects of augmented auditory feedback on psychomotor skill learning in precision shooting. *Journal of Sport and Exercise Psychology* 26, 2 (2004), 306–316.

[32] John W Krakauer. 2006. Motor learning: its relevance to stroke recovery and neurorehabilitation. *Current opinion in neurology* 19, 1 (2006), 84–90.

[33] Joseph J LaViola Jr, Ernst Kruijff, Ryan P McMahan, Doug Bowman, and Ivan P Poupyrev. 2017. *3D user interfaces: theory and practice.* Addison-Wesley Professional.

[34] Judith Ley-Flores, Eslam Alshami, Aneesha Singh, Frédéric Bevilacqua, Nadia Bianchi-Berthouze, Ophelia Deroy, and Ana Tajadura-Jiménez. 2022. Effects of pitch and musical sounds on body-representations when moving with sound. *Scientific reports* 12, 1 (2022), 1–20.

[35] I Scott MacKenzie. 1992. Fitts' law as a research and design tool in human-computer interaction. *Human-computer interaction* 7, 1 (1992), 91–139.

[36] I. Scott MacKenzie and Poika Isokoski. 2008. Fitts' Throughput and the Speed-Accuracy Tradeoff. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Florence, Italy) *(CHI '08)*. Association for Computing Machinery, New York, NY, USA, 1633–1636. https://doi.org/10.1145/1357054.1357308

[37] I Scott MacKenzie and Aleks Oniszczak. 1998. A comparison of three selection techniques for touchpads. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM Press/Addison-Wesley Publishing Co., 336–343.

[38] Alexander Marquardt, Ernst Kruijff, Christina Trepkowski, Jens Maiero, Andrea Schwandt, André Hinkenjann, Wolfgang Stuerzlinger, and Johannes Schöning. 2018. Audio-Tactile Proximity Feedback for Enhancing 3D Manipulation. In *24th Symposium on Virtual Reality Software and Technology (VRST '18)*. Article 2, 10 pages. https://doi.org/10.1145/3281505.3281525

[39] R Randall McKnight, Christian A Pean, J Stewart Buck, John S Hwang, Joseph R Hsu, and Sarah N Pierrie. 2020. Virtual reality and augmented reality—translating surgical training into surgical technique. *Current Reviews in Musculoskeletal Medicine* (2020), 1–12.

[40] Helen C. Miles, Serban R. Pop, Simon J. Watt, Gavin P. Lawrence, and Nigel W. John. 2012. A review of virtual environments for training in ball sports. *Computers & Graphics* 36, 6 (2012), 714 – 726. https://doi.org/10.1016/j.cag.2012.04.007

[41] Alec Moore, Marwan Kodeih, Anoushka Singhania, Angelina Wu, Tassneen Bashir, and Ryan McMahan. 2019. The importance of intersection disambiguation for virtual hand techniques. In *2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, 310–317.

[42] Alec G Moore, John G Hatch, Stephen Kuehl, and Ryan P McMahan. 2018. VOTE: A ray-casting study of vote-oriented technique enhancements. *International Journal of Human-Computer Studies* 120 (2018), 36–48.

[43] A. K. Mutasim, W. Stuerzlinger, and A. U. Batmaz. 2020. Gaze Tracking for Eye-Hand Coordination Training Systems in Virtual Reality. Extended Abstract. , 9 pages. https://doi.org/10.1145/3334480.3382924

[44] W Todd Nelson, Lawrence J Hettinger, James A Cunningham, Bart J Brickman, Michael W Haas, and Richard L McKinley. 1998. Effects of localized auditory information on visual target detection performance using a helmet-mounted display. *Human factors* 40, 3 (1998), 452–460.

[45] NIKE-SPARQ. 2011. Nike Introduces SPARQ Sensory Performance (SSP). Retrieved July 10, 2018 from https://www.youtube.com/watch?v=c1kVvJqBVx0

[46] Takayuki Nozawa, Erwin Wu, and Hideki Koike. 2019. Vr ski coach: Indoor ski training system visualizing difference from leading skier. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 1341–1342.

[47] Joviša M Obrenović, Vladimir D Nešić, and Milkica Nešić. 1996. The reaction time in relation to the modality of stimulation. *Facta universitatis-series: Physical Education* 1, 3 (1996), 85–90.

[48] George Papagiannakis, Panos Trahanias, Eustathios Kenanidis, and Eleftherios Tsiridis. 2018. Psychomotor surgical training in virtual reality. In *The adult hip-master case series and techniques*. Springer, 827–830.

[49] David R Perrott, John Cisneros, Richard L Mckinley, and William R D'Angelo. 1996. Aurally aided visual search under virtual and free-field listening conditions. *Human factors* 38, 4 (1996), 702–715.

[50] Duc-Minh Pham and Wolfgang Stuerzlinger. 2019. Is the pen mightier than the controller? a comparison of input devices for selection in virtual and augmented reality. In *25th ACM Symposium on Virtual Reality Software and Technology*. 1–11.

[51] Dmitri Poltavski and David Biberdorf. 2015. The role of visual perception measures used in sports vision programmes in predicting actual game performance in Division I collegiate hockey players. *Journal of sports sciences* 33, 6 (2015), 597–608.

[52] Quotronics Limited. 2019. BATAK Fitness training solutions. Retrieved July 10, 2018 from https://www.batak.com/

[53] Louis Rosenberg and Scott Brave. 1996. Using force feedback to enhance human performance in graphical user interfaces. In *Conference companion on Human factors in computing systems*. ACM, 291–292.

[54] Roland Sigrist. 2011. Visual and auditory augmented concurrent feedback in a complex motor task. *Presence* 20, 1 (2011), 15–32.

[55] Roland Sigrist, Georg Rauter, Robert Riener, and Peter Wolf. 2013. Augmented visual, auditory, haptic, and multimodal feedback in motor learning: a review. *Psychonomic bulletin & review* 20, 1 (2013), 21–53.

[56] Jason Sterkenburg, Steven Landry, and Myounghoon Jeon. 2017. Influences of visual and auditory displays on aimed movements using air gesture controls. Georgia Institute of Technology, 1–5.

[57] Jason Sterkenburg, Steven Landry, and Myounghoon Jeon. 2019. Design and evaluation of auditory-supported air gesture controls in vehicles. *Journal on Multimodal User Interfaces* 13, 2 (2019), 55–70.

[58] Walter Sturm and Klaus Willmes. 1991. Efficacy of a reaction training on various attentional and cognitive functions in stroke patients. *Neuropsychological Rehabilitation* 1, 4 (1991), 259–280.

[59] Robert J. Teather and Wolfgang Stuerzlinger. 2011. Pointing at 3D Targets in a Stereo Head-Tracked Virtual Environment. In *Symposium on 3D User Interfaces (3DUI '11)*. 87–94. https://doi.org/10.1109/3DUI.2011.5759222

[60] Robert J Teather and Wolfgang Stuerzlinger. 2014. Visual aids in 3D point selection experiments. In *Proceedings of the 2nd ACM symposium on Spatial user interaction*. 127–136.

[61] Kelvin Valentino, Kevin Christian, and Endra Joelianto. 2017. Virtual reality flight simulator. *Internetworking Indonesia Journal* 9, 1 (2017), 21–25.

[62] Jacob O Wobbrock, Leah Findlater, Darren Gergle, and James J Higgins. 2011. The aligned rank transform for nonparametric factorial analyses using only anova procedures. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 143–146.

[63] Frank TJM Zaal and Reinoud J Bootsma. 2011. Virtual reality as a tool for the study of perception-action: The case of running to catch fly balls. *Presence* 20, 1 (2011), 93–103.

[64] Mihaela A Zahariev and Christine L MacKenzie. 2003. Auditory, graphical and haptic contact cues for a reach, grasp, and place task in an augmented environment. In *Proceedings of the 5th international conference on Multimodal interfaces*. 273–276.