Effects of Different Auditory Feedback Frequencies in Virtual Reality 3D Pointing Tasks

Anil Ufuk Batmaz *
Simon Fraser University
BC, Canada

Wolfgang Stuerzlinger †
Simon Fraser University
BC, Canada

Abstract

Auditory error feedback is commonly used in 3D Virtual Reality (VR) pointing experiments to increase participants' awareness of their misses. However, few papers describe the parameters of the auditory feedback, such as the frequency. In this study, we asked 15 participants to perform an ISO 9241-411 pointing task in a distributed remote experiment. In our study, we used three forms of auditory feedback, i.e., C4 (262 Hz), C8 (4186 Hz) and none. According to the results, we observed a speed-accuracy trade-off for the C8 tones compared to C4 ones: subjects were slower, and their throughput performance decreased with the C8 while their error rate decreased. Still, for larger targets there was no speed-accuracy trade-off, and subjects were only slower with C8 tones. Overall, the frequency of the feedback had a significant impact on the user's performance. We thus suggest that practitioners, developers, and designers report the frequency they used in their VR applications.

Index Terms: Human-centered computing—Human Computer Interaction (HCI); Human-centered computing—Virtual Reality; Human-centered computing—Pointing;

1 Introduction

One of the purposes of 3D pointing studies in Virtual Reality (VR) or Augmented Reality (AR) is to understand how human performance changes with different input devices or methods, while they use a given visual display system [21]. Most of these studies use effective throughput to compare user input performance [21–23].

In such studies, sound is frequently used as a feedback mechanism, along with haptic and visual feedback [14, 29, 30]. Human hearing is sensitive to frequencies between 20 Hz to 20 kHz [19]. The upper and lower limits vary with different factors, including age, diseases, medications and even the environment [26], with 8 kHz considered to be audible to everyone (absent medical conditions) until 60 years of age. In the original Fitts' task, subjects performed the "selection" by tapping with a stylus on metal plates [13] while simultaneously seeing the target and feeling the impact on the fingertips. Moreover, hitting these metal plates also yielded auditory feedback, which helped subjects to perceive their successful task execution. Recent work in 3D pointing used auditory feedback to provide a stimulus to the participants for key events, such as when they "miss" a target [2, 5, 6, 27, 28]. However, researchers typically do not share detailed information on the type of auditory feedback they used in their work.

Recent studies revealed that when tones with different pitch are used for auditory error feedback, subjects' performance can be significantly impacted. Batmaz et al. [4] showed that when continuous tones with a higher pitch, i.e., C8, are used as auditory error feedback in a steering, subjects get uncomfortable as they hear this sound all the time. Thus, instead of focusing both on their speed and precision, participants typically prioritized making fewer mistakes [3, 4]. This work already demonstrated to some degree that it is possible to change or direct user performance by varying the pitch of auditory feedback, at least for continuous actions. Moreover, the same work showed that individuals can balance the speed-precision trade-off when they are exposed to C4 tones, i.e., middle C, where they also achieved their optimal performance [3, 4].

The above-mentioned approach [4] assessed user performance with continuous auditory feedback. Yet, in most 3D VR pointing studies, subjects hear auditory error feedback only when they made an error, i.e., when the subject performed a selection action outside of the target. The auditory feedback for such a miss needs to be short, since the participant typically already started moving the cursor to the next target. Also, Batmaz et al. [4] used a steering task which does not permit the use of throughput, a measure that combines of time, precision, and accuracy.

In this work, we investigated the following research question: *Is it possible to affect motor performance through middle and high pitches for auditory error feedback in Fitts' law pointing tasks?*

To answer these questions, we conducted a remote 3D VR pointing study using an ISO 9241-411 task [17]. Results showed an outcome similar to previous work, where the participants slowed down to reduce their error rate with high-pitch auditory error feedback [3, 4]. We hope that the results of this study encourage experimenters, researchers, and developers to report the parameters of the auditory feedback they used in their systems.

2 Previous Work

In this section, we discuss previous work on Fitts' law, effective throughput, and auditory error feedback.

2.1 Fitts' Law

Fitts' law models rapid aimed human movements [13]. Mackenzie's formulation for the Fitts' task is shown in Equation 1:

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

In Equation 1, movement time represent the time spent between two consecutive targets until the selection occurs. A is the distance between the two target centers and W is the width of a target. The coefficients a and b in Equation 1 are found by linear regression. ID represents the index of difficulty, which specifies the precision needed to execute the task [17].

2.2 Effective throughput

In this study, we follow the ISO 9241-411:2012 standard [17] and define the effective throughput as:

$$Throughput = \left(\frac{ID_e}{Movement\,Time}\right) \quad (2)$$

In Equation 2, movement time represents the task **execution time** and ID_e is the effective index of difficulty. According to ISO 9241-411:2012 [17], ID_e is the "user **precision** achieved in accomplishing

*e-mail: abatmaz@sfu.ca
†e-mail: w.s@sfu.ca, http://vvise.iat.sfu.ca

a task" [17]. In this paper, we also analyzed the effective index of difficulty, which expresses the precision of the participants, and is calculated as [17]:

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

In Equation 3, A_e represents the the effective distance which is the real distance traversed to execute the task, i.e., the distance between two selection points. W_e, on the other hand, is the effective target width, calculated as $W_e = 4.133 \times SD_x$, where SD_x is the standard deviation of selection coordinates along the task axis. SD_x expresses the **accuracy** of the task execution [22,23], as defined in MacKenzie and Isokoski's work [22]: we first measure the distance between each selection point and the target center and then calculate the standard deviation of these distances. Similar to the definition of precision in the ISO 9241-411 document, we consider the SD_x as the accuracy achieved by the users during the task execution.

When we look at Equation 2, we can see that the throughput based on effective measures combines time (movement time), precision (ID_e), and accuracy SD_x into a single equation, which incorporates all elements of the speed-accuracy trade-off. We analyze the effects of different forms of auditory feedback on user performance through measuring the throughput of the participants.

2.3 Auditory feedback

In the original Fitts' task, participants had to tap onto a metal plate, which provided a positive auditory feedback that the selection succeeded [13]. When the subjects missed the target, they tapped the stylus onto the table surface surrounding the target, which provided a different form of auditory feedback for a pointing error.

The effects of auditory feedback and its effect on user performance has been studied in various tasks in the past, e.g., [30]. Kotinnen et al. [20] studied a shooting task and showed that user performance can be increased through auditory feedback by mapping the frequency of the sound to the deviations in the trajectory. Another study on rowing-type movements showed that mapping the characteristics of human movements to the frequency and timbre of the sound can assist subjects in following a desired movement trajectories [29]. Other similar studies, such as [12,15,18], also used auditory feedback to improve the user performance by mapping key events to features of the played audio sound.

The effects of auditory feedback have also been analyzed for pointing tasks [1,34,35]. Their results showed that auditory cues can be useful to improve user performance in pointing tasks.

A recent study used continuous auditory error feedback during a VR steering task and changed the frequency of the sound depending on the deviation from the ideal path [3,4]. In that study, the authors choose a range of frequencies based on the piano keyboard, ranging between C1 and C8. The results showed that sound frequency had an impact on user performance and that pitch can affect the speed-accuracy trade-off in steering tasks in VR. The findings of that paper suggest that an optimal speed-precision trade-off is achieved with a C4 tone (262 Hz) and that participants pay more attention to their accuracy when they are exposed to C8 tones (4186Hz), which also increased their execution time [4]. However, this study did not test discrete auditory error feedback and did not analyze the speed-accuracy trade-off with throughput.

3 Motivation and Hypothesis

Our main goal for this study was to understand how pitch in audio feedback affects user performance during a 3D pointing task. From previous work, we know that error feedback can improve user performance, especially auditory error feedback [30, 34, 35]. On the other hand, a recent study showed that user performance can vary with different sound frequencies: subjects reached an optimal speed-precision trade-off at a middle C4 frequency, but with a higher C8 pitch they prioritized reducing the number of errors, which increased execution time [3, 10]. However, this previous work used a steering task with continuous auditory feedback. Thus, the impact of discrete auditory error feedback on the speed-precision trade-off of a pointing task, where the sound is audible only briefly after each error, is not yet established. In this study, we examine the following hypothesis: **user performance is negatively impacted with high pitch error feedback in 3D VR pointing studies.**

4 User Study

4.1 Subjects

We recruited 15 right-handed subjects (8 male and 7 female) with an average age of 29.05 $\pm$4.26 from a participant pool at the local university. Due to COVID-19, the experiment was conducted remotely. We only collected data from participants that had a computer able to run Steam VR on Windows computers. Six participants used a HTC Vive, four a HTC Vive Pro, three an Oculus Quest, and two participants used an Oculus Rift. We asked participants to wear headphones if their VR Head Mounted Display (HMD) did not have built-in speakers. Before starting the experiment, we also asked subjects to adjust the interpupillary distance of their headsets. One participant reported 0-2 hours of daily computer usage, another one 4-6 hours, six 6-8 hours, four 8-10 hours and three participant more than 10 hours. Two participants reported playing 3D mobile games 5-10 hours weekly and the rest 0-5 hours. Eight participants reported 0-5 hours of weekly computer game playing, four 5-10 hours, two 10-20 hours and one participant more than 20 hours.

4.2 Procedure

The experimenter connected remotely to the participants' computer through a video-conference application. Through this link, the experimenter monitored the progress of the study during the experiment, but did not record the participant's screen(s).

After the demographic questionnaire, the experimenter explained and demonstrated the experiment procedure to each participant.

In the virtual environment, subjects were placed at the center of a room with pictorial depth cues, as shown in Fig. 1. In this study, subjects performed a IS0 9241:411-2012 [17] experiment with eleven spherical targets for each round of trials. These targets were placed 40 cm away at the eye level of the subjects.

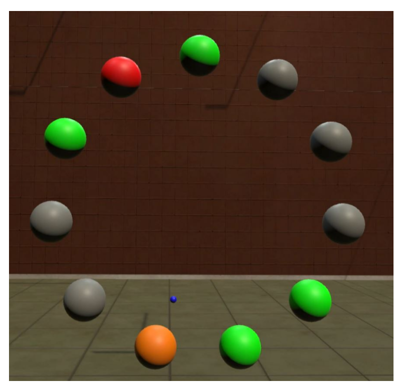


Figure 1: Experimental scene. Successfully selected targets are shown in green, missed targets in red, idle targets in grey, and the next target in orange. The cursor is shown in blue.

We used virtual hand selection technique and placed the 1 cm blue cursor 3 cm above the VR controller for it. The majority of the state-of-the art VR controllers are designed to be used with a power grip, which encouraged us to offset the cursor along the local axis of the VR controllers' main handle. For a selection, subjects had to place the cursor inside the target currently shown in orange in Fig. 1.

To confirm a selection, we asked subjects to press the space bar on the keyboard with their non-dominant hand to mitigate the effects of the "Heisenberg" effect [11].

The first target was randomly chosen by the software and participants executed the task either in a clockwise or counter-clockwise direction. We asked subjects to be "as fast and as precise as possible" when selecting targets.

When the cursor was inside a target, we changed the target color to blue to give visual feedback through highlighting [32]. If the participant selected the target (through pressing the spacebar) while the cursor was inside of the target, we changed the target color to green and registered the selection as a "hit". Otherwise, we changed the target color to red, played an error sound and registered the selection as a "miss". For data analysis, we used the center of the cursor position in 3D as the selection point.

When we played the error sound for a miss, we used three different forms of auditory feedback. One condition for the **auditory feedback** used a **C4** tone (262Hz), which is also known as the middle C on a piano. The second used a **C8** (4186 Hz), i.e., the highest C on a piano. In the third condition, which served as the baseline, we did not play any sound when the subjects made en error. We chose the above-mentioned frequencies based previous work [3, 4], where the speed-accuracy trade off was observed for the high frequency, but not for middle one. This choice also enables us to compare the results of this work with previous research that reported the used frequencies [10]. We counterbalanced the order of the three auditory feedback conditions with a Latin square.

We used three **target sizes** and two **target distances** comprising 6 unique IDs, see below. We chose target sizes and distances to enable us to compare our findings with other studies, such as [10]. The software randomly chose the ID for each round of trials.

During the experiment, we asked subjects to fix their computer volume to 40% and, if their VR HMD did not include built-in speakers, to wear headphones. Overall, the experiment lasted about 15 minutes for each participant.

4.3 Experimental Design

To analyze the effects of the different forms of auditory feedback, we used a two-way within-subjects design with three **auditory feedback** (AF_3 = No sound feedback, C4 = 262 Hz, C8 = 4186 Hz) conditions and 6 **ID**s (ID_6) comprising a AF_3 x ID_6 design. We used 6 unique IDs between 2.19 and 4.14 based on 2 target distances (TD_2 = 12.5, 25 cm) and 3 target sizes (TS_3 = 1.5, 2.5, and 3.5 cm). In total, we collected data from AF_3 x ID_6 x 11 times x 15 subjects = 2970 trials. For each, we collected participants' execution time (seconds), error rate (%), throughput (bits per second or BPS), SD_x (cm), and ID_e (bits).

For the analysis of the results, we used SPSS 24. We used Skewness (S) and Kurtosis (K) to analyze the normality of the dependent variables. We considered the data as normally distributed if the S and K were within ± 1.5 [16, 24]. When the data was not normal, we used the Aligned Rank Transform (ART) [33]. All the Repeated Measures (RM) ANOVA analysis results are shown in Table 1. We used the Bonferroni method for post-hoc analysis. The figures present the means and standard error of the means of the dependent variables.

5 User Study Results

5.1 Time Results

The time dependent variable was not normally distributed (S = 1.22, K = 2.30), so we used ART. The RM-ANOVA results are shown in Table 1 and Fig. 2(a). According to these results, subjects were slower with the C8 tone compared to the C4.

5.2 Error Rate Results

The error rate dependent variable was normally distributed (S = 1.11, K = 0.68). The RM-ANOVA results are shown in Table 1 and Fig. 2(b). According to these results, subjects made fewer errors with C8 auditory feedback compared to C4.

Table 1: RM ANOVA Results

	Auditory Feedback	ID	Auditory Feedback & ID
Time	$F(2, 28) = 5.05$ $p < 0.05$ η^2=0.265	$F(5, 70) = 31.070$ $p < 0.001$ η^2=0.689	$F(10, 140) = 4.628$ $p < 0.001$ $\eta^2 = 0.248$
Error rate	$F(1.305, 18.27) = 4.675$ $p < 0.05$ η^2=0.250	$F(5, 70) = 13.023$ $p < 0.001$ η^2=0.482	$F(10, 140) = 1.775$ $p < 0.001$ $\eta^2 = 0.113$
Throughput	$F(2, 28) = 6.052$ $p < 0.001$ η^2=0.302	$F(5, 70) = 11.120$ $p < 0.001$ η^2=0.443	$F(10, 140) = 0.715$ $p = 0.709$ $\eta^2 = 0.049$
SD_x	$F(2, 28) = 3.962$ $p < 0.05$ η^2=0.221	$F(5, 70) = 21.499$ $p < 0.001$ η^2=0.606	$F(10, 140) = 0.750$ $p < 0.001$ η^2=0.051
ID_e	$F(2, 28) = 3.707$ $p < 0.05$ η^2=0.209	$F(5, 70) = 85.143$ $p < 0.001$ η^2=0.859	$F(10, 140) = 1.389$ $p = 0.191$ η^2=0.090

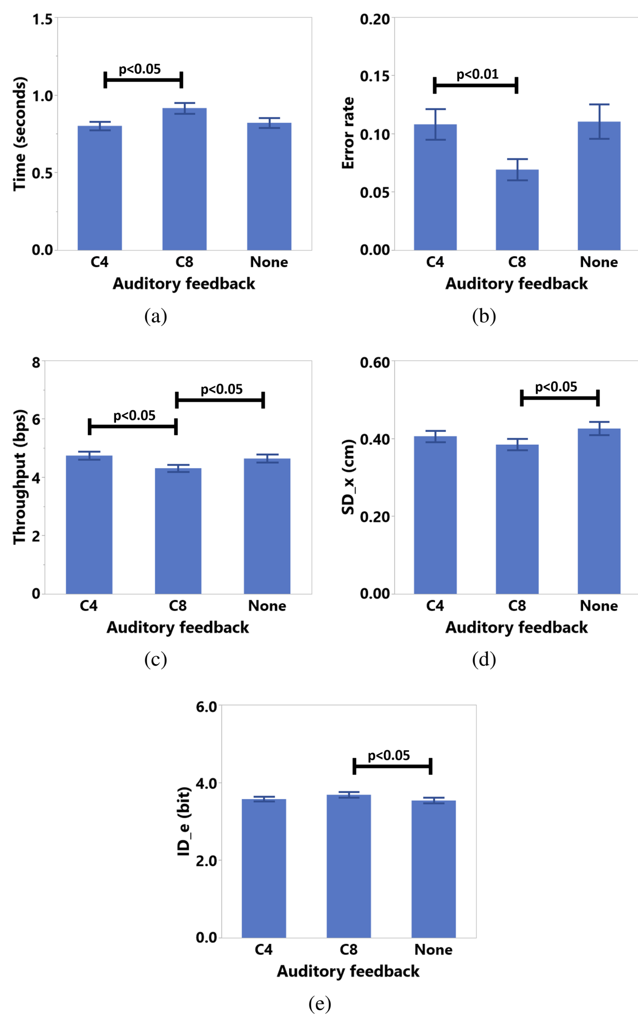


Figure 2: Results for auditory feedback for (a) time, (b) error rate, (c) throughput, (d) standard deviation, and (e) effective index of difficulty.

5.3 Throughput Results

The throughput dependent variable was normally distributed (S = 0.54, K = 0.22). The RM-ANOVA results are shown in Table 1 and Fig. 2(c). According to this measure, subjects' performance decreased with C8 tones compared to C4 or no auditory feedback.

5.4 Standard Deviation Results

The standard deviation dependent variable was normally distributed (S = 0.42, K = 0.09). The RM-ANOVA results are shown in Table 1 and Fig. 2(d). According to these results, subjects' accuracy decreased with C8 tones compared to no auditory feedback.

5.5 Effective Index of Difficult Results

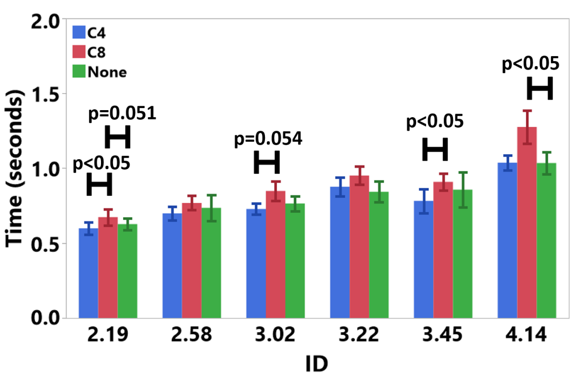
The effective index of difficulty dependent variable was normally distributed (S = 0.47, K = 0.05). The RM-ANOVA results are shown in Table 1 and Fig. 2(e). According to these results, the precision of the subjects increased with C8 auditory feedback compared to no auditory feedback.

5.6 Interaction Results

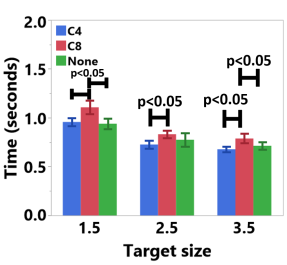
We found significant interactions between auditory feedback and ID for time, error rate, and SD_x, Table 1, and thus decided to further investigate these findings. Since the ID is the combination of the target size (W) and target distance (A), we analyzed our results with a three-way within-subjects ANOVA for auditory feedback, target size, and target distance, i.e., 3_{AF} x 3_{TS} x 3_{TD}.

5.6.1 Detailed analysis of time

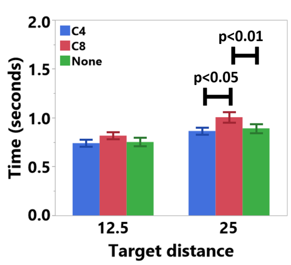
According to the results in Table 1, there is a significant interaction between auditory feedback and ID for time. These results are shown in Fig. 3(a). When we further analyzed these results, we found a significant interaction between auditory feedback and target size (F(4,56) = 2.56, $p < 0.05$, $\eta^2 = 0.155$), as shown in Fig. 3(b), and auditory feedback and target distance (F(2,28) = 2.944, $p < 0.05$, $\eta^2 = 0.174$), as shown in Fig. 3(c). According to these results, subjects were slower with C8 tones for all three target sizes. Subjects were also slower with C8 tones for the 25 cm target distance.



(a)



(b)

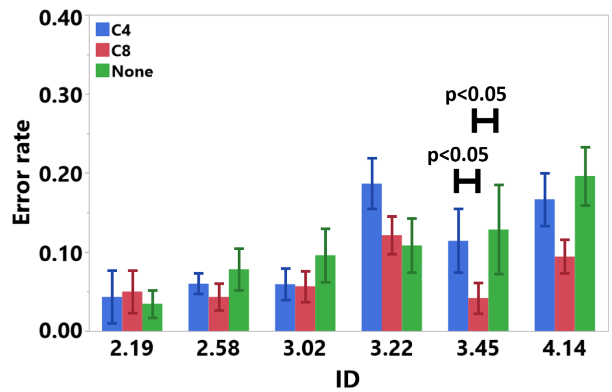


(c)

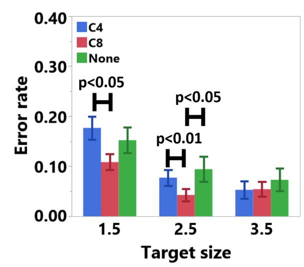
Figure 3: Detailed time analysis for interactions of (a) auditory feedback and ID, (b) auditory feedback and target size, and (c) auditory feedback and target distance.

5.6.2 Detailed analysis of error rate

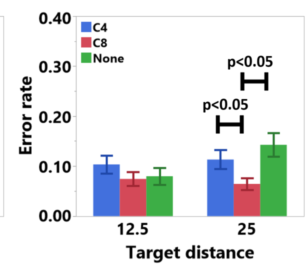
According to Table 1, we found a significant interaction between auditory feedback and ID for error rate. These results are shown in Fig. 4(a). When we further analyzed these results, we did not find a significant interaction between auditory feedback and target size (F(2.553, 35.747) = 1.324, p = 0.272, $\eta^2 = 0.086$), as shown in Fig. 4(b). Thus, we did not further analyze these results. However, we found a significant interaction between auditory feedback and target distance (F(2,28) = 4.891, $p < 0.05$, $\eta^2 = 0.259$), as shown in Fig. 4(c). According to these results, subjects made fewer errors with C8 tones for targets at 25 cm.



(a)



(b)



(c)

Figure 4: Detailed error rate analysis for interaction of (a) auditory feedback and ID, (b) auditory feedback and target size, and (c) auditory feedback and target distance.

5.6.3 Detailed analysis of SD_x

According to the results in Table 1, there is a significant interaction between auditory feedback and ID for SD_x. These results are shown in Fig. 5(a). When we further analyzed these results, we did not find a significant interaction between auditory feedback and target size (F(4, 56) = 1.324, p = 0.893, $\eta^2 = 0.019$), as shown in Fig. 5(b), and auditory feedback and target distance (F(2,28) = 1.798, p = 0.184, $\eta^2 = 0.114$), as shown in Fig. 5(c). Since both detailed interaction analyses yielded no significant result, we do not further analyze and discuss these results.

5.7 Fitts' law analysis

Fitts' law analysis based on Equation 1 revealed that the movement time can be modeled as MT = 0.312 + 0.21 ID, $R^2 = 0.88$ for C4 tones, MT = -0.00025 + 0.29 ID, $R^2 = 0.91$ for C8, and MT = 0.19 + 0.19 ID, $R^2 = 0.97$ for no auditory feedback, as shown in Fig. 6.

6 Discussion

In this work, we analyzed user performance in a 3D VR pointing experiment with the virtual hand interaction technique and three

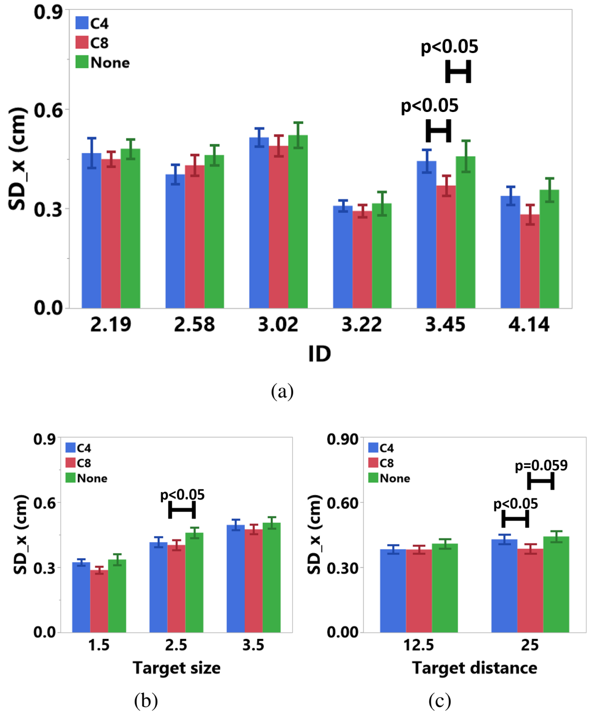


Figure 5: Detailed standard error analysis for interaction of (a) auditory feedback and ID, (b) auditory feedback and target size, and (c) auditory feedback and target distance.

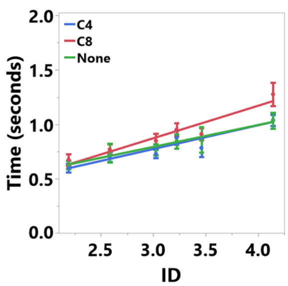


Figure 6: Fitts' law analysis results for the auditory feedback conditions.

different forms of auditory feedback.

Our results on time, error rate, and throughput show that when a high-frequency pitch is used as auditory error feedback, user performance significantly decreased in terms of time and throughput. These results support our hypothesis, that **user performance is negatively impacted with high pitch error feedback in 3D VR pointing studies**. In contrast, a higher pitch significantly reduced the error rate, and accuracy and precision were significantly increased. This finding is not in line with our hypothesis and we speculate that this is a consequence of the speed-accuracy trade-off inherent in Fitts' Law studies.

We speculate that, as in previous work [3, 4, 10], subjects were trying to be careful not to make mistakes so they could avoid hearing the C8 error feedback. Thus, they prioritized the error rate over execution time and also sacrificed their overall pointing throughput. We also surmise that subjects found this frequency (slightly) irritating, and were actively trying to avoid this sound.

Our more detailed analysis of the interactions showed that subjects were slower with C8 tones for all target sizes and with targets at larger distances, i.e., 25 cm. When we looked at the detailed analysis, we saw that error rate of the participants did not decrease with the 3.5 cm target size for C4 frequency. In other words, while the participants were slower with wide targets, their error rate did not decrease with the C8 tones compared to other forms of auditory feedback. Specifically, we did not observe any speed-accuracy trade-off for wide targets, and the C8 pitch only made participants slower with the 3.5 cm sized targets.

These results show not only that auditory feedback has an impact on user performance, but also document that its features, such as pitch of the tone, have an impact. We thus suggest that researchers, developers, and practitioners report the frequency of their auditory feedback to increase the reproducability of their work. Also, when the findings of two different studies that are otherwise very similar do not match, we suggest that researchers consider different forms of auditory feedback as a potential reason.

7 Limitations

Even though we analyzed 3D pointing user performance in a VR environment, we only investigated the virtual hand selection technique. Although the virtual hand technique is one of the most used interaction techniques, there are other techniques, such as ray casting. Still, we do not expect major differences in terms of outcomes, unless the pointing technique is prone to other limitations, such as jitter [6–9].

Due to COVID-19, we had to conduct the experiment in a remote, distributed manner. We followed the suggestions by Steed et al. [31]. The only requirement to run the experiment was to have access to a VR HMD and a PC running Steam VR software. Thus, participants used the hardware available in their home, which meant they used different VR HMDs and graphics cards.As we still observed consistent outcomes, we believe that collecting data with different VR systems increased the external validity of our results. Since the number of headsets used in this work was not evenly distributed, we did not include the VR headsets as a dependent variable; as the low number of individual headsets does not support detailed analysis. We acknowledge that the use of a variety of headsets added more noise to the collected data. Still, we did not observe any major differences in the collected data from different headsets and systems, and so are fairly confident that distributing the experiment did not affect our main results.

The sound frequencies used in this study are well within the human hearing range [25]. Before the experiment, the experimenter asked for verbal confirmation that participants could hear the error sounds. However, due to restrictions imposed by COVID-19, we could not verify participants' perception in a more objective manner.

Unlike previous work on steering tasks [3], we did not investigate all different tones that a keyboard can produce, i.e., the full range from C1 to C8. Thus, the results in this paper only support the previous findings within a small range of frequencies, but still extend the results for pointing tasks. The effect of the full range of pitch in auditory error feedback still needs to be further investigated in 3D pointing studies.

8 Conclusion & Future Work

In this paper, we performed a VR pointing study with the ISO 9241-411 [17] task with three forms of auditory feedback, i.e., a C4 tone (262 Hz), a C8 (4186 Hz), and no audio. The results showed that higher pitch for auditory error feedback, e.g., a C8, can reduce the error rate of the subjects but also increases their execution time and decreases the throughput. In other words, when we used a higher pitch, we observed a speed-accuracy trade-off in the results. Thus, we recommend that software developers, practitioners, and researcher carefully choose the auditory feedback that they will play during their study and also to report the parameters for their auditory feedback in their publications to enhance comparability of results. Moreover, we also recommend using middle frequency sounds, such as a C4, for auditory error feedback to reach the optimal speed-accuracy trade-off in 3D pointing studies.

In the future, we want to extend this work to different interaction techniques and other features of sound, such as the volume, length, timbre, timing, and quality. Furthermore, the results here only apply for discrete auditory error feedback, so the findings here could also be extended to continuous auditory error feedback during VR pointing experiments.

REFERENCES

[1] M. Akamatsu, I. MacKenzie, and T. Hasbroucq. A comparison of tactile, auditory, and visual feedback in a pointing task using a mouse-type device. *Ergonomics*, 38(4):816—827, April 1995. doi: 10.1080/00140139508925152

[2] M. D. Barrera Machuca and W. Stuerzlinger. Do stereo display deficiencies affect 3D pointing? In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, p. LBW126. ACM, 2018.

[3] A. U. Batmaz. *Speed, precision and grip force analysis of human manual operations with and without direct visual input*. PhD thesis, Strasbourg, 2018.

[4] A. U. Batmaz, M. de Mathelin, and B. Dresp-Langley. Different sound pitch effects on motor performance of individuals in head-mounted virtual reality. In *Perception*, vol. 48, pp. 21–21, 2019.

[5] A. U. Batmaz, M. D. B. Machuca, D. Pham, and W. Stuerzlinger. Do head-mounted display stereo deficiencies affect 3D pointing tasks in AR and VR? In *Conference on Virtual Reality and 3D User Interfaces*, VR '19, pp. 585–592. IEEE, Mar 2019. doi: 10.1109/VR.2019.8797975

[6] A. U. Batmaz, A. K. Mutasim, and W. Stuerzlinger. Precision vs. power grip: A comparison of pen grip styles for selection in virtual reality. In *Conference on Virtual Reality and 3D User Interfaces, Abstracts and Workshops, Workshop on Novel Input Devices and Interaction Techniques at VR*, NIDIT '20, pp. 23–28. IEEE, Mar 2020. doi: 10.1109/VRW50115.2020.00012

[7] A. U. Batmaz, M. R. Seraji, J. Kneifel, and W. Stuerzlinger. No jitter please: Effects of rotational and positional jitter on 3D mid-air interaction. In *Future Technologies Conference*, vol. AISC 1289 of *FTC '20*. Springer, Nov 2020. doi: 10.1007/978-3-030-63089-8_52

[8] A. U. Batmaz and W. Stuerzlinger. The effect of rotational jitter on 3D pointing tasks. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, CHI EA '19. ACM, May 2019. doi: 10.1145/3290607.3312752

[9] A. U. Batmaz and W. Stuerzlinger. Effects of 3D rotational jitter and selection methods on 3D pointing tasks. In *Workshop on Novel Input Devices and Interaction Techniques at VR*, NIDIT '19, pp. 1687–1692. IEEE, Mar 2019. doi: 10.1109/VR.2019.8798038

[10] A. U. Batmaz and W. Stuerzlinger. The effect of pitch in auditory error feedback for fitts' tasks in virtual reality training systems. In *Conference on Virtual Reality and 3D User Interfaces*, VR '21. IEEE, Mar 2021.

[11] D. Bowman, C. Wingrave, J. Campbell, and V. Q. Ly. Using pinch gloves (tm) for both natural and abstract interaction techniques in virtual environments, 2001.

[12] E. O. Boyer, B. M. Babayan, F. Bevilacqua, M. Noisternig, O. Warusfel, A. Roby-Brami, S. Hanneton, and I. Viaud-Delmon. From ear to hand: the role of the auditory-motor loop in pointing to an auditory source. *Frontiers in computational neuroscience*, 7:26, 2013.

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

[14] M. Funk, J. Heusler, E. Akcay, K. Weiland, and A. Schmidt. Haptic, auditory, or visual? towards optimal error feedback at manual assembly workplaces. In *Proceedings of the 9th ACM International Conference on PErvasive Technologies Related to Assistive Environments*, pp. 1–6, 2016.

[15] A. Godbout and J. E. Boyd. 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.

[16] J. F. Hair Jr, W. C. Black, B. J. Babin, and R. E. Anderson. Multivariate data analysis, 2014.

[17] ISO 9241-411:2012. Ergonomics of human-system interaction – part 411: Evaluation methods for the design of physical input devices. ISO, May 2012.

[18] G. Jakus, K. Stojmenova, S. Tomažič, and J. Sodnik. A system for efficient motor learning using multimodal augmented feedback. *Multimedia Tools and Applications*, 76(20):20409–20421, 2017.

[19] M. Kleiman-Weiner and J. Berger. The sound of one arm swinging: a model for multidimensional auditory display of physical motion. In *International Conference on Auditory Display*. Georgia Institute of Technology, 2006.

[20] N. Konttinen, K. Mononen, J. Viitasalo, and T. Mets. The effects of augmented auditory feedback on psychomotor skill learning in precision shooting. *Journal of Sport and Exercise Psychology*, 26(2):306–316, 2004.

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

[22] I. S. MacKenzie and P. Isokoski. Fitts' throughput and the speed-accuracy tradeoff. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '08, p. 1633–1636. Association for Computing Machinery, New York, NY, USA, 2008. doi: 10.1145/1357054.1357308

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

[24] P. Mallery and D. George. SPSS for windows step by step: a simple guide and reference. *Allyn, Bacon, Boston,*, 2003.

[25] B. Masterton, H. Heffner, and R. Ravizza. The evolution of human hearing. *The Journal of the Acoustical Society of America*, 45(4):966–985, 1969.

[26] J. H. Mills and J. A. Going. Review of environmental factors affecting hearing. *Environmental health perspectives*, 44:119–127, 1982.

[27] A. K. Mutasim, W. Stuerzlinger, and A. U. Batmaz. Gaze tracking for eye-hand coordination training systems in virtual reality. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, CHI EA '20. ACM, Apr 2020. doi: 10.1145/3334480.3382924

[28] D. Pham and W. Stuerzlinger. Is the pen mightier than the controller? a comparison of input devices for selection in virtual and augmented reality. In *25th Symposium on Virtual Reality Software and Technology*, VRST '19. ACM, Nov 2019. doi: 10.1145/3359996.3364264

[29] R. Sigrist. Visual and auditory augmented concurrent feedback in a complex motor task. *Presence*, 20(1):15–32, 2011. doi: 10.1162/pres_a_00032

[30] R. Sigrist, G. Rauter, R. Riener, and P. Wolf. Augmented visual, auditory, haptic, and multimodal feedback in motor learning: a review. *Psychonomic bulletin & review*, 20(1):21–53, 2013.

[31] A. Steed, F. R. Ortega, A. S. Williams, E. Kruijff, W. Stuerzlinger, A. U. Batmaz, A. S. Won, E. S. Rosenberg, A. L. Simeone, and A. Hayes. Evaluating immersive experiences during covid-19 and beyond. *Interactions*, 27(4):62–67, Jul 2020. doi: 10.1145/3406098

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

[33] J. O. Wobbrock, L. Findlater, D. Gergle, and J. J. Higgins. The aligned rank transform for nonparametric factorial analyses using only ANOVA procedures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pp. 143–146. ACM, New York, NY, USA, 2011. doi: 10.1145/1978942.1978963

[34] M. A. Zahariev and C. L. MacKenzie. 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*, ICMI '03, p. 273–276. Association for Computing Machinery, New York, NY, USA, 2003. doi: 10.1145/958432.958481

[35] M. A. Zahariev and C. L. Mackenzie. Auditory contact cues improve performance when grasping augmented and virtual objects with a tool. *Experimental brain research*, 186(4):619–627, 2008.