

Exploring the Use of Auditory Feedback as a Guide for 3D Drawing in Extended Reality

E. Gruy¹  and F. Berthaut² 

¹CRISAL, University of Lille, France

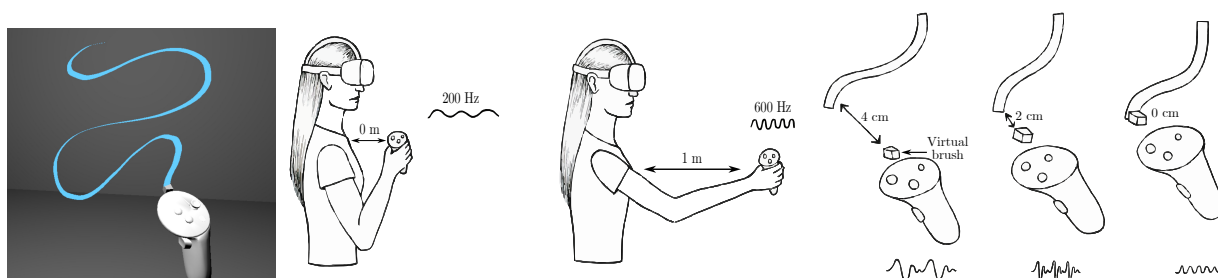


Figure 1: View of the virtual environment from the headset and illustration of the different auditory feedback presented in this article. From left to right: view from the headset when drawing, changes in frequency for the DEPTH feedback and changes in amplitude modulation for the PROXIMITY feedback.

Abstract

3D drawing (or sketching) in Extended Reality (XR) is more difficult to master than 2D drawing. It is performed in mid-air and does not benefit from a physical surface, has more degrees of freedom, and is often dependent on Head-Mounted Displays that contain inaccurate stereoscopic rendering (due to the vergence-accommodation conflict). Many approaches were explored in the scientific literature in order to compensate for these issues, such as visual guides, the use of haptic feedback, or beautifying techniques. However, very few focus on the use of auditory feedback as a possible guide for increasing spatial awareness, and thus accuracy, while drawing inside a 3D space, although sound has proven to be a useful additional feedback in several 3D interactive contexts. In this paper, we explored several auditory feedback (based on the depth of the controller and the proximity to other strokes in the canvas) to try and see if it improves the accuracy of drawings in XR. To do so, we conducted an experiment with 21 participants, mostly novices in drawing. The results show that the use of the proposed auditory feedback do not have an influence on the accuracy of the drawings, but do have an influence on the participants' perceived accuracy and the confidence in their ability to perform the task.

CCS Concepts

• **Computing methodologies** → **Mixed / augmented reality; Virtual reality;** • **Human-centered computing** → **Usability testing; Sound-based input / output;**

1. Introduction

3D drawing (or sketching), used in many contexts such as architectural design, industrial design, art production, installations and performance, now often relies on Extended Reality (XR) technologies, more specifically on Head-Mounted Displays (HMD), as they allow a user to be fully immersed inside a 3D space within which drawings can be freely manipulated and explored. However, these stereoscopic displays have limitations. The rendered environment can not fully replicate the physical world, as the displays create a

conflict between vergence and accommodation [HGAB08]). This can lead to inaccurate spatial cues, and therefore alter depth perception inside the 3D environment [BMS19, BMPS19]. In the case of 3D drawing, where spatial cues are an important component in the interaction, this can negatively influence the quality of the drawings.

3D drawing with XR controllers also induce several challenges. First, 3D drawings are performed in mid-air rather than on a surface, as it is the case for 2D drawing, which decreases the strokes

accuracy [AKA*17]. Also, it requires higher sensorimotor and cognitive (specifically visual-spatial cognition [TC24]) demands from a user, due to the increase in degrees of freedom [WIMB10]. Some specificities, such as the spatial ability of a user, can also affect the shape likeness of the drawings [BMSA19].

Prior work have explored ways to compensate for these issues and improve 3D drawing. They range from visual guides [MSA19, TPBM*22], to the use of haptic feedback [EBMS*20, KZL07, DGK*20], beautifying algorithms [MAS*18, YDSG21, YAS*21], or different input/interaction techniques [DSK18, JZF*21]. Very few have, however, looked into the use of auditory feedback to guide the drawings by increasing spatial awareness, although it has proven to be useful in other 3D interactive contexts [LG05, GWFL22, MK97].

Sound can have an influence on one's body representation, as well as their emotional state [LFAS*22, SS20]. It can enhance motor learning [EFS*16] (especially in people who have a musical background [LMM*22]), or influence the execution of a movement [BBG*20, LFAS*22, BBB*13]. In tasks such as, for example, 3D navigation in virtual environments [LG05], improving effective throughput [BYLS22], or in trajectory-based finger gestures [GKL*18], auditory feedback has shown to be effective in improving one's performance. Moreover, some studies have found that auditory guidance can increase presence [RBRG21], and that a higher sense of presence can have beneficial effects on performance [GLRS20]. When a task complexity gets higher, auditory feedback can also serve as an additional information source (thus using two sensory channels) for a user to rely on, as a backup [MCM*16]. It is especially interesting for 3D drawing, since different shape complexities can be achieved, for example comparing a single square or cube to a detailed 3D object.

In addition to that, auditory feedback for 3D drawing can have other advantages. It does not overload the visual channel, which could, in the case of complex drawings, be an advantage in comparison to visual guides. Unlike haptic feedback devices, it allows for the use of hands or alternative controllers which do not provide vibrotactile feedback. Finally, it is not intrusive nor constraining, as it does not modify the appearance of strokes, which is the case with beautifying algorithms. The latter is important since in the case of expressive drawing, users might not appreciate to lose control over their work, although this can be the opposite for architectural design or other disciplines that necessitate precise and/or geometrical shapes. When comparing potential sonification strategies for 2D and 3D drawing, the main differences between them lie in the difficulties posed by 3D drawing, since the guidance is mostly used to compensate for aspects of drawing that do not exist within a 2D space (for example depth perception issues and plane deviations when connecting strokes). Therefore, this article aims at exploring different auditory feedback to improve the accuracy of 3D drawing, by giving users additional spatial cues. This could be the first step in designing a non-constraining, multimodal system that uses auditory feedback in addition to the visual feedback, with applications that are either expressiveness-oriented or precision-oriented.

1.1. Contributions

Our contributions are two-fold:

1. We propose two techniques for guiding 3D drawing through auditory feedback on the depth and proximity of the virtual brush.
2. We evaluate their impact on the user experience and provide insights on the use of auditory feedback for drawing in XR.

2. Related Work

This paper delves into themes such as 3D assisted mid-air drawing and audio guidance in 3D interaction in XR. The next subsections detail the related work for each topic.

2.1. 3D Assisted Mid-air Drawing

In order to assist/guide 3D drawing in XR, several techniques have been previously explored. Firstly, specific input or interaction techniques have been designed on several occasions. Dudley *et al.* [DSK18] propose an Augmented Reality (AR) application aimed at industrial design that works bare-handed. Three input techniques are available, one freehand, and the other two depending on control points. Another application that uses a specific interaction technique is HandPainter [JZF*21]. It relies on a hand-based physical proxy to switch between 3D mid-air drawing to 2D drawing on a surface. Here, one hand serves as a canvas, and the index of the other one is used as a pen, to limit the inaccuracies induced by mid-air drawing. There is also a beautifying algorithm applied to the strokes.

Other projects rely on beautifying techniques, such as Multiplanes [MAS*18], ScaffoldSketch [YDSG21] and CASSIE [YAS*21]. Multiplanes beautifies the strokes either in real-time or with trigger points, and also provides some visual guides. ScaffoldSketch comprises two modes with different algorithms to alter the strokes: one more specifically designed for drawing lines (*i.e.*, the scaffolding), Scaffold, and the other for drawing curves, Sketch. Finally, CASSIE uses 3D sketching to create a mesh as a curve network, with it being beautified without changing the original strokes too much. All the projects listed above are made with industrial design or geometrical sketching in mind (if we omit HandPainter), and are not really suited for expressive drawing.

It is also possible to find projects that explore visual guides as a way to improve the shape-likeness of drawings without altering the expressiveness of the strokes. For example, Smart3DGuides [MSA19] gives the user indications regarding the depth and the orientation of the controller to reduce the mistakes in sub-action planning. Turkmen *et al.* [TPBM*22] also explore the influence of discrete and continuous visual guides on drawing accuracy. However, visual guides would probably work best in environments where the drawings are simple (Turkmen *et al.* used simple shapes for the experiment and Smart3DGuides is made for novices), otherwise it might overload the visual channel.

Finally, a number of research has been conducted on the use of haptic feedback. VRSketchIn [DGK*20] employs a tracked pen and tablet to perform 3D drawing on a constrained 2D surface, hence providing passive haptic feedback. VRSketchPen [EBMS*20] presents two haptic modalities, with either pneumatic force feedback to simulate contact on a surface, or vibrotactile feedback to simulate textures. It is designed to be non-constraining

and thus preserves the expressiveness of drawing. Drawing on Air [KZL07] is designed for 3D curve drawing in mid-air and uses a SensAble Phantom force feedback device to perform the interaction, with two techniques available, using either one or two hands. These projects, however, require specific equipment (whether it is for the device in itself or its tracking) that can be expensive and/or complicated to get outside of a research context. Auditory feedback, on the other hand, only requires speakers, making it more widely available.

Many of the projects listed above are designed with architectural/industrial design or geometrical drawing in mind (because of the constraints it implies), in opposition to a more expressive/artistic approach. Also, few are non-constraining, non-invasive and easy to set up without requiring specific equipment. This is what we aim to do in this article, by guiding 3D drawing with an auditory feedback.

2.2. Audio Guidance for 3D Interaction in XR

Auditory feedback has been used in several occasions as a way to enhance 3D interaction in XR. For example, Reynaert *et al.* [RBRG21] look into the effect of rhythm on fatigue in mid-air gestures in Virtual Reality (VR), by testing different sound conditions. They show that the sound can have an effect on the perceived fatigue, difficulty, and on the sense of presence. Other than rhythm, Gao *et al.* [GWFL22] study spatial auditory cues for guidance tasks in immersive virtual environments. They specifically target elevation, as the sound localization rendering has a limited accuracy.

Additionally, auditory feedback is often considered as a way to compensate for another sense, typically vision, by giving a user additional spatial cues. For instance, May *et al.* [MSWW19] explore several types of auditory feedback for object targeting in XR with limited visibility. Their experiment showed improvements in both performance and subjective workload. Guarese *et al.* [GPR*24] study sonification methods in AR, this time without a visual feedback, for vertical guidance. They also detail different scenarios in which this technology could be beneficial. In a similar manner, Lokki and Gröhn [LG05] find that 3D navigation in a virtual environment is possible with only auditory cues, although the best results are obtained with both the visual and auditory feedback. Finally, for AR drawing with a narrow field of view, Xu *et al.* [XLHT22] use auditory feedback to help prevent shape deviations when drawing in mid-air with a mobile device. Their technique significantly reduces the effort and mental demand required by the task. This last project is the most similar to us, although the contexts are different: we use a HMD and not a mobile device, effectively removing the field of view issue. This leads to different interaction techniques with the system, and thus different needs in terms of auditory feedback.

3. Experiment

To test if the use of auditory feedback can increase one's precision when drawing in 3D, we designed an experiment, which is described in the following subsections.

This experiment used a Meta Quest 2 HMD and one controller

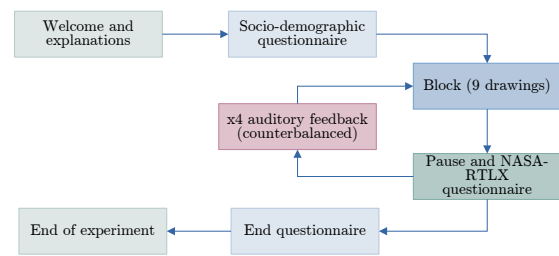


Figure 2: Detailed experimental procedure for each participant, with counterbalanced auditory feedback blocks (balanced latin square) between participants to avoid fatigue/learning effects.

(left or right depending on the participants' dominant hand) for drawing and navigating through the virtual environment. The 3D drawing application was designed with the Godot game engine version 4.4.1, as it was easy to add functionalities according to our needs. The strokes were ribbon-shaped, fully opaque, a pre-defined and non modifiable colour, and unshaded, as seen in Figure 1.

The audio part of the experiment was designed using Puredata, with the IVMI_builder Godot plugin [Ber25] to deploy directly on the HMD. This way, the whole of the application could be run on the HMD without having to rely on an additional computer to display the auditory feedback to the participants. We therefore used the HMD speakers rather than headphones. The experiment took place at a lab, in a silent room to avoid auditory disturbances, and usually lasted around an hour. The experimental procedure is shown in Figure 2.

We chose to conduct the experiment in VR instead of Mixed Reality (MR), as VR tends to suffer from a distorted egocentric depth perception [KW08, RVH13], which is less the case with MR devices [JSS*08, SSE15].

3.1. Hypotheses

Based on the literature, we made the two following hypotheses:

- H1: Using an auditory feedback to guide the movements will increase the strokes accuracy ([XLHT22, GWFL22]). More specifically, the DEPTH feedback will improve the planarity of the drawings on the XY-plane, and the PROXIMITY feedback will increase the matching of two lines and reduce their degree of deviation.
- H2: With increased complexity in the shapes to reproduce, users will rely more on both feedback (visual and auditory) at the same time ([MCM*16]).

3.2. Auditory Feedback

As stated in the introduction, 3D drawing in XR can induce several challenges, including an altered depth perception and inaccurately rendered spatial cues. On this basis, two auditory feedback types (DEPTH and PROXIMITY) were tested during this experiment, both separately and in combination with one another. Therefore, the participants were confronted to 4 different conditions:

1. **NO SOUND:** No sound was produced. This condition was the baseline.
2. **DEPTH:** The auditory feedback (a sine wave) was linked to the position of the controller along the Z-axis relative to the participant, *i.e.*, in front of them. The closer the controller was to the participant's body, the lower was the frequency, and vice versa (see Figure 1). We chose pitch to represent the depth, since it is the most used parameter in sonification studies [DB13], and has shown to improve the accuracy of 3D drawing in a different context [XLHT22].
3. **PROXIMITY:** The auditory feedback (a sine wave at a 440 Hz frequency with amplitude modulated by a Low Frequency Oscillator (LFO) ranging from 1Hz to 15Hz) was linked to the proximity with other strokes in the environment. If the controller passed the proximity threshold, the sound was activated, and the closer the participant got to a stroke, the faster the amplitude modulation became, until the sound was continuous (meaning they were touching the strokes, see Figure 1). The application determined the closest point to the controller by calculating the distances with each point of a stroke until it found the smallest (*i.e.*, iterating over the points of each stroke and computing the Euclidean distance to the controller, which did not induce performance drops due to the strokes' simple geometry). We chose amplitude modulation to represent the proximity to other strokes, since it has shown to have a positive effect on object targeting [MSWW19], and could therefore help to compensate for the higher visual-spatial cognitive demands of 3D drawing, more specifically for connecting strokes.
4. **DEPTH+PROXIMITY:** A combination of the two conditions above. The frequency is then mapped to the distance to the participant, while the amplitude modulation depends on the proximity to other strokes.

After adjusting the frequency range during the conception phase of the experiment, the DEPTH frequency was set between 200 and 600 Hz, as the auditory feedback can alter the participants' performances [BYLS22], and therefore a frequency that is too high could be considered annoying and negatively impact the results. As the participant is placed at the center of the virtual environment (at a $(0, y, 0)$ point, with y being their height while sitting), and considering a maximum arm extension of around 0.65 meter, the mapping function for the DEPTH is the following: $frequency = 200 + (dist / -0.65 * 400)$, with $dist$ representing the controller's distance on the Z-axis relative to the participant's position (Godot's coordinate system being right-handed).

The PROXIMITY threshold was set to a distance that was quite low (approximately 4 cm inside the VE) as to not confuse the participants, since there were no visual cues to signal which point they were the closest to and they could only rely on the auditory feedback. Too big of a threshold could have become ambiguous if many strokes were drawn, as the auditory feedback would have been activated at a distance were the closest point would be impossible to visually identify. The amplitude was modulated by a sine wave for which the frequency was defined by $frequency = (15 - (min_dist / 0.04) * 15) (min_dist \text{ being the closest point})$, and the modulation gain as $gain = (output + 1) / 2 * 0.5 + 0.5$ (with output being the output of the sine wave).

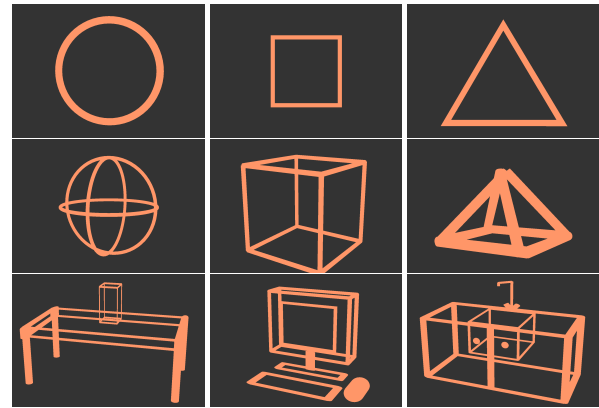


Figure 3: All the drawings participants had to reproduce. From top left to bottom right: circle, square, triangle, sphere, cube, pyramid, table, computer, kitchen sink.

Finally, the audio was not permanent and only activated when touching (without the need to press/start drawing) the drawing button of the controller, to allow the participants to take breaks. The volume was calibrated during the design phase of the experiment so that it was loud enough with the headset, and participants were told they could lower the volume a bit if it was uncomfortable.

The above auditory feedback were chosen because they were clear (pure sine waves) and different enough from one another that it was simple to differentiate them. However, one could imagine applying similar modulations (*e.g.*, changes in pitch and amplitude) using richer sound such as instrumental sounds generated with samples, granular synthesis or physical modelling.

3.3. Experiment Design

During the experiment, the participants had to complete 4 blocks composed of 9 drawings. Each drawing had to be reproduced according to a given model. The drawings had different complexities: three were simple 2D shapes (a circle, a triangle and a square), three were simple 3D shapes (a sphere, a pyramid and a cube) and the last three were more complex 3D shapes (a kitchen sink, a computer and a table with a vase) (see Figure 3). The model was visible when pressing a button and disappeared when releasing it. When it was visible, it was placed around one meter in front of the participants, slightly above them, to discourage them from trying to retrace it. In addition to that, it was not possible to draw when the model was visible, and vice versa. The participants did not have the possibility to erase their strokes, as a way to encourage them to stay focused on the task.

We decided to test out different shape complexities, in 2D and 3D, to see if the participants relied more on the audio feedback when confronted to more complex shapes (in addition to the visual feedback), and if this had an impact on the results. Each block had a different auditory feedback, selected using a balanced latin square (therefore, participants did not have the auditory feedback in the same order). The participants had a training phase at the beginning of a block, where they could experiment with the system and the

auditory feedback, before moving on with the drawings. They had no time constraints overall, and were told to take their time to be as precise as possible (since the instructions can have an effect on the results [PGA*16]). However, they were still timed throughout the experiment.

In order to limit the bias due to movements which would influence spatial perception [KLB*98, BMSA19] and would be difficult to control, participants had to complete the experiment while sitting on a chair and were asked to move as little as possible (with the exception of looking at the models). Between each block, participants were asked to remove the HMD for a 2 minutes break, to avoid developing arm fatigue and to answer a questionnaire.

3.4. Collected Data and Evaluation

The data collected during the experiment, and the subsequent evaluation strategies, are described in the following sections. It is to note that we decided not to measure the spatial ability of the participants [KH01], since we want to see if the auditory feedback has an improvement on the accuracy of the drawing no matter the base level of the participant, and hence their spatial ability.

3.4.1. Socio-demographic Data

The participants were asked their age, gender, dominant hand, XR experience, the frequency at which they drew on a regular basis, and a self evaluation of their drawing level.

3.4.2. NASA-RTLX

In between blocks, the participants had to answer a raw NASA-RTLX (NASA-RTLX) questionnaire [HS88], to be able to perform a comparison between the auditory feedback at the end. We did not use the weighted factors, since we are not interested in obtaining a global score at the end, but rather want to look at each factor individually. We settled on the NASA-RTLX and not the System Usability Scale (SUS) [B*96], since we wish to investigate how the auditory feedback impacts the users' task load, and not the perceived usability of the system.

3.4.3. Drawing Accuracy

All the 36 drawings performed by each participant were saved as GLTF objects to be evaluated afterwards. To do so, we used a simple 3D objects viewer that allowed zooms and rotations of the drawings. For the evaluation framework [MIKS24], we chose to base ours on the one designed by Wiese *et al.* [WIMB10] (line straightness, matching of two lines, degree of deviation, corrective movements). We, however, added supplemental evaluation categories (planarity and resemblance to the model), detailed on Table 1. We settled for a manual evaluation framework, because some models were not suited for an automatic review process (typically the complex 3D shapes, since deviations from the model did not always imply a low drawing quality).

As per Wiese *et al.* [WIMB10], each drawing was evaluated according to the different categories by the first author (without hearing to the auditory feedback). Each category was marked on a 1 to 3 scale, with 1 being good quality and 3 being bad quality. The

category scores (except for the resemblance to the model) were averaged according to the amount of elements present on the drawing (strokes, connections between the strokes), with each element having its own score, to obtain a more nuanced rating. For the XY, YZ and ZX planes, as 2D drawings are only planar on one dimension, the scores for the XY, YZ and ZX dimension for the 3D drawings were averaged into one plane value. This way, the planarity scores were aligned between 2D and 3D drawings. To get the global drawing accuracy, all the scores were summed up into one value, that could therefore range from 6 (very good quality) to 18 (very bad quality).

3.4.4. Drawing Completion Time

The time taken by the participants to complete the drawings (from the first drawn stroke until they moved on to the next drawing) was collected in order to evaluate if there were differences between the auditory feedback.

3.4.5. End Questionnaire

At the end of the session, the participants were asked several questions about how they perceived the auditory and the visual feedback in relation to the complexity of the drawings to reproduce (*i.e.*, how much they relied on the auditory feedback, how much they ignored it, how much they relied on the visuals). They were also asked how understandable was the auditory feedback, and how enjoyable and unpleasant was the auditory aspect of the application. Finally, they had to rank all the auditory feedback according to what they favoured.

3.5. Participants

A total of 21 participants (14 males, 6 females, 1 non-binary) aged 21 to 29 (mean age = 24.33 years old, SD = 1.91) agreed to take part in the experiment. Only one was left-handed, and 18 out of the 21 had already used XR technologies (*i.e.*, HMDs) before (with the average XR experience being 2.94/5). When asked how often they drew on a daily basis (on a 1 to 5 scale, with 1 being never and 5 being very often), the average answer was 1.81/5 (with the highest given score being 4), and the average drawing level of the participants (on a 1 to 5 scale, with 1 being novice and 5 being expert) was 1.9/5 (with the highest given score being 3). The participants took part on a voluntary basis, after giving their informed consent, and were not paid for their participation. The experiment was approved by the ethics committee of the authors' institution.

3.6. Results

All statistical analyses were conducted using JASP v.0.19.3.

3.6.1. Drawing Accuracy

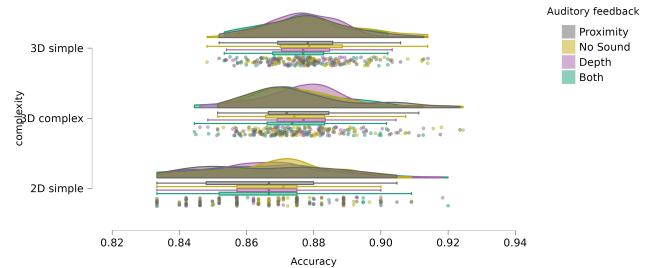
The experiment resulted in a 2 (DEPTH) x 2 (PROXIMITY) x 3 (COMPLEXITY) factorial design. In order to analyse the effect of the three factors on drawing accuracy, we first checked the normal distribution of the data using a Shapiro-Wilk test, and then applied a Box-Cox function ($x = (x^\lambda - 1)/\lambda$) with $\lambda = -1$) to normalize the set of data. We then performed a repeated measures ANOVA

Table 1: Additional evaluation categories, based on Wiese et al. evaluation framework [WIMB10].

Category	Description of dimensions of a category	Exemplary anchor sketches
Planarity (XY, YZ and ZX axis)	1. The plane is flat (the strokes are mostly aligned when looking from the profile).	
	2. The plane deviates slightly from its desired placement (there are some waves when looking at the profile).	
	3. The plane deviates heavily from its desired placement (there are numerous waves when looking at the profile).	
Resemblance to the model	1. The shape is very similar to the model.	
	2. The shape deviates from the model, some elements are misplaced (example: there are missing strokes).	
	3. The shape deviates heavily from the model, a lot of elements are misplaced.	

(with a Greenhouse-Geisser sphericity correction) because the data was normally distributed. A statistically significant main effect was found for COMPLEXITY ($F(1.55, 31.1) = 20.747, p < 0.001, \eta_p^2 = 0.509$), but not for DEPTH ($F(1.0, 20.0) = 2.053, p = 0.167, \eta_p^2 = 0.093$), or PROXIMITY ($F(1.0, 20.0) = 0.746, p = 0.398, \eta_p^2 = 0.036$). In order to gain a better understanding of the effect of both variables, we also performed a Bayesian ANOVA [LVW15] which confirmed a very strong evidence in favour of an effect of COMPLEXITY ($BF_{10} = 22348.986$), and strong evidence in favour of an absence of effect of DEPTH ($BF_{10} = 0.396$), and PROXIMITY ($BF_{10} = 0.336$).

If we look at the evaluation categories individually, we also find evidence of an effect of COMPLEXITY (line straightness: $F(1.6, 31.8) = 38.771, p < 0.001, \eta_p^2 = 0.660$, matching of two lines: $F(1.4, 28.9) = 21.227, p < 0.001, \eta_p^2 = 0.515$, degree of deviation: $F(1.4, 28.1) = 13.118, p < 0.001, \eta_p^2 = 0.396$, XY-planarity: $F(1.6, 31.6) = 20.231, p < 0.001, \eta_p^2 = 0.503$), but not of DEPTH (line straightness: $F(1.0, 20.0) = 0.153, p = 0.699, \eta_p^2 = 0.008$, matching of two lines: $F(1.0, 20.0) = 0.332, p = 0.571, \eta_p^2 = 0.016$, degree of deviation: $F(1.0, 20.0) = 1.116, p = 0.303, \eta_p^2 = 0.053$, XY-planarity: $F(1.0, 20.0) = 2.288, p = 0.146, \eta_p^2 = 0.103$), or PROXIMITY (line straightness: $F(1.0, 20.0) = 2.403, p = 0.137, \eta_p^2 = 0.107$, matching of two lines: $F(1.0, 20.0) = 5.250, p = 0.033, \eta_p^2 = 0.208$, degree of deviation: $F(1.0, 20.0) = 1.280, p = 0.271, \eta_p^2 = 0.060$, XY-planarity: $F(1.0, 20.0) = 0.098, p = 0.757, \eta_p^2 = 0.005$).

**Figure 4:** Raincloud plot of the accuracy results for each drawing complexity, comparing the auditory feedback conditions.

3.6.2. NASA-RTLX

A Shapiro-Wilk test revealed that the data was not normally distributed. After normalizing the data using a Box-Cox function ($x = \ln x$), we performed a Friedman test. Out of the 6 factors, we observed a statistically significant effect of AUDITORYFEEDBACK for the Performance ($X^2 = 11.757, p = 0.008$) and Effort ($X^2 = 8.345, p = 0.039$) factors. Conover's Post-hoc tests with Bonferroni correction revealed only a statistically significant difference between the DEPTH and DEPTH+PROXIMITY conditions for the Effort factor ($p = 0.022$).

3.6.3. Drawing Completion Time

An ANOVA was performed on the time spent drawing. It revealed a statistically significant main effect of AUDITORYFEED-

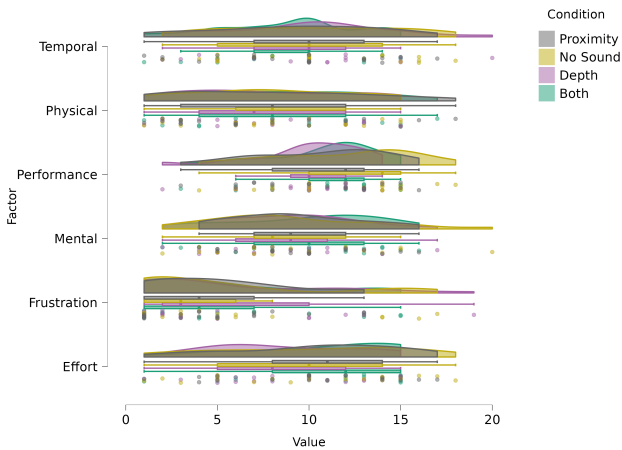


Figure 5: Raincloud plot of the NASA-RTLX dimensions, comparing the auditory feedback conditions.

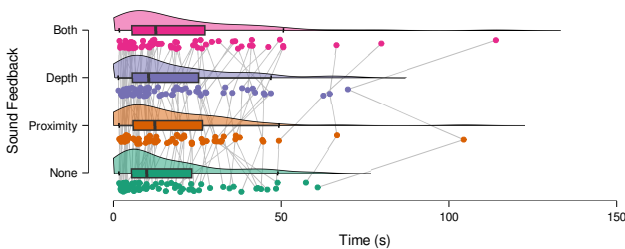


Figure 6: Raincloud plot of the drawing time, averaged across the three model complexities, comparing the auditory feedback conditions.

BACK ($F(1.54, 32.35) = 4.263, p = 0.031$) and COMPLEXITY ($F(1.05, 22.06) = 107.117, p < 0.001$). Post-hoc tests with Bonferroni correction did not reveal any statistically significant differences between AUDITORYFEEDBACK values. In order to refine this analysis, we performed a Bayesian ANOVA, which also revealed moderate evidence for an effect of AUDITORYFEEDBACK ($BF_{10} = 4.413$) and decisive evidence for an effect of COMPLEXITY ($BF_{10} = 1.8 * 10^{14}$). Furthermore, post-hoc tests revealed strong evidence for differences between the DEPTH+PROXIMITY feedback, and the other types of feedback (with NO SOUND $BF_{10} = 17.179$, with PROXIMITY $BF_{10} = 11.302$, with DEPTH $BF_{10} = 5.115$), with DEPTH+PROXIMITY leading to a longer drawing time. These results are depicted in Figure 6.

3.6.4. End Questionnaire

Results for the questionnaire are shown in Figure 7. They show clear differences in preferences due to COMPLEXITY and AUDITORYFEEDBACK, on the usage of feedback. The overall appreciation of auditory feedback however remains limited.

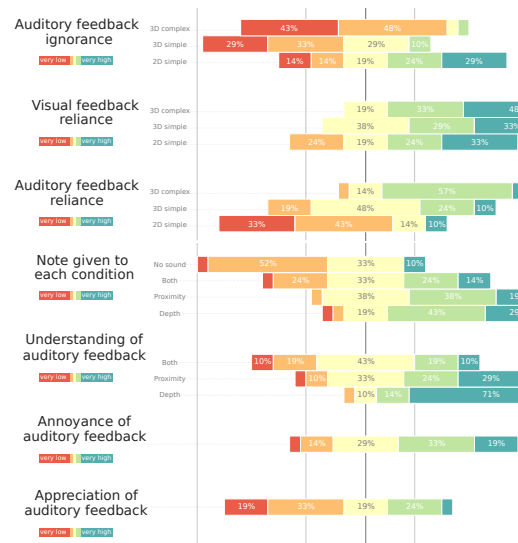


Figure 7: Results of the end questionnaire highlighting perceived differences between drawing complexities and auditory feedback conditions.

4. Discussion

Looking at the previous results on drawing accuracy, task load, drawing time and questionnaire, we can derive a number of insights on the use of auditory feedback for guiding 3D drawing.

4.1. Auditory Feedback Improves Perceived but not Actual Performance

The results that were obtained during the experiment tend to show that the chosen audio parameters did not have an effect on the drawing accuracy (see Figure 4), whether it was DEPTH improving the XY-planarity, or PROXIMITY improving the matching of two lines and reducing the deviations, thus denying hypothesis H1. However, there are still interesting elements to note. First, the NASA-RTLX results indicate that the auditory feedback had a significant effect on the Performance factor. This means that even though the accuracy did not change between the auditory feedback, the participants perceived differences: they felt like they performed the best with the DEPTH condition, and that they performed the worse with the NO SOUND condition. This goes hand in hand with the participants' impressions that were gathered during the experiment. Those who started with an auditory feedback were unsettled when reaching the NO SOUND condition (P19: *It feels so weird without sound now, like it's missing something*), often commenting on how badly they thought they were going to perform with it (P21: *I'll try my best, but it might look terrible*). One participant even commented that not having audio feedback felt like an important sense was taken from you. Overall, participants felt more comfortable in their ability to perform the task when they had an auditory feedback, although this did not affect the statistical results of the drawing accuracy. The NASA-RTLX results also show that with the DEPTH condition, less efforts were required to complete the drawings.

A few elements could explain why the accuracy results were not

significant. With the DEPTH condition, as the sound was continuous, it was sometimes difficult for the participants to notice when there were slight variations, and thus a plane deviation / strokes not connecting. As for the PROXIMITY condition, it often led to corrective movements, that were sometimes large. Thus, there was either straighter lines but with a disconnect of strokes, or lines correctly connecting, but inducing a shape deviation because of the corrective movements. A way of compensating for these issues could include distinct note changes and/or additional visual cues.

4.2. Auditory Feedback Reliance and Time Spent on the Task

As we can see on Figure 7, participants did rely more on the auditory feedback when the drawing complexity increased, thus confirming H2. For the 2D simple shapes, as the additional spatial cues were not crucial to perform the drawings, the auditory feedback was mostly ignored, but this was not the case anymore for the 3D simple shapes, and especially the 3D complex shapes. It could be linked to the visual channel occlusion, as way more strokes were present on the canvas, and there was a need for additional spatial cues to correctly perform the task. It could also be because the more complex the model became, the more visuomotor coordination was required to execute the movements, and the auditory feedback was helpful in this instance. So although the auditory feedback did not impact the drawing accuracy, participants used it as guidance regardless. This ties in with previous studies, which found that, in an audio-visual task, users relied more on the auditory feedback for backup as the complexity increased [MCM*16].

Another insight pertains to the cognitive load (Mental demand factor of the NASA-RTLX). The results are not significant, but that could be explained by the two different profiles encountered among the participants. They either felt that the use of auditory feedback reduced their cognitive load, because they could rely on it to facilitate the drawing process (especially with complex shapes) (P11: *It's less efforts because you can rely on the sound*), or that, conversely, it increased it, as they had to listen to the audio to draw and it required a lot of focus (P13: *You have to pay attention to more things*). However, we do not know what could induce these cognitive differences. Our results finally indicate that added auditory feedback lead to longer completion times (see Figure 6). This finding correlates with previous research on the effect of sound guides and rhythm on user experience [RBRG21], which suggested that sound increased perceived fatigue by having participants focus more on the performed gestures. A similar effect could be the reason for our results, with an increased focus leading to slower gestures.

4.3. Preferences in Auditory Feedback

The ranks given by the participants for each auditory feedback (Figure 7) show that DEPTH was the preferred feedback, and also that it was the most understandable, followed by the PROXIMITY condition. The DEPTH+PROXIMITY condition was ranked as the hardest to understand, since it is the addition of two different auditory feedback and required more focus (the Mental demand was rated as being higher in this condition, see Figure 5). Consequently, it was less appreciated than the other two. One issue that could explain this rating is that the proximity threshold was low, and

if participants were doing fast gestures, they had trouble hearing the oscillation over the depth feedback. Increasing the proximity threshold, and adding a visual cue indicating the point they are the closest to, could probably resolve the issue. We can note that the NO SOUND condition was on average rated the lowest by the participants, meaning they found this condition to be the least interesting. About the global perception of the auditory feedback, it was found to be quite enjoyable and only mildly annoying, which was rather surprising considering it is pure sine waves. Only one participant actively complained about it being very irritating, the others staying more nuanced (for example stating that it was fine for how long it lasted, but a prolonged exposition would become unpleasant). Therefore, exploring more complex audio parameters (*i.e.*, changes in timbre, harmonics, musical patterns...), or the same parameters on more complex sounds such as audio files played through granular synthesis, could probably be perceived as more enjoyable, and even elevate the drawing experience to a complete audio-visual experience (as it can be seen with 3D musical drawing [NKOM18, VSN20, BMPHS15, GB24]).

4.4. Limitations

Some limitations to this study might include the small amount of audio features tested, as we only had auditory feedback mapped to the depth of the controller and its proximity to other strokes. Additionally, the study was performed in a controlled environment, where the task was imposed onto the participants, preventing the drawings from being expressive (plus they were only geometrical shapes). Even if we had significant statistical results linking auditory feedback to increased drawing accuracy, nothing indicates that the system would have been usable outside of a controlled experiment (*i.e.*, for expressive drawing). That would require additional studies. Also, the participants who took part in the experiment were mostly novices in drawing. If we had a majority of more advanced level participants, this could have had an impact on the results. The study could be replicated with a small number of expert users. It is also worth noting that we did not evaluate spatial ability nor asked for the musical background of the participants, although this can impact either their 3D drawing or motor-learning capabilities. Future work should investigate how these parameters could affect auditory feedback guidance in the context of 3D drawing.

5. Conclusion

In this paper, we explored the use of auditory feedback for guiding 3D drawing in XR, with the goal of increasing the accuracy and compensating for visual depth perception issues. The results show that although the auditory feedback had no significant effect on the accuracy of the drawings, it had an influence on the participants' confidence and their perceived performance. They highlight the tendency to rely on another sensory modality when the drawing complexity increases. Further work will focus on different types of auditory feedback, for example with more complex sounds, and look into the combination of spatial cues from different feedback (visual or haptic, for example indicating the proximity to other strokes with a visual indication, or using vibrotactile feedback) in addition to the auditory feedback, as a way to increase drawing accuracy.

References

- [AKA*17] ARORA R., KAZI R. H., ANDERSON F., GROSSMAN T., SINGH K., FITZMAURICE G.: Experimental evaluation of sketching on surfaces in vr. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2017), CHI '17, Association for Computing Machinery, p. 56435654. doi: 10.1145/3025453.3025474. 2
- [B*96] BROOKE J., ET AL.: Sus-a quick and dirty usability scale. *Usability evaluation in industry* 189, 194 (1996), 4–7. 5
- [BBB*13] BOYER E. O., BABAYAN B. M., BEVILACQUA F., NOIS-TERNIG M., WARUSFEL O., ROBY-BRAMI A., HANNETON S., VIAUD-DELMON I.: From ear to hand: the role of the auditory-motor loop in pointing to an auditory source. *Frontiers in computational neuroscience* 7 (2013), 26. doi:10.3389/fncom.2013.00026. 2
- [BBG*20] BOYER E. O., BEVILACQUA F., GUIGON E., HANNETON S., ROBY-BRAMI A.: Modulation of ellipses drawing by sonification. *Experimental Brain Research* 238 (2020), 1011–1024. 2
- [Ber25] BERTHAUT F.: GDPD / IVMI-BUILDER: A Libre Software Framework for Extended Reality Musical Instruments and Sonic Installations. In *Proceedings of the 19th Linux Audio Conference* (Lyon, France, June 2025). 3
- [BMPHS15] BERTHAUT F., MARTINEZ PLASENCIA D., HACHET M., SUBRAMANIAN S.: Reflets: Combining and Revealing Spaces for Musical Performances. In *New Interfaces for Musical Expression (NIME)* (Baton Rouge, United States, May 2015). 8
- [BMPS19] BATMAZ A. U., MACHUCA M. D. B., PHAM D. M., STUERZLINGER W.: Do head-mounted display stereo deficiencies affect 3d pointing tasks in ar and vr? In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)* (2019), pp. 585–592. doi: 10.1109/VR.2019.8797975. 1
- [BMS19] BARRERA MACHUCA M. D., STUERZLINGER W.: The effect of stereo display deficiencies on virtual hand pointing. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2019), CHI '19, Association for Computing Machinery, p. 114. doi:10.1145/3290605.3300437. 1
- [BMSA19] BARRERA MACHUCA M. D., STUERZLINGER W., ASENTE P.: The effect of spatial ability on immersive 3d drawing. In *Proceedings of the 2019 Conference on Creativity and Cognition* (New York, NY, USA, 2019), C&C '19, Association for Computing Machinery, p. 173186. doi:10.1145/3325480.3325489. 2, 5
- [BYLS22] BATMAZ A. U., YU K., LIANG H.-N., STUERZLINGER W.: Improving effective throughput performance using auditory feedback in virtual reality. In *Proceedings of the 2022 ACM Symposium on Spatial User Interaction* (New York, NY, USA, 2022), SUI '22, Association for Computing Machinery. doi:10.1145/3565970.3567702. 2, 4
- [DB13] DUBUS G., BRESIN R.: A systematic review of mapping strategies for the sonification of physical quantities. *PloS one* 8, 12 (2013), e82491. 4
- [DGK*20] DREY T., GUGENHEIMER J., KARLBAUER J., MILO M., RUKZIO E.: Vrsketchin: Exploring the design space of pen and tablet interaction for 3d sketching in virtual reality. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2020), CHI '20, Association for Computing Machinery, p. 114. doi:10.1145/3313831.3376628. 2
- [DSK18] DUDLEY J. J., SCHUFF H., KRISTENSSON P. O.: Bare-handed 3d drawing in augmented reality. In *Proceedings of the 2018 Designing Interactive Systems Conference* (New York, NY, USA, 2018), DIS '18, Association for Computing Machinery, p. 241252. doi:10.1145/3196709.3196737. 2
- [EBMS*20] ELSAYED H., BARRERA MACHUCA M. D., SCHAARSMIDT C., MARKY K., MÜLLER F., RIEMANN J., MATVIENKO A., SCHMITZ M., WEIGEL M., MÜHLHÄUSER M.: Vrsketchpen: Unconstrained haptic assistance for sketching in virtual 3d environments. In *Proceedings of the 26th ACM Symposium on Virtual Reality Software and Technology* (New York, NY, USA, 2020), VRST '20, Association for Computing Machinery. doi:10.1145/3385956.3418953. 2
- [EFS*16] EFFENBERG A. O., FEHSE U., SCHMITZ G., KRUEGER B., MECHLING H.: Movement sonification: Effects on motor learning beyond rhythmic adjustments. *Frontiers in neuroscience* 10 (2016), 219. 2
- [GB24] GRUY E., BERTHAUT F.: MagneTip: Reintroducing a Physical Interaction Loop for 3D Musical Drawing in Extended Reality. In *New Interfaces for Musical Expression* (Utrecht (Netherlands), Netherlands, Sept. 2024). 8
- [GKL*18] GAO B., KIM H., LEE H., LEE J., KIM J.-I.: Effects of continuous auditory feedback on drawing trajectory-based finger gestures. *IEEE Transactions on Human-Machine Systems* 48, 6 (2018), 658–669. doi:10.1109/THMS.2018.2850329. 2
- [GLRS20] GRASSINI S., LAUMANN K., RASMUSSEN SKOGSTAD M.: The use of virtual reality alone does not promote training performance (but sense of presence does). *Frontiers in psychology* 11 (2020), 1743. 2
- [GPR*24] GUARESE R., PRETTY E., RENATA A., POLSON D., ZAMBETTA F.: Exploring audio interfaces for vertical guidance in augmented reality via hand-based feedback. *IEEE Transactions on Visualization and Computer Graphics* 30, 5 (2024), 2818–2828. doi: 10.1109/TVCG.2024.3372040. 3
- [GWFL22] GAO Z., WANG H., FENG G., LV H.: Exploring sonification mapping strategies for spatial auditory guidance in immersive virtual environments. *ACM Trans. Appl. Percept.* 19, 3 (Sept. 2022). doi:10.1145/3528171. 2, 3
- [HGAB08] HOFFMAN D. M., GIRSHICK A. R., AKELEY K., BANKS M. S.: Vergence–accommodation conflicts hinder visual performance and cause visual fatigue. *Journal of vision* 8, 3 (2008), 33–33. 1
- [HS88] HART S. G., STAVELAND L. E.: Development of nasa-tlx (task load index): Results of empirical and theoretical research. In *Advances in psychology*, vol. 52. Elsevier, 1988, pp. 139–183. 5
- [JSS*08] JONES J. A., SWAN J. E., SINGH G., KOLSTAD E., ELLIS S. R.: The effects of virtual reality, augmented reality, and motion parallax on egocentric depth perception. In *Proceedings of the 5th Symposium on Applied Perception in Graphics and Visualization* (New York, NY, USA, 2008), APGV '08, Association for Computing Machinery, p. 914. doi:10.1145/1394281.1394283. 3
- [JZF*21] JIANG Y., ZHANG C., FU H., CANNVÒ A., LAMBERTI F., LAU H. Y. K., WANG W.: Handpainter - 3d sketching in vr with hand-based physical proxy. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2021), CHI '21, Association for Computing Machinery. doi:10.1145/3411764.3445302. 2
- [KH01] KOZHEVNIKOV M., HEGARTY M.: A dissociation between object manipulation spatial ability and spatial orientation ability. *Memory & cognition* 29, 5 (2001), 745–756. 5
- [KLB*98] KLATZKY R. L., LOOMIS J. M., BEALL A. C., CHANCE S. S., GOLLEDGE R. G.: Spatial updating of self-position and orientation during real, imagined, and virtual locomotion. *Psychological science* 9, 4 (1998), 293–298. 5
- [KW08] KLEIBER M., WINKELHOLZ C.: Distortion of depth perception in virtual environments using stereoscopic displays: quantitative assessment and corrective measures. In *Stereoscopic Displays and Applications XIX* (2008), vol. 6803, SPIE, pp. 124–134. 3
- [KZL07] KEEFE D., ZELEZNIK R., LAIDLAW D.: Drawing on air: Input techniques for controlled 3d line illustration. *IEEE transactions on visualization and computer graphics* 13, 5 (2007), 1067–1081. doi: 10.1109/TVCG.2007.1060. 2, 3
- [LFAS*22] LEY-FLORES J., ALSHAMI E., SINGH A., BEVILACQUA F., BIANCHI-BERTHOUE N., DEROY O., TAJADURA-JIMÉNEZ A.: Effects of pitch and musical sounds on body-representations when moving with sound. *Scientific reports* 12, 1 (2022), 2676. doi:10.1038/s41598-022-06210-x. 2

- [LG05] LOKKI T., GROHN M.: Navigation with auditory cues in a virtual environment. *IEEE MultiMedia* 12, 2 (2005), 80–86. doi:10.1109/MMUL.2005.33. 2, 3
- [LMM*22] LIU W., MAGALHAES M. A., MACKAY W. E., BEAUDOUIN-LAFON M., BEVILACQUA F.: Motor variability in complex gesture learning: Effects of movement sonification and musical background. *ACM Trans. Appl. Percept.* 19, 1 (Jan. 2022). doi:10.1145/3482967. 2
- [LVW15] LY A., VERHAGEN A. J., WAGENMAKERS E.-J.: Harold jeffreys default bayes factor hypothesis tests: Explanation, extension, and application in psychology. *Journal of Mathematical Psychology* 72 (08 2015). doi:10.1016/j.jmp.2015.06.004. 6
- [MAS*18] MACHUCA M. D. B., ASENTE P., STUERZLINGER W., LU J., KIM B.: Multiplanes: Assisted freehand vr sketching. In *Proceedings of the 2018 ACM Symposium on Spatial User Interaction* (New York, NY, USA, 2018), SUI '18, Association for Computing Machinery, p. 3647. doi:10.1145/3267782.3267786. 2
- [MCM*16] METATLA O., CORREIA N. N., MARTIN F., BRYAN-KINNS N., STOCKMAN T.: Tap the shapetones: Exploring the effects of cross-modal congruence in an audio-visual interface. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2016), CHI '16, Association for Computing Machinery, p. 10551066. doi:10.1145/2858036.2858456. 2, 3, 8
- [MIKS24] MACHUCA M. D. B., ISRAEL J. H., KEEFE D. F., STUERZLINGER W.: Toward more comprehensive evaluations of 3d immersive sketching, drawing, and painting. *IEEE Transactions on Visualization and Computer Graphics* 30, 8 (2024), 4648–4664. doi:10.1109/TVCG.2023.3276291. 5
- [MK97] MEREU S. W., KAZMAN R.: Audio enhanced 3d interfaces for visually impaired users. *SIGCAPH Comput. Phys. Handicap.*, 57 (Jan. 1997), 1015. doi:10.1145/250025.250029. 2
- [MSA19] MACHUCA M. D. B., STUERZLINGER W., ASENTE P.: Smart3dguides: Making unconstrained immersive 3d drawing more accurate. In *Proceedings of the 25th ACM Symposium on Virtual Reality Software and Technology* (New York, NY, USA, 2019), VRST '19, Association for Computing Machinery. doi:10.1145/3359996.3364254. 2
- [MSWW19] MAY K. R., SOBEL B., WILSON J., WALKER B. N.: Auditory displays to facilitate object targeting in 3d space. In *Proceedings of the 2019 International Conference on Auditory display* (2019), vol. 8, pp. 155–162. 3, 4
- [NKOM18] NAKAGAWA R., KOMATSUBARA R., OTA T., MITSUMORI T.: Drawing sound in mr space: A multi-user audiovisual experience in mixed reality space. In *Proceedings of the Virtual Reality International Conference-Laval Virtual* (2018), pp. 1–4. 8
- [PGA*16] PARSEIHIAN G., GONDRE C., ARAMAKI M., YSTAD S., KRONLAND-MARTINET R.: Comparison and evaluation of sonification strategies for guidance tasks. *IEEE Transactions on Multimedia* 18, 4 (2016), 674–686. doi:10.1109/TMM.2016.2531978. 5
- [RBRG21] REYNAERT V., BERTHAUT F., REKIK Y., GRISONI L.: The effect of rhythm in mid-air gestures on the user experience in virtual reality. In *Human-Computer Interaction INTERACT 2021: 18th IFIP TC 13 International Conference, Bari, Italy, August 30 September 3, 2021, Proceedings, Part III* (Berlin, Heidelberg, 2021), Springer-Verlag, p. 182191. doi:10.1007/978-3-030-85613-7_13. 2, 3, 8
- [RVH13] RENNER R. S., VELICHKOVSKY B. M., HELMERT J. R.: The perception of egocentric distances in virtual environments - a review. *ACM Comput. Surv.* 46, 2 (Dec. 2013). doi:10.1145/2543581.2543590. 3
- [SS20] STANTON T. R., SPENCE C.: The influence of auditory cues on bodily and movement perception. *Frontiers in psychology* 10 (2020), 3001. 2
- [SSE15] SWAN J. E., SINGH G., ELLIS S. R.: Matching and reaching depth judgments with real and augmented reality targets. *IEEE Transactions on Visualization and Computer Graphics* 21, 11 (2015), 1289–1298. doi:10.1109/TVCG.2015.2459895. 3
- [TC24] TUNG Y.-H., CHANG C.-Y.: How three-dimensional sketching environments affect spatial thinking: A functional magnetic resonance imaging study of virtual reality. *PLOS ONE* 19, 3 (03 2024), 1–19. doi:10.1371/journal.pone.0294451. 2
- [TPBM*22] TÜRKMEN R., PFEUFFER K., BARRERA MACHUCA M. D., BATMAZ A. U., GELLERSEN H.: Exploring discrete drawing guides to assist users in accurate mid-air sketching in vr. In *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2022), CHI EA '22, Association for Computing Machinery. doi:10.1145/3491101.3519737. 2
- [VSN20] VALER R., SCHRAMM R., NEDEL L.: Musical brush: Exploring creativity in an ar-based tool combining music and drawing generation. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)* (2020), pp. 634–635. doi:10.1109/VRW50115.2020.00168. 8
- [WIMB10] WIESE E., ISRAEL J. H., MEYER A., BONGARTZ S.: Investigating the learnability of immersive free-hand sketching. In *Proceedings of the Seventh Sketch-Based Interfaces and Modeling Symposium* (Goslar, DEU, 2010), SBIM '10, Eurographics Association, p. 135142. 2, 5, 6
- [XLHT22] XU H., LYU F., HUANG J., TU H.: Applying sonification to sketching in the air with mobile ar devices. *IEEE Transactions on Human-Machine Systems* 52, 6 (2022), 1352–1363. doi:10.1109/THMS.2022.3186592. 3, 4
- [YAS*21] YU E., ARORA R., STANKO T., BÆRENTZEN J. A., SINGH K., BOUSSEAU A.: Cassie: Curve and surface sketching in immersive environments. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2021), CHI '21, Association for Computing Machinery. doi:10.1145/3411764.3445158. 2
- [YDSG21] YU X., DIVERDI S., SHARMA A., GINGOLD Y.: ScaffoldsSketch: Accurate industrial design drawing in vr. In *The 34th Annual ACM Symposium on User Interface Software and Technology* (New York, NY, USA, 2021), UIST '21, Association for Computing Machinery, p. 372384. doi:10.1145/3472749.3474756. 2