An Artists' Perspectives on Natural Interactions for Virtual Reality 3D Sketching

Richard Rodriguez
Colorado State University
Fort Collins, USA
richir@colostate.edu

Brian T. Sullivan
Colorado State University
Fort Collins, USA
brian.t.sullivan@colostate.edu

Mayra Donaji Barrera Machuca
Dalhousie University
Halifax, CA
mbarrera@dal.ca

Anil Ufuk Batmaz
Concordia University
Montreal, CA
ufuk.batmaz@concordia.ca

Cyane Tornatzky
Colorado State University
Fort Collins, USA
cyane.tornatzky@colostate.edu

Francisco R. Ortega
Colorado State University
Fort Collins, USA
fortega@colostate.edu

Figure 1: A dog drawn in VR by a participant using OpenBrush.

ABSTRACT

Virtual Reality (VR) applications like OpenBrush offer artists access to 3D sketching tools within the digital 3D virtual space. These 3D sketching tools allow users to "paint" using virtual digital strokes that emulate real-world mark-making. Yet, users paint these strokes through (unimodal) VR controllers. Given that sketching in VR is a relatively nascent field, this paper investigates ways to expand our understanding of sketching in virtual space, taking full advantage of what an immersive digital canvas offers. Through a study conducted with the participation of artists, we identify potential methods for natural multimodal and unimodal interaction techniques in 3D sketching. These methods demonstrate ways to incrementally improve existing interaction techniques and incorporate artistic feedback into the design.

CHI '24, May 11–16, 2024, Honolulu, HI, USA

ACM ISBN 979-8-4007-0330-0/24/05
https://doi.org/10.1145/3613904.3642758

CCS CONCEPTS

• **Human-centered computing** → **Gestural input**; **Interactive systems and tools**; **User studies**.

KEYWORDS

Virtual Reality, Gestures, Speech, Multimodal Interaction, 3D Sketching

ACM Reference Format:

Richard Rodriguez, Brian T. Sullivan, Mayra Donaji Barrera Machuca, Anil Ufuk Batmaz, Cyane Tornatzky, and Francisco R. Ortega. 2024. An Artists' Perspectives on Natural Interactions for Virtual Reality 3D Sketching. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24), May 11–16, 2024, Honolulu, HI, USA.* ACM, New York, NY, USA, 20 pages. https://doi.org/10.1145/3613904.3642758

1 INTRODUCTION

3D Sketching in Virtual Reality (VR) is a relatively new medium of artistic expression that allows users to create strokes in 3D space. Sketching in VR allows artists to experiment with mark making off the 2D plane; it allows them to explore how their bodies move in relation to that 3D space; and it allows artists to interact with technology that interfaces with the 3D "canvas". Artists have already taken to the new 3D sketching medium and have created a wide variety of innovative works. An early example of such explorations is the "Final Spin" from Jen Zen presented at the SIGGRAPH 2000 Art Gallery [115], where Jen Zen used 3D sketching to create human figures. Modern VR Head-Mounted Displays (HMDs) have increased the use of 3D sketching by artists who use this medium in live performances where they sketch immersive paintings that the audience can experience. For example, the artist Anna Zhilyaeva made a 3D interpretation of "Liberty Leading the People" in the Louvre Museum in Paris [116]. Another example is Aura Garden [90], where the audience collaborated to create light sketches in VR. Despite the artistic possibilities that this new medium offers, studies in modern 3D sketching systems have primarily focused on the creation of different types of strokes within the sketching program and controller unimodal interaction for naive users (see Table 1). Although artists have taken to VR sketching to creative cutting-edge works, there is a need for more research to understand and address the specific requirements of artists in this context.

This paper examines suggestions from trained artists on their ideas around novel input methods for sketching in VR. We examine the potential of alternative solutions for 3D sketching in VR that involve natural user interactions with unimodal input methods such as speech, gestures, gaze, and/or a multimodal combination of these. By understanding the needs of artists in a VR sketching application, we are able to make recommendations so that sketching applications can be designed to meet the artists' needs. Our goal is to identify novel ways to improve the artist experience when using 3D sketching systems, so they can express themselves better within the virtual space.

Natural user interactions [76, 99] offer several advantages that could improve artist experiences when using VR HMDs. For example, users can engage with virtual environments (VE) more naturally, mimicking real-world actions and communication, which reduces the cognitive load and provides a more immersive and intuitive interaction paradigm [111]. Moreover, multimodal interactions, which combine multiple channels, such as gestures, speech, pen, gaze, and touch, provide further advantages to communicate and interact with computers. For example, the ability to include multiple input channels allows tasks to be tailored to individual preferences and abilities, promoting accessibility [44]. The resiliency offered by multimodal interaction also increases the system's robustness, ensuring a more reliable interaction even in challenging conditions (e.g., higher workload) [80, 106, 107]. Overall, multimodal interaction amplifies the sense of presence and agency in VR, fostering deeper engagement and enabling a wider range of users to navigate natively and interact within immersive digital spaces seamlessly.

Despite the advantages of using natural user and multimodal interactions, few commercial 3D sketching systems support them. A common problem with these systems, like OpenBrush [35], is that they have to work on the constraints of commercial VR HMDs, e.g., using a controller as an input, or cater to specific populations, e.g., GravitySketch built by designers, for designers motto [92] who want a fast prototyping tool in VR and collaborate with others. However, prior studies involving artists have revealed that to facilitate 3D sketching effectively, it is essential to provide artists with suitable tools [56]. For example, when an artist draws in 3D, they are not just working on a flat surface like they experience in traditional 2D sketching and must try to convey depth, perspective, and complex spatial relationships simultaneously. For example, artist James R. Eads [73] creates portals to imaginary universes, syncing his strokes to the subtle beat of music. Viewers can walk through each portal and experience his vision, but also hear sounds that pulse through strokes drawn using Tilt Brush. While selecting sketching tools, changing colors, or adjusting brush sizes might not resemble the mental model of the users familiar with 2D sketching, the versatility of VR allows users to interact with multiple modes of mid-air interaction, which adds more dimension to users' creative process. This may allow them to express their ideas more comprehensively and with greater nuance.

When applying natural user interactions to 3D sketching in VR, it might be possible to make the user experience feel more intuitive, seamless, and similar to how artists interact with the physical world by mimicking how they work with physical materials [87]. In VR, for example, artists could use gestures, voice commands, or stylus input based on their preferences, reducing the need to learn complex menus and commands. Moreover, multimodal interactions can simplify the user experience by allowing artists to choose the interaction method that feels most natural to them. When interaction methods mimic real-world actions, the mapping between user intent and system response becomes more intuitive and requires less training in the new environment, e.g., using gestures that resemble physical actions or voice commands that directly describe what you want. Furthermore, artists are not limited to one type of input device, so they can adapt their approach based on what feels most natural and effective for each stage of their creative process. Combining gestures, voice commands, and different input devices would allow them to capture their ideas fully, translating their creative vision into a more accurate digital representation [93].

To examine how these interactions can help artists, we conducted semi-structured interviews [60] with artists from the local university's art program to explore how natural and multimodal interaction techniques might improve the 3D sketching process in VR. Semi-structured interviews are centered around a topic, where the interviewer asks open-ended questions, and the interviewee's answer reflects their personal experience, allowing the interviewer

Table 1: Overview of previous studies that have focused on unimodal interactions and multimodal interactions.

Paper	Participant Type	Modalities	Modality Count	Interaction Type	Software Tested	Sketching Activity	Focus and Conclusion
[31]	Designers	Pen, Pen+Tablet	2	U	Custom	Open sketch	Study focused on usability and concluded that VR is not optimized for sketching.
[51]	Designers	Gesture, Multitouch	2	U	Custom	Open sketch	Focused on usability and concluded that interaction in VR is still challenging.
[52]	Naive users	Pen, Pen+Tablet, Gesture	3	U	Custom	Open sketch, Single stroke	Focused on the app performance. Participants provided a usability report. Authors agreed that the components of VR sketching should be explored.
[86]	Naive users	Gesture	1	U	Custom	Single stroke	Explored one natural user interaction and focused on usability of the application.
[11]	Naive users	Gesture	1	U	Custom	Copy model, Open sketch, Single stroke	Study focused on the usability of the software through one unimodal interaction.
[110]	Naive users	Gesture, Speech+ Gesture	2	U, M	Custom	Copy model	Study focused on the usability of the software thorugh one unimodal and one multimodal interaction.
[56]	Artists	Gesture	1	U	Custom	Open sketch	Explored the usability of VR sketching space and tools for artists. Need more tools to support artists.
Our paper	**Artists**	**Bimanual, Controller, Gaze, Pen, Speech, Controller+ Gesture, Gesture+ Gaze, Gesture+ Speech**	**8**	**U, M**	**Commercial**	**Open sketch**	**Explores novel tools that can be beneficial to artist, natural user interactions, whether unimodal or multimodal, that artists prefer when interacting with VR sketching applications, and the usability of commercial VR sketching applications.**

Legend:
U - Unimodal
M - Multimodal

to gain a deeper understanding from the interviewee's perspective. In the context of our study, this approach enabled us to understand their perspective on 3D sketching and tools. This is particularly relevant, considering that artists may anticipate experiencing a seamless transition from 2D sketching into a 3D sketching system. The truth is different, as depth perception [12] and dependency on spatial abilities [13] are issues that plague users inside an immersive 3D VE. These are just a few of the issues that affect all users, regardless of background, when using VR. Therefore, to improve existing 3D sketching systems, we collected data from the artists' perspectives, to inform us of their needs and offer recommendations.

In this paper, we extend previous work on multimodal interaction [62, 106, 107] from simple [106, 107] to complex 3D environments [117] by proposing novel multimodal interaction techniques for 3D sketching. We also extend previous work on the advantages of using different input types simultaneously [31, 51, 52] to incorporate the artist's perspective. Our results suggest implementing multimodal and unimodal natural interaction techniques in future 3D sketching applications to create a more comprehensive and immersive experience for artistic users. The real-world usability of these proposed interaction techniques and features can be studied and evaluated to refine the techniques further. Our findings can aid in designing 3D sketching and other VR art applications. They may be incorporated into other domains of VR, such as annotation in immersive analytics or work in architecture and interior design. Our contributions are:

- A study, using semi-structured interviews, that asks artists their opinions on using natural unimodal interactions and adding multimodal interactions to 3D sketching systems. We found that the way artists interact varies from one individual to another, therefore, having additional unimodal interactions and adding multimodal interactions to 3D sketching will allow artists to use natural interactions that they are used to.
- A study on how artists evaluate the usability of a commercial 3D sketching system. We found that Open Brush [35] was rated above average by most of the artists for the two tasks that were assigned during the study.
- Recommendations for developers and designers of future 3D sketching applications and possibly other applications that have properties in common (e.g., annotation in VR). These recommendations include other unimodal and multimodal interactions to cater to the needs of each artist. Other recommendations include tools that were suggested by the participants which will aid in minimizing the artist's workflow.

2 RELATED WORK

2.1 3D Sketching

3D sketching as formally defined by Arora et al. [6] is *"a type of technology-enabled sketching where: (i) the physical act of mark making is accomplished off-the-page in a 3D, body-centric space, (ii) a computer-based tracking system records the spatial movement of the drawing implement, and (iii) the resulting sketch is often displayed in this same 3D space, e.g., via the use of immersive computer displays, as in virtual and augmented realities (VR and AR)"* (Arora et al. [6], p. 149). This way of sketching is flexible and fast [102], and is intuitive for 3D input [49, 98]. Due to these advantages, several companies have released applications that enable users to sketch and design in 3D such as Tilt Brush [39] (now open-source OpenBrush [35]), Gravity Sketch [92], and Quill [66]. These examples of commercial 3D sketching software have made 3D sketching available in various disciplines, including art, modeling, filmmaking, architecture, visualization design & research, medicine, and cultural heritage [100].

Despite the advantages of 3D sketching, correctly positioning a stroke in 3D space is challenging, as users are affected by high sensorimotor [105] and cognitive [13, 79] demands, depth perception issues in stereo displays [12, 16, 17], and the absence of physical support [7]. Previous work has studied the control and ergonomic aspects of sketching in mid-air [7, 57] and the learnability of 3D sketching [13, 105] to identify the cause(s) of these positioning inaccuracies. Other work has studied the advantages and disadvantages of 3D sketching as a medium for creativity and design by comparing it against pen-and-paper [47]. Finally, previous work has also studied how 3D sketching affects the act of ideation [32, 112]. Here, we aim to understand the expectations of artists for 3D sketching, focusing on their needs for alternate unimodal interactions and multimodal interactions within 3D sketching.

2.2 Interaction Techniques for 3D Sketching

Since the early 1990s, previous research has proposed multiple novel interaction devices and techniques for 3D sketching [10]. The devices include pens [33, 46, 85, 101] and physical surfaces [31, 52] that aim to provide a surface to draw on. For example, Elsayed et al. [33] demonstrated that active haptic feedback reduces errors in VR when a physical surface is not present. The interaction techniques, like virtual surfaces [8, 11, 61], beautification [11, 34] or novel metaphors to create strokes [50, 52, 86] aim to reduce the sensorimotor and cognitive demands of 3D sketching. For example, Barrera Machuca et al. [11], via Multiplanes, empowered participants by assisting them in sketching with snapping and beautification of strokes, which reduced the participants' cognitive and sensorimotor demands. Finally, another approach uses visual guides to improve the user's shape accuracy [14, 41, 97, 113]. One limitation of these approaches is that they mostly focus on unimodal interactions that use either controllers, gestures, or gaze.

We were able to identify previous work focusing on unimodal and limited multimodal interactions for 3D sketching systems (see Table 1). Some previous research outside of VR focuses on the effect of multimodal interactions on creativity [118] or user experience [110]. There has also been a lot of work that uses multimodal interactions for 3D modeling using CAD systems [20, 25, 74, 91, 96]. For example, participants in the study conducted by Wolf et al. [110] reported that using multimodal interactions, instead of unimodal interactions, allowed them to feel a higher state of presence. Another example is VR-CAD, in which Bourdot et al. [20] reported that using natural interactions allowed the participants to intuitively manage CAD objects, minimizing complications that are commonly expected of CAD applications. The advantages of using multimodal interactions for design in VR include: a higher sense of flow, higher intuitive use and lower mental workload, and a higher sense of presence [110]. They also provide similar creativity levels to unimodal interactions [118]. Due to the advantages they provide, it is important to understand how to add or incorporate multimodal interactions in 3D sketching systems.

2.3 Multimodal Interaction

Multimodal interactions are the combination of multiple input types like gestures, speech, pen, gaze, and touch. The combination of these inputs can have three properties: synchronous versus asynchronous, symmetric versus asymmetric, and dependent versus independent [63]. Synchronous interaction is one where the user can perform multiple interactions at the same time, whereas, in an asynchronous interaction, the actions do not need to happen at the same time. An example of synchronous interaction would be selecting an object while using speech to tell the system to change colors at the same time. With asynchronous interaction, one could select an object and then give a spoken instruction to change the color after the selection was made, but not at the same time. Symmetrical interactions are usually bimanual in nature, and the actions on one hand mirror what the other hand is doing; asymmetrical interactions do not have to mirror what the other hand is doing and thus act independently of each other. An example of symmetrical interaction is when one is painting a mirror image, like painting the wings of a butterfly with both hands. Similarly. asymmetrical interaction is, e.g., when one is painting and one hand has a menu palette and the other has a brush, so both hands are performing different tasks. A dependent interaction is one where an interaction depends on the other to accomplish a task, such as hands working in tandem with one hand controlling the color palette while the other hand controls the brush. In contrast, the interactions involved do not rely on each other in an independent interaction. An independent interaction could be when both hands can act as brushes and each can be used to draw, regardless of one another.

Researchers have continued investigating multimodal interaction since the work from Bolt [19]. Another important work by Hinckley et al. looked at Pen+Touch and described what type of interactions were possible [45]. Multiple studies have also examined multimodal gesture and speech inputs using mid-air gestures [5, 24, 43, 64, 71, 106]. For example, using gesture elicitation, Williams et al. [106] showed that multimodal interactions are essential to interact with augmented reality (AR) HMDs in a natural way. Yet, some have examined only a subset of gestures, such as 2D gestures (e.g., multitouch) [69, 84] or paddling gestures [48]. While multiple studies using gesture + speech interactions have been studied, they have concentrated in 2D environments or 3D environments using desktop displays, with less work in AR/VR [82, 107]. It is possible to find multimodal interaction examples, such as Internet of Things home controls [54], 3D computer-aided design in a 2-dimensional environment [58], and web browsing on televisions [71, 75, 108]. For example, Wittorf et al. [108] found that users preferred certain mid-air gestures when interacting with a wall-display. The larger amount of work has been in multimodal gesture and speech fusion and recognition [19, 24, 53, 81], although some of them have used limited gesture sets [26] or limited speech dictionaries [64]. Overall, the research conducted thus far has tested input feasibility and human adaptability and created more intuitive and discoverable interaction sets [109], yet the type of inputs are limited, without clear transferability to more complex applications.

3 MOTIVATION AND RESEARCH QUESTIONS

In previous studies, researchers collaborated with designers to evaluate the usability of novel VR interaction systems [4, 98]. Similarly, previous research also included designers in exploring new techniques or input devices for sketching in VR [31, 51, 59, 89]. For

example, Drey et al. [31] did a usability walk-through with six participants to understand the design space between 2D (pen on a tablet) and 3D input (6 Degrees-of-Freedom (DOF) pen) for 3D sketching. Yet, few works have focused on the experiences of artists when using 3D sketching systems [40, 55]. For example, Keefe et al. [55] studied collaboration and visualization in VR sketching and found that the sketching system lacked the tools needed for artists to capture their intended designs. Also, to our knowledge, there have been no studies where the artists evaluated commercially available VR sketching applications (see Table 1). By filling this gap, we aim to ensure that new tools align with the creative processes and expectations of artists, which will allow this community to be more active in the space.

Our research follows previous approaches to understanding adding input types to a 3D sketching system. First, this paper explores novel ways to use the current tools available in commercial software using unimodal and multimodal natural interaction, e.g., using any combination of gestures, speech, eye gaze, pen, and controller. Second, the paper aims to identify new tools that could be added to a 3D sketching system and that can benefit from these novel input methods. Using perspectives from artists, we investigate the following research questions:

- **RQ1** What tools of commercial 3D sketching systems help artist in their sketching process?
- **RQ2** What natural multimodal interactions can 3D sketching systems add to help artist in their sketching process?
- **RQ3** How do artists perceive the usability of commercial 3D sketching systems?

While RQ1 investigates the identification of the tools of commercial 3D sketching systems that help artists in their sketching process, RQ2 explores new unimodal and multimodal natural interactions for current tools and new features for current commercial systems that fulfill their needs. Finally, RQ3 examines usability deficiencies of 3D sketching systems from the perspective of artists. Artists and designers have distinct priorities. Unlike designers who emphasize performance and speed, artists concentrate on the creative process and achieving the final result. By identifying novel unimodal and multimodal natural interactions, designers of future 3D sketching systems can create better tools considering various use cases and go beyond using a controller as an input method.

4 USER STUDY

4.1 Methodology

Participants. For the study, thirteen participants (8 females and 5 males) studying art at the local university were recruited. Their ages ranged between 20 and 28 (M = 22.4, SD = 2.4). Eleven participants had previously used AR/VR before. Seven indicated that they had their vision corrected, two through the use of glasses, one through contact lenses, and the other four did not specify. Two participants had previously experimented with Tilt Brush (now called OpenBrush) in VR. Two participants were double majors (computer science and fine art), while the rest were specifically fine art majors. Our study was limited to artists, because they have experience working in fine arts from their classes and studio practice, giving the study a population closer to that of established artists as compared to naive users. All participants were either enrolled in or recently graduated from a Bachelor of Fine Art (BFA) degree. The BFA program requires foundational class work that includes coursework in drawing, painting, sculpture, and digital media. Additionally, the pre-survey questionnaire allowed for the participant to volunteer information, such as specific applications they had worked with (e.g., ZBrush, Maya, Blender, Cinema 4D, Autodesk 3DS Max), but none of our participants volunteered that information.

Equipment. The 3D sketching program was run on an Alienware Aurora R14 desktop equipped with an AMD Ryzen 9 5900 12-core processor running at 3.0 GHz, with a total of 32 GB of system RAM and an NVIDIA Ge-Force RTX 3080 with 26 GB of onboard memory with the GPU running at 1710 MHz. The desktop ran Microsoft Windows 11 Home (version 10.0.22621, build 22621). The participants used an HTC Vive Pro Eye with two controllers and two lighthouses to access the 3D Sketching system. Finally, a Go-Pro Hero 7 with a 128 GB memory card was used to record the interaction of the participants during the study. For the 3D sketching application, we used a fork of OpenBrush v2.3.0 [35]. OpenBrush was run through Unity 3D, version 2019.4.25f1 (as recommended by the contributors). As the participants drew in the 3D Sketching program, their drawings were recorded by Unity's Recorder to capture the participant's perspective from the HMD.

Procedure. Upon arrival, each participant followed a series of tasks, mentioned here and in Figure 4. The participant first signed three forms: a vision attestation form, the consent form, and a pre-survey questionnaire (including their demographics and any prior VR experience). They were informed of what a semi-structured interview is, and that this study uses semi-structured interviews to collect data. The participant then watched a video tutorial that showed the basics of using OpenBrush in VR [1]. After the video, the participant was fitted with the HMD and controllers to repeat the basic operations they had just seen in the OpenBrush video tutorial, allowing them to practice by replicating what they had just watched. When the tutorial was finished, the participant removed the HMD and controllers. Next, the participant watched another video that explained the different types of unimodal and multimodal natural interactions and their categories [2]. After the second video and before starting the study, the researchers allowed the participants to ask any questions, but none of the participants had questions on the procedure.

For the first part of the study (Phase One), the participant was fitted with the HMD/controllers and was tasked with drawing a 3-dimensional dog using any tool available in the 3D sketching application. The participant had a 2-meter by 2.1-meter rectangular space, free of obstacles, in order to sketch freely in OpenBrush. Phase One ended after 10 minutes, at which time the participant was given the choice to take a 2-minute break or continue directly to Phase Two. In Phase Two, the participant was fitted with the HMD/controllers (if the 2-minute break was taken) and was tasked with drawing a ground, a path, and a tree. Just as in Phase One, they were allowed to use any tools they liked but had a 15-minute

[1]The OpenBrush tutorial can be seen at https://youtu.be/XqrwfRKjv7U.
[2]The Unimodal and Multimodal video can be seen at https://youtu.be/zxeCdDaPk-8.

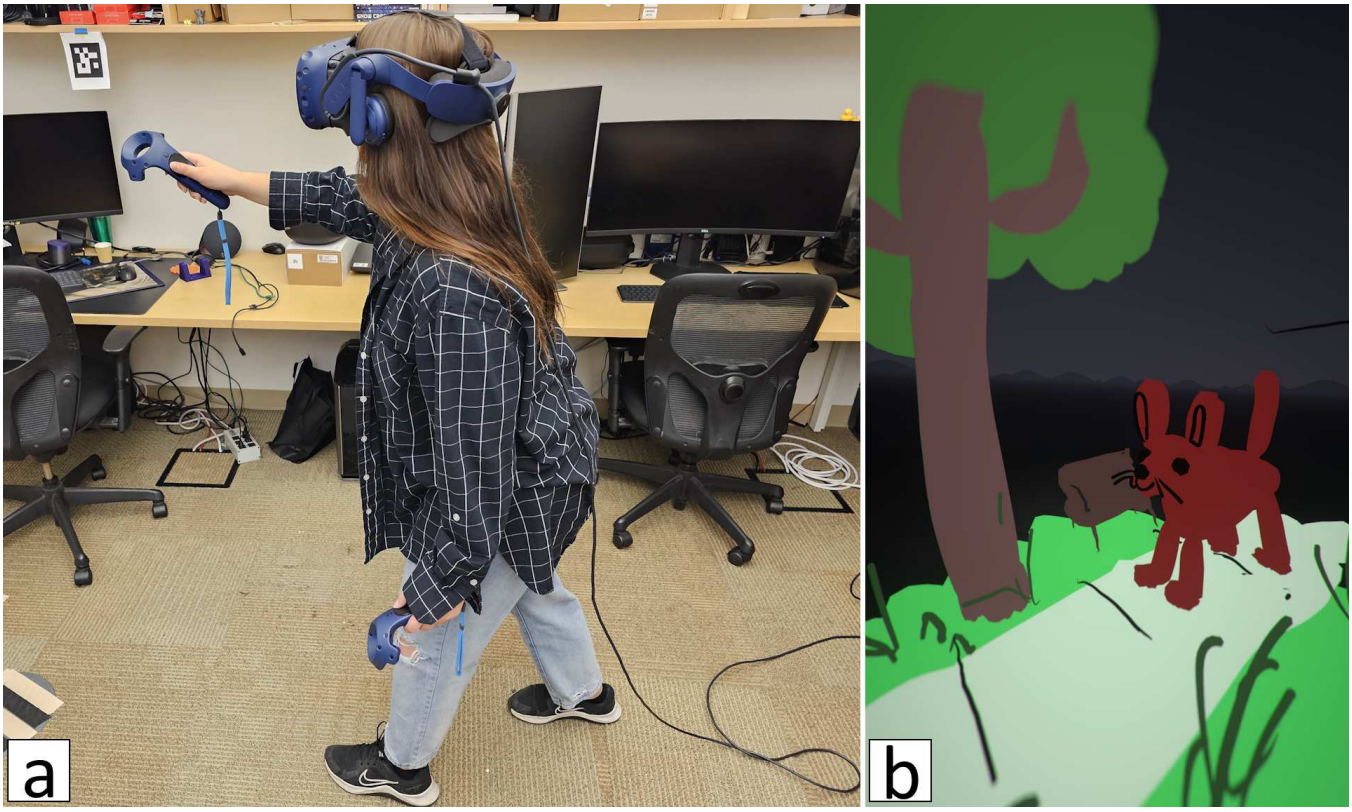

Figure 2: The person shown in a) was not a participant, to preserve anonymity, but is shown in a similar pose displayed by P4. b) Shows the drawing made by P4.

time limit. Participants could add additional constructs to the scene as long as they had drawn the ground, the path, and the tree, and the time limit had not been reached. Some of the completed works can be observed on the right-hand side of Figure 2 and Figure 3 with the corresponding participants appearing on the left. After removing the HMD, the participant was asked to complete a System Usability Scale (SUS) [9, 65]. Afterward, a post-study interview (see the supplementary materials for the interview questions) was done, where the participants told the researchers about their experiences. Finally, participants were offered class credit or a $20 Amazon gift card for their time. In total, the entire study lasted around 57 minutes.

Data Collection. Each participant's movements in the physical space during their sketching session were recorded using a Go-Pro camera (Figure 2 and Figure 3). The camera was fixed with an overview of the sketching area. The recording of the session began after participants watched the video tutorial about the multimodal interaction. After the video tutorial, we asked participants if they had any questions about the video and the task, but none of the participants had questions. Then, participants were asked to follow the "think-out-loud" method [77] while sketching to help understand their thought processes as they drew specific elements. Moreover, the study researcher periodically asked the participants if multimodal interaction techniques would assist with the participant's current task. This occurred every time the participant switched to something "new" or after asking for assistance about how to navigate the system. If the participants asked for assistance, the researchers provided verbal help to resolve the issue and followed up by asking if an alternative interaction technique could have aided in accomplishing that task or prevented the issue. Sometimes, participants did not have a response to follow-up questions. To allow researchers to examine the participants' actions while suggesting other unimodal or multimodal interactions, the screen of the PC running OpenBrush was recorded.

Following a completed participant session, the video/audio recordings from the Go-Pro camera were synchronized with the headset recordings from Unity. This allowed a simultaneous analysis of the participants' real-world motions and what they saw in the virtual world. The audio recordings were also automatically transcribed using the Microsoft Word Web App's transcription feature [67]. Each of the two authors reviewed half of the transcripts to fix transcription mistakes. When necessary, corrections were made to the transcriptions using the Go-Pro recording.

Data Analysis. Following an approach inspired by Braun and Clarke [22, 23], this study uses researcher reflexivity as a pillar of the thematic analysis. Because of this epistemological and ontological position, researchers were able to avoid measuring inter-coder agreement. The agreement poses the existence of a researcher "bias"

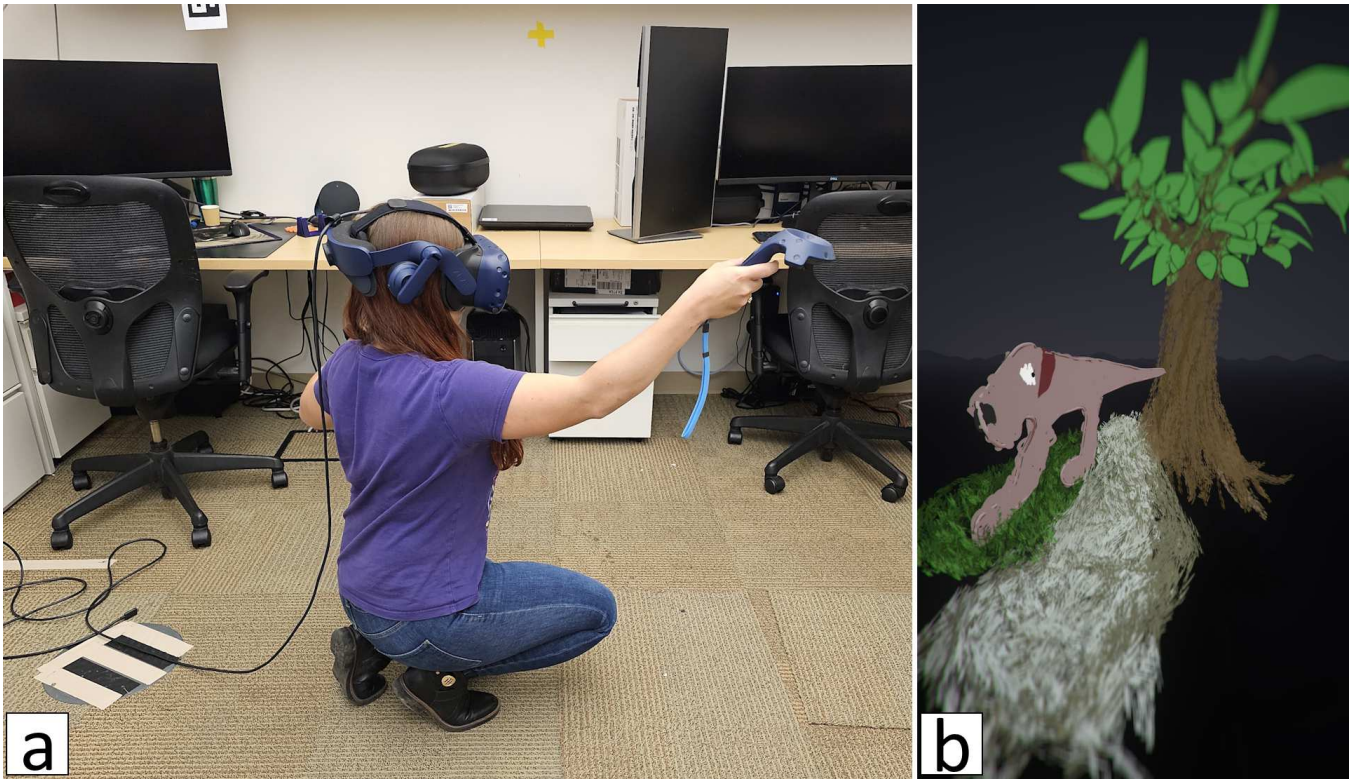

Figure 3: The person shown in a) was not a participant, to preserve anonymity, but is shown in a similar pose displayed by P13. b) Shows the drawing made by P13.

and tries to minimize it (as well as performing consensus coding), anchored in the belief that there is an objective way of coding and that this objective method is more desirable. Instead, researchers in this study recognize the situated nature of coding and its inherent partiality and subjectivity [28].

Three researchers conducted a qualitative analysis of the interviews. Two of these researchers (or, more specifically, coders) ran the user study and were familiar with the data. The third coder had previous thematic analysis experience and helped the lead coders through the process. Two coders were male, and one was a female. Two coders had undergraduate degrees in Fine Arts, either in animation and digital art or in film/cinema production. One had formal training in drawing and sketching, and the other had over 14 years of experience in 2D art. The third coder did not have formal training in drawing or sketching.

Of the two researchers who led the user study, one coder was assigned seven interview transcripts, and the other was assigned six. Each transcript was assigned to the coder who originally conducted the interview. This assignment leverages familiarity with the data as key to analysis [18, 21]. The two coders used a template with columns for transcript excerpts, codes, and comments. The coders were further familiarized with the data by re-reading their transcripts and taking notes. Individually and inductively, they coded their transcripts to create a system to encode the data while keeping a list of this encoded data and their descriptions to track their own process. Then, they shared the coded data and discussed the construction of themes. The themes were refined in conversations among two coders who conducted the study and then proposed to the third coder for further discussion. For this final part, the third coder participated in the discussions and helped define the final themes. The lead coders met five times and three additional times with the third coder. Ultimately, the themes were proposed to the rest of the team for further discussion.

5 FINDINGS

5.1 Qualitative Findings using Thematic Analysis

The analysis characterizes artists' expectations about what features 3D sketching applications should have. The questions focus on identifying the tools of commercial 3D sketching systems that help artists in their sketching process (RQ1) or integrating unimodal and multimodal natural interactions into 3D sketching systems (RQ2). Recognizing that the participants are art students at Colorado State University, our results account for this population, which has specific cultural expectations of design tools [94, 103]. Our data indicates a familiarity with complex desktop tools, yet not enough experience with 3D sketching. The research indicates that the participants (i.e., artists) wanted to improve current features and add other input modalities. It also shows that artists expect 3D sketching systems to have more features than other design tools.

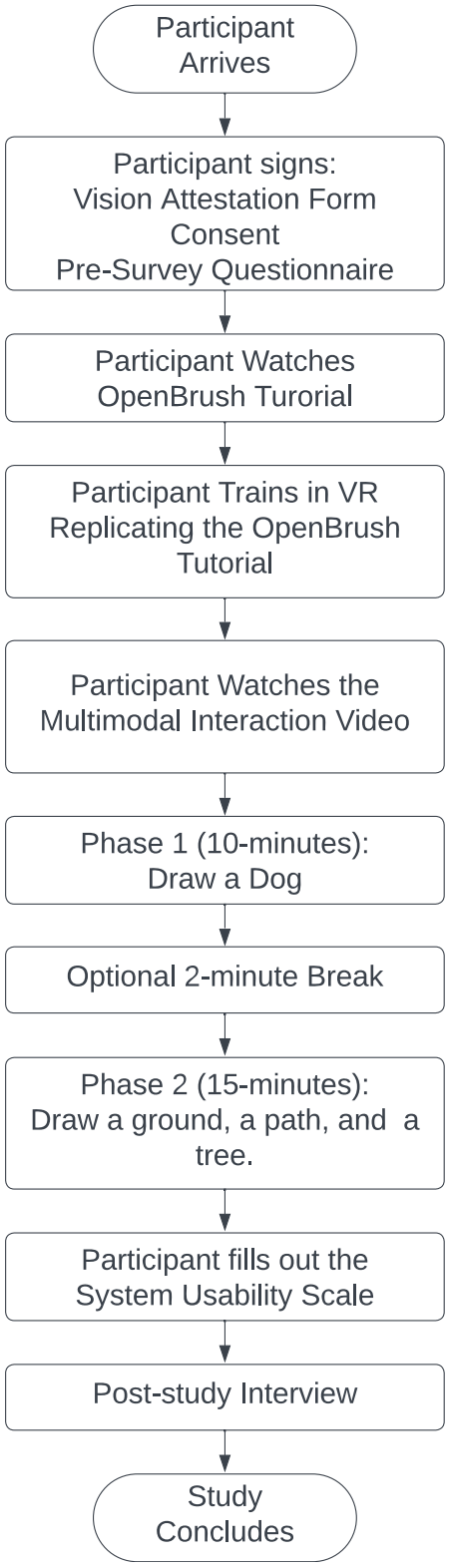

Figure 4: Each participant followed the same script, pictured above, throughout the study.

The following section further develops the paper's themes to describe the requested features and input modalities artists suggested for 3D sketching systems.

5.1.1 Alternative Modalities to Current Features. Our data indicates that the study's participants identified the need to remap current tools in the 3D sketching system tested (OpenBrush) to novel input methods. This remapping does not modify the existing functionality of the tool but rather a way to control it. We grouped these suggestions into three main categories: brush, object interaction, and menu (Figure 5.) The Brush category includes any interaction that affects the brush style. The Object Interaction category includes any action that selects or manipulates the object/stroke of the drawing. Finally, the Menu category includes choosing a tool or doing an action from a menu.

Brush. In 3D sketching systems, the brush tool is fundamental for users to create new strokes by moving the VR controller in space. Interestingly, most participants did not mention changing the input method to draw strokes. Only two participants suggested other ways to create strokes. P10 mentioned that a physical, real-world pen would be a useful interaction method to accomplish the same functionality as the controller. P2 mentioned using a gesture plus the controller to redraw strokes by selecting a stroke and adding vertices to it. P2 described this as, *"adjust it [...] like grab [...] certain [...] parts of it like I can grab this middle part like by selecting it and [...] use my hands [...] to like stretch it in the way that I want it to look."* This interaction is known as redrawing [6], and is present in applications such as Adobe Illustrator [1] and Adobe Photoshop [2].

One important aspect of the brush tool is its characteristics of a stroke drawn by moving the controller. In most 3D sketching systems, these characteristics control a stroke's color, texture, and width. Users change brush's characteristics via settings found in a menu that sits on the opposite controller's virtual menu system in the 3D space. For artists, access to changing the brush's settings could be improved through gestures. Yet, among the participants, there was no consensus on which gestures to use. P12 suggested natural gestures like swiping left or right, *"if there is a type of motion where I can just like maybe like swipe like a certain way to like just like change brushes."* On the other hand, P8 suggested wrist movements, describing, *"maybe a wrist flick to be able to change between the two brushes."*

Object Interaction. Unlike traditional 2D sketching with pen-and-paper, 3D strokes exist as objects in space that the user can manipulate (e.g., translate, rotate, and scale). Users can also manipulate other objects inside the environment, like drawing guides. Most 3D sketching systems allow users to manipulate these objects using one- or two-handed interactions with the controllers. Interacting with objects is an important task for artists, whether moving the object or affecting it by changing its properties. Participants suggested manipulating objects with other input modalities, such as gesture, speech, gaze, or bimanual interaction.

For unimodal input methods, participants who suggested using gestures mentioned the need for more natural interactions with the hand. One example of this is P2, who said that if it were possible to *"grab this middle part like by selecting it and like use my hands or something like that to like stretch it in the way that I want it to look."*

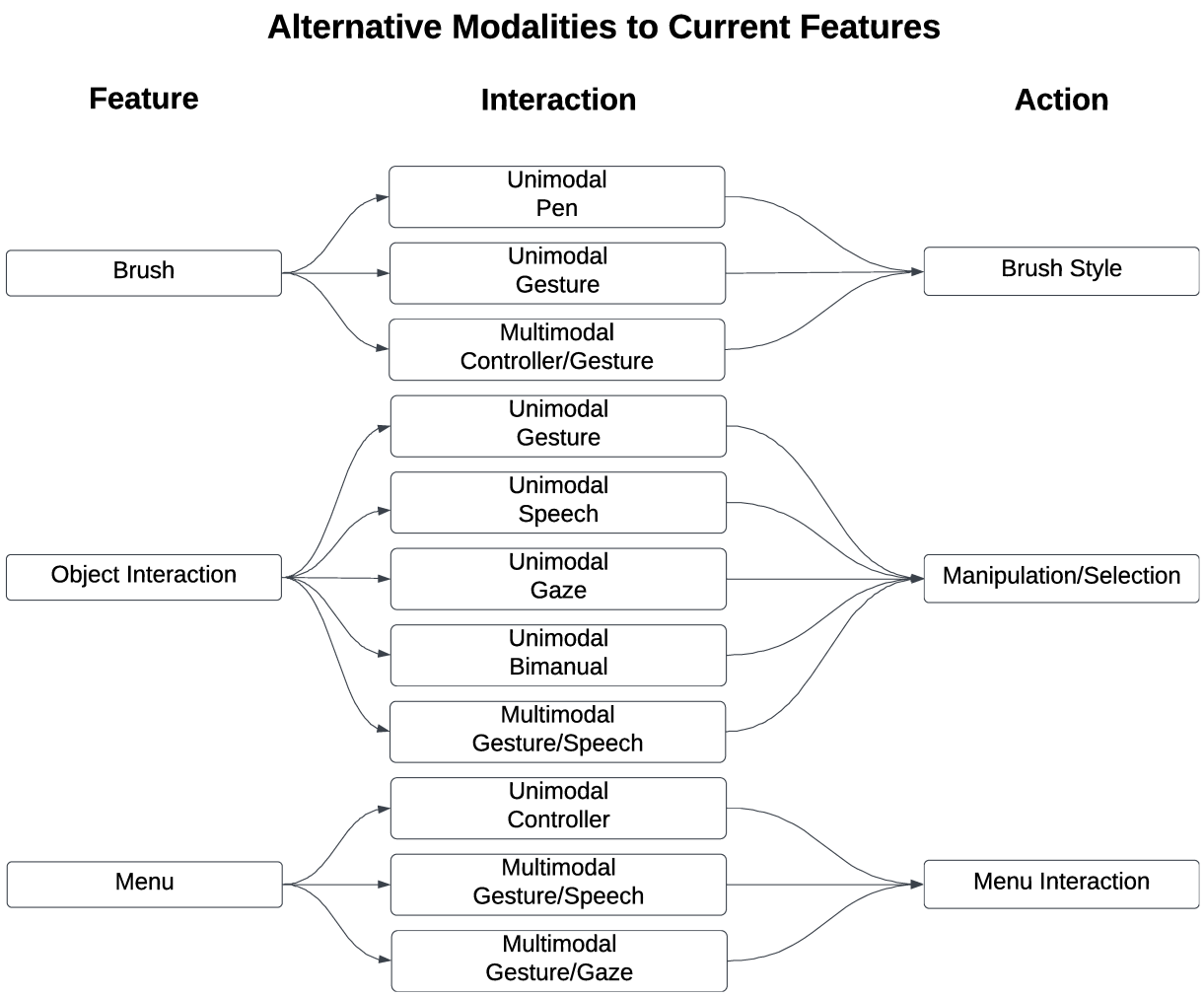

Figure 5: The diagram shows how the main categories (left column) could potentially be remapped into other modalities (middle column) and the effect that it would have on an action (right column).

P2 stated that this method would be preferable to using a controller to scale the stroke. Other participants also wanted to use their hands, but in a bimanual interaction. For example, P3 mentioned that if *"you could kind of use both hands to, like, grow a selection around something from a distance."* The participants suggested other input modalities, like speech and gaze, to make the interaction faster. For example, P13 wanted the ability to use speech to *"select everything and all of the dots I've drawn,"* and P7 mentioned *"if I was looking there and I could just kind of grow a selection where I was looking."*

The participants also suggested multimodal interactions for object manipulation. Examples of proposed multimodal interactions include merging gesture and speech. An example of this is P2's suggestion to use gesture and speech to delete strokes, *"I could probably like point at it and like tell it to erase it."* Also, while attempting to select strokes, P1 mentioned that gesture and gaze would be a good way to manipulate strokes, *"I feel like that would be a gaze with [...] my hand gesture."*

Menu. Accessing the menu is important to reveal all the tools available to participants. The menu allow users to modify the properties or characteristics of the strokes in the 3D environment, like changing colors, textures, or brush width. P9 and P13 suggested extending the current way to switch between tools or properties; P9 wanted to continue using the controller to alternate between tools by *"double click[ing] on a button to go back to [the] previous tool."* Similarly, P13 did not want to switch to a different input modality but instead wanted to use a different combination on the controller to switch colors. P13 demonstrated such action to the researchers by tapping on the controller trackpad. While both participants preferred the controller for the current unimodal input, their methods for switching between tools differed slightly.

Other participants felt comfortable using multimodal inputs to interact in the environment. P2 wanted to use a combination of gesture and speech to erase strokes in the environment. In using gesture followed immediately by the verbal command *"tell it to erase it,"* P2 hoped to minimize accessing the menu multiple times -

one time to perform a selection, and the second to access the erase feature from the menu. In contrast, P1 wanted to minimize the time needed to access the menu when duplicating strokes. Duplicating strokes involves selecting the strokes that will be duplicated, followed by another menu command to duplicate them. P1 hoped to save time by looking at the strokes that needed to be selected. Then, while doing a circular motion on the controller with the *"hands and then I used the gesture right here"* to duplicate the strokes. Both participants wanted to save time by minimizing the number of times they needed to access the menu to perform common tasks. Accessing the menu multiple times would have distracted the participants, but multimodal inputs could have allowed them to focus on the task at hand.

5.1.2 Proposed Features. Some of the participants' suggestions on new functionalities are not currently available in OpenBrush. We also examined various tools and 3D drawing software available in the market, including Open Brush, Gravity Sketch, ShapesXR [29], Paint 3D [68], Paint.Net [30], Photoshop, and Blender [37] (Table 2), and could only identify one solution that met the suggestions of the participants in Blender, which provides basic functionality for manipulating objects [36] in VR. We grouped these suggestions in five main categories, creation, manipulations, menu, selection, and animation (Figure 6 and Figure 7), and discussed them in detail below. The creation category is for creating objects, other than strokes, in the environment. The manipulations category allows the participant to alter the appearance of a stroke by splitting it, sculpting it, moving it, or erasing it from the environment. The beautification feature takes a non-straight line and ties all the points together into a perfect line. The proposed menu category would provide access to a menu or a set of sequential commands. The selection category would allow selection through other input modalities, such as speech, and grouping of multiple strokes via the controller. The animation category proposes a simulation that is composed of interactions between objects, and this simulation keeps repeating.

Creation. While users can manipulate objects via the standard translation, rotation, and scaling, adding additional details, such as texture, is not a feature that is currently available in the application. P1 would have preferred to alter a selected stroke to reflect a particular aesthetic vision. P1 wanted to create a specific texture, but could not do so due to the current limitation of the software. Another aspect was that 7 participants were interested in turning strokes that resembled a shape into a perfect geometric shape. Artists commonly use applications, such as Adobe Photoshop and Blender, to create geometric shapes from drawings. In Notability [38] for the iPad, this feature is known as perfect shapes, where the application, based on a machine learning model attempts to approximate the shape that the user is drawing and creates a perfect shape, replacing the user's drawing. This technique is also known as beautification. The approach for beautification differed slightly among the participants who proposed the feature. P4 suggested using speech to generate a 3-dimensional flat circle, not a sphere, by saying, *"large circle, or something like that."* On the other hand, P11 wanted to use speech to generate objects, but in this case, P11 wanted to generate full 3D shapes, such as a sphere or a cube. Furthermore, P11 wanted to be as specific as possible on where the 3D shape had to go by saying *"I want this on [...] the Z plane or the Y plane."* While the requests were similar, generating the requested shapes differed slightly. In contrast, P9 was interested in generating custom shapes. P9 wanted to generate fur on the side of the dog by issuing the verbal command, *"generate [fur] all over the surface."*

Participants P8 and P11 (who use digital drawing applications) were interested in not only generating shapes but also filling the surface created by strokes or filling the volume. P8 and P11 agreed that filling the surface created by strokes was important. They differed in the object that was being filled. While painting the grass, P8 suggested a *"fill feature so I could [...] connect a line here and then use a paint bucket to fill this all green would be interesting."* In contrast, P11 wanted to perform the same function but to fill the surface of a pre-made shape. In extending P8's request, P12 wanted to fill the surface of any surface, regardless of the number of strokes that the object was made of. One observation is that the three participants (i.e., P8, P11, and P12) wanted to use only speech for the fill feature. However, P13 wanted a similar function by using gestures. When attempting to fill the volume of an object, P13 mentioned that *"you could like make the shapes [...] come in filled"* by gesturing towards the object. While speech and gesture were the most common inputs, the preferred unimodal input was speech. Interestingly, two participants, P1 and P9, mentioned being assisted by artificial intelligence (AI), such as P1, after drawing a dog, wanted *"kind of AI generated to give you this."*

Although the 3D sketching application allows participants to use their dominant hands to draw, it is limited by not allowing both hands to select strokes or draw. P8 would have liked to spread both arms to select all strokes that appeared between them from the headset's perspective. Instead of using both hands to control the selection, P4 wanted to use the non-dominant hand to control the size of the stroke being drawn by the current brush. In the current system, the stroke size can be controlled by the dominant hand by swiping left or right on the controller trackpad but not by the opposite controller. In contrast, P3 wanted to be more involved in the drawing by using both hands (bimanual) to draw independently. While there was a disagreement on how they would use both hands to affect their drawing, the participants mentioned they would have benefited from using bimanual interaction to advance their drawings.

Manipulations. Artists may start with mental images of what they envision, but they may modify their visions as the drawing progresses. In order to allow for modification, participants proposed manipulating strokes using a set of inputs that includes beautification, stroke splitting, sculpting, moving, and erasing features. The beautification of shapes was previously mentioned, but one participant wanted the beautification of single lines. P2 wanted to turn a stroke into a straight line by speaking *"make the line straight"* through the microphone (i.e., speech). P6 found it difficult to create a flat surface to draw the path and thus wanted the controller to have the ability to create a flat surface in the environment. P8 wanted to use straight lines. Unlike P2, however, P8 did not want a stroke to be beautified into a straight line, but rather wanted the application to draw a straight line.

In 2D, adjusting a stroke could be done by splitting it or removing part of it. In the tested application, a stroke can be removed or left as-is, but it cannot be split. P7 mentioned that erasing *"the whole*

Table 2: A comparison of the features across various commercially available desktop, VR, and a hybrid desktop/VR application.

Application	Detailing	Filling	Generating	Beautification	Stroke Splitting	Sculpting	Moving	Erasing	Shortcut	Tool Selection	Menu	Grouping	Selection	Animation
Open Brush	No	No	No	No	No	No	Yes ●	Yes ●	Yes ●	Yes ●	Yes ●	No	Yes ●	No
Gravity Sketch	Yes	No	Yes	No	No	No	Yes ●	Yes ●	Yes ●	Yes ●	Yes ●	No	Yes ●	No
ShapesXR	No	No	Yes	No	No	No	Yes ●	Yes ●	No	Yes ●	Yes ●	Yes	Yes ●	Yes ▲
Paint 3D	Yes	Yes	Yes	No	Yes	No	Yes ●	Yes	Yes	Yes	Yes	Yes	Yes	No
Paint.Net	No	Yes	Yes	No	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
Photoshop	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Blender ⊛	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Legend:
● Unimodal via the VR controller only.
▲ Limited animation when hovering over an object.
⊛ This application is a desktop system but, offers limited VR support.

VR Applications
Desktop Applications
Desktop/VR Application

stroke and not just like individual parts of the stroke" was inefficient, as the participant would need to account for additional time to create new strokes by having to erase the current stroke, then creating two additional strokes to give the appearance of a split stroke. To resolve that, P2 suggested splitting a stroke by saying *"pull it apart"* while using a gesture, issuing a verbal command by saying *"split this line*," or using a slicing gesture on the stroke.

Some branches of fine arts, like sculpting or even painting, can require artists to use their hands when working with clay or clay-like materials. P10 and P11, who enjoy sculpting, would like to see sculpting offered in future releases of OpenBrush. P10 wanted to use pre-made geometric shapes with the volume inside them filled to *"just start kind of like sculpting"* from the outside and working towards the inside. When asked if there was a preference between drawing and sculpting, P10 responded by saying that using hands for *"sculpting [...] would probably be even more preferable."* It is clear that the participants were trying to associate previous knowledge from real-life sculpting to sculpting in VR.

Finally, six participants wanted better control of the strokes or an alternate way to remove them. In the current version of OpenBrush, to select a stroke, the user has to make contact with the controller and the stroke. Instead of walking to a stroke to select it with the controller and then move it to another position, P11 wanted to *"point at something and say like or just like being able to point to something and grab it,"* as in using ray-cast pointing to select strokes that were far away. P11 also wanted to use ray-cast pointing to highlight an object to either verbally tell the application to select it or grab it with the controller and then move it to a more suitable location. Similarly, P4 wanted to be able to erase a stroke by just *"point[ing] at it and like tell it to erase it."* In the case of these two participants, a multimodal interaction would have been suitable to accomplish their goal.

Menu. As each participant had taken at least one digital art class, they had experience using application interface menus. Although some applications on the desktop support accessing menus via speech, the tested VR 3D sketching application did not. P11 wanted to access the tools in the menu employing speech by merely *"say[ing] the name"* of the shortcut corresponding to the menu. From the participant's view, a shortcut, just like the shortcuts found on popular applications like Adobe Photoshop, allows the participant to reach a tool or an action by skipping several menus, thus saving time. When painting on a 2D digital canvas like Procreate [88] on an iPad, an artist can use a side palette to test out the brush size and color before using it to digitally draw with. While the tested application allows the participant to change the size of the controller by swiping left or right, P8 suggested a different method to access the tool by pressing on the controller trackpad rather than swiping left or right. The reasoning behind this, as P8 explained, is *"to make that be a part of the trackpad, because it is a little bit choppy."* As P8 was swiping on the controller, the location of the controller in the VR environment was constantly drifting. At the same time, P8 suggested removing the menu on the non-dominant hand. The head rotation required to look at the non-dominant menu hand and select a different tool was described as distracting. P8's reason follows: *"when I have to stop and find this button, I mean it is not that hard to find, but some way that you could swipe up on the trackpad and open a menu would be, I think, a little bit more efficient."* A pop-up menu close to the dominant (or drawing) controller would have been more efficient by minimizing the time needed to rotate the head.

Selection. An important aspect of 3D systems, such as OpenBrush, is the ability to select specific strokes or a group of strokes. Selecting strokes allows the user to erase or duplicate a single stroke or multiple strokes, which minimizes the time the user has to spend to erase or duplicate them. P3 would have liked to select strokes by using a bimanual interaction, like a T-pose, where the distance between the hands hands would indicate the range of the desired selection. Another way the same participant wanted to do a stroke selection was by using speech. P4, P7, and P9 agreed on using speech to select all the strokes in the environment by saying *"select all."* P8 suggested two different methods: using a dedicated button on the controller, which P9 agreed on, or using a combination of speech and gesture. Stroke selection would *"probably use gaze*," according to P12, who was asked which modality of interaction would be preferred for selecting strokes. P13 felt that speech would be

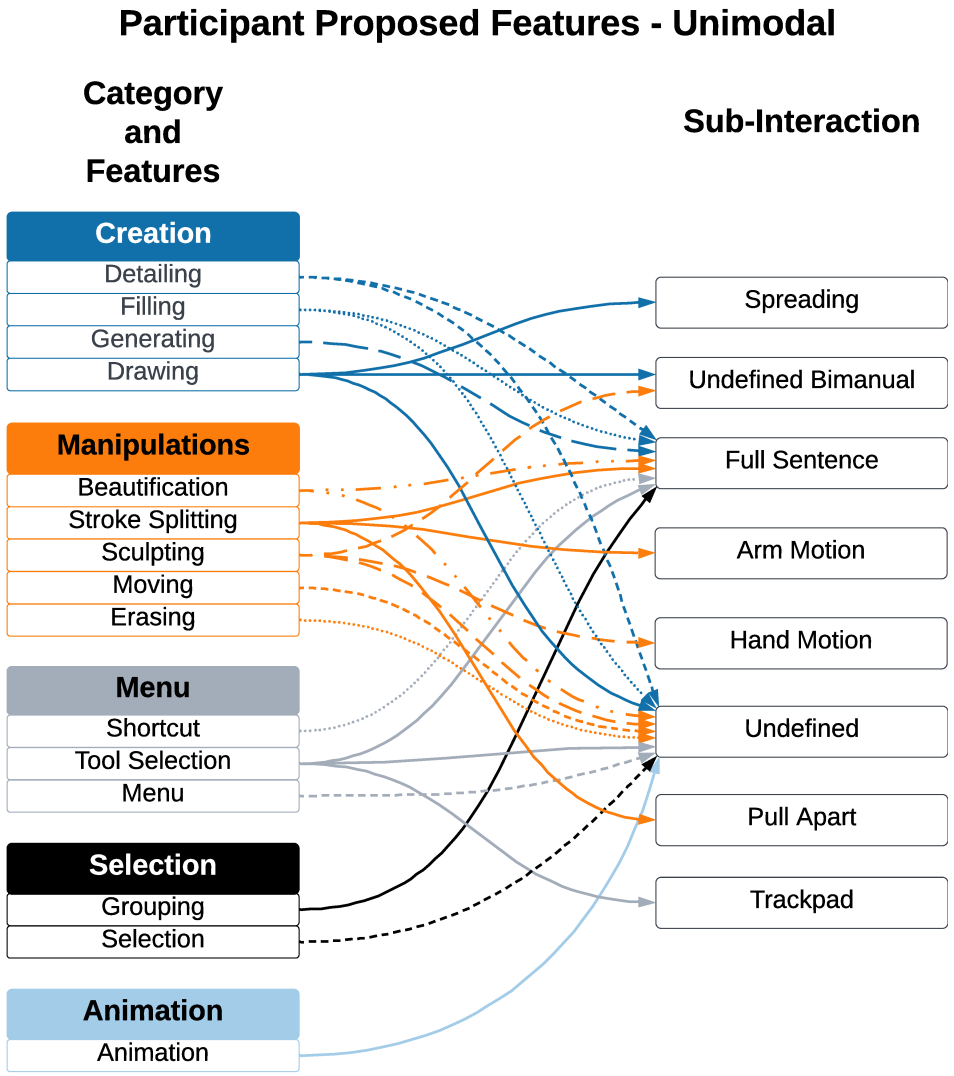

Figure 6: Diagram showing the proposed unimodal features (left column) and the sub-interactions (right column) they map into. The sub-interactions labeled undefined (i.e.: undefined bimanual, undefined) refer to an specific interaction where the participant did not mention how to accomplish the given interaction. For example, for drawing, the participant mentioned they wanted to use a bimanual interaction but did not specify the hand/arm movements they would use, therefore, the participant did not define it.

useful in selecting all the strokes by echoing the command, *"select everything,"* which would group all the strokes in the environment.

Animation. While the tested application (OpenBrush) allows participants to showcase their creative side, animation is not supported. Some brush effects perform an animation as part of their texture, but the participant does not have any control over this animation. P1 wanted to create a custom animation that kept repeating itself: the effect of lightning coming out of bubbles. While this could not be created, due to the limitation of the software, P1 said that it *"would be nice"* if that feature existed.

Multimodal Features. Multimodal interaction refers to an interaction that involves two or more input modalities being used to accomplish a task in the system (see Figure 7). For example, a participant may want to point to a stroke and say delete. For selection, P8 was the only one that suggested using a combination of speech and gesture. When grouping the features into common categories, it was found that participants in our study mostly proposed multimodal interaction techniques for creation tasks.

Multimodal Creation. Participants proposed specific features for filling shapes or objects with colors or textures and generating

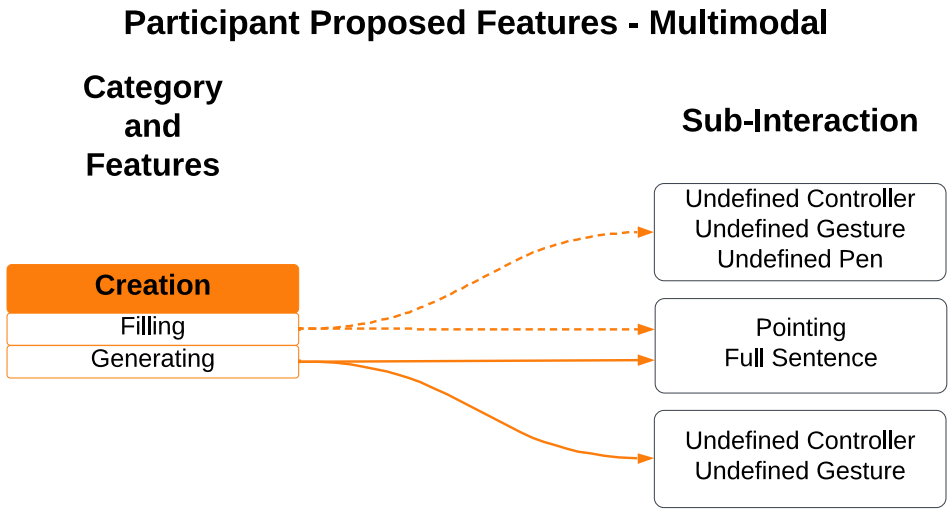

Figure 7: Diagram showing the proposed multimodal features (left column) and the sub-interactions (right column) they map into. The sub-interactions labeled undefined (i.e.: undefined controller, undefined gesture) refer to an specific interaction where the participant did not mention how to accomplish the given interaction. For example, for filling, the participant mentioned they wanted to use the controller but did not specify which button or combinations of buttons to use on the controller, therefore, the participant did not define it.

shapes and objects. As with the unimodal case of this feature category, these features were grouped under "Creation" since they would involve creating additional content in the VE. Unlike the Creation category for unimodal interactions, however, no detailing or drawing features were proposed for use with multimodal interaction techniques.

Filling. Participants also expressed the desire for OpenBrush to allow them to fill the inside or surface of an object or shape. Although some proposed techniques for accomplishing this involved unimodal interactions, others proposed multimodal interaction techniques. P2 proposed multimodal interaction technique, to point at an existing object and then use speech to fill it with color or texture. To fill in a tree, for instance, P2 described, *"Pointing at it, telling it [...], 'Fill this tree up with green.' "* This approach entailed drawing some kind of outline to indicate the tree, which P2 said could possibly mean drawing the wireframe for the object. It was not clear whether P2 meant creating a wireframe mesh, as is found in 3D modeling, or simply drawing an outline of the object and then specifying that it should be filled. Filling shapes/objects was also proposed by P10 to be accomplished through the coordinated use of the controller, a pen, and gesture. This interaction technique was focused primarily on texture and would involve selecting the drawn outline of a shape/object with the controller and then using the gesture and pen in undefined ways to fill the object with a desired texture.

Generating. During drawing tasks, participants wanted to be able to generate objects and shapes in OpenBrush. As described previously, some of the proposed interaction techniques for this desired feature only involved unimodal interactions. Other proposed interaction techniques for generating shapes and objects involved multiple modalities working in tandem. This sometimes involved a combination of full-sentence speech and pointing. When asked if an alternative interaction technique could help create the ground, P2 wanted to *"Point at, like say, two points [...] and say, 'Make a square.' "* P2 further elaborated this proposed interaction technique by pointing to two separate points, such as the opposite corners of a square, followed by the verbal command to make a square, and the system will use those 2 points as a reference and create a square. Meanwhile, for such 3D objects as cylinders, P2 said that pointing at two points could specify the top and bottom of the object. Further details in defining the dimensions of the shapes and objects were not provided by P2. P2 also proposed generating more complex objects at a specified location by pointing and simply saying to generate this. One example given was to *"...point at, like, a certain point within, like, the bark of the tree and [...] tell it to sprout a branch.' "* Alternatively, P13 proposed using a combination of controller, gesture, and full-sentence speech to generate shapes. This interaction technique would use speech to say, as P13 described, *"Make me a circle,"* and then gesture could be used to specify where to place the shape/object while the controller would be used to control the other attributes of the shape/object, such as the size.

Because many of the proposed multimodal interaction techniques involved speech commands, implementing these interactions would involve accurate speech recognition that can also incorporate the context provided by the other interaction techniques. For instance, when pointing at an object and using speech to fill it with color, the system will need to recognize what object is being pointed at and connect that to the spoken instructions. Due to some aspects of the proposed interaction techniques being vaguely described by participants, future work would also involve identifying what kinds of gesture, controller, or pen actions would be necessary to make

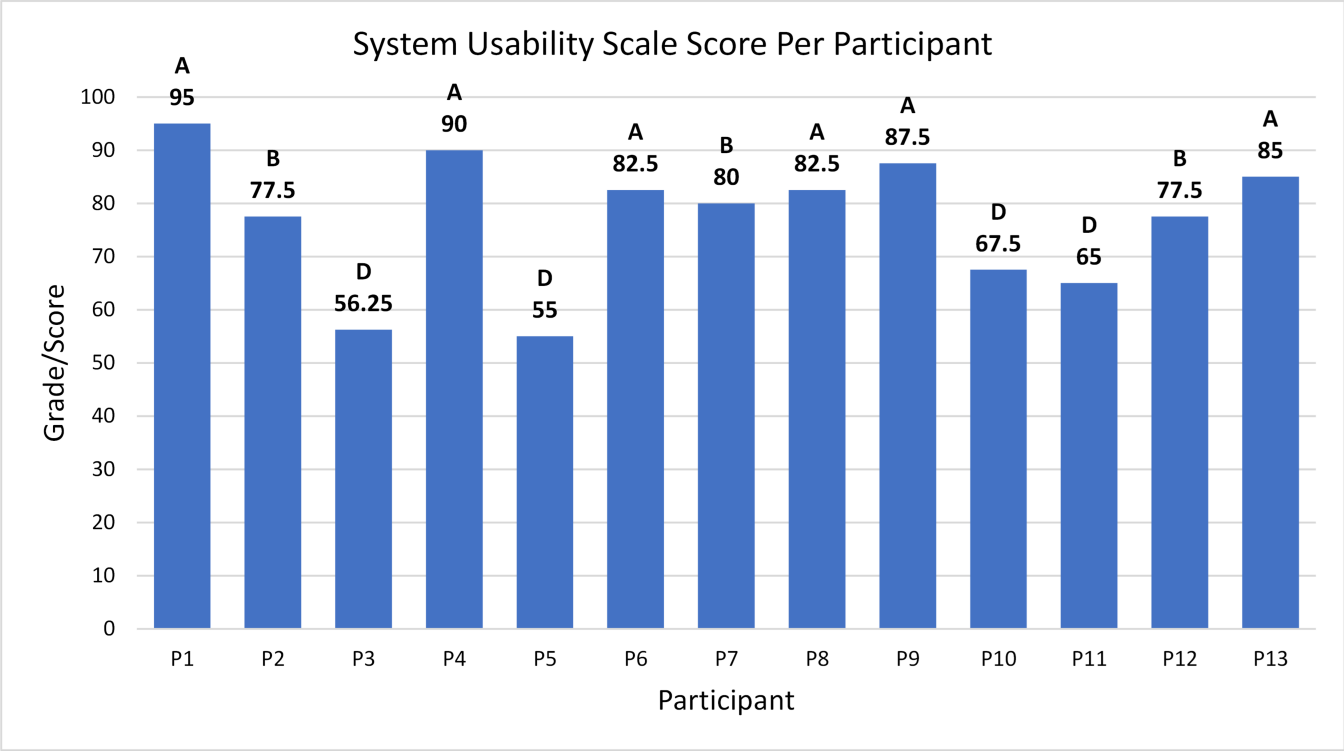

Figure 8: The calculated System Usability Scale (SUS) score and grade per participant (M=77.01, SD=12.5).

these multimodal interaction techniques effective and satisfying for users.

5.2 Quantitative Findings using the System Usability Scale

Following the need to run 3D sketching evaluations with stable systems [15], the usability of the 3D sketching system was evaluated by having participants answer the SUS questionnaire. A post-study interview was conducted to get participants' opinions on the current 3D sketching system. (See the supplementary materials for the interview questions.) Participants rated the system positively, as shown in Figure 8. The overall average score was 77.01 (SD = 12.5), corresponding to letter grade B, showing the system's usability is above average. When looking at the individual statements, participant statements indicate usability. One participant's evaluation with an average score was 4.15 (SD = 0.99) stated, *"I think that I would like to use this system frequently."* Another evaluation with an average rating of 4.08 (SD = 0.76) declared, *"I thought the system was easy to use."* Both respondents scored the system above average. The positive rating of the system, which earned it the letter grade B, is supported further by additional participant interviews. P5, for example, commented, *"I love this, this is great!"* P1 mentioned, *"That is nice, kind of very satisfying."* Similarly, in feedback about the application, P2 said that *"it is very nice."*

6 DISCUSSION

In this study, we asked artists their opinions about the tools in current commercial 3D sketching systems, and they evaluated the usability of those tools. We also ran a user study to allow trained artists to share their ideas around novel input methods for 3D sketching that are natural and multimodal. In addition, artists were asked about their perceptions of the usability of commercial 3D sketching systems.

6.1 Adding novel tools

RQ1 was about *the tools of commercial 3D sketching systems help artist in their sketching process.* Our participants identified four elements, Menu, Shape Creation, Manipulations, and Animation, of the commercial system study that did not meet their needs or could be improved. Here we describe our findings for each one of them:

Menu. One feature not available in the sketching application is the ability to maximize the display space by removing the menu on the non-dominant hand. This menu could potentially distract the artists and hide important parts of the sketch when the artist is doing a visual search. Moacdieh et al. [70] studied the effects of cluttering the display, which ultimately affects the performance. It has been shown that having a higher field of view (FOV) [83] leads to better performance. Therefore, manually hiding the menu when artists are not using it maximizes the FOV, which would allow for a better visual search, and increase the performance of the artist

when sketching in VR. Another requested feature was a preview pane to test different combinations of brushes, brush sizes, and colors. Just like the menu on the non-dominant hand, this preview pane should only be shown at the artist's request. Removing it when not necessary would maximize the FOV.

Shape Creation. One of the desired features requested by artists was the ability to create basic geometric shapes. Gravity Sketch and ShapesXR (Table 2) are VR applications that already provide this functionality, which is similar to creating shapes on desktop applications (Table 2). Basic 2D shape creation in VR has previously been explored by Barrera Machuca et al. [11] by using beautification. Two artists also wanted to generate shapes with the assistance of AI. Chen et al. [27] have explored using natural language to generate colored 3D shapes, as well as composite 3D shapes, such as tables and chairs. Incorporating such technologies would allow artists to create basic geometric shapes.

Manipulations. The artists also wanted to manipulate the strokes by turning them into straight lines or splitting them. Turning a non-straight stroke into a straight line has previously been demonstrated by the use of beautification [11]. This turns a sequence of dots that look almost straight into a straight line. On the other hand, Jiang et al. [52] showed that splitting a stroke can be done via a cut operation. The sketching application has a feature called Snip that breaks the stroke, however, this function affects the stroke not only at the location where the split occurs, but also the shapes of the newly split strokes. The cut operation only affects the stroke at the site where it is cut, leaving the rest of the shape of the two strokes unaffected. One artist wanted to manipulate objects closer by hand-sculpting them. Currently, Blender (Table 2) supports basic sculpting in VR, but for finer control and details, users still have to launch the desktop counterpart of Blender to finalize those sculptures.

Animation. The sketching application includes some brushes with animations. Several applications, such as ShapesXR (Table 2), also provide limited animation when hovering over an object. Although only one artist suggested creating animations, it would be worthwhile if a future release allows artists to create a basic animation that extends beyond the animated brushes that are already included in the application.

All of these suggested tools share a commonality: they resemble features found in other desktop and VR applications used for 3D modeling. Our results suggest that new 3D sketching applications should include tools that artists are already familiar with from other software, as the artists expect to find similar tools across software.

6.2 Adding unimodal and multimodal interactions

RQ2 concerned identifying *which natural multimodal interactions 3D sketching systems can add to help the artist in their sketching process.* Interestingly, we found that our participants mostly focused on tools not related to the process of sketching in a broad sense, but on the interactions that help manipulate the strokes such as object interaction, selection, and manipulation. We also found that when proposing multimodal interactions, our participants only proposed interactions with two different input methods, e.g., gesture and speech or controller and gesture, but no more.

Brush. As an alternative to using the VR controller to control the brush, only two artists mentioned the pen. One would think that since all the participants are artists, that they would elect to go with the pen as the preferred input device. When we analyzed the demographics of two artists who suggested a pen, we noticed that one is in their senior year and the other had recently graduated. The artist who had just graduated used VR extensively and had worked with a team to create a VR game. Because of their familiarity with HMDs and their different uses, this artist chose the pen because they wanted the most appropriate tool for interacting in VR as a sketching medium. The second artist chose the pen because they focused on digital drawing and pixel art as their primary form of artwork. These artists' daily activities and type of artistic practice affected their choice of using a brush.

Object Interaction. We observed a relationship between artists who wanted to use bimanual interaction versus artists who wanted to use gestures. A possible explanation for the choice of interaction may depend on the number of hours that each of these participants spends in front of a computer as opposed to those artists who spend the majority of their time on a controller. The artist who suggested a bimanual interaction spends considerably more time on a computer, using both hands to manage the keyboard and mouse. In contrast, the artist who predominately uses the controller must make broader gestures due to both hands being tied up with the controller.

Menu. For an alternative way to access menu, we looked at the two most prominent modalities requested by the artists to access the menu: controller and speech. The artist who chose the controller is accustomed to using similar types of input devices, such as game controllers, therefore, it may have been a natural preference influenced by their experience in gaming. The artist who chose speech selected this interaction as a personal preference, possibly related to phone use, wherein users can use voice activation to engage with their phones. What we can gather from these artists is that accessing the menu through other modalities is a personal choice that may be influenced by other familiar technologies in their environments.

Manipulations. Interestingly, there were two artists who wanted to split a stroke by pulling it apart. Both artists had experienced VR before and both were in a similar age group. The first artist works on sculptural art projects using both hands, therefore, using both hands to pull apart the stroke could be a natural translation from the physical world. The second artist primarily works with desktop applications, possibly influencing their choice to select using a bimanual interaction for pulling a stroke apart. We hypothesize that this artist may want to pull apart a stroke as they predominately use their hands to do separate operations when on the keyboard and mouse. For deleting strokes, two participants wanted to use gesture and speech combined. We hypothesize that using both gesture and speech to specifically choose a stroke and use speech to indicate what must be done with it could be a simplified process for this artist.

Selection. For selecting strokes, 7 artists suggested speech. Upon checking pre-survey demographics, we noticed that they all have previous experience with VR systems. We hypothesize that because VR applications often come with audio-visual tutorials, but not an in-depth written manual, speech is a more natural response in the environment—allowing the application to process the verbal command on their behalf.

6.3 System usability scale (SUS)

RQ3 was about examining *the artists' opinion on the usability of commercial 3D sketching systems.* Our findings show a high usability score for the sketching application (Figure 8). However, there were four artists who rated the system with a 'D' or below average. The pre-survey questionnaire revealed that these artists played computer games for less than 6 hours on a weekly basis. We speculate that for people who are accustomed to playing games, it is relatively easy to switch from one game controller to another, or a VR controller for that matter [3]. Also, the only way that artists were able to directly interact with the sketching application was through the VR controllers. Using this unfamiliar input technology might have been challenging for them, leading to frustration and a lower usability score.

Table 3: This table illustrates the preferred modality per feature among the participants. The features stroke splitting and erasing were the only two categories with ties.

Category	Feature	Preferred Interaction
Creation	Detailing	Bimanual
	Filling	Speech
	Generating	Speech
	Drawing	Bimanual
Manipulations	Beautification	Controller
	Stroke Splitting	Controller, Gesture
	Sculpting	Gesture
	Moving	Speech
	Erasing	Gesture, Gesture+Speech
Menu	Shortcut	Controller
	Tool Selection	Speech
	Menu	Controller
Selection	Grouping	Controller
	Selection	Speech
Animation	Animation	Gaze

7 RECOMMENDATIONS TO ADD NOVEL TOOLS AND UNIMODAL/MULTIMODAL NATURAL INTERACTIONS TO 3D SKETCHING SYSTEMS

Based on the interviews and the feedback received, we have created recommendations for adding novel tools for future sketching application.

7.1 Adding novel tools

During the interviews, the artists made suggestions for tools that could help them during their sketching process. One tool would be a disappearing menu and a preview pane. These should only appear at the request of the artist, and can be tied to an on-and-off mechanism, e.g., a toggle button on the controller or a spoken command, such as "show me the menu".

The artists also mentioned the generation of primitive shapes, which would help them save time for sketching. Two artists mentioned the possibility of being assisted by AI. Currently, large language models, such as ChatGPT by OpenAI [78], can be integrated into game engines [95]. Such similar software can help artists as novel tools.

Developers can implement beautification techniques, such as those displayed in Multiplanes [11] to estimate the shape that the artist is attempting to draw. Methods such as Nestor [42] and GCN [104], which uses a neural network, have also been shown to be successful. For the splitting function, as in HandPainter [52], a stroke splitting can break a mesh into two watertight pieces, and two new objects should be added as nodes to the scene graph.

Allowing artists to implement their animation can be challenging. Developers can create empty animation objects that can be used by the artist to assign modified object properties, and these can be saved in the timeline as they are added. By letting the artists play the animation, they can see how the object changes based on the modified properties.

7.2 Adding unimodal and multimodal natural interactions

During our study, the only way to sketch in the application was through the VR controllers. Every artist has a personalized style, and the sketching application should be able to accommodate that style. As it was suggested by the participants, other unimodal interactions should be made available to cater to each artist's style. Additionally, some participants suggested adding multimodal interactions. Therefore, we recommend adding other input modalities, such as gesture and speech, to help the artists navigate the software in a more natural way. By adding this additional functionality the participant will be able to use the software in a natural way.

Additionally, the hardware used in this study exists to meet general VR user needs. While the current hardware allows 3D sketching to be done, designing controllers that are more specifically oriented toward sketching and other artistic applications could offer additional options for users to interact effectively with the system. Designing hardware options in varying sizes could also accommodate users with different hand sizes and different ranges of mobility. By increasing the hardware options along with the software capabilities, users can direct more conscious effort towards working and creating and less on interacting with the tools.

Having multiple modalities available for interaction provides a versatile toolkit for expressing users' ideas. Users can start with quick, fluid gestures to lay down the basic structure of their sketches. They could then use speech or controller functions to refine their ideas. This encourages experimentation and exploration as users are not confined to specific tools or techniques. This freeform approach can lead to novel design concepts.

8 LIMITATIONS AND FUTURE WORK

Semi-structured interviews allowed the discovery of ways to improve 3D sketching interactions from 13 artists. However, when inspecting the preferences in Table 3, a theme of unimodal interaction is present in most of the features. Bimanual only appears in detailing and drawing features, and controller appears a few more times (which can be bimanual). However, multimodal interaction appears only in the moving feature.

We speculate that the reason why the multimodal interaction was not always preferred or present is that our participants had limited exposure to the VR/AR sketching application. 11 of the 13 participants had used a VR/AR headset before, 2 of those 11 participants mentioned using the headset for a very short period of time. Most of them had not used or had limited time working in 3D sketching. Most of the work in producing gestures has been observed using elicitation studies, e.g., [99, 109, 114], but not using the approach we took. As expressed earlier, elicitation studies have a number of limitations in complex systems such as 3D sketching. Therefore, the next step of our research is to provide HMDs to artists and ask them to use 3D sketching over a series of weeks. In addition, during this period, we plan to provide a series of weekly videos describing different methods about unimodal and multimodal interaction to familiarize artists with modalities. Then, the artists would be invited for a study when they have mastered 3D sketching. This would allow us to combine their art expertise with the experience they have gathered in the 3D environment. We hope to find that the approach improves the participants' familiarity with the different types of interaction modalities that they can propose. The second study will also increase the times that were given in this experiment and provide a multi-session approach, which is similar to the production methodology suggested by Morris et al. [72].

Another option is to seek a larger set of artist from different places in a future study, given that our study had participants from the local university and in the age range of 20 to 28, which limits the feedback received to those of younger adults only and with a certain type of experience.

9 CONCLUSION

In this paper, we proposed new ways to interact with 3D sketching systems in VR, using one or more input modalities from the artists' point of view with semi-structured interviews. The study involved interviewing 13 artists, on which we performed a thematic analysis. We identified alternative modalities to current features and proposed new features that mimic features available on desktop software, which might give artists an advantage. The suggestions about input methods made by the artists in the study were informed in part by their coursework in drawing and sculpture. The suggestions were also influenced by the participants' computer use and familiarity with gaming. We also gathered insights from artists experienced in traditional physical creation to explore ways in which developers could enhance the intuitiveness of unimodal and multimodal interactions in VR sketching. If developers can create more opportunities for artists to interact with VR sketching more naturally, it could allow artists to interact more seamlessly with their creations. It could also lead to greater adoption of VR sketching within the artistic community and those in the greater creative VR world.

We also provide recommendations for future 3D sketching systems in VR to make it easier for artists to transition from a desktop to an immersive 3D sketching environment. Some participants mentioned that implementing these new features will make the system efficient for people to interact in the VE, rather than trying to overcome the system's limitations. While 3D sketching systems in VR are great for casual users, our recommendations are based on artists' perspectives. Thus, they are geared toward ensuring that future artists are productive and efficient when sketching in 3D VEs.

ACKNOWLEDGMENTS

We like to thank and acknowledge the funding support from the National Science Foundation via grants NSF 2327569, 2238313, 2223432, 2037417, and 1948254; support from the Defense Advanced Research Projects Agency via grant DARPA HR00112110011; and support from the Office of Naval Research via grants ONR N00014-21-1-2949 and ONR N00014-21-1-2580. We also want to extend our appreciation to Adam Coler for the valuable assistance in proofreading and suggesting edits.

REFERENCES

[1] Adobe Inc. 2023. Adobe Illustrator. https://www.adobe.com/products/illustrator.html

[2] Adobe Inc. 2023. Adobe Photoshop. https://www.adobe.com/products/photoshop.html

[3] Serefraz Akyaman and Ekrem C. Alppay. 2021. A Critical Review of Video Game Controller Designs, In Game + Design Education. *Game+ Design Education: Proceedings of PUDCAD 2020* 13, 311–323. https://doi.org/10.1007/978-3-030-65060-5_25

[4] Sang-Gyun An, Yongkwan Kim, Joon H. Lee, and Seok-Hyung Bae. 2017. Collaborative Experience Prototyping of Automotive Interior in VR with 3D Sketching and Haptic Helpers. In *Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Oldenburg, Germany) *(AutomotiveUI '17)*. Association for Computing Machinery, New York, NY, USA, 183–192. https://doi.org/10.1145/3122986.3123002

[5] Dimitra Anastasiou, Cui Jian, and Desislava Zhekova. 2012. Speech and Gesture Interaction in an Ambient Assisted Living Lab. In *Proceedings of the 1st Workshop on Speech and Multimodal Interaction in Assistive Environments*, Dimitra Anastasiou, Desislava Zhekova, Cui Jian, and Robert Ross (Eds.). Association for Computational Linguistics, Jeju, Republic of Korea, 18–27. https://aclanthology.org/W12-3504

[6] Rahul Arora, Mayra Donaji Barrera Machuca, Philipp Wacker, Daniel F. Keefe, and Johann Habakuk Israel. 2022. Introduction to 3D Sketching. In *Interactive Sketch-Based Interfaces and Modelling for Design*, Alexandra Bonnici and Kenneth P. Camilleri (Eds.). River Series in Document Engineering, New York, USA, Chapter 6, 151–178. https://doi.org/10.1201/9781003360650

[7] Rahul Arora, Rubaiat Habib Kazi, Fraser Anderson, Tovi Grossman, Karan Singh, and George Fitzmaurice. 2017. Experimental Evaluation of Sketching on Surfaces in VR. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 5643–5654. https://doi.org/10.1145/3025453.3025474

[8] Rahul Arora, Rubaiat Habib Kazi, Tovi Grossman, George Fitzmaurice, and Karan Singh. 2018. SymbiosisSketch: Combining 2D & 3D Sketching for Designing Detailed 3D Objects in Situ. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC Canada). ACM, Montreal, Canada, 1–15. https://doi.org/10.1145/3173574.3173759

[9] Aaron Bangor, Philip T. Kortum, and James T. Miller. 2008. An Empirical Evaluation of the System Usability Scale. *International Journal of Humanâ€“Computer Interaction* 24, 6 (2008), 574–594. https://doi.org/10.1080/10447310802205776

[10] Mayra Donaji Barrera Machuca, Rahul Arora, Philipp Wacker, Daniel F. Keefe, and Johann Habakuk Israel. 2022. Interaction Devices and Techniques for 3D Sketching. In *Interactive Sketch-Based Interfaces and Modelling for Design*, Alexandra Bonnici and Kenneth P. Camilleri (Eds.). River Series in Document Engineering, New York, USA, Chapter 8, 229–249. https://doi.org/10.1201/9781003360650

[11] Mayra Donaji Barrera Machuca, Paul Asente, Wolfgang Stuerzlinger, Jingwan Lu, and Byungmoon Kim. 2018. Multiplanes: Assisted Freehand VR Sketching. In *Symposium on Spatial User Interaction* (Berlin, Germany) *(SUI '18)*. Association for Computing Machinery, New York, NY, USA, 36–47. https://doi.org/10.1145/3267782.3267786

[12] Mayra Donaji Barrera Machuca and Wolfgang Stuerzlinger. 2019. The Effect of Stereo Display Deficiencies on Virtual Hand Pointing. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, Article 207, 14 pages. https://doi.org/10.1145/3290605.3300437

[13] Mayra Donaji Barrera Machuca, Wolfgang Stuerzlinger, and Paul Asente. 2019. The Effect of Spatial Ability on Immersive 3D Drawing. In *Proceedings of the 2019 on Creativity and Cognition* (San Diego, CA, USA) *(C&C '19)*. Association for Computing Machinery, New York, NY, USA, 173–186. https://doi.org/10.1145/3325480.3325489

[14] Mayra Donaji Barrera Machuca, Wolfgang Stuerzlinger, and Paul Asente. 2019. Smart3DGuides: Making Unconstrained Immersive 3D Drawing More Accurate. In *25th ACM Symposium on Virtual Reality Software and Technology* (Parramatta, NSW, Australia) *(VRST '19)*. Association for Computing Machinery, New York, NY, USA, Article 37, 13 pages. https://doi.org/10.1145/3359996.3364254

[15] Mayra Donaji Barrera Barrera Machuca, Johann Habakuk Israel, Daniel F. Keefe, and Wolfgang Stuerzlinger. 2023. Toward More Comprehensive Evaluations of 3D Immersive Sketching, Drawing, and Painting. *IEEE Transactions on Visualization and Computer Graphics* (2023), 1–18. https://doi.org/10.1109/TVCG.2023.3276291

[16] Anil Ufuk Batmaz, Mayra Donaji Barrera Machuca, Duc Minh Pham, and Wolfgang Stuerzlinger. 2019. 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)*. IEEE, IEEE, Osaka, Japan, 585–592. https://doi.org/10.1109/VR.2019.8797975

[17] Anil Ufuk Batmaz, Mayra Donaji Barrera Machuca, Junwei Sun, and Wolfgang Stuerzlinger. 2022. The Effect of the Vergence-Accommodation Conflict on Virtual Hand Pointing in Immersive Displays. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 633, 15 pages. https://doi.org/10.1145/3491102.3502067

[18] Cindy M. Bird. 2005. How I stopped dreading and learned to love transcription. *Qualitative inquiry* 11, 2 (2005), 226–248. https://doi.org/10.1177/107780040427341

[19] Richard A. Bolt. 1980. "Put-That-There": Voice and Gesture at the Graphics Interface. In *Proceedings of the 7th Annual Conference on Computer Graphics and Interactive Techniques* (Seattle, Washington, USA) *(SIGGRAPH '80)*. Association for Computing Machinery, New York, NY, USA, 262–270. https://doi.org/10.1145/800250.807503

[20] Patrick Bourdot, Thomas Convard, Flavien Picon, Mehdi Ammi, Damien Touraine, and Jean-Marc Vézien. 2010. VR–CAD integration: Multimodal immersive interaction and advanced haptic paradigms for implicit edition of CAD models. *Computer-Aided Design* 42, 5 (2010), 445–461. https://doi.org/10.1016/j.cad.2008.10.014 Advanced and Emerging Virtual and Augmented Reality Technologies in Product Design.

[21] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101. https://doi.org/10.1191/1478088706qp063oa

[22] Virginia Braun and Victoria Clarke. 2014. What can "thematic analysis" offer health and wellbeing researchers? *International journal of qualitative studies on health and well-being* 9, 1 (2014), 26152. https://doi.org/10.3402/qhw.v9.26152

[23] Virginia Braun and Victoria Clarke. 2019. Reflecting on reflexive thematic analysis. *Qualitative research in sport, exercise and health* 11, 4 (2019), 589–597. https://doi.org/10.1080/2159676X.2019.1628806

[24] Sébastien Carbini, Lionel Delphin-Poulat, Laurence Perron, and Jean E. Viallet. 2006. From a Wizard of Oz Experiment to a Real Time Speech and Gesture Multimodal Interface. *Signal Process.* 86, 12 (dec 2006), 3559–3577. https://doi.org/10.1016/j.sigpro.2006.04.001

[25] Damien Chamaret and Paul Richard. 2008. Multimodal Prop-Based Interaction with Virtual Mock-up: CAD Model Integration and Human Performance Evaluation. In *Proceedings of the 2008 ACM Symposium on Virtual Reality Software and Technology* (Bordeaux, France) *(VRST '08)*. Association for Computing Machinery, New York, NY, USA, 259–260. https://doi.org/10.1145/1450579.1450642

[26] Edwin Chan, Teddy Seyed, Wolfgang Stuerzlinger, Xing-Dong Yang, and Frank Maurer. 2016. User Elicitation on Single-Hand Microgestures. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 3403–3414. https://doi.org/10.1145/2858036.2858589

[27] Kevin Chen, Christopher B. Choy, Manolis Savva, Angel X. Chang, Thomas Funkhouser, and Silvio Savarese. 2019. Text2Shape: Generating Shapes from Natural Language by Learning Joint Embeddings. In *Computer Vision – ACCV 2018*, C. V. Jawahar, Hongdong Li, Greg Mori, and Konrad Schindler (Eds.). Springer International Publishing, Cham, 100–116. https://doi.org/10.48550/arXiv.1803.08495

[28] Victoria Clarke and Virginia Braun. 2021. Thematic analysis: a practical guide. *Thematic Analysis* (2021), 1–100.

[29] Shapes Corp. 2023. ShapesXR. https://www.shapesxr.com/

[30] dotPDN LLC. 2023. Paint.net. https://www.getpaint.net/

[31] Tobias Drey, Jan Gugenheimer, Julian Karlbauer, Maximilian Milo, and Enrico Rukzio. 2020. 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* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3313831.3376628

[32] Philip Ekströmer, Jens Wängdahl, and Renee Wever. 2018. Virtual Reality Sketching for Design Ideation. , 9 pages.

[33] Hesham Elsayed, Mayra Donaji Barrera Machuca, Christian Schaarschmidt, Karola Marky, Florian Müller, Jan Riemann, Andrii Matviienko, Martin Schmitz, Martin Weigel, and Max Mühlhäuser. 2020. VRSketchPen: Unconstrained Haptic Assistance for Sketching in Virtual 3D Environments. In *26th ACM Symposium on Virtual Reality Software and Technology* (Virtual Event, Canada) *(VRST '20)*. Association for Computing Machinery, New York, NY, USA, Article 3, 11 pages. https://doi.org/10.1145/3385956.3418953

[34] Michele Fiorentino, Giuseppe Monno, Pietro A. Renzulli, and Antonio E. Uva. 2003. 3D Sketch Stroke Segmentation and Fitting in Virtual Reality. In *International conference on the Computer Graphics and Vision*. GraphiCon'2003, Moscow, Russia, 8.

[35] Icosa Foundation. 2023. Open Brush. https://openbrush.app/.

[36] The Blender Foundation. 2002. Blender 3.0: Virtual Reality. https://wiki.blender.org/wiki/Reference/Release_Notes/3.0/Virtual_Reality

[37] The Blender Foundation. 2023. Blender. http://www.blender.org/

[38] Ginger Labs, Inc. 2023. Notability. https://notability.com/

[39] Google. 2023. Tilt Brush. https://www.tiltbrush.com/.

[40] Jen Grey. 2002. Human-computer interaction in life drawing, a fine artist's perspective. In *Proceedings Sixth International Conference on Information Visualisation*. IEEE, London, UK, 761–770. https://doi.org/10.1109/IV.2002.1028866

[41] Tovi Grossman, Ravin Balakrishnan, Gordon Kurtenbach, George Fitzmaurice, Azam Khan, and Bill Buxton. 2002. Creating Principal 3D Curves with Digital Tape Drawing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Minneapolis, Minnesota, USA) *(CHI '02)*. Association for Computing Machinery, New York, NY, USA, 121–128. https://doi.org/10.1145/503376.503398

[42] Nate Hagbi, Oriel Bergig, Jihad El-Sana, and Mark Billinghurst. 2011. Shape Recognition and Pose Estimation for Mobile Augmented Reality. *IEEE Transactions on Visualization and Computer Graphics* 17, 10 (2011), 1369–1379. https://doi.org/10.1109/TVCG.2010.241

[43] Alexander G. Hauptmann. 1989. Speech and Gestures for Graphic Image Manipulation. *SIGCHI Bull.* 20, SI (mar 1989), 241–245. https://doi.org/10.1145/67450.67496

[44] Julia Himmelsbach, Markus Garschall, Sebastian Egger, Susanne Steffek, and Manfred Tscheligi. 2015. Enabling Accessibility through Multimodality? Interaction Modality Choices of Older Adults. In *Proceedings of the 14th International Conference on Mobile and Ubiquitous Multimedia* (Linz, Austria) *(MUM '15)*. Association for Computing Machinery, New York, NY, USA, 195–199. https://doi.org/10.1145/2836041.2836060

[45] Ken Hinckley, Koji Yatani, Michel Pahud, Nicole Coddington, Jenny Rodenhouse, Andy Wilson, Hrvoje Benko, and Bill Buxton. 2010. Pen + Touch = New Tools. In *Proceedings of the 23nd Annual ACM Symposium on User Interface Software and Technology* (New York, New York, USA) *(UIST '10)*. Association for Computing Machinery, New York, NY, USA, 27–36. https://doi.org/10.1145/1866029.1866036

[46] Philipp P. Hoffmann, Hesham Elsayed, Max Mühlhäuser, Rina R. Wehbe, and Mayra Donaji Barrera Machuca. 2023. ThermalPen: Adding Thermal Haptic Feedback to 3D Sketching. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI EA '23)*. Association for Computing Machinery, New York, NY, USA, Article 474, 4 pages. https://doi.org/10.1145/3544549.3583901

[47] Samory Houzangbe, Dimitri Masson, Sylvain Fleury, David Antonio Gómez Jáuregui, Jeremy Legardeur, Simon Richir, and Nadine Couture. 2022. Is virtual reality the solution? A comparison between 3D and 2D creative sketching tools in the early design process. *Frontiers in Virtual Reality* 3 (2022), 958223. https://doi.org/10.3389/frvir.2022.958223

[48] Sylvia Irawati, Scott Green, Mark Billinghurst, Andreas Duenser, and Heedong Ko. 2006. An Evaluation of an Augmented Reality Multimodal Interface Using Speech and Paddle Gestures. In *Proceedings of the 16th International Conference on Advances in Artificial Reality and Tele-Existence* (Hangzhou, China) *(ICAT'06)*. Springer-Verlag, Berlin, Heidelberg, 272–283. https://doi.org/10.1007/11941354_28

[49] Johann Habakuk Israel, Eva Wiese, Magdalena Mateescu, Christian Zöllner, and Rainer Stark. 2009. Investigating three-dimensional sketching for early conceptual design—Results from expert discussions and user studies. *Computers & Graphics* 33, 4 (2009), 462–473. https://doi.org/10.1016/j.cag.2009.05.005

[50] Bret Jackson and Daniel F. Keefe. 2016. Lift-Off: Using Reference Imagery and Freehand Sketching to Create 3D Models in VR. *IEEE Transactions on Visualization and Computer Graphics* 22, 4 (2016), 1442–1451. https://doi.org/10.1109/TVCG.2016.2518099

[51] Hans-Christian Jetter, Roman Rädle, Tiare Feuchtner, Christoph Anthes, Judith Friedl, and Clemens N. Klokmose. 2020. "In VR, Everything is Possible!": Sketching and Simulating Spatially-Aware Interactive Spaces in Virtual Reality. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–16. https://doi.org/10.1145/3313831.3376652

[52] Ying Jiang, Congyi Zhang, Hongbo Fu, Alberto Cannavò, Fabrizio Lamberti, Henry Y.K. Lau, and Wenping Wang. 2021. HandPainter - 3D Sketching in VR with Hand-Based Physical Proxy. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 412, 13 pages. https://doi.org/10.1145/3411764.3445302

[53] Michael Johnston, Philip R. Cohen, David McGee, Sharon L. Oviatt, James A. Pittman, and Ira Smith. 1997. Unification-Based Multimodal Integration. In *ACL '98/EACL '98: Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics and Eighth Conference of the European Chapter of the Association for Computational Linguistics* (Madrid, Spain) *(ACL '98/EACL '98)*. Association for Computational Linguistics, USA, 281–288. https://doi.org/10.3115/976909.979653

[54] Runchang Kang, Anhong Guo, Gierad Laput, Yang Li, and Xiang 'Anthony' Chen. 2019. Minuet: Multimodal Interaction with an Internet of Things. In *SUI '19: Symposium on Spatial User Interaction* (New Orleans, LA, USA) *(SUI '19)*. Association for Computing Machinery, New York, NY, USA, Article 2, 10 pages. https://doi.org/10.1145/3357251.3357581

[55] Daniel F. Keefe, Daniel Acevedo, Jadrian Miles, Fritz Drury, Sharon M. Swartz, and David H. Laidlaw. 2008. Scientific Sketching for Collaborative VR Visualization Design. *IEEE Transactions on Visualization and Computer Graphics* 14, 4 (2008), 835–847. https://doi.org/10.1109/TVCG.2008.31

[56] Daniel F. Keefe, David B. Karelitz, Eileen L. Vote, and David H. Laidlaw. 2005. Artistic collaboration in designing VR visualizations. *IEEE Computer Graphics and Applications* 25, 2 (2005), 18–23. https://doi.org/10.1109/MCG.2005.34

[57] Daniel F. Keefe, Robert Zeleznik, and David Laidlaw. 2007. Drawing on Air: Input Techniques for Controlled 3D Line Illustration. *IEEE Transactions on Visualization and Computer Graphics* 13, 5 (2007), 1067–1081. https://doi.org/10.1109/TVCG.2007.1060

[58] Sumbul Khan and Bige Tunçer. 2019. Gesture and speech elicitation for 3D CAD modeling in conceptual design. *Automation in Construction* 106 (2019), 102847. https://doi.org/10.1016/j.autcon.2019.102847

[59] Yongkwan Kim and Seok-Hyung Bae. 2016. SketchingWithHands: 3D Sketching Handheld Products with First-Person Hand Posture. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology* (Tokyo, Japan) *(UIST '16)*. Association for Computing Machinery, New York, NY, USA, 797–808. https://doi.org/10.1145/2984511.2984567

[60] Eleanor Knott, Aliya H. Rao, Kate Summers, and Chana Teeger. 2022. Interviews in the social sciences. *Nature Reviews Methods Primers* 2, 1 (15 Sep 2022), 73. https://doi.org/10.1038/s43586-022-00150-6

[61] Kin C. Kwan and Hongbo Fu. 2019. Mobi3DSketch: 3D Sketching in Mobile AR. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–11. https://doi.org/10.1145/3290605.3300406

[62] Joseph J. LaViola Jr., Sarah Buchanan, and Corey Pittman. 2014. *Multimodal Input for Perceptual User Interfaces*. John Wiley & Sons, Ltd, , Chapter 9, 285–312. https://doi.org/10.1002/9781118706237.ch9

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

[64] Minkyung Lee and Mark Billinghurst. 2008. A Wizard of Oz Study for an AR Multimodal Interface. In *Proceedings of the 10th International Conference on Multimodal Interfaces* (Chania, Crete, Greece) *(ICMI '08)*. Association for Computing Machinery, New York, NY, USA, 249–256. https://doi.org/10.1145/1452392.1452444

[65] James R. Lewis. 2018. The System Usability Scale: Past, Present, and Future. *International Journal of Humanâ€“Computer Interaction* 34, 7 (2018), 577–590. https://doi.org/10.1080/10447318.2018.1455307

[66] SmoothStep LLC. 2023. Quill. http://quill.art/.

[67] Microsoft. 2023. Transcribe Your Recordings: OneNote for Microsoft 365, Word Web App. https://support.microsoft.com/en-us/office/transcribe-your-recordings-7fc2efec-245e-45f0-b053-2a97531ecf57. https://support.microsoft.com/en-us/office/transcribe-your-recordings-7fc2efec-245e-45f0-b053-2a97531ecf57

[68] Microsoft Corp. 2023. Paint 3D. https://apps.microsoft.com/detail/9NBLGGH5FV99/

[69] Christophe Mignot, Claude Valot, and Noëlle Carbonell. 1993. An Experimental Study of Future "natural" Multimodal Human-Computer Interaction. In *INTERACT '93 and CHI '93 Conference Companion on Human Factors in Computing Systems* (Amsterdam, The Netherlands) *(CHI '93)*. Association for Computing Machinery, New York, NY, USA, 67–68. https://doi.org/10.1145/259964.260075

[70] Nadine Moacdieh and Nadine Sarter. 2015. Display Clutter: A Review of Definitions and Measurement Techniques. *Human Factors* 57, 1 (2015), 61–100. https://doi.org/10.1177/0018720814541145

[71] Meredith R. Morris. 2012. Web on the Wall: Insights from a Multimodal Interaction Elicitation Study. In *Proceedings of the 2012 ACM International Conference on Interactive Tabletops and Surfaces* (Cambridge, Massachusetts, USA) *(ITS '12)*. Association for Computing Machinery, New York, NY, USA, 95–104. https://doi.org/10.1145/2396636.2396651

[72] Meredith R. Morris, Andreea Danielescu, Steven Drucker, Danyel Fisher, Bongshin Lee, M. C. Schraefel, and Jacob O. Wobbrock. 2014. Reducing Legacy Bias in Gesture Elicitation Studies. *Interactions* 21, 3 (may 2014), 40–45. https://doi.org/10.1145/2591689

[73] Beckett Mufson. 2017. https://www.vice.com/en/article/8qvwzk/james-r-eads-tiltbrush-vr-mushroom-forest

[74] Vishwas G. Nanjundaswamy, Amit Kulkarni, Zhuo Chen, Prakhar Jaiswal, Sree S. S., Anoop Verma, and Rahul Rai. 2013. Intuitive 3D Computer-Aided Design (CAD) System With Multimodal Interfaces. In *ASME 2013 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference (International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Vol. Volume 2A: 33rd Computers and Information in Engineering Conference)*. ASME, Portland, Oregon, USA, V02AT02A037. https://doi.org/10.1115/DETC2013-12277

[75] Michael Nebeling, Alexander Huber, David Ott, and Moira C. Norrie. 2014. Web on the Wall Reloaded: Implementation, Replication and Refinement of User-Defined Interaction Sets. In *Proceedings of the Ninth ACM International Conference on Interactive Tabletops and Surfaces* (Dresden, Germany) *(ITS '14)*. Association for Computing Machinery, New York, NY, USA, 15–24. https://doi.org/10.1145/2669485.2669497

[76] Michael Nielsen, Moritz Störring, Thomas B. Moeslund, and Erik Granum. 2004. A Procedure for Developing Intuitive and Ergonomic Gesture Interfaces for HCI. In *Gesture-Based Communication in Human-Computer Interaction*, Antonio Camurri and Gualtiero Volpe (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 409–420. https://doi.org/10.1007/978-3-540-24598-8_38

[77] Gary M. Olson, Susan A. Duffy, and Robert L. Mack. 2018. Thinking-out-loud as a method for studying real-time comprehension processes. In *New methods in reading comprehension research*. Routledge, London, UK, 253–286. https://doi.org/10.4324/9780429505379

[78] Inc. Open AI. 2023. https://platform.openai.com/

[79] Alfred Oti and Nathan Crilly. 2021. Immersive 3D sketching tools: Implications for visual thinking and communication. *Computers & Graphics* 94 (2021), 111–123. https://doi.org/10.1016/j.cag.2020.10.007

[80] Sharon L. Oviatt. 1997. Multimodal Interactive Maps: Designing for Human Performance. *Hum.-Comput. Interact.* 12, 1 (mar 1997), 93–129. https://www.tandfonline.com/doi/abs/10.1080/07370024.1997.9667241

[81] Sharon L. Oviatt, Antonella DeAngeli, and Karen Kuhn. 1997. Integration and Synchronization of Input Modes during Multimodal Human-Computer Interaction. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems* (Atlanta, Georgia, USA) *(CHI '97)*. Association for Computing Machinery, New York, NY, USA, 415–422. https://doi.org/10.1145/258549.258821

[82] Tran Pham, Jo Vermeulen, Anthony Tang, and Lindsay MacDonald Vermeulen. 2018. Scale Impacts Elicited Gestures for Manipulating Holograms: Implications for AR Gesture Design. In *Proceedings of the 2018 Designing Interactive Systems Conference* (Hong Kong, China) *(DIS '18)*. Association for Computing Machinery, New York, NY, USA, 227–240. https://doi.org/10.1145/3196709.3196719

[83] Eric D. Ragan, Doug A. Bowman, Regis Kopper, Cheryl Stinson, Siroberto Scerbo, and Ryan P. McMahan. 2015. Effects of Field of View and Visual Complexity on Virtual Reality Training Effectiveness for a Visual Scanning Task. *IEEE Transactions on Visualization and Computer Graphics* 21, 7 (2015), 794–807. https://doi.org/10.1109/TVCG.2015.2403312

[84] Sandrine Robbe. 1998. An Empirical Study of Speech and Gesture Interaction: Toward the Definition of Ergonomic Design Guidelines. In *CHI 98 Conference Summary on Human Factors in Computing Systems* (Los Angeles, California, USA) *(CHI '98)*. Association for Computing Machinery, New York, NY, USA, 349–350. https://doi.org/10.1145/286498.286815

[85] Hugo Romat, Andreas Fender, Manuel Meier, and Christian Holz. 2021. Flashpen: A High-Fidelity and High-Precision Multi-Surface Pen for Virtual Reality. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*. IEEE, Lisboa, Portugal, 306–315. https://doi.org/10.1109/VR50410.2021.00053

[86] Enrique Rosales, Chrystiano Araújo, Jafet Rodriguez, Nicholas Vining, Dongwook Yoon, and Alla Sheffer. 2021. AdaptiBrush: Adaptive General and Predictable VR Ribbon Brush. *ACM Trans. Graph.* 40, 6, Article 247 (dec 2021), 15 pages. https://doi.org/10.1145/3478513.3480511

[87] Ayshwarya Saktheeswaran, Arjun Srinivasan, and John Stasko. 2020. Touch? Speech? or Touch and Speech? Investigating Multimodal Interaction for Visual Network Exploration and Analysis. *IEEE Transactions on Visualization and*

Computer Graphics 26, 6 (2020), 2168–2179. https://doi.org/10.1109/TVCG.2020.2970512

[88] Savage Interactive Pty Ltd. 2023. Procreate. https://procreate.com/ipad/

[89] Bahar Sener and Owain Pedgley. 2008. Novel Multimodal Interaction for Industrial Design. In *Human Computer Interaction*, Ioannis Pavlidis (Ed.). IntechOpen, Rijeka, Chapter 13, 195–214. https://doi.org/10.5772/6300

[90] Jinsil H. Seo, Michael Bruner, and Nathanael Ayres. 2018. Aura Garden: Collective and Collaborative Aesthetics of Light Sculpting in Virtual Reality. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI EA '18)*. Association for Computing Machinery, New York, NY, USA, 1–6. https://doi.org/10.1145/3170427.3177761

[91] Anirudh Sharma, Sriganesh Madhvanath, Ankit Shekhawat, and Mark Billinghurst. 2011. MozArt: A Multimodal Interface for Conceptual 3D Modeling. In *Proceedings of the 13th International Conference on Multimodal Interfaces* (Alicante, Spain) *(ICMI '11)*. Association for Computing Machinery, New York, NY, USA, 307–310. https://doi.org/10.1145/2070481.2070538

[92] Gravity Sketch. 2023. GravitySketch. https://www.gravitysketch.com/.

[93] Chengyu Su, Chao Yang, Yonghui Chen, Fupan Wang, Fang Wang, Yadong Wu, and Xiaorong Zhang. 2021. Natural multimodal interaction in immersive flow visualization. *Visual Informatics* 5, 4 (2021), 56–66. https://doi.org/10.1016/j.visinf.2021.12.005

[94] Koun-Tem Sun, Hsin-Te Chan, and Kuan-Chien Meng. 2010. Research on the application of virtual reality on arts core curricula. In *5th International Conference on Computer Sciences and Convergence Information Technology*. IEEE, Seoul, South Korea, 234–239. https://doi.org/10.1109/ICCIT.2010.5711063

[95] Unity Technologies. 2023. https://unitynlp.readthedocs.io/en/latest

[96] Daniela Trevisan, Felipe Carvalho, Alberto Raposo, Carla Freitas, and Luciana Nedel. 2010. Supporting the design of multimodal interactions: a case study in a 3D sculpture application. In *Proceedings of the XII Symposium on Virtual and Augmented Reality*. DIMAP-UFRN, Natal, Brazil, 153–162.

[97] Rumeysa Türkmen, Ken Pfeuffer, Mayra Donaji Barrera Machuca, Anil Ufuk Batmaz, and Hans Gellersen. 2022. 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 Orleans, LA, USA) *(CHI EA '22)*. Association for Computing Machinery, New York, NY, USA, Article 276, 7 pages. https://doi.org/10.1145/3491101.3519737

[98] Sander Van Goethem, Jouke Verlinden, Regan Watts, and Stijn Verwulgen. 2021. User Experience Study on Ideating Wearables in VR. In *Proceedings of the Design Society: 23rd International Conference on Engineering Design (ICED21)*, Vol. 1. Cambridge University Press, Gothenburg, Sweden, 3339–3348. https://doi.org/10.1017/pds.2021.595

[99] Santiago Villarreal-Narvaez, Jean Vanderdonckt, Radu-Daniel Vatavu, and Jacob O. Wobbrock. 2020. A Systematic Review of Gesture Elicitation Studies: What Can We Learn from 216 Studies?. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference* (Eindhoven, Netherlands) *(DIS '20)*. Association for Computing Machinery, New York, NY, USA, 855–872. https://doi.org/10.1145/3357236.3395511

[100] Philipp Wacker, Rahul Arora, Mayra Donaji Barrera Machuca, Daniel F. Keefe, and Johann Habakuk Israel. 2022. 3D Sketching Application Scenarios. In *Interactive Sketch-Based Interfaces and Modelling for Design*, Alexandra Bonnici and Kenneth P. Camilleri (Eds.). River Series in Document Engineering, New York, USA, Chapter 9, 241–261. https://doi.org/10.1201/9781003360650

[101] Philipp Wacker, Oliver Nowak, Simon Voelker, and Jan Borchers. 2019. ARPen: Mid-Air Object Manipulation Techniques for a Bimanual AR System with Pen & Smartphone. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3290605.3300849

[102] Philipp Wacker, Adrian Wagner, Simon Voelker, and Jan Borchers. 2018. Physical Guides: An Analysis of 3D Sketching Performance on Physical Objects in Augmented Reality. In *Proceedings of the Symposium on Spatial User Interaction* (Berlin, Germany) *(SUI '18)*. Association for Computing Machinery, New York, NY, USA, 25–35. https://doi.org/10.1145/3267782.3267788

[103] Yu-Han Wang and Marco Ajovalasit. 2020. Involving cultural sensitivity in the design process: a design toolkit for Chinese cultural products. *International Journal of Art & Design Education* 39, 3 (2020), 565–584. https://doi.org/10.1111/jade.12301

[104] Xin Wei, Ruixuan Yu, and Jian Sun. 2020. View-GCN: View-Based Graph Convolutional Network for 3D Shape Analysis. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, Seattle, WA, USA, 1847–1856. https://doi.org/10.1109/CVPR42600.2020.00192

[105] Eva Wiese, Johann Habakuk Israel, Achim Meyer, and Sara Bongartz. 2010. Investigating the Learnability of Immersive Free-Hand Sketching. In *Seventh Sketch-based Interfaces and Modeling Symposium* (Annecy, France) *(SBIM '10)*. Eurographics Association, Goslar, DEU, 135–142. https://dl.acm.org/doi/10.5555/1923363.1923387

[106] Adam S. Williams, Jason Garcia, and Francisco R. Ortega. 2020. Understanding Multimodal User Gesture and Speech Behavior for Object Manipulation in Augmented Reality Using Elicitation. *IEEE Transactions on Visualization and Computer Graphics* 26, 12 (2020), 3479–3489. https://doi.org/10.1109/TVCG.2020.3023566

[107] Adam S. Williams and Francisco R. Ortega. 2020. Understanding Gesture and Speech Multimodal Interactions for Manipulation Tasks in Augmented Reality Using Unconstrained Elicitation. *Proc. ACM Hum.-Comput. Interact.* 4, ISS, Article 202 (nov 2020), 21 pages. https://doi.org/10.1145/3427330

[108] Markus L. Wittorf and Mikkel R. Jakobsen. 2016. Eliciting Mid-Air Gestures for Wall-Display Interaction. In *Proceedings of the 9th Nordic Conference on Human-Computer Interaction* (Gothenburg, Sweden) *(NordiCHI '16)*. Association for Computing Machinery, New York, NY, USA, Article 3, 4 pages. https://doi.org/10.1145/2971485.2971503

[109] Jacob O. Wobbrock, Meredith R. Morris, and Andrew D. Wilson. 2009. User-Defined Gestures for Surface Computing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Boston, MA, USA) *(CHI '09)*. Association for Computing Machinery, New York, NY, USA, 1083–1092. https://doi.org/10.1145/1518701.1518866

[110] Erik Wolf, Sara Klüber, Chris Zimmerer, Jean-Luc Lugrin, and Marc E. Latoschik. 2019. "Paint That Object Yellow": Multimodal Interaction to Enhance Creativity During Design Tasks in VR. In *2019 International Conference on Multimodal Interaction* (Suzhou, China) *(ICMI '19)*. Association for Computing Machinery, New York, NY, USA, 195–204. https://doi.org/10.1145/3340555.3353724

[111] Xiaozhe Yang, Pei-Yu Cheng, Xin Liu, and Sheng-Pao Shih. 2023. The impact of immersive virtual reality on art education: A study of flow state, cognitive load, brain state, and motivation. *Education and Information Technologies* (07 2023), 1–20. https://doi.org/10.1007/s10639-023-12041-8

[112] Xiaozhe Yang, Lin Lin, Pei-Yu Cheng, Xue Yang, Youqun Ren, and Yueh-Min Huang. 2018. Examining creativity through a virtual reality support system. *Educational Technology Research and Development* 66, 5 (2018), 1231–1254. https://doi.org/10.1007/s11423-018-9604-z

[113] Ya-Ting Yue, Xiaolong Zhang, Yongliang Yang, Gang Ren, Yi-King Choi, and Wenping Wang. 2017. WireDraw: 3D Wire Sculpturing Guided with Mixed Reality. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 3693–3704. https://doi.org/10.1145/3025453.3025792

[114] IonuŢ-Alexandru ZaiŢi, Ştefan-Gheorghe Pentiuc, and Radu-Daniel Vatavu. 2015. On free-hand TV control: experimental results on user-elicited gestures with Leap Motion. *Personal and Ubiquitous Computing* 19, 5 (01 Aug 2015), 821–838. https://doi.org/10.1007/s00779-015-0863-y

[115] Jen Zen. 2000. https://digitalartarchive.siggraph.org/artwork/jen-zen-jen-grey-final-spin/

[116] Anna Zhilyaeva. 2018. https://superrare.com/artwork-v2/liberty-23312

[117] Xiaoyan Zhou. 2023. Designing Navigation Tool for Immersive Analytics in AR. In *2023 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. IEEE, Shanghai, China, 975–976. https://doi.org/10.1109/VRW58643.2023.00330

[118] Chris Zimmerer, Erik Wolf, Sara Wolf, Martin Fischbach, Jean-Luc Lugrin, and Marc E. Latoschik. 2020. Finally on Par?! Multimodal and Unimodal Interaction for Open Creative Design Tasks in Virtual Reality. In *Proceedings of the 2020 International Conference on Multimodal Interaction* (Virtual Event, Netherlands) *(ICMI '20)*. Association for Computing Machinery, New York, NY, USA, 222–231. https://doi.org/10.1145/3382507.3418850