The Effect of Interface Types and Immersive Environments on Drawing Accuracy and User Comfort

Qianyuan Zou*
The University of Auckland, Auckland, New Zealand

Huidong Bai†
The University of Auckland, Auckland, New Zealand

Zhuang Chang‡
The University of Auckland, Auckland, New Zealand

Zirui Xiao§
RingNet Pte. Ltd., Singapore

Suizi Tian¶
University College London, London, United Kingdom

Henry Been-Lirn Duh‖
The Hong Kong Polytechnic University, Hong Kong

Allan Fowler**
The University of Auckland, Auckland, New Zealand

Mark Billinghurst††
The University of Auckland, Auckland, New Zealand

Abstract

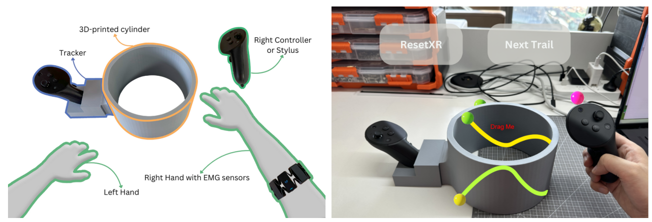
In this research, we investigate the effectiveness of asymmetric interactions (HandStylus, HandController, and TwoHands) in Augmented Reality (AR), Virtual Reality (VR), and Extended Reality (XR) for 3D digital drawing overlaying on physical and virtual objects. We evaluate the input accuracy and fatigue of these object-based 3D drawing experiences using quantitative measurements and further explore the correlation between these outcomes with subjective questionnaires. We found significant independence between environments and interface types, which considerably influence the performance and usability of 3D immersive drawing. We noted discrepancies between users' subjective experiences and objective performance. Specifically, although AR drawing on physical objects provides superior accuracy and minimal muscle fatigue due to tangible feedback, and the TwoHands interaction offers the highest precision, the subjective results show the reverse outcome. Based on these findings, we propose design recommendations and discuss directions for future research in immersive drawing environments.

Index Terms: Mixed reality, virtual reality, immersive drawing, content creation, fatigue, surface electromyography

1 Introduction

Using immersive technologies for design is growing in popularity. This trend has caused a merging of traditional design methods with advanced digital technologies, resulting in a noticeable blending of Virtual Reality (VR) and Augmented Reality (AR) applications within the design workflow. AR/VR tools are characterized by intuitive spatial interaction, immediate visual feedback, and highly synergistic compatibility with design workflows. Tools such as Tilt Brush [11] or Gravity Sketch [12] are particularly useful in the initial stages of artistic workflows, where immersive experiences enable designers to sketch and preview their concepts within a spatial context quickly.

There are two critical aspects of immersive design workflows: Environment and Interface Types. In these workflows, transitioning information across environments, such as importing reference images from the real world to VR, can be complex. Devices like the Apple Vision Pro, which allow seamless switching between VR and AR without removing the headset, simplify this process. On the other hand, while AR provides tangible feedback and real-world visibility, AR devices typically have a smaller field of view (FOV) than VR devices. The latest video see-through Mixed Reality (MR) headsets are anticipated to overcome these limitations by offering a real-world view without sacrificing the FOV.



(a) (b)

Figure 1: The prototype system setup (a) and illustration of user's view with HandController interaction in an XR environment (b).

Interface Types also play critical roles in design workflow. In a conventional design, users often hold a different tool in each hand to accomplish design tasks. For example, during the computer-aided design process using a desktop computer, users employ a mouse or stylus with their right hand while simultaneously using a keyboard with their left hand. This tool-based asymmetric interaction has been demonstrated to be superior for VR sketching [49]. Headsets also support using distinct tools in each hand, enhancing the efficient use of diverse tools from the real world in virtual settings.

Based on these two key features, our research investigates how the Environment or Interface Types can affect the usability and performance of precise object-based immersive drawing. We built a proof-of-concept system Figure 1a to assess the following research question: *How do different immersive environments and interface types affect the performance and usability of immersive 3D object-based content generation tasks?*

Our main innovations and contributions are:

- We discussed the influence of tangible feedback and physical spatial constraint across AR/VR/XR sketching experiences;
- We collected physiological signals for fatigue analysis by comparing them with subjective questionnaires;
- We evaluated the effect of different immersive environments and interface types on performance and usability for object-based 3D immersive drawing.

2 Related work

2.1 Design Tasks in Immersive Environments

Extended Reality (XR) technology, including VR and AR, integrates real and virtual environments to enhance user interaction and immersion[34], it allows users to interact with real-world objects, such as tools on their desks, which encourages the integration of familiar digital design tools like mice, keyboards, and graphics

*e-mail: qzou984@aucklanduni.ac.nz
†e-mail: huidong.bai@auckland.ac.nz
‡e-mail: zcha621@aucklanduni.ac.nz
§e-mail: zirui.xiao@ring.lin
¶e-mail: ucju178@ucl.ac.uk
‖e-mail: henry.duh@polyu.edu.hk
**e-mail: allan.fowler@auckland.ac.nz
††e-mail: mark.billinghurst@auckland.ac.nz

979-8-3315-1647-5/24/$31.00 ©2024 IEEE
DOI 10.1109/ISMAR62088.2024.00099

tablets into XR settings [9]. This diminishes the universality of the traditional two-controller VR input method, often leading users to switch between different input tools for various contexts, such as transitioning between XR and VR environments. To address this, some researchers have proposed bringing traditional devices into VR environments. For example, Bai et al.[3] showcased the ability to stream data from a mobile phone into VR, allowing users to interact with phone applications seamlessly within VR. Similarly, Meta's Horizon Workroom [27] integrates physical desks, keyboards, and computers into VR, enabling users to use a real keyboard and screen that corresponds with their virtual equivalents. Horvat et al.[17] pointed out that in the design workflow, there are instances where content from the real world and virtual world needs to interact. For example, sketching in VR often requires references from the real world, and after the sketch is completed, it needs to be exported to a real-world application for preview. They argue that there needs to be more research addressing the seamless connection between different realities in the design workflow.

Little research has explored transition effects along the reality-virtuality continuum [38], incorporating VR prototyping in the design workflow. Gruenefeld et al.[13] introduced a rapid prototyping system that enables users to swiftly transition between AR and VR realities and accomplish interface prototyping within the design workflow. However, their study primarily used in-depth interviews to evaluate the usability of design workflows within various immersive environments. The sole objective data they gathered was the time users took to complete a prototype in different settings, which is insufficient for a comprehensive analysis. Furthermore, they acknowledged that their research did not incorporate the tools typically used by users in real-world design workflows. Aside from this paper, limited research has supported design prototyping across the entire reality-virtuality continuum.

2.2 Immersive Object-based Content Creation

Immersive object editing encompasses object translation, scale and rotation [37]. Except for select studies dedicated to input hardware research[24, 46], this functionality is predominantly facilitated by the inherent input modes of VR systems, such as using handheld controllers or hand gesture input [17]. Corresponding studies have explored how different input devices or Interface Types can be employed for content generation. For instance, Drey et al.[10] discussed using a drawing tablet in VR environments for sketching. Zou et al. [48] explored the impact of interacting with different real-world tools in each hand for sketching in virtual environments. Xu et al. [42] proposed using gesture input to expedite the generation of preset 3D geometric objects as drawing references. However, due to hardware limitations, most studies predominantly addressed content creation within a singular reality dimension. As XR Head-Mounted Displays (HMDs) mature and the demand for XR design workflows increases, a significant research gap has emerged in Interface Types for content creation and 3D object editing within XR environments. Existing research on integrating drawing tasks into XR is sparse. Our study addresses this by focusing on deconstructing creation and editing into object-based drawing tasks. Participants will perform 3D object editing tasks in various immersive environments (AR, VR, XR) using our prototype Interface Types, allowing us to assess whether XR environments, which combine real-world visibility with virtual-world flexibility, and offer advantages over exclusive AR or VR settings.

2.3 Input Modalities

In VR interactions, symmetric Interface Types with identical input tools like standard VR controllers or hand gesture input are common. These systems, exemplified by the Meta Ques [28] and Hololens [29], offer stability and versatility [49]. Conversely, real-world design tasks often involve using different tools in each hand.Jerlad [21] suggests that replicating these real-world experiences in VR can enhance user immersion and experience. Previous studies [49, 42, 45, 2, 31, 41] investigate various asymmetric interactions, such as using gesture input alongside VR controllers or stylus and combining smartphones with VR controllers for tasks like 3D modelling and object editing. These studies highlight the potential of asymmetric interactions to improve user experience by mirroring real-world usage patterns, although they may not match the efficiency of symmetric interactions due to hardware constraints. However, the majority of these research studies heavily depend on subjective surveys to assess task load and usability. We argue that incorporating objective measurements for usability, such as physiological sensing, is lacking in this research and could provide a more comprehensive understanding of the user experience.

2.4 User Comfort in Immersive Drawing

User comfort refers to how users feel when using a product or system, with respect to the ease of use, convenience, and overall experience. It includes factors such as ergonomics, usability, effectiveness, and how well the technology meets the user's needs and preferences. Ensuring user comfort is essential for a positive user experience and can influence satisfaction, efficiency, and productivity [25]. Immersive drawing involves extensive movement of muscles in the arms, wrists, and fingers. Arm muscle fatigue inevitably occurs during this process, possibly affecting the task load [22, 44, 26]. However, many immersive VR content creation studies predominantly rely on subjective surveys, such as the NASA-TLX [14], to measure task load. This method alone may not provide a robust measurement. There is also some previous research that has utilized vision-based workload evaluation methods to quantify arm fatigue during mid-air interactions [20, 16]. Vision-based systems often require careful calibration and specific setup conditions, which can be time-consuming and complex. On the other hand, physiological sensors such as electromyographic (EMG) devices can objectively measure muscle activity. Combined with traditional subjective questionnaires, they might yield a more comprehensive understanding of the task load and user comfort.

There have been preliminary attempts in previous research. For example, Penumudi et al.[32] examined EMG signals in the neck of users in a virtual environment, and Iqbal et al. [19] assessed arm fatigue in simple VR interaction tasks. However, these studies fall short of bridging the significant research gap concerning enhancing subjective questionnaires with physiological signals to address their lack of objectivity in measuring task load. Penumudi, et al.'s study, involved users in relatively static positions passively receiving information in a virtual setting, whereas the interaction tasks in Iqbal et al.'s research were overly simplistic compared to authentic VR drawing tasks, diverging significantly from typical user scenarios.

In summary, effective immersive design workflows require integrating AR/VR with the physical world to ensure seamless content transitions. Our literature review highlights a promising approach incorporating physical tools or replicas within XR environments, aiming to establish a flexible input system [30, 8]. Prior research advocates for including real-time objective measures, such as EMG, to provide a more thorough evaluation of immersive systems [6]. Our study aims to explore the usability and performance of different interface types in handling intricate 3D object-based drawings within different immersive environments. User evaluations will encompass conventional subjective surveys and quantifiable indicators like physiological signals and input accuracy.

3 System Design and Implementation

We developed the prototype system shown in Figure 1a. The Meta Quest Pro, selected for its dual VR and AR/XR capabilities through video see-through technology, was employed as our operational HMD. The default Meta Quest Pro controller was used as an in-

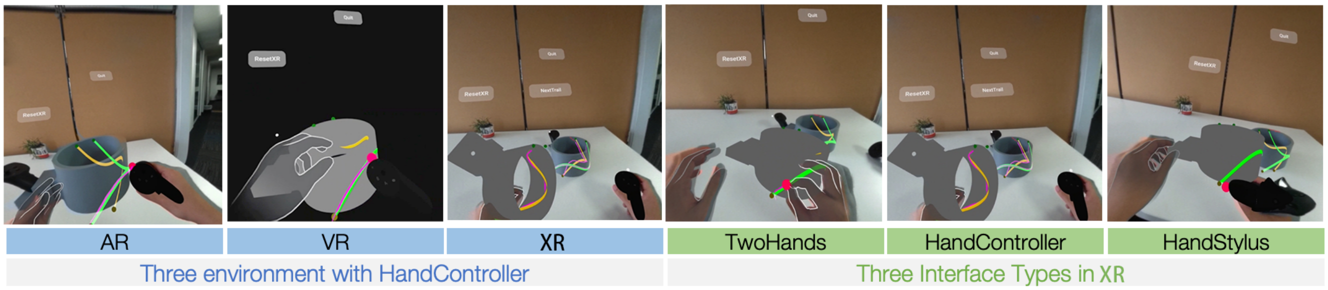


Figure 2: We have three environments and three interface types in this study. On the left (Blue), we show three environments in one interface type; On the right (Green), we show three interface types in one environment.

put device, enabling it to function as a stylus input when held in the opposite orientation. With this setup, we can maintain consistency in the weight of the controller and stylus. We used an eight-channel EMG sensor (Myo Band) [35] to monitor muscle fatigue in the user's dominant forearm, targeting key muscles like the brachioradialis and flexor carpi radialis. By being positioned near the elbow, the sensor effectively captured EMG signals during the experiment [35]. Meta's built-in gesture recognition function was employed for gesture input. A second Quest Pro controller was adapted to track the position of the target object—a 3D-printed cylinder with a 17.7 cm outer diameter, 15.8 cm inner diameter, 0.95 cm wall thickness, and a height of 10 cm. This particular size was chosen for the target due to the following reasons: 1) The barrel shape with open ends minimizes obstructions, improving the Quest Pro's tracking stability. 2) The cylinder is sized to allow hand and controller movements within its walls, mimicking real-life spatial constraints. 3) The barrel shape simulates real-world drawing conditions where designers often interact with curved surfaces, such as porcelain vases. This setup provides realistic feedback that both supports and limits the drawing process as previously mentioned, effectively representing the typical challenges designers face when working with three-dimensional objects.

Based on the review of related research, our hypotheses were: ***H1:*** *Different immersive environments significantly affect drawing accuracy and user comfort;* ***H2:*** *Different interface types significantly affect drawing accuracy and user comfort;* ***H3:*** *There is a significant interaction effect between interface types and the immersive environment;* ***H4:*** *There is a significant correlation between objective measures and subjective assessment of drawing accuracy and user comfort.*

4 User Study

4.1 Experimental Design

The primary goal of this study was to evaluate how different environments and interaction modalities influence user performance and experience, using both subjective questionnaires and objective data. We aimed to assess the effect on drawing accuracy, muscle fatigue, system suitability, mental workload, and simulator sickness. The study examined two factors: Environment and Interface Types, each with three levels. Environments were categorized as AR, VR, and XR, while Interface Types included HandController, HandStylus, and TwoHands. In AR, users can view the real cylinder along with the target and the drawn curves. In VR, users can see a completely virtual cylinder along with both the target and the drawn curves. In XR, users can see the real and virtual cylinders simultaneously, as well as the target and the curves they draw. The screenshots are shown in Figure 2.

4.2 Participants

A total of 36 participants (24 male, 12 female) from the university campus were recruited for the study. Age distribution was as follows: 24 participants between 21 and 30 years and 12 participants between 31 and 40 years. Participants' experience with VR varied, with 36% rating their familiarity as 4/5, 25% as 5/5, 17% as 3/5, 17% as 2/5, and 5% as 1/5, resulting in an average experience score of 3.6/5. In terms of VR sketching skills, 33% rated themselves as 1/5, 28% at 2/5, 25% at 3/5, 11% again at 3/5, and 3% at 5/5, with an average skill level of 2.2 out of 5, where 5 indicates very experienced. All the participants' dominant hand is the right hand. Participants were divided into three groups to experience different environments (AR, VR, XR) and used all three interaction modalities (HandController, HandStylus, TwoHands). This design allowed for a within-subjects approach in terms of interaction modalities and a between-subjects approach in terms of environments. The testing sequence for different interface types was randomized across participants using a Latin square design to mitigate learning effects.

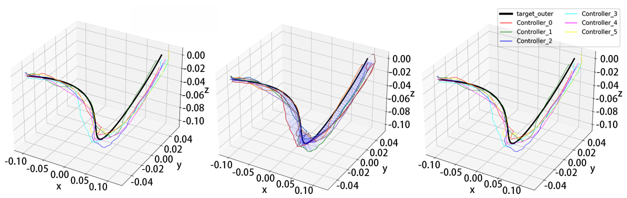
4.3 Tasks and Data Logging

The participants' primary task involved tracing predetermined target curves on a hollow cylinder's inner and outer surfaces. Inspired by Chen [6] and Zou [49], we use Equation equ:curve to generate our target curve, consisting of a sinusoidal curve rotating around the central axis of the hollow cylinder. This rotation was modulated by a linear function along the cylinder's axial direction, ensuring that the curve started at the bottom and ended at the top of the hollow cylinder.

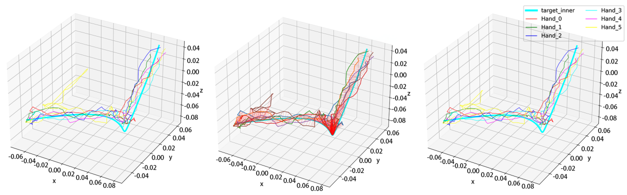
$$\begin{cases} x = r \cdot \cos\theta, \\ z = r \cdot \sin\theta, \\ y = A \cdot \sin(2\pi f t + \varphi) + A\,(2t-1), \\ t = \frac{i}{n-1}, \quad where \quad n > 1, i \in [0, n) \end{cases} \tag{1}$$

Parameters: n is the total number of points on the curve, set to 1000 in our study. i is the index of points on the curve within $[0, 1000)$. The radius r has inner $r = 7.9\,cm$ and outer $r = 8.85\,cm$. θ is the phase in the x, z plane with outer $\theta = \pi(t+1)$ and inner $\theta = \pi(t+1.2)$. The amplitude of the wave A is set to $5\,cm$, which is equal to half of the hollow cylinder's height because the cylinder's coordinate origin is in the geometry centre of the hollow cylinder. The frequency f of the sinusoidal curve is set to 1Hz, and initial phase φ is 0.

To begin drawing in VR or XR, subjects used their left hand to pinch a red "Drag Me" tag, creating an object replica; this feature is absent in the AR mode. They then picked up a Pink sphere, serving as the tip of the drawing tool, and touched a Yellow sphere to start drawing. The task involved drawing lines from the Yellow to a Green sphere, following a highlighted path. Participants were instructed to draw around two cylinder layers, starting with the outer layer. Each stroke had to be completed in one attempt without corrections or repetitions. After finishing two strokes, participants pressed the 'Next Trial' button. The drawing task consisted of six trials, evaluating drawing ability across different modalities. Performance metrics included stroke accuracy relative to the path and the time taken for each stroke. The illustration of the XR environment's user interaction is depicted in Figure 1b.



(a) Outer curves drawn by a user using TwoHands in the AR environment, based on the raw curves (Left), the processing (Middle) and the filtered paths (Right).



(b) Inner curves drawn by a user using TwoHands in the AR environment based on the raw curves (Left), the processing (Middle) and the filtered paths (Right).

Figure 3: The procedures of filtering noise from original data.

The screenshots of three environments with HandController interaction are shown in Figure 2. In the AR environment, participants interacted with a real cylinder, tracing predetermined patterns with a tool in their right hand while their left hand held the cylinder. The traced lines immediately appeared on the cylinder. In the VR environment, participants manipulated a virtual cylinder, tracing patterns in a similar manner. The lines appeared on the virtual cylinder in real-time. After tracing, they transitioned to the AR environment to view these lines on a real cylinder. In the XR environment, participants could see both real and virtual cylinders to trace patterns. The lines they traced appeared simultaneously on both types of cylinders, merging the real and virtual elements seamlessly.

4.4 Measurements and Analysis

Given that our task emphasizes precision over speed, we informed participants before the experiment to focus primarily on the accuracy of their drawings. While we encouraged them to complete the task as fast as possible, we explicitly noted no time limits.

Input Accuracy: Based on the studies by Chen [6] and Zou [49], to calculate the accuracy of the curves generated by the users, we computed the Euclidean distance between each sampled point and the target curve, as well as the mean value of all sampled points. Our research modified the vortex-like curves into spatial curves that adhered to the inner and outer walls of the cylinder to increase the complexity of the curves. Similarly, we also removed sampling points whose Euclidean distance from the target curve exceeded five centimeters to enhance the reliability and stability of the data. Figure 3 shows all the outer and inner wall curves drawn by a user using the TwoHands interface in the AR environment with original and filtered curves. In this study, 1296 trials of input data were recorded and analysed. We sampled 1000 times the spatial coordinates of the pink ball (representing the brush) held in the user's right hand.

EMG Signal: We used the Python programming language and the signal processing and machine learning libraries of Scipy for the EMG data preprocessing and extraction pipeline. The raw data, data processing flowchart and source code are available on GitHub[1]. The EMG signals captured by the Myo armband are nonstationary and multi-component, ranging from -128 to 128 activation units, which reflect the amplified muscle electrical potentials. These signals are sampled at a rate of 200 Hz. Like other EMG devices, the Myo armband's signals may include artifacts from power lines, motion, instrumental and environmental noise, crosstalk, and electrocardiographic interference, often influenced by the sensor's placement on the moving arm [1, 47, 7, 40]. Effective preprocessing is essential to mitigate these artifacts and extract reliable features for fatigue monitoring. Considering the spectral characteristics of the aforementioned artefacts and that Myo's sampling rate is 200 Hz, we applied a fourth-order Butterworth band-pass filter with a range of $[20 \sim 90]$ Hz to minimize the artefacts. A notch filter around 50 Hz was also used to eliminate power line noise. Each channel was divided by its maximum value to normalize the data. The sEMG signals were split into shorter chunks [18] to extract muscle fatigue features in the time and frequency domains.

Regarding fatigue monitoring, extensive research has identified effective sEMG features that can reflect various aspects of muscle activity [39]. Muscular fatigue manifestation is characterized by increased temporal features of sEMG and decreased spectral and temporal-spectral features[43, 36]. Thus, we utilized the root mean square (RMS) feature, a temporal property, to monitor the state of muscular fatigue. We perform our analysis by first extracting RMS features from the entire session and from 0.2-second windows with 10% overlap. RMS data from eight EMG channels is then averaged to form a single representative feature. A 2nd-order polynomial is fitted to these features to assess trends in muscular fatigue. A positive trend in time-domain features and a negative trend in spectral features indicate increasing fatigue. We exclude sessions if the maximum amplitude is less than ± 20 activation units or if no trend changes in signals are observed, enhancing the reliability of our data and filtering out low signal-to-noise ratios.

Subjective Questionnaires: At the end of each session, the participants were asked to complete several subjective questionnaires. We used the Simulator Sickness Questionnaire [5] for measuring motion sickness and the NASA-TLX Questionnaire [15] for measuring mental and physical load. We also measured the system's usability using the System Usability Scale (SUS) [4]. After completing all three Interface Types, participants commented on open questions in a post-experiment questionnaire.

5 Results

A two-way ANOVA analyzed the recorded dataset, using Environment and Interface types as independent variables at a 95% confidence level, employing Tukey's Honestly Significant Difference (HSD) for post-hoc analysis when necessary [33]. This allowed for the examination of both main and interaction effects. The evaluation of Accuracy and Fatigue was segmented into inner and outer wall sections due to the differing spatial positions of the target curves and muscle groups involved in the drawing process. Details are as follows.

5.1 Objective Data

Accuracy Outer: The results of the outer wall Input Accuracy are shown in Figure 4a. The main effect of the environment on accuracy was found to be statistically significant, $F(2,633) = 17.23$, $p < 0.0001$. Similarly, the main effect of the Interface Type on accuracy was also statistically significant, $F(2,633) = 9.19$, $p < 0.0001$. The interaction effect between environment and Interface Type was not statistically significant, $F(4,633) = 0.058$, $p = 0.994$.

[1] https://github.com/FrostyAlien/VIP_MRTK3

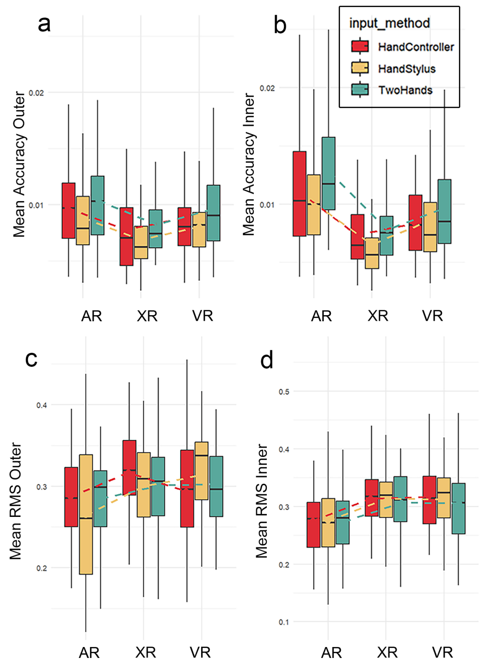
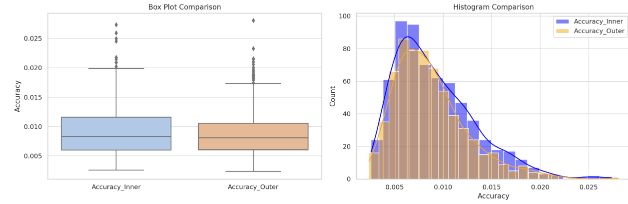


Figure 4: Two–way interaction effects on objective measurements in accuracy and fatigue.

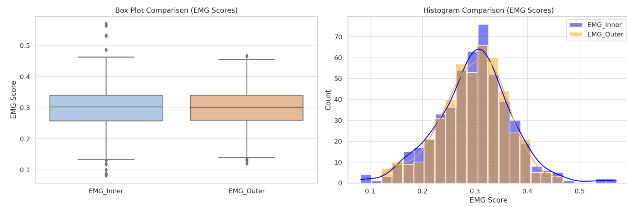
This implies that the effect of one factor on accuracy does not depend significantly on the other factor. The post-hoc results suggested Accuracy was significantly higher in AR compared to both XR ($p < 0.0001$) and VR ($p = 0.0493$), with XR showing the lowest accuracy. VR also outperformed XR ($p = 0.0018$). Regarding the Interface Types, there was a significant decrease in accuracy between the HandStylus and HandController ($p = 0.0046$); moreover, using TwoHands was significantly more accurate than using the HandStylus ($p = 0.0001$). The accuracy of using TwoHands and HandController was not statistically different.

Accuracy Inner: The results of the outer wall Input Accuracy are shown in Figure 4b. The main effect of the environment on accuracy was found to be statistically significant, $F(2,629) = 64.927$, $p < 0.0001$. Similarly, the main effect of the Interface Type on accuracy was also statistically significant, $F(2,629) = 11.035$, $p < 0.0001$. The interaction effect between environment and Interface Type was not statistically significant, $F(4,629) = 1.39$, $p = 0.235$. This implies that the effect of one factor on accuracy does not depend significantly on the other factor. In terms of environment, post-hoc results show significant differences between VR and AR (*mean diff* = -0.0024, $p < 0.001$), and VR and XR (*mean diff* = 0.00167, $p < 0.001$). The difference between XR and AR was also significant (*mean diff* = -0.00408, $p < 0.001$). In terms of Interface Types, significant differences were noted between TwoHands and HandStylus (*mean diff* = 0.0016, $p < 0.001$), and TwoHands and HandController (*mean diff* = 0.0011, $p = 0.004$), with TwoHands consistently showing higher accuracy scores. No significant difference was detected between HandStylus and HandController (*mean diff* = -0.0005, $p = 0.342$).

RMS Slope Outer: The results are shown in Figure 4c. There is a significant effect of the environment on RMS values $F(2,457) = 8.740$, $p = 0.000188$, indicating a statistically significant difference between the environments. However, the Interface Type did not significantly affect the RMS values $F(2,457) = 0.384$, $p = 0.682$. Moreover, the interaction between the environment and Interface Type was insignificant $F(4,457) = 1.961$, $p = 0.099$. The posthoc



(a) Comparison Between Accuracy Inner and Outer.



(b) Comparison Between RMS Value Inner and Outer.

Figure 5: Objective data of Inner and Outer layer comparison.

analysis reveals statistically significant mean differences between XR and AR (*mean diff* = 0.0264, $p = 0.00038$) and VR and AR (*mean diff* = 0.0244, $p = 0.00286$), but no significant difference between VR and XR (*mean diff* = -0.00195, $p = 0.959$). The results suggest that XR and VR show statistically significant higher muscular fatigue than AR, but there's no significant difference between XR and VR.

RMS Slope Inner: The results are shown in Figure 4d. The results indicated a significant effect of the environment on EMG scores (F(2, 487) = 15.156, $p < .0001$), but no significant effects were found for the Interface Type (F(2, 487) = 0.509, $p = 0.601$) or the interaction between environment and method (F(2, 487) = 0.062, $p = 0.993$). The post-hoc results indicated statistically significant differences in RMS value between XR and AR (*mean diff* = 0.0363, $p < .0001$) and VR and AR (*mean diff* = 0.0366, $p < .0001$), but no significant difference between VR and XR (*mean diff* = 0.0003, $p < .0001$). The results suggest that the AR environment causes less muscular fatigue than XR and VR in drawing tasks on the Inner layer.

Inner vs Outer: We specifically compared the performance between the object's inner and outer across all conditions to study the influence caused by the spatial constraint during the interaction. The results of the paired t-test revealed a statistically significant difference between Accuracy Inner and Accuracy Outer (*mean diff* = 0.000423, t-statistic = 2.2155, $p = 0.0271$), indicating higher accuracy for the Inner Layer compared to the Outer layer (See Figure 5a). In contrast, the paired t-test comparing RMS Inner and RMS Outer showed no statistically significant difference (*mean diff* = 0.0019, t-statistic = 0.4399, $p = 0.6602$), suggesting comparable levels between the Inner and Outer RMS value (See Figure 5b).

5.2 Subjective Data

System Usability Scale: The results of the SUS are shown in Figure 6a. The results showed a significant effect of interface type on System Usability Scores (F(2, 99) = 3.619, $p = 0.0304$). However, the environment did not significantly affect the System Usability Scores (F(2, 99) = 0.802, $p = 0.4511$). Moreover, the interaction between environment and interface type was not significant (F(4, 179) = 0.112, $p = 0.9779$). The results of Tukey's posthoc analysis showed that the HandStylus and HandController differed significantly (*mean diff* = -12.64, p = 0.0227), while the comparisons between TwoHands and HandController (*mean diff* = -6.74, p = 0.3281), and TwoHands and HandStylus (*mean diff* = 5.90, p = 0.4237), did not show significant differences. The results

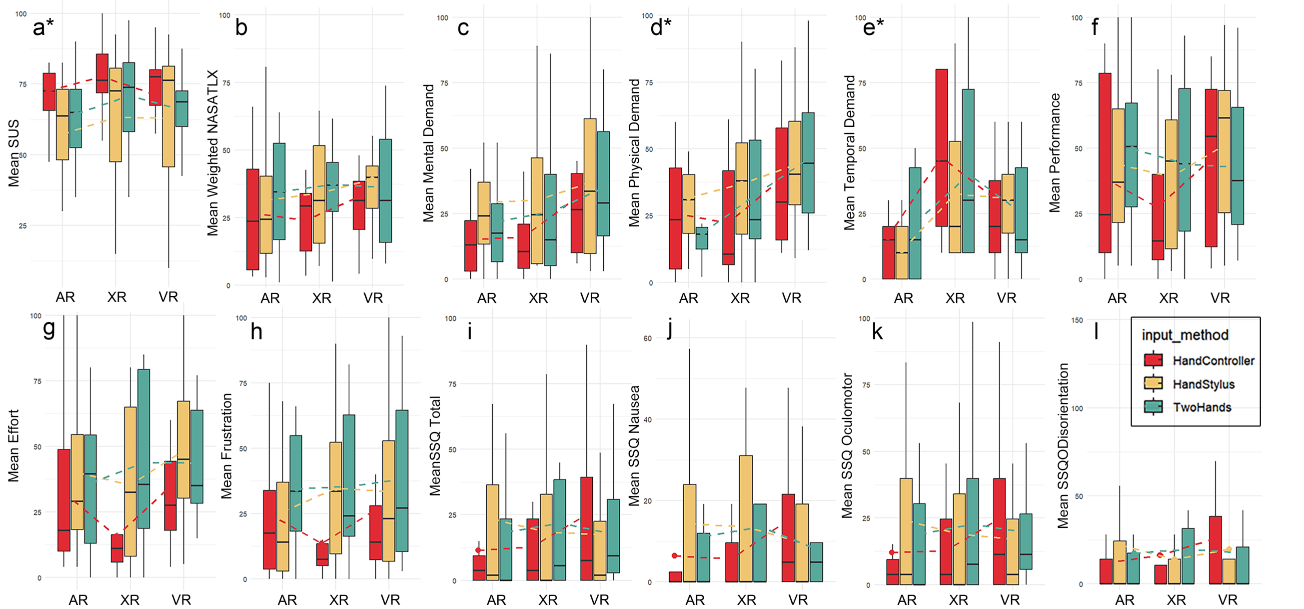


Figure 6: Analysis of two–way interaction effects on measured subjective results from different questionnaires (Asterisks have been added to the plots that indicate the posthoc analysis conducted).

suggest that the type of Interface Type significantly affects the usability scores in immersive environments, particularly HandStylus was significantly more preferred by participants than the HandController, while there were no significant usability differences between TwoHands and either the HandController or the HandStylus.

NASA Task Load Index: We separately computed the weighted NASA-TLX and its subset to have a better insights of perceived task load.

Weighted NASA-TLX - The two-way ANOVA results indicated no significant factors influencing the total Weighted NASA-TLX score. Details are as follows: Environment (F(2, 99) = 0.536, p = 0.587), Interface Type (F(2, 99) = 1.473, p = 0.234), The interaction between the Environment and the Interface Type (F(4, 99) = 0.185, p = 0.946) (See Figure 6b).

Mental Demand - The Two-Way ANOVA results indicated no significant factors influencing the total Mental Demand. Details are as follows: Environment (F(2, 99) = 2.550, p = 0.083), Interface Type (F(2, 99) = 1.969, p = 0.145), The interaction between the Environment and the Interface Type (F(4, 99) = 0.125, p = 0.973) (See Figure 6c).

Physical Demand - The Two-Way ANOVA results indicated the Environment significantly influenced the total Physical Demand. Details are as follows: Environment (F(2, 99) = 5.43, p = 0.006), Interface Type (F(2, 99) = 1.23, p = 0.298), The interaction between the Environment and the Interface Type (F(4, 99) = 0.7, p = 0.593). The post-hoc results indicated VR exhibits significantly lower Physical Demand scores compared to AR (*mean diff* = 17.667, $p < 0.05$). However, the comparison between VR vs XR (*mean diff* = 12.361, $p = 0.068$), and XR vs AR (*mean diff* = 5.305, $p = 0.6$) did not show a significant difference (See Figure 6d).

Temporal Demand - The Two-Way ANOVA results indicated the Environment significantly influenced the total Temporal Demand. Details are as follows: Environment (F(2, 99) = 7.962, p = 0.000622), Interface Type (F(2, 99) = 0.6, p = 0.551), The interaction between the Environment and the Interface Type (F(2, 99) = 0.564, p = 0.689). The post-hoc results indicated XR exhibits significantly higher Temporal Demand scores compared to AR (*mean diff* = 23.611, $p < 0.001$). However, the comparison between VR vs XR (*mean diff* = -10.833, $p = 0.166$), and VR vs AR (*mean diff* = 12.77778, $p = 0.084$) did not show a significant difference (See Figure 6e).

Performance - The Two-Way ANOVA results indicated no significant factors influencing the total Performance. Details are as follows: Environment (F(2, 99) = 0.951, p = 0.39), Interface Type (F(2, 99) = 0.773, p = 0.464), The interaction between the Environment and the Interface Type (F(4, 99) = 0.379, p = 0.823) (See Figure 6f).

Effort - The Two-Way ANOVA results indicated no significant factors influencing the total Effort. Details are as follows: Environment (F(2, 99) = 1.283, p = 0.282), Interface Type (F(2, 99) = 2.804, p = 0.065), The interaction between the Environment and the Interface Type (F(4, 99) = 0.666, p = 0.617) (See Figure 6g).

Frustration - The Two-Way ANOVA results indicated no significant factors influencing the total Effort. Details are as follows: Environment (F(2, 99) = 0.29, p = 0.749), Interface Type (F(2, 99) = 2.325, p = 0.103), The interaction between the Environment and the Interface Type (F(4, 99) = 0.407, p = 0.803) (See Figure 6h).

Simulator Sickness Quesionnaire: The Two-Way ANOVA results indicated no significant factors influencing the total Simulator Sickness score. Details are as follows: Environment (F(2, 99) = 0.075, p = 0.928), Interface Type (F(2, 99) = 0.143, p = 0.867), The interaction between the Environment and the Interface Type (F(2, 99) = 0.488, p = 0.745)(See Figure 6i). We listed the sub-scale score details as below.

Nausea - No significant factors influence the Nausea score. Details are as follows: Environment (F(2, 99) = 0.003, p = 0.997), Interface Type (F(2, 99) = 0.352, p = 0.704), The interaction between the Environment and the Interface Type (F(2, 99) = 0.834, p = 0.506)(See Figure 6j).

Oculomotor Disturbance - No significant factors influencing the Oculomotor Disturbance score. Details are as follows: Environment (F(2, 99) = 0.079, p = 0.924), Interface Type (F(2, 99) = 0.233, p = 0.792), The interaction between the Environment and the Interface Type (F(2, 99) = 0.505, p = 0.732)(See Figure 6k).

Disorientation - No significant factors influencing the Disorientation score. Details are as follows: Environment (F(2, 99) = 0.162, p = 0.850), Interface Type (F(2, 99) = 0.004, p = 0.996), The interaction between the Environment and the Interface Type (F(2, 99) = 0.259, p = 0.903) (See Figure 6l).

User Preference: The post-study questionnaire results regarding user preferences are shown in Figure 7. The results indicated that HandController was the most preferred in AR at 66.67%, in VR at 41.67%, and in XR at 41.67%. This consistency highlights the HandController's dominant role as the interaction of choice, amassing an overall preference score of 150.00%. Conversely, the HandStylus, while less favoured in AR at 25.00%, showed a significant increase in preference in VR, matching the HandController at 41.67%, and observed a moderate preference in XR at 33.33%. This interaction type garnered an overall preference of 100.00%. The TwoHands interaction revealed a more varied pattern, being least favoured across technologies, with scores often falling below those of the HandController and HandStylus, tallying an overall preference of 50.00%.

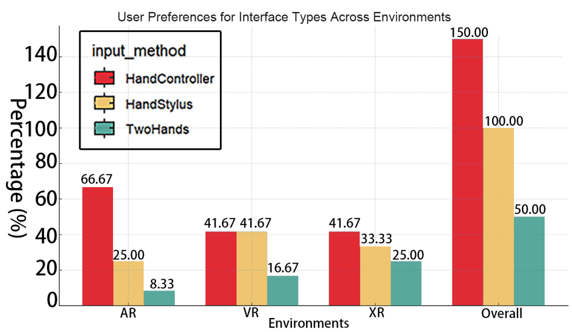


Figure 7: User preferences for three interface types across three environments.

5.3 Result Highlights

The highlights of the results are listed below: **Accuracy Outer**: AR outperformed VR and XR. TwoHands interface outperformed HandStylus; **Accuracy Inner**: AR was more accurate than both VR and XR. VR outperformed XR. TwoHands interface excelled over both HandStylus and HandController. **RMS Outer**: XR exhibited higher RMS than AR; **RMS Inner**: XR exhibited higher RMS than AR. VR showed higher RMS compared to XR. **Physical Demand**: AR requires higher physical demand than VR; **Temporal Demand**: XR was perceived more temporal demand than AR. **SUS**: The HandController outperformed HandStylus. **User preference**: Across all environments, the HandController is the most preferred, followed by the HandStylus, while the TwoHands is the least preferred.

6 Discussion

6.1 Key Findings

Based on the results of the user study, **H1** was accepted, **H2** was partially accepted, **H3** was rejected, and **H4** was partially accepted.

H1 was accepted since we observed significant effects of immersive environments on drawing accuracy and user comfort from both subjective and objective perspectives. Specifically, in ***AR environment***: Users demonstrated significantly higher accuracy in drawing tasks in AR compared to those in VR and XR, particularly when real objects effectively provided tangible feedback. Nonetheless, the physical dimensions of these real objects could limit their effectiveness for drawing tasks. For example, the cylinder used in the task had an inner diameter just large enough to fit the user's hand and controller, which restricted the range of hand movements during drawing. As a result, this limitation might lead users to perceive drawing in AR as more physically demanding than in VR. This perception contrasts with the objective data on physical demand, measured by RMS values, where the findings suggested the opposite. Furthermore, the perceived temporal demand in AR was lower than in XR despite informing participants pre-experiment that the task prioritized accuracy over speed and that time would not be measured. This difference between objective data and subjective user perceptions merits further exploration, as it suggests that traditional questionnaires, which rely on users' subjective feelings and memory, may not always be reliable. In ***VR environment***: Objective results indicate higher accuracy in drawing curves on the inner wall in a VR environment compared to an XR environment. However, VR also leads to greater muscular fatigue. We hypothesize that the lack of real-world visual references in VR compels users to concentrate more on the virtual lines, potentially improving accuracy at the expense of increased fatigue. Notably, these objective findings were not mirrored by the subjective questionnaires. In ***XR environment***: Objective results demonstrate that drawing in an XR environment leads to greater fatigue and lower accuracy than in an AR environment. We found that subjectively, users perceive the temporal demand in XR to be significantly higher than in AR. As mentioned previously, although users were informed that they need not be concerned about the time taken to complete tasks, their perception of task duration may still be influenced by the fatigue experienced during drawing. Considering that the performance of XR and VR and the usability perceived by users are relatively similar, one reason for this may be the absence of tangible feedback in our experimental setup, resulting in higher fatigue and lower accuracy than AR. Given the strong correlation between fatigue levels and accuracy in this experiment, the impact of fatigue on drawing accuracy merits further investigation.

H2 was partially accepted since we observed significant effects of interface types only on drawing accuracy from both subjective and objective perspectives. Objective results indicate that the various interface types tested do not show significant differences in fatigue experienced during use. However, the TwoHands interface consistently demonstrated higher accuracy than the HandStylus and HandController across all environments. Despite the higher objective accuracy of TwoHands, subjective questionnaire results demonstrated that TwoHands was the least preferred interface type among users in all environments. We believe the higher accuracy in the Two Hands condition, compared to the HandController and HandStylus conditions, is largely due to the bulkiness of the controller or stylus. This bulk creates a gap between the stylus tip and the target curve, an issue absent when drawing with bare hands. This difference is especially notable when drawing the inner walls of objects, where the constrained diameter leads to significantly better accuracy with bare hands than with interfaces involving a controller. However, drawing with bare hands lacks the tangible feedback provided by holding a controller, subjectively making this method less preferred by users. Based on verbal feedback from users, it was found that the majority preferred to hold a tool while drawing, as this aligned with their habitual practices in the real world. For instance, AR group P09 said: *I think the main reason the Pen is best is because not only are we most used to it, but the grip also feels the most steady*. While VR group P11 said: *The tip of the stylus and pen is easy to use when drawing the strokes*. These insights underscore the importance of familiarity and physical comfort in the user experience of interface tools.

H3 was rejected since the study results revealed no significant interaction effect between Environment and Interface Type, suggesting that each factor influences usability and performance directly, independent of the influence of the other factor. This independence indicates that, in this case, the effects of Environment and Interface Type on usability or performance are not contingent upon each other. These conclusions are supported by users' subjective experiences and corroborated by objective data.

H4 was partially accepted.We observed discrepancies between some subjective questionnaire results and the objective data from sensors, which we discuss in the following section. We attribute these differences mainly to the fact that subjective questionnaires rely on participants' memories and perceptions. As posited by Oberauer and Kliegl [30], longer tasks increase the likelihood of distractions that can disrupt memory encoding and retrieval. To counter the limitations of traditional questionnaires, we recommend complementing them with real-time data collection using physiological sensors and embedded software. This multi-modal approach could reduce memory interference and retrospective bias, thus improving the accuracy and ecological validity of user experience evaluations.

6.2 Observation

During the study and post-study interviews, we observed that many participants verbally expressed difficulties while drawing the inner wall curves of the cylinder in the AR condition. For instance, P03 from the AR group mentioned in the post-study interview: *It's very difficult to draw in the inner layer.* Similarly, P08 stated: *The calibration of the target strokes is not perfectly matching with the physical world; I think it influenced the haptic feedback, making me feel difficult to draw*. We believe the challenges users exhibit are related to the physical constraints posed by the inner wall. While this constraint objectively enhances drawing accuracy and reduces fatigue, the unnatural posture it requires may lead participants to perceive the AR condition as more demanding.

In addition, We suggest that this tangible feedback also influenced participants' choice of interface type to some extent. Although the objective data showed that TwoHands exhibited superior accuracy when drawing the inner wall, users subjectively rated TwoHands as the least favourable interface type across all environments. We infer that despite the cylinder's inner wall size limiting the performance of the controller and stylus, participants still preferred more stable control devices. This aligns with conclusions drawn from previous studies [49].

6.3 Design Recommendations

For designers: Incorporating tangible user interface elements into AR/VR/XR content creation interface is highly recommended to enhance user experience. Tangible feedback is essential in object-based content creation tasks as it significantly improves accuracy, aligns subjective perceptions with objective measures, and reduces fatigue by facilitating more natural interactions. Understanding the effects of tangible feedback on user performance and perception across different immersive environments and interface types is crucial. Thus, integrating tangible interfaces can optimize user engagement and effectiveness in augmented, virtual, and mixed-reality applications.

For researchers: It is worth exploring a broader array of assessment tools that can capture the complete, real-time user experience. Given the observed discrepancies between subjective questionnaires and objective data, it is advisable to develop integrated assessment methods. These methods should effectively bridge user perceptions and real-time performance data. Incorporating techniques such as real-time physiological monitoring or eye-tracking could offer deeper insights into user experience and performance, thereby enhancing the comprehensiveness of evaluations in user studies.

6.4 Limitation and Direction for Future work

This study identifies several limitations. First, while a correlation between user fatigue and drawing accuracy was observed, the nature of this relationship remains unclear due to the experiment design. Our study design does not explicitly aim to investigate the correlation between hand fatigue and drawing accuracy, which typically requires manual manipulation of hand fatigue (e.g., having participants wear hand weights and perform the same drawing task). Instead, our study focuses on exploring the performance of various interface types across different immersive environments. Hence, our experiment does not include controller weight as a controlled variable. Future research should further explore this relationship to clarify the direct effects between fatigue and accuracy. The study design may also introduce variability due to individual differences. We used a two-way ANOVA design. This hybrid design balanced clean experimental isolation to avoid the learning effect with practical usability measurements [23]. Although it increases data analysis complexity and potential variability, it was necessary to capture the nuanced interplay of the studied factors. Future research could employ more common approaches, such as randomizing experimental conditions, rather than omitting comparisons.

Second, the study used the Meta Quest Pro controllers as the handheld controllers and styluses for the HandController and HandStylus interfaces, ensuring that both tools were of equal weight. However, feedback from users indicated that holding the Quest Pro controller in reverse, as required for the stylus input, felt unnatural. This awkward grip likely influenced users' subjective experiences, as reflected in the SUS results, though it was not apparent in the objective data. Therefore, the impact of grip posture on performance and usability when using the same controllers of the same shape warrants further investigation.

Third, in our current study, while we deconstructed several fundamental operations of the drawing task and combined them into a highly representative task, it remains somewhat simplistic compared to real-world drawing scenarios. Future research could consider engaging users in more complex and long-term drawing tasks, coupled with real-time objective data sampling and analysis. This approach may provide a more comprehensive understanding of immersive drawing performance and its influencing factors.

Fourth, the use of a black background in the VR environment aimed to minimize visual distractions and focus users on the drawing task, simulating typical VR applications like TiltBrush. However, this may have contributed to higher accuracy in AR environments. The inconsistency in trajectory visualizations—partially occluded in VR and MR but unobstructed in AR—also requires further investigation to understand its impact on accuracy.

7 Conclusion

Our study combined objective measurements with subjective assessments to thoroughly investigate the effects of various immersive environments and interface types on performance and usability within immersive 3D object-based drawing tasks. Using a two-way factorial design with 36 participants, our findings reveal that environment and interface type independently affect performance and usability, with no significant interaction effects. Our research also indicated a disparity between objective data and subjective assessments, pointing to potential limitations in traditional subjective questionnaires. Specifically, the AR environment demonstrated consistently higher accuracy and reduced muscle fatigue, yet users reported it as having a higher physical demand. Similarly, the TwoHands interface showed objectively superior performance but was subjectively rated as the least preferred by participants.

In future studies, we plan to build upon our current research by incorporating additional physiological sensors, such as Galvanic Skin Response (GSR) and vision-based sensors, to measure cognitive load. We also intend to increase the complexity and duration of the drawing tasks. Through these enhancements, we aim to explore the relationships among cognitive load, fatigue levels, and operational precision in 3D tasks more objectively.

References

[1] M. Abduo and M. Galster. Myo Gesture Control Armband for Medical Applications. 2015. Publisher: University of Canterbury. 4

[2] R. Arisandi, M. Otsuki, A. Kimura, F. Shibata, and H. Tamura. Virtual handcrafting: Building virtual wood models using tooldevice. *Proceedings of the IEEE*, 102(2):185–195, 2014. 2

[3] H. Bai, L. Zhang, J. Yang, and M. Billinghurst. Bringing full-featured mobile phone interaction into virtual reality. *Computers & Graphics*, 97:42–53, 2021. 2

[4] A. Bangor, P. T. Kortum, and J. T. Miller. An empirical evaluation of the system usability scale. *Intl. Journal of Human–Computer Interaction*, 24(6):574–594, 2008. 4

[5] P. Bimberg, T. Weissker, and A. Kulik. On the usage of the simulator sickness questionnaire for virtual reality research. In *2020 IEEE conference on virtual reality and 3D user interfaces abstracts and workshops (VRW)*, pp. 464–467. IEEE, 2020. 4

[6] C. Chen, M. Yarmand, Z. Xu, V. Singh, Y. Zhang, and N. Weibel. Investigating input modality and task geometry on precision-first 3d drawing in virtual reality. In *2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 384–393. IEEE, 2022. 2, 3, 4

[7] R. H. Chowdhury, M. B. Reaz, M. A. B. M. Ali, A. A. Bakar, K. Chellappan, and T. G. Chang. Surface electromyography signal processing and classification techniques. *Sensors*, 13(9):12431–12466, 2013. 4

[8] G. Christou, E. Lai-Chong Law, W. Green, and K. Hornbaek. Reality-based interaction evaluation methods and challenges. *International Journal of Human-Computer Studies*, 69(1):1–2, 2011. doi: 10.1016/j.ijhcs.2010.11.001 2

[9] Y. Dan, Z. Shen, Y. Zhu, and L. Huang. Using mixed reality (mr) to improve on-site design experience in community planning. *Applied Sciences*, 11(7), 2021. doi: 10.3390/app11073071 2

[10] T. Drey, J. Gugenheimer, J. Karlbauer, M. Milo, and E. Rukzio. 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*, pp. 1–14, 2020. 2

[11] Google. Tilt brush. https://www.tiltbrush.com/. Accessed: 2024-07-24. 1

[12] Gravity Sketch. Gravity sketch. https://www.gravitysketch.com/. Accessed: 2024-07-24. 1

[13] U. Gruenefeld, J. Auda, F. Mathis, S. Schneegass, M. Khamis, J. Gugenheimer, and S. Mayer. Vrception: Rapid prototyping of cross-reality systems in virtual reality. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, pp. 1–15, 2022. 2

[14] S. G. Hart. Nasa-task load index (nasa-tlx); 20 years later. In *Proceedings of the human factors and ergonomics society annual meeting*, vol. 50, pp. 904–908. Sage publications Sage CA: Los Angeles, CA, 2006. 2

[15] S. G. Hart and L. E. Staveland. Development of nasa-tlx (task load index): Results of empirical and theoretical research. In *Advances in psychology*, vol. 52, pp. 139–183. Elsevier, 1988. 4

[16] J. D. Hincapié-Ramos, X. Guo, P. Moghadasian, and P. Irani. Consumed endurance: a metric to quantify arm fatigue of mid-air interactions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, p. 1063–1072. Association for Computing Machinery, New York, NY, USA, 2014. doi: 10.1145/2556288.2557130 2

[17] N. Horvat, S. Kunnen, M. Štorga, A. Nagarajah, and S. Škec. Immersive virtual reality applications for design reviews: Systematic literature review and classification scheme for functionalities. *Advanced Engineering Informatics*, 54:101760, 2022. 2

[18] S. Inam, S. Al Harmain, S. Shafique, M. Afzal, A. Rabail, F. Amin, and M. Waqar. A brief review of strategies used for emg signal classification. In *2021 International Conference on Artificial Intelligence (ICAI)*, pp. 140–145. IEEE, 2021. 4

[19] H. Iqbal, S. Latif, Y. Yan, C. Yu, and Y. Shi. Reducing arm fatigue in virtual reality by introducing 3d-spatial offset. *IEEE Access*, 9:64085–64104, 2021. 2

[20] S. Jang, W. Stuerzlinger, S. Ambike, and K. Ramani. Modeling cumulative arm fatigue in mid-air interaction based on perceived exertion and kinetics of arm motion. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, p. 3328–3339. Association for Computing Machinery, New York, NY, USA, 2017. doi: 10.1145/3025453.3025523 2

[21] J. Jerald. *The VR book: Human-centered design for virtual reality*. Morgan & Claypool, 2015. 2

[22] J. J. LaViola Jr, E. Kruijff, R. P. McMahan, D. Bowman, and I. P. Poupyrev. *3D user interfaces: theory and practice*. Addison-Wesley Professional, 2017. 2

[23] J. Lazar, J. H. Feng, and H. Hochheiser. *Research methods in human-computer interaction*. Morgan Kaufmann, 2017. 8

[24] N. Li, T. Han, F. Tian, J. Huang, M. Sun, P. Irani, and J. Alexander. Get a grip: Evaluating grip gestures for vr input using a lightweight pen. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1–13, 2020. 2

[25] C.-H. Lu, C.-L. Wu, M.-Y. Weng, W.-C. Chen, and L.-C. Fu. Context-aware energy saving system with multiple comfort-constrained optimization in m2m-based home environment. *IEEE Transactions on Automation Science and Engineering*, 14(3):1400–1414, 2017. doi: 10.1109/TASE.2015.2440303 2

[26] N. Maitlo, Y. Wang, C. P. Chen, L. Mi, and W. Zhang. Hand-gesture-recognition based text input method for ar/vr wearable devices. *arXiv e-prints*, pp. arXiv–1907, 2019. 2

[27] Meta. Horizon workrooms. https://forwork.meta.com/nz/horizon-workrooms/. Accessed: 2024-07-24. 2

[28] Meta. Meta quest. https://www.meta.com/nz/quest/. Accessed: 2024-07-24. 2

[29] Microsoft. Microsoft hololens. https://www.microsoft.com/en-us/hololens. Accessed: 2024-07-24. 2

[30] K. Oberauer and R. Kliegl. A formal model of capacity limits in working memory. *Journal of memory and language*, 55(4):601–626, 2006. 2, 8

[31] M. Otsuki, K. Sugihara, A. Toda, F. Shibata, and A. Kimura. A brush device with visual and haptic feedback for virtual painting of 3d virtual objects. *Virtual Reality*, 22:167–181, 2018. 2

[32] S. A. Penumudi, V. A. Kuppam, J. H. Kim, and J. Hwang. The effects of target location on musculoskeletal load, task performance, and subjective discomfort during virtual reality interactions. *Applied ergonomics*, 84:103010, 2020. 2

[33] D. G. Pereira, A. Afonso, and F. M. Medeiros. Overview of friedman's test and post-hoc analysis. *Communications in Statistics-Simulation and Computation*, 44(10):2636–2653, 2015. 4

[34] J. Ratcliffe, F. Soave, N. Bryan-Kinns, L. Tokarchuk, and I. Farkhatdinov. Extended reality (xr) remote research: a survey of drawbacks and opportunities. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI '21. Association for Computing Machinery, New York, NY, USA, 2021. doi: 10.1145/3411764.3445170 1

[35] S. Rawat, S. Vats, and P. Kumar. Evaluating and exploring the myo armband. In *2016 International Conference System Modeling & Advancement in Research Trends (SMART)*, pp. 115–120. IEEE, 2016. 3

[36] D. Roman-Liu, T. Tokarski, and K. Wójcik. Quantitative assessment of upper limb muscle fatigue depending on the conditions of repetitive task load. *Journal of Electromyography and Kinesiology*, 14(6):671–682, 2004. 4

[37] M. F. Shiratuddin and W. Thabet. Utilizing a 3d game engine to develop a virtual design review system. 2011. 2

[38] R. Skarbez, M. Smith, and M. C. Whitton. Revisiting milgram and kishino's reality-virtuality continuum. *Frontiers in Virtual Reality*, 2:647997, 2021. 2

[39] J. Sun, G. Liu, Y. Sun, K. Lin, Z. Zhou, and J. Cai. Application of surface electromyography in exercise fatigue: A review. *Frontiers in Systems Neuroscience*, 16, 2022. 4

[40] D. C. Toledo-Pérez, J. Rodríguez-Reséndiz, R. A. Gómez-Loenzo, and J. Jauregui-Correa. Support vector machine-based emg signal classification techniques: A review. *Applied Sciences*, 9(20):4402, 2019. 4

[41] A. Withana, M. Kondo, Y. Makino, G. Kakehi, M. Sugimoto, and

M. Inami. Impact: Immersive haptic stylus to enable direct touch and manipulation for surface computing. *Computers in Entertainment (CIE)*, 8(2):1–16, 2010. 2

[42] X. Xu, Y. Zhou, B. Shao, G. Feng, and C. Yu. Gesturesurface: Vr sketching through assembling scaffold surface with non-dominant hand. *IEEE Transactions on Visualization and Computer Graphics*, 5 2023. doi: 10.1109/TVCG.2023.3247059 2

[43] H. A. Yousif, A. Zakaria, N. A. Rahim, A. F. B. Salleh, M. Mahmood, K. A. Alfarhan, L. M. Kamarudin, S. M. Mamduh, A. M. Hasan, and M. K. Hussain. Assessment of muscles fatigue based on surface emg signals using machine learning and statistical approaches: a review. In *IOP conference series: materials science and engineering*, vol. 705, p. 012010. IOP Publishing, 2019. 4

[44] F. Zhang, S. Chu, R. Pan, N. Ji, and L. Xi. Double hand-gesture interaction for walk-through in vr environment. In *2017 IEEE/ACIS 16th International Conference on Computer and Information Science (ICIS)*, pp. 539–544. IEEE, 2017. 2

[45] L. Zhang, W. He, H. Bai, Q. Zou, S. Wang, and M. Billinghurst. A hybrid 2d–3d tangible interface combining a smartphone and controller for virtual reality. *Virtual Reality*, 27(2):1273–1291, 2023. 2

[46] L. Zhang, Y. Liu, H. Bai, Q. Zou, Z. Chang, W. He, S. Wang, and M. Billinghurst. Robot-enabled tangible virtual assembly with coordinated midair object placement. *Robotics and Computer-Integrated Manufacturing*, 79:102434, 2023. 2

[47] Y. Zheng and X. Hu. Interference removal from electromyography based on independent component analysis. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 27(5):887–894, 2019. 4

[48] Q. Zou, H. Bai, L. Gao, A. Fowler, and M. Billinghurst. Asymmetric interfaces with stylus and gesture for vr sketching. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 968–969, 2022. doi: 10.1109/VRW55335.2022.00335 2

[49] Q. Zou, H. Bai, L. Gao, G. A. Lee, A. Fowler, and M. Billinghurst. Stylus and gesture asymmetric interaction for fast and precise sketching in virtual reality. *International Journal of Human–Computer Interaction*, pp. 1–18, 2024. 1, 2, 3, 4, 8