Comparing Spatial Interaction Modalities for 2D-Widgets in Productivity Applications in Virtual Reality

Ying-Chao Tung
University of Washington
Seattle, WA, USA
yct56@cs.washington.edu

Stephen DiVerdi
Adobe Research
San Francisco, CA, USA
diverdi@adobe.com

Aaron Hertzmann
Adobe Research
San Francisco, CA, USA
hertzmann@adobe.com

Jacob O. Wobbrock
Information School | DUB Group
University of Washington
Seattle, WA, USA
wobbrock@uw.edu

ABSTRACT

Virtual reality applications often rely on standard 2D widgets such as buttons and sliders, but each application presents them differently. We identify a two-axis design space for 2D widgets in VR: the widget placement can be world-, view-, or hand-stabilized, and the input mechanism can be mouse, raycast, touch, or gaze. We also include a table-stabilized placement which puts widgets in the 2D plane of the user's physical desk, enabling passive-haptic feedback via touch input. We conduct a study with six representative modalities: *table touch*, *hand touch*, *world raycast*, *world touch*, *world mouse*, and *world gaze*. Results for button selection and slider manipulation show that *hand touch*, *table touch*, and *world raycast* are most successful, while users prefer *table touch* and *hand touch*. A subsequent qualitative pilot study of digital artists using *table touch* and *hand touch* for 2D drawing finds that *table touch* is perceived as more precise, but user preference varies with ergonomic factors.

CCS CONCEPTS

• **Human-centered computing → Interaction techniques**;

KEYWORDS

virtual reality, 3D user interface, 2D widgets

1 INTRODUCTION

The advent of consumer virtual reality (VR) hardware has caused a resurgence of interest in 3D user interface design. Most content is entertainment—gaming and video consumption—but a few productivity applications have emerged and shown tremendous promise, focusing on media such as 3D painting [22], modeling [63], and 360 video production [62]. As users spend more time in VR and do more work there, it raises the question of how can VR best support productivity applications. For example, what would a word processor or 2D graphic design tool look like in VR? Each VR application currently presents its own conventions for how to access, organize, and interact with mostly 2D UI widgets, drawing heavily from existing 3DUI literature. Conversely, a major factor in the success of desktop GUIs is the established and reliable set of expectations around how basic UI elements—buttons, sliders, panels—behave. Therefore, the adoption of VR productivity applications is predicated on the open question: what is the best way to interact with 2D widgets in VR?

Figure 1: A participant doing 2D sketching in VR using tabletop interaction with an office desk. At left is the view of the user at their desk, and at right is the user's view in the HMD.

We believe it is most useful for practitioners to understand the methods that are employed commercially today. 3D user interfaces have already been surveyed in academic literature [3, 7, 8, 27], but the most successful of these techniques have yet to become adopted by commercial VR applications. One explanation is that industry lags behind academia. We suggest instead that academic studies of individual interaction techniques lack the context that comes with productivity software. Application interaction modalities must perform well, but must also be general enough to work across a variety of tools, and should heavily emphasize ease of use (e.g. Apple famously retained one-button mice for many years). As an example, SQUAD [45] may be the state of the art 3D selection interface, but using it in a real application requires the user to suffer a cognitive load for each 3D selection task. One result of this hypothesis is that the interaction modalities employed in commercial applications are not simply behind the academic literature, but are a particular subset that are simultaneously accurate, general, flexible, and easy to use. This makes it interesting to survey this specific class of interfaces as a special case.

In this paper, we investigate the space of input modalities for interacting with 2D UI widgets, specifically buttons and sliders, as well as a simple 2D drawing task, for a user seated at an office desk. We survey the types of input modalities currently in use in available VR applications, such as casting rays to static 2D windows or physically touching hand-held panels. We include in the survey an input modality with passive haptic feedback that is available with existing commercial harware. The modality leverages the user's physical desk to provide passive-haptic feedback for widgets placed on the tabletop. Based on the survey, we propose a design space that separates the placement of the widgets from the widget activation mechanism. Specifically, widgets can be placed in the world, on the user's hand, in the user's view, or on the table, and they can be activated by touch, gaze direction, ray-casting, or the mouse. We also discuss potential trade-offs and examples in this design space.

Motivated by our design space, we perform two studies. First, we compare performance in button selection and slider manipulation tasks across the different input modalities. Second, we use the results to design a simple 2D VR drawing tool. We conduct a pilot expert review of a drawing task to compare tradeoffs in this space and receive qualitative feedback. While we do not explicitly evaluate our modalities in an augmented reality (AR) context, we believe our results would be applicable in that context as well.

The contributions of this work are: (1) a design space that organizes the common input modalities in VR productivity applications; (2) empirical results from a study of 2D widget interaction and drawing in virtual reality; (3) empirical results from a pilot expert review of our most successful input modalities.

2 RELATED WORK

User interfaces for 3D applications have been studied extensively [8]. Early work explored the issues posed by 3D interaction for user interface design [36]. Subsequent work proposes new 3D widgets that support the types of interactions supported by 2D WIMP interfaces, but more suited to the 3D environment. For example, many variants of pull-down menus have been developed [13], including for data-gloves [9] or mobile phones [49]. Other research focuses on the core 3D interaction tasks of object selection and manipulation. Virtual hand and virtual pointer techniques have been developed and compared [7, 69], and more recently an extensive survey of derivative techniques was presented [3]. While these custom 3D interfaces have been successful academically, that success has not thus far translated into commercial software.

An alternate strategy is to translate standard 2D GUI widgets into 3D. Early work put 2D windows directly into AR [16] and evaluated interaction with buttons and click-and-drag type gestures [52]. Surveys of 3D interaction techniques have repeatedly made the claim that 2D WIMP interfaces are poorly suited to 3D VR experiences, because of ergonomic factors, poor mapping of degrees of freedom, and loss of presence or immersion [31, 39]. However, our experience shows that 2D widgets are quite common in commercial VR applications. We believe this is because many user tasks in VR are not in fact 3D tasks (e.g., place a 3D object), but 2D tasks situated in 3D scenes (e.g., turn on a light or adjust a dimmer), and 2D GUI widgets are well-suited for these tasks.

Within the design space of 3D interfaces, there are a number of recurring interaction modalities that are relevant to current commercial VR applications and thus our work.

Raycasting refers to laser pointer input where the hand's pose casts a ray into the scene to intersect with interactive objects. Beyond simple raycasting, previous work has explored bendable [79] or scaled [1] rays to improve selection. Raycasting has also been augmented with additional input modalities, such as a touchpad [46]. Finally while most raycasting is used to intersect with world-stabilized objects, it can also be used for hand-stabilized widgets [15].

Conversely, *touch* input is a virtual hand metaphor that requires moving the input device to physically intersect with a 3D object. In contrast to raycasting, touch requires more motion but can have higher accuracy [55]. Similar to raycasting, touch interfaces have also been improved upon, as in the *Go-Go* technique [68].

Perhaps the simplest input modality is to use the standard desktop *mouse* as a 2D input that selects whatever the cursor occludes in the view (i.e. casting a ray from the eye through the cursor into the scene). Mouse input can be used for widgets that are view-stabilized [62] or world-stabilized [6, 29], and may even outperform 3D inputs for some 3D tasks [5]. As a mouse cursor is normally a 2D icon, adapting it to stereo viewing requires considering its stereo disparity and how it interacts with occluding geometry [74].

When more capable input hardware is not available, *gaze* direction can achieve a level of interactivity, as in Google Cardboard type VR devices. With a single button, gaze can be used to control navigation [71] or even to type [90]. If eye-tracking is available, then eye-gaze can provide further expressiveness [67].

While data gloves were popular in the past, optical hand pose tracking systems are now being developed [82]. Hand tracking and gestures are commercially-available in HoloLens [59] and Leap Motion [48]. Free hand gestures can be used for direct 3D selection and manipulation [51] or navigation of view-stabilized menus [14]. As commercial HMDs do not come with built-in gesture input sensors today, we do not consider this modality in our evaluation.

A common criticism of 3D interfaces is the lack of haptic feedback, and one solution is to rely on passive haptics from an available physical *table*. Grossman and Widgor surveyed 3D interactions on the tabletop and categorized prior works by display and input spaces [27]. On the one hand, tabletop multitouch displays can support 2D tasks in the surface plane with AR displays or virtual environments [2, 18, 30, 34]. On the other hand, tabletop touch displays also enable 3D tasks above the surface of the table [11, 47, 54, 78, 86], or even with hybrid gestures [4]. Alternatively, a tracked, hand-held tablet can provide a similar haptic experience [9, 81].

Finally, while not immersive, large displays (e.g., wall-sized) present similar usability problems. A number of studies have compared different inputs including mouse, raycast, touch, and gaze for different tasks and types of screens [43, 44, 60, 77]. Raycasting has also been augmented with cameras [42] or mobile phones [41], or even bimanual hand gestures [89]. While this work informs our design, it does not directly answer questions about input modality performance in virtual reality.

3 INTERACTION MODALITIES

We propose a design space that is informed by the current practice in VR productivity applications. We also include a modality with haptic feedback that is currently available with existing hardware.

3.1 Commercial Practice

To ground our investigation in tools for consumer VR users, we survey techniques currently used in VR productivity applications.

Figure 2: Using Medium [63] for sculpting. The UI is situated in the 3D environment and is activated by a ray cast from the right hand. At left, the in-headset view, and at right, a user making a selection.

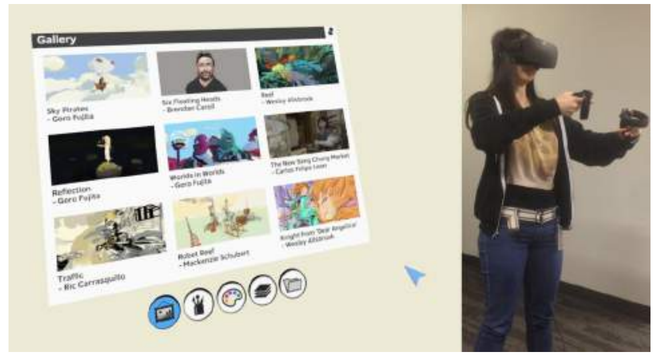

Figure 3: Using Quill [65] for painting. The UI panel follows the left hand and is activated by touches from the right hand. At left, the in-headset view, and at right, a person making a selection.

The modalities used in these applications must balance the goals of being easily accessible, easy to learn, and effective, while integrating well with the rest of the application and interface.

Hardware. In VR productivity applications, the user is situated within a 3D environment and controls the environment with different hardware input devices. In addition to head-tracked gaze direction, the user interacts through hand-held controllers, such as a positionally-tracked controller (e.g., Oculus Touch [64]). Orientation-only controllers (e.g., Google Daydream [25]) can be used, often mimicking positional input. Conventional input devices can also be used when resting on a table, such as the keyboard and mouse. Button-only inputs are sometimes used (e.g., the Oculus Remote [64]) for conventional menu selection, but productivity tasks require more sophisticated controllers.

Laser Pointing. When a positionally-tracked controller is available, perhaps the most common input modality is to select by casting a ray from the user's controller to a UI element that is fixed in world space, akin to using a laser pointer (Figure 2). This interface is widely used in VR applications, including the default home interfaces in Oculus [64] and Daydream [25], as well in many games and other applications.

Palette Touch and Cast. In 3D painting and sculpting interfaces, the user may navigate throughout the world, so there is no fixed position where interface elements may be located. Instead, menu elements are attached to the user's subdominant hand (Figure 3). In Quill [65] and Gravity Sketch [26], the user then either "touches" menu elements with the other hand, as an oil painter would hold a paint palette in one hand and touch the paint brush to it with the other hand. In other modeling tools, including TiltBrush [22], Medium [63], and Blocks [23], raycasting is used from the dominant controller to the palette. This can occasionally cause confusion: the dominant controller is used both for painting/modeling operations and for menu selection, depending on whether the controller is pointing at the palette. The controller mode can unexpectedly switch to palette selection during modeling. Blocks mitigates this by limiting the ray-casting distance.

Quill, TiltBrush, and Gravity Sketch also allow menus to be "pinned," converting them to a world-space alignment.

Panel Touch. Applications such as SoundStage [33] use a direct physical metaphor, placing UI elements in the 3D scene and activating them with direct hand touch. This may be the most physically intuitive strategy, but it comes with the physical cost of difficulty activating elements that are beyond arm's reach.

Some applications that use palette touch, including Quill [65], TiltBrush [22] and Gravity Sketch [26], optionally allow interface windows to be pinned to 3D world locations, creating the panel touch experience. This allows users to save time by placing menus convenient for their 3D workspace, but the panels need to be manually moved when navigating to a new 3D view.

Games in which the player is situated in a virtual environment often include buttons, knobs, dials, or other controls in the world. Examples include Please Don't Touch Anything [19], I Expect You to Die [73], and Rick and Morty Virtual Rick-ality [66]. Lone Echo [72] includes hand-, view-, and world-stabilized interface elements. In these cases, the choice of interaction is driven by storytelling and immersion goals rather than productivity.

Mouse and Keyboard. Some software, such as BigScreen [6] and Virtual Desktop [29], approaches VR productivity from the opposite direction: displaying existing desktop applications directly in VR, using the mouse and keyboard. These applications benefit from the much larger apparent display space of VR, but otherwise use conventional input and output mechanisms. The Vremiere 360 video-editing system [62] also uses mouse and keyboard input, to leverage users' comfort with existing input metaphors (Figure 4).

Gaze Control. Some hardware (e.g., Google Cardboard [24]), provides no position input other than the viewer's gaze direction. In this case, the user must look at the target widget and then either wait a fixed duration, or press the only button on the device (Figure 5). As this method is quite slow, it is generally not used when a controller is available. This is used, for example, in the New York Times [83], Jaunt [40], and Gala 360 [80] viewers. Note that gaze direction is determined by the direction of the user's head, since eye tracking is not widely available.

Figure 4: Using Vremiere [62] for video editing. The UI panel is affixed to the view and is activated by the mouse and keyboard. At left, the in-headset view, and inset, a person using the interface.

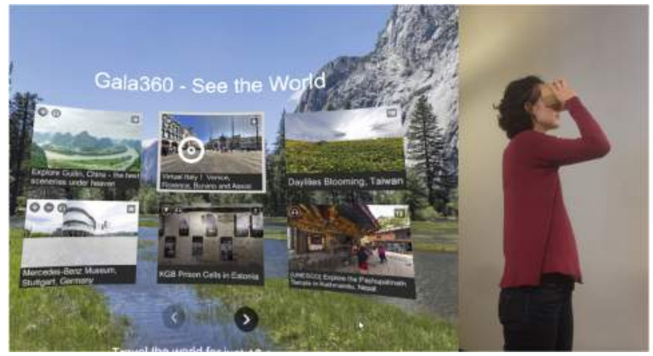

Figure 5: Using Gala 360 [80] for viewing. The UI is oriented statically around the headset and is activated by rotating the headset to center on the desired element and either waiting a fixed duration or using a physical button. At left, the in-headset view, and at right, a person using the interface.

3.2 Table-Based Input

Compared to *raycast* or *panel touch*, interactions with passive haptic feedback augments touch interactions and enables more precise interface manipulation [9, 12, 28, 35]. The table-based interaction, one of the common interactions with passive haptic feedback, has been explored with different configurations, such as AR or virtual environments [4, 11, 18, 56, 85].

For the specific case of productivity applications, users will often be seated at a computer desk for reasons of ergonomics and comfort. Moreover, the presence of the physical table and the existing tracking hardware make the table-based interaction available for commercial VR applications. Thus, we include table-based input into our design space and evaluation to understand its usability with the current commercial hardware. Technically this is a subclass of the panel touch metaphor described earlier, where panels are statically positioned in the scene, but the table-based interaction require the UI widgets to be positioned in the plane of the desk, and activated by the direct touch.

We prototype the interaction with Oculus Touch [64] controllers. The user must calibrate the height and extent of the space; this is similar to the room calibration currently required by VR devices. Calibration is performed by placing the controllers at opposite corners of the free space on the table.

Some tasks, such as 2D drawing, require fine precision control. The Oculus Touch controller uses a *power grip* [61] in which the handle is grasped between the palm and fingers, making fine-scale control difficult. We propose an alternative *precision grip* [61] in which the handle is grasped between the thumb and index finger and the butt of the controller is the selection point (Figure 6). We also attach felt pads on the controllers to make drawing easier. A proper stylus would be preferable but we focus on commercially available hardware.

3.3 Interaction Design Space

We observe that interfaces can be described along two separate axes constituting a design space: Widgets may be **placed** in the scene stabilized with respect to the *world*, *view*, *hand*, or *table*, and widgets may be **activated** by *raycast*, *gaze*, *touch*, or *mouse*.

Existing interfaces can be categorized by selecting a widget placement and an activation mechanism, for example Vremiere [62] uses *view mouse*, whereas Quill [65] uses *hand touch*. Table 1 illustrates this design space. We observe several combinations that are not previously tested. Some combinations do not make sense. For example, it is impossible to control a view-stabilized UI by gaze direction. *Hand mouse* and *view raycast* are similarly unlikely. Other combinations are simply less popular, e.g., at least one *hand raycast* interface does exist [15] but it does not have obvious benefits.

We now describe the widget placement categories of our design space more concretely.

World. Widgets are placed on 2D planes that are positioned with respect to the world coordinate system. As the user moves and turns, the widgets will go in and out of view like real objects in the environment.

Within this category, widgets may be stabilized with respect to either the *virtual world* or the *real world*. Virtual World alignment is used when widgets represent "physical" buttons in a virtual world, whereas real world alignment is used in Quill and TiltBrush for "pinned" menus. Real-World and Virtual-World Stabilization behave the same when the user moves normally around in their workspace; they are different when the user teleports or scales the world.

View. Widgets are overlaid on top of the rendered 3D scene, positioned with respect to the user's view coordinates and oriented to always face the viewer. This creates a "heads up display" where the wigets are always visible regardless of the user's pose.

Hand. Widgets are attached to the user's tracked, non-dominant hand. The user can move their hand to control the visibility of the widgets, and interacting with them requires a bimanual gesture.

Table. Widgets are placed in the 2D plane of the physical desk in the user's workspace. The widgets behave the same as in the *world* condition, with the additional affordance that the feel of the real desk will help the user maintain their frame of reference.

Now we describe the activation techniques of our design space.

Raycast. The controller casts a ray which is intersected with the widget's 2D plane to compute a simulated 2D mouse coordinate. The controller trigger is used as a primary mouse button.

Figure 6: Gripping the Oculus Touch controller. At left is the *precision grip* we use for *table touch*. At right is the *power grip* we use for the *world raycast*, *world touch*, and *hand touch* modalities.

Gaze. A ray is cast from the headset's position, in the direction of the headset orientation (the center of the user's view), which yields a 2D coordinate as in *raycast*. A single button on the headset is used as a primary mouse button.

Touch. A point on the controller is used to compute 3D proximity to widgets in the scene, and if when it crosses a distance threshold, a mouse down event is simulated, until the distance goes back above that threshold and a mouse up event is simulated.

Mouse. The desktop mouse controls a view-stabilized 2D cursor in the users field of view. A ray is implicitly cast from the user's eye through the cursor's 2D position into the scene to intersect with widgets and compute a 2D coordinate in the widget plane. The mouse's primary button is used.

4 QUANTITATIVE EVALUATION

To evaluate the performance of and users' preference for different UI modalities, we conducted a user study with two basic UI manipulation tasks: button selection and slider manipulation.

We recruited 16 able-bodied participants (9 male, 7 female), aged 21 to 45 ($M = 28$, $SD = 6.88$). Three of the participants had experience using non-gaming applications in VR. Six had never ever experienced VR. Participants were paid for their participation.

Participants used our customized experiment software with Oculus Rift headset and Oculus Touch controllers on an Micro-Star International (MSI) Dominator Pro 2017 laptop with a GeForce GTX 980 graphic card and Windows 10 OS running in front of a standard office table. The Oculus tracking sensors were positioned on another desk next to the study table to ensure the tracking system could sense the controller's and headset's movement across the entire table (Figure 7).

4.1 Procedure

Each study session consisted of three stages.

First, participants were given an introduction to each UI modality. The experimenter then showed the user two tasks and explained the procedure of the study. Each participant had at most three minutes to try with each UI modality with a UI menu containing multiple buttons and sliders (Figure 9).

Second, participants were required to perform button selection tasks first and then slider manipulation tasks with six UI modalities in a random order. Each task was performed five times before moving to the next task.

	World	View	Hand	Table
Raycast	Common [25, 64]	Difficult	Unused	Unused
Gaze	Limited [40, 80, 83]	Difficult	Unused	Unused
Touch	Common [19, 33, 66, 73]	Unused	Common [22, 23, 26, 63, 65]	Evaluated
Mouse	Common [6, 29]	Limited [62]	Difficult	Unused

Table 1: Summary of our design space. Activation methods are shown on the vertical axis, and placements along the horizontal axis. Cited examples are described in the text. Cells are colored as follows: Common : Frequently used in current applications, Limited : Used rarely, such as in cases of limited controllers, Unused : Not currently used for productivity, but could be, Difficult : Would be difficult or impossible to use, and Evaluated : Not currently used for productivity, but was evaluated in this paper.

Third, the experimenter asked the participant questions revised from an existing questionnaire about computer system usability [50].

4.2 Tasks

In order to better study 2D UI interaction techniques, we can decompose user interactions into basic motions, using what Shneiderman called *Widget-Level* decomposition [76]. This approach looks at widgets that are defined in the system and decomposes them based on their implementation. Extracted from common productivity applications in VR, two distinct types of actions—discrete and continuous—are evaluated with buttons and sliders in our testbed.

In the *button selection* task, the participant was shown a 3×3 grid of buttons and a *start button* (Figure 8 left). The participant was required to trigger the *start button* before each trial to provide a consistent starting location.

In the *slider manipulation* task, the participant was asked to drag the value of a slider to the target value shown on a neighboring text field (Figure 8 right). The goal was to measure how precisely the user can position the slider in a short amount of time. Each trial was therefore limited to five seconds, even if by that point the user could not achieve the target value. The five-second duration is drawn from Andujar and Argelaguet [1].

To reduce the influence of confounding factors in our study, we set the visual size of the UI widget to be constant across all interaction modalities. We use a unit of measure called a *dmm* (distance-independent millimeter) [58]. The *dmm* simplifies the traditional user-centered visual size measurement, which is defined as the vertical and horizontal angles the object occupies in the user's field of view. In accordance with the minimum size recommendation, we set the size of the button in our study to 64 $dmm \times 64$ dmm with 16 *dmm* padding. For the slider, we set the interactive area as 300 $dmm \times 64$ dmm, which users reported as the minimum size they

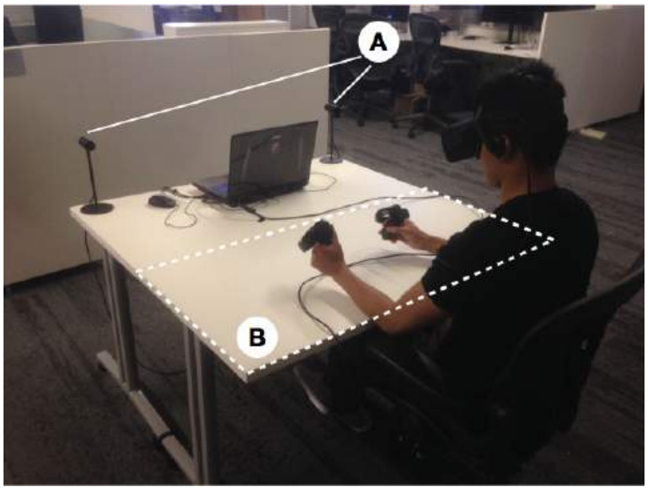

Figure 7: The study environment. Two Oculus position tracking sensors (A) are spaced apart to ensure good coverage of the entire interaction space (B). The participant wears an Oculus Rift headset and holds Oculus Touch controllers.

could manipulate comfortably in our pilot study. Due to the time limit of the study, we only evaluated with minimum size widgets, which can be the lower bound reference of human performance with different modalities. Different factors, such as sizes, distances, or other forms of widgets, might be an interesting direction for further evaluation in the future.

Note that we did not include any Fitts' Law test [17, 70] since the interaction area (the size of the widget visually perceived by users) was dynamically varied by users with *hand touch* modality. With *hand touch* modality, users can control the position of UI wdgets by their nondominant hands. We did not dynamically change the size of the widget based on the distance between the user and the widget since there was no such configuration in the prior works or existing commercial applications in VR.

4.3 UI Modalities

Drawing from our proposed design space, we compare the performance of six UI modalities on two 2D widget manipulation tasks. For each participant, we measured their arm length and desk height with a quick calibration procedure. The configuration for each UI modality is as follows.

World raycast: the UI widgets are placed one arm's length in front of the participant to keep them out of reach. The ray is visualized by a line emanating from the controller, and a dot at the widget intersection point (Figure 9A).

World gaze: the UI widgets are placed the same as in *world raycast*. The ray is cast from the user's head position, so only the dot at the widget intersection point (in the center of the field of view) is visualized (Figure 9B).

World mouse: the UI widgets are placed the same as in *world raycast*. We also add a background canvas to help participants feel like the cursor was moving on a virtual screen rather than floating in mid-air. The cursor is visualized as a small 2D arrow (Figure 9C).

World touch: the UI widgets are placed between 0.50 to 0.75 arm's lengths from the user, which can be adjusted by the user before starting the task. The distance range was defined since widgets were placed too close or too far caused fatigue rapidly. The controller touch point is visualized as a small 3D triangle (Figure 9D).

Hand touch: the UI widgets are placed on top of the participant's non-dominant hand tilted up 45° from the up axis of the handheld controller. The controller touch point is the same as in *world touch* (Figure 9E).

Table touch: the UI widgets are placed on a virtual table with height equal to the physical table in front of the participant. The controller touch point is at the base of the controller handle (Figure 9F).

4.4 Design and Analysis

The study was a factorial within-subjects design with the following factors and levels:

- *Modality* (six techniques, in alphabetical order): hand touch, table touch, world gaze, world mouse, world raycast, world touch.
- *Trial*: 1-5
- *Participant*: 1-16

Modality was modeled as a fixed effect. *Trial* was nested within *Modality* and modeled as a random effect. *Participant* was also modeled as a random effect.[1]

With 16 participants performing five trials using each of six input modalities, we generated 480 data points for the button selection task and 480 data points for the slider manipulation task. Subsequent interviews yielded responses on six Likert scales for each input modality, or 576 responses in all. Finally, participants ranked each modality for overall preference, producing 96 rank data points.

Assumptions for normality and homogeneity of variance were tested using the Shapiro-Wilk [75] and Brown-Forsythe [10] tests, respectively. As our data was found unsuitable for parametric analysis, or was ordinal or rank-based in nature, nonparametric tests were used in all analyses. Specifically, the Aligned Rank Transform procedure was used for nonparametric analyses of variance [38, 57, 87]. Multinomial and binomial tests were used for analyzing counts. Post hoc pairwise comparisons were corrected using Tukey's correction [84] or Holm's sequential Bonferroni procedure [37]. Statistical tests were carried out in R using, among others, the `ARTool`, `lme4`, `lsmeans`, and `XNomial` packages.

5 RESULTS

We present the results of our user study and the Likert and preference responses obtained in our subsequent interviews.

5.1 Button Selection Task

The button selection task examined button selection times for the six different interaction modalities. The means and standard deviations for each modality are shown in Table 2.

[1]Fixed effects are those whose levels are chosen explicitly and are of explicit interest. Random effects are those whose levels are not of interest in themselves, per se, but are sampled randomly from a larger population about which we wish to generalize. See Frederick [20] for further explanation, and Littell et al. [53] for a worked example.

Figure 8: User study tasks. At left is an in-headset view of the button selection task. One of nine buttons randomly turns red for each trial. At right is an in-headset view of the slider manipulation task. The target value is shown on the top of the slider with the bigger font size and bold style. To the right side of the slider is its current value.

Input Modality	Mean Time (s)	Stdev Time (s)
Hand touch	1.057	0.366
World raycast	1.082	0.553
World touch	1.132	0.489
World gaze	1.258	0.602
Table touch	1.298	1.540
World mouse	1.846	1.034

Table 2: Means and standard deviations for button selection times (in seconds) by input modality, in ascending order. Lower is better.

Unsurprisingly, with time as a measure, the response violates a Shapiro-Wilk test for normality ($W = 0.58$, $p < .0001$). It also violates the Brown-Forsythe test for homogeneity of variance ($F_{5,474} = 4.81$, $p < .001$). Accordingly, the nonparametric Aligned Rank Transform procedure was used. There was a significant effect of Modality on button selection time ($F_{5,24} = 11.11$, $p < .0001$). Post hoc pairwise comparisons using Tukey's correction reveal that all input modalities were significantly faster than *world mouse* ($p < .05$), and that *world raycast* was significantly faster than *world gaze* ($p < .05$).

5.2 Slider Manipulation Task

The slider manipulation task examined how close to a target value (1–100) a slider could be positioned within a five second window. The means and standard deviations for the distances from these target values are shown in Table 3.

As with the button selection task, normality was violated ($W = 0.34$, $p < .0001$). (Although in this case, homogeneity of variance was not violated ($F_{5,474} = 1.82$, n.s.).) Nonparametric analysis was therefore chosen. There was a significant effect of Modality on slider distance ($F_{5,24} = 6.34$, $p < .001$). Post hoc pairwise comparisons using Tukey's correction show that *world raycast* was significantly more accurate than *world mouse* ($p < .05$), but that *hand touch* and *world touch* were significantly less accurate than *world mouse* ($p < .01$). Also, *table touch* was significantly more accurate than *world touch* ($p < .01$).

We conducted a further analysis of the slider task by examining the number of trials that were "correct," i.e., when the slider was placed on the target value exactly. The number of correct trials of 80 total trials for each input modality is shown in Table 3. A multinomial test indicates that the input modalities produced significantly different correctness counts ($p < .01$). Follow-up binomial tests for each modality, corrected with Holm's sequential Bonferroni procedure, indicate that *world mouse* and *world touch* deviated significantly from the other correctness counts ($p < .05$), the former being much more correct and the latter being much less correct.

Input Modality	Mean Distance	Stdev Distance	Correct ($n = 80$)
Table touch	1.125	2.852	38
World raycast	1.138	1.220	31
World mouse	1.788	5.305	55
Hand touch	1.863	3.393	26
World gaze	2.063	6.643	34
World touch	3.525	8.344	21

Table 3: Means and standard deviations for slider distances from target values by input modality, in ascending order. Lower is better. Also shown is the number of correct trials where the target value was selected exactly (out of a possible $n = 80$).

5.3 Likert and Preference Responses

In our subsequent interviews, we asked participants to rate each input modality on six Likert scales ranging from 1 to 5. Figure 10 shows average Likert responses. As the Likert responses were ordinal, we analyzed them with the nonparametric Aligned Rank Transform procedure. (We also confirmed each result with the nonparametric Friedman test [21], and statistical conclusions agreed in all cases.)

5.3.1 Ease of Use of Each Input Modality. Participants indicated, for each input modality, its general ease of use where 1 was "very difficult" and 5 was "very easy." On average, *table touch* was rated easiest ($M = 4.50$, $SD = 0.73$) and *world gaze* most difficult ($M = 2.94$, $SD = 1.24$). An omnibus test shows a significant effect of Modality on perceived ease of use ($F_{5,75} = 6.46$, $p < .0001$). Post hoc pairwise comparisons using Tukey's correction indicate that *world gaze* was significantly more difficult than, in descending order of ease, *table touch*, *world raycast*, *world mouse*, and *hand touch* ($p < .05$). Also, *world touch* was significantly more difficult to use than *table touch* ($p < .05$).

5.3.2 Ease of Button Selection Task. Participants also indicated, for each input modality, its ease for completing the button selection task. All modalities received average scores above 4.0 except *world gaze* ($M = 3.75$, $SD = 0.58$), again rated most difficult. *World touch* was rated easiest on average ($M = 4.81$, $SD = 0.40$). An omnibus test shows a significant effect of Modality on perceived ease of use for button selection ($F_{5,75} = 9.08$, $p < .0001$). Post hoc pairwise comparisons using Tukey's correction show that *world gaze* was significantly more difficult than, in descending order of ease, *world touch*, *hand touch*, *table touch*, and *world raycast* ($p < .001$).

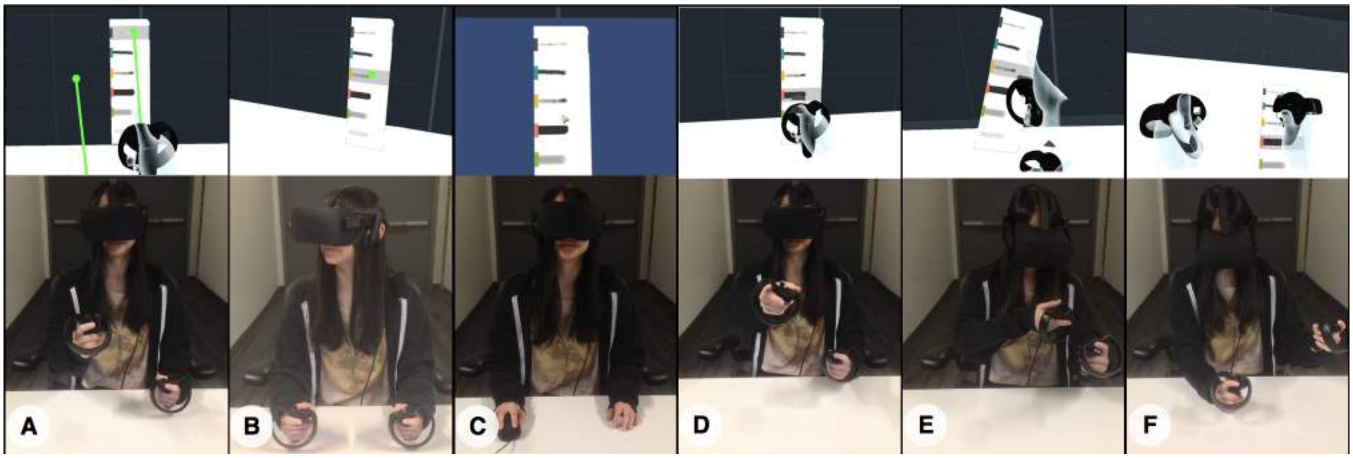

Figure 9: A study participant selecting the brush tool with six interaction modalities during the practice session. From left to right, A) *world raycast*, B) *world gaze*, C) *world mouse* D) *world touch*, E) *hand touch*, F) *table touch*.

5.3.3 Ease of Slider Manipulation Task. Participants rated, for each input modality, its ease for performing the slider manipulation task. As for general ease of use, *table touch* was rated easiest for the slider manipulation task ($M = 4.25$, $SD = 0.93$). Once again, *world gaze* was rated most difficult ($M = 2.19$, $SD = 1.11$). In general, ratings were below 4.0 except for *table touch*, indicating this was a more difficult task. An omnibus test reveals a significant effect of Modality on perceived ease of use for the slider manipulation task ($F_{5,75} = 9.69$, $p < .0001$). Post hoc pairwise comparisons using Tukey's correction show that *table touch* was significantly easier than, in descending order of ease, *hand touch*, *world touch*, and *world gaze* ($p < .05$). Furthermore, *world mouse* and *world raycast* were significantly easier than *world gaze* ($p < .01$). Finally, *world mouse* was also easier than *world touch* ($p < .05$).

5.3.4 Ease of Learning. Participants also rated input modalities for ease of learning. All modalities except *world gaze* received average ratings above 4.0, with *world mouse* being rated as easiest to learn ($M = 4.88$, $SD = 0.34$). Again, *world gaze* was rated most difficult ($M = 3.50$, $SD = 1.26$). An omnibus test indicates a significant effect of Modality on perceived ease of learning ($F_{5,75} = 5.85$, $p < .001$). Post hoc pairwise comparisons using Tukey's correction indicate that *world gaze* was significantly more difficult than, in descending order of ease, *world mouse*, *table touch*, *world raycast*, and *world touch* ($p < .05$).

5.3.5 Fatigue vs. Comfort. Participants indicated, for each input modality, how fatiguing or comfortable it was to use, where 1 was "very tiring" and 5 was "very comfortable." On average, *table touch* was considered most comfortable ($M = 4.38$, $SD = 0.62$). In contrast, *world gaze* was most tiring ($M = 2.56$, $SD = 1.31$). An omnibus test shows a significant effect of Modality on perceived fatigue vs. comfort ($F_{5,75} = 8.90$, $p < .0001$). Post hoc pairwise comparisons using Tukey's correction reveal that *world gaze* and *world touch* were both significantly more fatiguing than, in descending order of comfort, *table touch*, *world mouse*, and *world raycast* ($p < .05$).

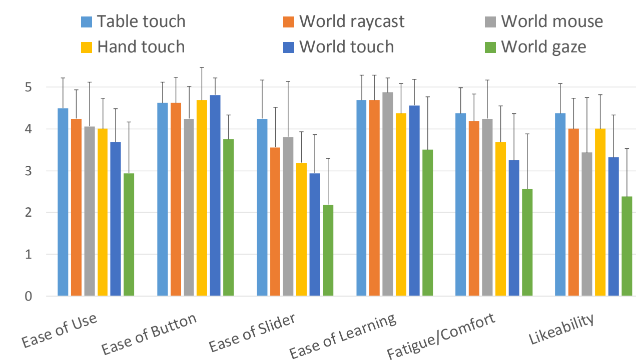

Figure 10: Mean ratings for each of six input modalities on six Likert scales (1–5) from 16 participants. Higher is better. Input modalities are ordered by descending overall ease of use, the leftmost scale. Error bars are +1 *SD*.

5.3.6 Likeability. Participants also rated, for each input modality, how much they liked it, where 1 was "did not like at all" and 5 was "liked it very much." Perhaps unsurprisingly, likeability responses generally comported with perceptions of ease and comfort. Specifically, *table touch* was most liked ($M = 4.38$, $SD = 0.72$) and *world gaze* was least liked ($M = 2.38$, $SD = 1.15$). Other input modalities receiving scores of 4.0 or better were *world raycast* and *hand touch*. An omnibus test indicates a significant effect of Modality on likeability ($F_{5,75} = 9.08$, $p < .0001$). Post hoc pairwise comparisons using Tukey's correction indicate that *world gaze* was significantly less liked than, in descending order of likeability, *table touch*, *world raycast*, *hand touch*, and *world mouse* ($p < .05$). Also, *table touch* was significantly more liked than *world mouse* and *world touch* ($p < .05$).

5.3.7 Input Modality Preference. Participants rank-ordered (1st–6th) their overall preference among input modalities. Each modality was uniquely ranked by each participant. We calculated an overall preference score for each modality by giving $P = 6 - R + 1$ points for each ranking R, multiplying those points P by the number of participants N who chose that rank, and then summing over the given modality. For example, *table touch* was ranked $R = 1st$ by $N = 7$ participants, so $P = (6 - 1 + 1) = 6$, and $P \times N = 6 \times 7 = 42$ points. Similarly, *world gaze* was ranked $R = 6th$ by $N = 9$ participants, for $1 \times 9 = 9$ points. Total points for a modality are the sum of its points for each ranking, and are shown in Table 4. The maximum score a modality could receive would be if all $N = 16$ participants ranked it their 1st favorite, for $6 \times 16 = 96$ points. The minimum would be if all participants ranked it their 6th favorite (i.e., least favorite), for $1 \times 16 = 16$ points.

Input Modality	Total Score (range 16–96)	Ranked first by?	Ranked last by?
Table touch	78	7	0
Hand touch	67	4	0
World raycast	64	1	1
World mouse	50	4	6
World touch	48	0	0
World gaze	29	0	9

Table 4: Ranked preference scores for each modality (range 16–96). Higher is better. See text for details on how scores were calculated. Also shown are the number of participants (of $N = 16$) who ranked each input modality first (their favorite) or last (their least favorite).

A multinomial test indicates that the total scores were indeed significantly different ($p < .001$). Follow-up binomial tests for each modality, corrected with Holm's sequential Bonferroni procedure, indicate that *table touch* and *world gaze* deviated significantly from the other scores ($p < .05$), the former being much more preferred and the latter being much less preferred.

6 DISCUSSION

The goal of our study is to understand users' performance and preference on different UI modalities when performing 2D tasks in VR. In this section, we discuss the results of questionnaires and interviews for each of our input modalities, respectively.

6.1 Implications for Touch-based Interaction

According to the overall preference scores, *table touch* is significantly preferred over other input modalities. Seven of 16 users rated it as their favorite to use for 2D widget manipulation tasks. Participants mentioned that it was easier and more natural to use due to the haptic feedback provided by the table surface. Moreover, during the user study, users could perform the task while resting their arms on the table which they found to be comfortable. Using table for appropriate arm rest was also supported by prior works [32, 56]. However, in our interview session, some participants said that looking down at the table to use the interface made them tired, which is why they preferred *hand touch* instead.

Compared to *world touch*, our results show that *hand touch* has not only better performance on the button and slider tasks, but also higher scores for preference and usability. With the same input method, the placement of the user interface affects the preference. Participants indicated that dynamic adjustment of the UI placement helps them. Moreover, users reported that *hand touch* was more intuitive because the UI was always readily available on the hand, which requires less effort than searching for the UI in the world.

Due to the lack of the haptic feedback and the freedom to place the user interface dynamically, *world touch* has the worst performance among direct touch interactions. Keeping the arm stretched out in the air also causes fatigue, which was reported by participants. This emphasizes the importance of providing users a way to dynamically position *world touch* interfaces. Users all agreed that *world touch* is easy for button selection, but hard for precise manipulation such as slider control.

6.2 Implications for Pointer-based Interaction

Unsurprisingly, *world gaze* is rated most difficult for both tasks, since the modality requires head movement. Seven of 16 participants indicated concerns about fatigue for *world gaze* in our interviews. Participants feel that people cannot use *world gaze* for UI manipulation tasks frequently or for extended durations, since the neck movement made them tired quickly in our study. Besides, participants mentioned that it is hard to focus on the trigger of the slider and check the value at the same time, which makes them feel even more tired. Since *world gaze* requires users to focus on the trigger of the slider while manipulating it, we suggest the layout of the slider design should consider this factor. A better interface would display the current value on the handle rather than adjacent to the slider.

Users reported that *world mouse* is the easiest to learn because every participant was familiar with traditional desktop input methods. According to the result of the slider manipulation task, *world mouse* is significantly more accurate than other modalities. Interestingly, *world mouse* has a polarizing effect on user preference. Four of 16 participants ranked it as their top choice, but six others ranked it last. In our interviews, familiarity and precision are the main reasons people preferred it. In contrast, participants with the opposite opinion indicated they think the mouse constrains freedom in 3D space. They were also concerned about being able to find the mouse on the desk while wearing the headset.

Currently, *world raycast* is the most common modality for real world applications. Compared to other pointer interactions, *world raycast* has higher likeability and preference scores. Although only one person ranked it as their favorite modality, the overall preference score placed *world raycast* third of six. In our study, people believed that *world raycast* was easy and comfortable to control with small wrist movements. Furthermore, due to our study's seated setting, one third of participants reported resting the sides of their palms or their wrists on the table to grain more precision, which is similar to observations of previous work [36].

6.3 Physical Considerations

We designed our study so that all input modalities are equally applicable to the chosen scenario to allow a fair comparison, but real applications have constraints that make some modalities more

Figure 11: Our VR 2D drawing app. The *table touch* modality is demonstrated. To the right of the canvas are the tool menu and the color palette.

or less appropriate, depending on, for example, whether the user will be seated at a desk or navigating a 3D virtual space.

Among Placement strategies, Hand stabilization works in the general case of unconstrained 3D environments when the user might be moving through the 3D space; World stabilization works better when navigation is relatively constrained; and Table stabilization works even better, but requires the presence of a table without 3D navigation. Study participants also found Hand stabilization to be fast for some seated tasks that did not require navigation. Among Activation strategies, Raycast is the most general strategy; Touch works better, but only for elements that are near the user in 3D; and Mouse also can work well but requires a table. Gaze is the least effective, and only used for cases where no positional or orientation controller is available.

7 EXPERT REVIEW PILOT STUDY

Per our hypothesis that studying interaction modalities in isolation does not thoroughly test their performance, we conduct an additional study within the context of a productivity application, where the primary user goal is to perform a creative task, and the 2D widget manipulation is a support task. Our hope is this will give a more accurate sense of the actual performance of our modalities under demanding real world conditions.

We conducted an expert review pilot study of artists using a 2D drawing tool in VR (Figure 11). The interview of the expert review focused on the experience of interacting UI widgets with *table touch* or *hand touch* input modalities during the creative process in VR. Specifically, we used 2D drawing application rather than 3D painting tool in the study since 3D painting tools require a time-consuming learning phase to be familiar with functionalities and features for users. Moreover, the potential of creating 2D drawing in mixed reality has been explored by DodecaPen [88].

We recruited two experts (professional digital artists) who have worked with digital content creation tools for at least five years. One had prior VR experience (a VR video player).

The study lasted about 40 minutes, including 10 minutes for introduction and practice in the VR 2D drawing application. After practicing, the participants used the application to create a drawing, prompted by an inspirational image. For precise control, the drawing canvas is always configured for *table touch* interaction. The tool menu and color palette are controlled and placed together according to one of two modalities depending on the task, either *table touch* or *hand touch*.

7.1 Results

Participants were interviewed about the usability and comfort of switching between the UI panel and the drawing canvas. Both participants responded positively about *table touch* for precise control.

P1: *Sliders and interface worked better as well as natural feel of touching table made picking things in interface easier*

P2: *The feedback is nice, you know you're on it. The precision and the movement were nice.*

For the usability of *hand touch*, experts believed it was only suitable for the discrete manipulation task.

P1: *Not sure about this particular use because the interaction requires more precision for sliders, but buttons and quick actions could be useful in this format.*

P2: *It's good to control with buttons. Dragging the slider is a little awkward.*

In the *table touch* condition, since the content canvas and the UI panel are both placed on the table, the UI panel is shifted to the right of the content canvas. Experts had some comments about the placement.

P1: *Required more full body movement to position my line of sight.*

P2: *Slightly awkward to move tool to right to pick features but not as difficult as hand touch*

The two experts' feedback implies that precision and ergonomics are two main factors for using 2D UIs in creation tools in VR. Although *table touch* can provide more precise controls, the placement of the UI panel affects the usability and comfort. We believe further study of UI placement will provide more insights for future designers of VR productivity applications.

8 CONCLUSIONS AND FUTURE WORK

In this work, we survey the input modalities employed in commercial VR productivity applications, and based on this survey, we propose a design space that organizes those applications. We extend the design space with, *table touch*, which uniquely utilizes a physical desk in front of a seated user to support desktop productivity tasks. We present a user study of discrete and continuous 2D widget manipulation tasks and find that *table touch*, *hand touch*, and *world raycast* were more successful and preferred by users. We also present an expert review pilot study of modalities in a VR creation tool application, and find that user preference varies with ergonomics factors, such as the placement of the drawing canvas and the UI panel.

For future work, we would like to apply *table touch* to 3D content creation tools such as 3D painting and sculpting to reduce fatigue and increase precision in those contexts. We would also like to see how our design space applies to augmented reality applications. Today, our results do not only help designers create better user experiences for their VR applications, but also provide new avenues for researchers to explore.

REFERENCES

[1] Carlos Andujar and Ferran Argelaguet. 2006. Friction surfaces: scaled ray-casting manipulation for interacting with 2D GUIs. In *Eurographics Symposium on Virtual*

Environments, Ming Lin and Roger Hubbold (Eds.). The Eurographics Association. https://doi.org/10.2312/EGVE/EGVE06/101-108
[2] Carlos Andujar, Marta Fairen, and Ferran Argelaguet. 2006. A Cost-effective Approach for Developing Application-control GUIs for Virtual Environments. In *3D User Interfaces (3DUI'06)*. 45–52. https://doi.org/10.1109/VR.2006.6
[3] Ferran Argelaguet and Carlos Andujar. 2013. A survey of 3D object selection techniques for virtual environments. *Computers & Graphics* 37, 3 (2013), 121–136. https://doi.org/10.1016/j.cag.2012.12.003
[4] Hrvoje Benko, Edward W. Ishak, and Steven Feiner. 2004. Collaborative mixed reality visualization of an archaeological excavation. In *Third IEEE and ACM International Symposium on Mixed and Augmented Reality*. 132–140. https://doi.org/10.1109/ISMAR.2004.23
[5] François Bérard, Jessica Ip, Mitchel Benovoy, Dalia El-Shimy, Jeffrey R. Blum, and Jeremy R. Cooperstock. 2009. Did "Minority Report" Get It Wrong? Superiority of the Mouse over 3D Input Devices in a 3D Placement Task. In *Proceedings of the 12th IFIP TC 13 International Conference on Human-Computer Interaction: Part II (INTERACT '09)*. Springer-Verlag, Berlin, Heidelberg, 400–414. https://doi.org/10.1007/978-3-642-03658-3_45
[6] BigScreen. 2016. BigScreen. http://bigscreenvr.com/ (version 0.19.0).
[7] Doug A. Bowman and Larry F. Hodges. 1997. An Evaluation of Techniques for Grabbing and Manipulating Remote Objects in Immersive Virtual Environments. In *Proceedings of the 1997 Symposium on Interactive 3D Graphics (I3D '97)*. ACM, New York, NY, USA, 35–ff. https://doi.org/10.1145/253284.253301
[8] Doug A. Bowman, Ernst Kruijff, Joseph J. LaViola, and Ivan Poupyrev. 2004. *3D User Interfaces: Theory and Practice*. Addison Wesley Longman Publishing Co., Inc., Redwood City, CA, USA.
[9] Doug A. Bowman and Chadwick A. Wingrave. 2001. Design and Evaluation of Menu Systems for Immersive Virtual Environments. In *Proceedings of the Virtual Reality 2001 Conference (VR'01)*. IEEE Computer Society, Washington, DC, USA, 149–156. http://dl.acm.org/citation.cfm?id=580521.835855
[10] Morton B. Brown and Alan B. Forsythe. 1974. Robust Tests for the Equality of Variances. *J. Amer. Statist. Assoc.* 69, 346 (1974), 364–367. http://www.jstor.org/stable/2285659
[11] Gerd Bruder, Frank Steinicke, and Wolfgang Stuerzlinger. 2013. *Touching the Void Revisited: Analyses of Touch Behavior on and above Tabletop Surfaces*. Springer Berlin Heidelberg, Berlin, Heidelberg, 278–296. https://doi.org/10.1007/978-3-642-40483-2_19
[12] Dane Coffey, Nicholas Malbraaten, Trung B. Le, Iman Borazjani, Fotis Sotiropoulos, Arthur G. Erdman, and Daniel F. Keefe. 2012. Interactive Slice WIM: Navigating and Interrogating Volume Data Sets Using a Multisurface, Multitouch VR Interface. *IEEE Transactions on Visualization and Computer Graphics* 18, 10 (Oct 2012), 1614–1626. https://doi.org/10.1109/TVCG.2011.283
[13] Raimund Dachselt and Anett Hübner. 2007. Three-dimensional menus: A survey and taxonomy. *Computers & Graphics* 31, 1 (2007), 53–65. https://doi.org/10.1016/j.cag.2006.09.006
[14] Matthew M. Davis, Joseph L. Gabbard, Doug A. Bowman, and Dennis Gracanin. 2016. Depth-based 3D gesture multi-level radial menu for virtual object manipulation. In *2016 IEEE Virtual Reality (VR)*. 169–170. https://doi.org/10.1109/VR.2016.7504707
[15] Gerwin de Haan, Eric J. Griffith, Michal Koutek, and Frits H. Post. 2006. Hybrid Interfaces in VEs: Intent and Interaction. In *Proceedings of the 12th Eurographics Conference on Virtual Environments (EGVE'06)*. Eurographics Association, Aire-la-Ville, Switzerland, Switzerland, 109–118. https://doi.org/10.2312/EGVE/EGVE06/109-118
[16] Steven Feiner, Blair MacIntyre, Marcus Haupt, and Eliot Solomon. 1993. Windows on the World: 2D Windows for 3D Augmented Reality. In *Proceedings of the 6th Annual ACM Symposium on User Interface Software and Technology (UIST '93)*. ACM, New York, NY, USA, 145–155. https://doi.org/10.1145/168642.168657
[17] Paul M. Fitts. 1954. The information capacity of the human motor system in contrJournal of Experimental Psychology: General. 47(1954), 6 (1954). https://doi.org/10.1037/h0055392
[18] Clifton Forlines, Daniel Wigdor, Chia Shen, and Ravin Balakrishnan. 2007. Direct-touch vs. Mouse Input for Tabletop Displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '07)*. ACM, New York, NY, USA, 647–656. https://doi.org/10.1145/1240624.1240726
[19] ForwardXP. 2016. Please, Don't Touch Anything. http://www.forwardxp.com/ (version 2.0.1).
[20] Brigitte N. Frederick. 1999. Fixed-, random-, and mixed-effects ANOVA models: A user-friendly guide for increasing the generalizability of ANOVA results. ed. (1999), 111–122. http://eric.ed.gov/?id=ED426098
[21] Milton Friedman. 1937. The Use of Ranks to Avoid the Assumption of Normality Implicit in the Analysis of Variance. *J. Amer. Statist. Assoc.* 32, 200 (1937), 675–701. http://www.jstor.org/stable/2279372
[22] Google. 2016. Tilt Brush. https://www.tiltbrush.com/ (version 14.1).
[23] Google. 2017. Blocks. https://vr.google.com/blocks/ (version 1.0.5).
[24] Google. 2017. Cardboard. https://vr.google.com/cardboard/ Accessed on 9-18-2017.
[25] Google. 2017. Daydream View. https://vr.google.com/daydream/ Accessed on 9-18-2017.
[26] Gravity Sketch. 2017. Gravity Sketch VR. https://www.gravitysketch.com/ Accessed on 9-18-2017.
[27] Tovi Grossman and Daniel Wigdor. 2007. Going Deeper: a Taxonomy of 3D on the Tabletop. In *Horizontal Interactive Human-Computer Systems, 2007. TABLETOP '07. Second Annual IEEE International Workshop on*. 137–144. https://doi.org/10.1109/TABLETOP.2007.18
[28] Jan Gugenheimer, David Dobbelstein, Christian Winkler, Gabriel Haas, and Enrico Rukzio. 2016. FaceTouch: Enabling Touch Interaction in Display Fixed UIs for Mobile Virtual Reality. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology (UIST '16)*. ACM, New York, NY, USA, 49–60. https://doi.org/10.1145/2984511.2984576
[29] Guy Godin. 2016. Virtual Desktop. http://www.vrdesktop.net/ (version 1.5.1).
[30] Martin Hachet, Benoit Bossavit, Aurélie Cohé, and Jean-Baptiste de la Rivière. 2011. Toucheo: Multitouch and Stereo Combined in a Seamless Workspace. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology (UIST '11)*. ACM, New York, NY, USA, 587–592. https://doi.org/10.1145/2047196.2047273
[31] Chris Hand. 1997. A Survey of 3D Interaction Techniques. *Computer Graphics Forum* 16, 5 (1997), 269–281. https://doi.org/10.1111/1467-8659.00194
[32] Jeffrey T. Hansberger, Chao Peng, Shannon L. Mathis, Vaidyanath Areyur Shanthakumar, Sarah C. Meacham, Lizhou Cao, and Victoria R. Blakely. 2017. Dispelling the Gorilla Arm Syndrome: The Viability of Prolonged Gesture Interactions. In *Virtual, Augmented and Mixed Reality*, Stephanie Lackey and Jessie Chen (Eds.). Springer International Publishing, Cham, 505–520. https://doi.org/10.1007/978-3-319-57987-0_41
[33] Hard Light Labs. 2017. SoundStage. http://www.soundstagevr.com/
[34] Alex Hill and Andrew Johnson. 2008. Withindows: A Framework for Transitional Desktop and Immersive User Interfaces. In *2008 IEEE Symposium on 3D User Interfaces*. 3–10. https://doi.org/10.1109/3DUI.2008.4476584
[35] Ken Hinckley, Randy Pausch, John C. Goble, and Neal F. Kassell. 1994. Passive Real-world Interface Props for Neurosurgical Visualization. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '94)*. ACM, New York, NY, USA, 452–458. https://doi.org/10.1145/191666.191821
[36] Ken Hinckley, Randy Pausch, John C. Goble, and Neal F. Kassell. 1994. A Survey of Design Issues in Spatial Input. In *Proceedings of the 7th Annual ACM Symposium on User Interface Software and Technology (UIST '94)*. ACM, New York, NY, USA, 213–222. https://doi.org/10.1145/192426.192501
[37] Sture Holm. 1979. A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics* 6, 2 (1979), 65–70. http://www.jstor.org/stable/4615733
[38] Suleiman Tashtoush James J. Higgins. 1994. An aligned rank transform test for interaction. 2 (1994), 201–211.
[39] Jacek Jankowski and Martin Hachet. 2013. A Survey of Interaction Techniques for Interactive 3D Environments. In *Eurographics 2013 - State of the Art Reports*, M. Sbert and L. Szirmay-Kalos (Eds.). The Eurographics Association. https://doi.org/10.2312/conf/EG2013/stars/065-093
[40] JauntVR. 2017. Jaunt VR. https://www.jauntvr.com/ (version 1.30.0).
[41] Seokhee Jeon, Jane Hwang, Gerard J. Kim, and Mark Billinghurst. 2006. Interaction Techniques in Large Display Environments Using Hand-held Devices. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology (VRST '06)*. ACM, New York, NY, USA, 100–103. https://doi.org/10.1145/1180495.1180516
[42] Hao Jiang, Eyal Ofek, Neema Moraveji, and Yuanchun Shi. 2006. Direct Pointer: Direct Manipulation for Large-display Interaction Using Handheld Cameras. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '06)*. ACM, New York, NY, USA, 1107–1110. https://doi.org/10.1145/1124772.1124937
[43] Ricardo Jota, Miguel A. Nacenta, Joaquim A. Jorge, Sheelagh Carpendale, and Saul Greenberg. 2010. A Comparison of Ray Pointing Techniques for Very Large Displays. In *Proceedings of Graphics Interface 2010 (GI '10)*. Canadian Information Processing Society, Toronto, Ont., Canada, Canada, 269–276. http://dl.acm.org/citation.cfm?id=1839214.1839261
[44] Ricardo Jota, João M. Pereira, and Joaquim A. Jorge. 2009. A Comparative Study of Interaction Metaphors for Large-scale Displays. In *CHI '09 Extended Abstracts on Human Factors in Computing Systems (CHI EA '09)*. ACM, New York, NY, USA, 4135–4140. https://doi.org/10.1145/1520340.1520629
[45] Regis Kopper, Felipe Bacim, and Doug A. Bowman. 2011. Rapid and accurate 3D selection by progressive refinement. In *IEEE Symposium on 3D User Interfaces (3DUI)*. 67–74. https://doi.org/10.1109/3DUI.2011.5759219
[46] André Kunert, Alexander Kulik, Christopher Lux, and Bernd Fröhlich. 2009. Facilitating System Control in Ray-based Interaction Tasks. In *Proceedings of the 16th ACM Symposium on Virtual Reality Software and Technology (VRST '09)*. ACM, New York, NY, USA, 183–186. https://doi.org/10.1145/1643928.1643969
[47] Paul Lapides, Ehud Sharlin, Mario C. Sousa, and Lisa Streit. 2006. The 3D Tractus: a three-dimensional drawing board. In *First IEEE International Workshop on Horizontal Interactive Human-Computer Systems (TABLETOP '06)*. 169–176. https://doi.org/10.1109/TABLETOP.2006.33

[48] Leap Motion. 2017. Leap Motion. https://www.leapmotion.com/ Accessed on 9-18-2017.
[49] Hyeongmook Lee, Dongchul Kim, and Woontack Woo. 2011. Graphical Menus Using a Mobile Phone for Wearable AR Systems. In *2011 International Symposium on Ubiquitous Virtual Reality*. 55–58. https://doi.org/10.1109/ISUVR.2011.23
[50] James R. Lewis. 1995. IBM Computer Usability Satisfaction Questionnaires: Psychometric Evaluation and Instructions for Use. *Int. J. Hum.-Comput. Interact.* 7, 1 (Jan. 1995), 57–78. https://doi.org/10.1080/10447319509526110
[51] Wanhong Lin, Lear Du, Carisa Harris-Adamson, Alan Barr, and David Rempel. 2017. *Design of Hand Gestures for Manipulating Objects in Virtual Reality*. Springer International Publishing, Cham, 584–592. https://doi.org/10.1007/978-3-319-58071-5_44
[52] Robert W. Lindeman, John L. Sibert, and James K. Hahn. 1999. Towards Usable VR: An Empirical Study of User Interfaces for Immersive Virtual Environments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '99)*. ACM, New York, NY, USA, 64–71. https://doi.org/10.1145/302979.302995
[53] Ramon C. Littell, P R Henry, and Clarence B. Ammerman. 1998. Statistical analysis of repeated measures data using SAS procedures. *Journal of animal science* 76 (1998), 1216–1231. https://doi.org/10.2527/1998.7641216x
[54] Paul Lubos. 2014. Touching the Third Dimension. In *Proceedings of the Ninth ACM International Conference on Interactive Tabletops and Surfaces (ITS '14)*. ACM, New York, NY, USA, 473–478. https://doi.org/10.1145/2669485.2669490
[55] Paul Lubos, Gerd Bruder, and Frank Steinicke. 2014. Analysis of direct selection in head-mounted display environments. In *2014 IEEE Symposium on 3D User Interfaces (3DUI)*. 11–18. https://doi.org/10.1109/3DUI.2014.6798834
[56] Paul Lubos, Gerd Bruder, and Frank Steinicke. 2015. Influence of Comfort on 3D Selection Task Performance in Immersive Desktop Setups. *Journal of Virtual Reality and Broadcasting* 12(2015), 2 (2015). https://doi.org/10.20385/1860-2037/12.2015.2
[57] Hossein Mansouri, Robert L. Paige, and James G. Surles. 2004. Aligned Rank Transform Techniques for Analysis of Variance and Multiple Comparisons. *Communications in Statistics - Theory and Methods* 33, 9 (2004), 2217–2232. https://doi.org/10.1081/STA-200026599 arXiv:http://dx.doi.org/10.1081/STA-200026599
[58] Chris Mckenzie and Adam Glazier. 2017. Google I/O 2017 - Designing Screen Interfaces for VR. https://www.youtube.com/watch?v=ES9jArHRFHQ
[59] Microsoft. 2017. HoloLens. https://www.microsoft.com/en-us/hololens Accessed on 9-18-2017.
[60] Brad A. Myers, Rishi Bhatnagar, Jeffrey Nichols, Choon Hong Peck, Dave Kong, Robert Miller, and A. Chris Long. 2002. Interacting at a Distance: Measuring the Performance of Laser Pointers and Other Devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '02)*. ACM, New York, NY, USA, 33–40. https://doi.org/10.1145/503376.503383
[61] John R. Napier. 1956. THE PREHENSILE MOVEMENTS OF THE HUMAN HAND. *Bone & Joint Journal* 38-B, 4 (1956), 902–913. arXiv:http://bjj.boneandjoint.org.uk/content/38-B/4/902.full.pdf http://bjj.boneandjoint.org.uk/content/38-B/4/902
[62] Cuong Nguyen, Stephen DiVerdi, Aaron Hertzmann, and Feng Liu. 2017. Vremiere: In-Headset Virtual Reality Video Editing. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 5428–5438. https://doi.org/10.1145/3025453.3025675
[63] Oculus. 2017. Medium. https://www.oculus.com/medium/ (version 1.2.2).
[64] Oculus. 2017. Oculus Rift. https://www.oculus.com/ Accessed on 9-18-2017.
[65] Oculus. 2017. Quill. https://www.oculus.com/story-studio/quill/ (version 1.2).
[66] Owlchemy Labs. 2017. Rick and Morty: Virtual Rick-ality. http://www.adultswim.com/games/pc-console/rick-and-morty-simulator-virtual-rick-ality/ (version 1.0.10).
[67] Thammathip Piumsomboon, Gun Lee, Robert W. Lindeman, and Mark Billinghurst. 2017. Exploring natural eye-gaze-based interaction for immersive virtual reality. In *2017 IEEE Symposium on 3D User Interfaces (3DUI)*. 36–39. https://doi.org/10.1109/3DUI.2017.7893315
[68] Ivan Poupyrev, Mark Billinghurst, Suzanne Weghorst, and Tadao Ichikawa. 1996. The Go-go Interaction Technique: Non-linear Mapping for Direct Manipulation in VR. In *Proceedings of the 9th Annual ACM Symposium on User Interface Software and Technology (UIST '96)*. ACM, New York, NY, USA, 79–80. https://doi.org/10.1145/237091.237102
[69] Ivan Poupyrev, Tadao Ichikawa, Suzanne Weghorst, and Mark Billinghurst. 1998. Egocentric Object Manipulation in Virtual Environments: Empirical Evaluation of Interaction Techniques. *Computer Graphics Forum* 17, 3 (1998), 41–52. https://doi.org/10.1111/1467-8659.00252
[70] Ivan Poupyrev, Suzanne Weghorst, Mark Billinghurst, and Tadao Ichikawa. 1997. A Framework and Testbed for Studying Manipulation Techniques for Immersive VR. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology (VRST '97)*. ACM, New York, NY, USA, 21–28. https://doi.org/10.1145/261135.261141
[71] Wendy Powell, Vaughan Powell, Phillip Brown, Marc Cook, and Jahangir Uddin. 2016. *Influence of navigation interaction technique on perception and behaviour in mobile virtual reality*. The University of Reading, 73–81.
[72] Ready At Dawn. 2017. Lone Echo. https://www.echo.games/ (version 3.121).
[73] Schell Games. 2017. I Expect You To Die. https://iexpectyoutodie.schellgames.com/ (version 1.0.108744).
[74] Leila Schemali and Elmar Eisemann. 2014. Design and evaluation of mouse cursors in a stereoscopic desktop environment. In *2014 IEEE Symposium on 3D User Interfaces (3DUI)*. 67–70. https://doi.org/10.1109/3DUI.2014.6798844
[75] Samuel S. Shapiro and Martin B. Wilk. 1965. An Analysis of Variance Test for Normality (Complete Samples). *Biometrika* 52, 3/4 (1965), 591–611. http://www.jstor.org/stable/2333709
[76] Ben Shneiderman, Catherine Plaisant, Maxine Cohen, and Steven Jacobs. 2009. *Designing the User Interface: Strategies for Effective Human-Computer Interaction* (5th ed.). Addison-Wesley Publishing Company, USA.
[77] Garth Shoemaker, Anthony Tang, and Kellogg S. Booth. 2007. Shadow Reaching: A New Perspective on Interaction for Large Displays. In *Proceedings of the 20th Annual ACM Symposium on User Interface Software and Technology (UIST '07)*. ACM, New York, NY, USA, 53–56. https://doi.org/10.1145/1294211.1294221
[78] Frank Steinicke, Hrvoje Benko, Antonio Krüger, Daniel Keefe, Jean-Baptiste de la Riviére, Ken Anderson, Jonna Häkkilä, Leena Arhippainen, and Minna Pakanen. 2012. The 3rd Dimension of CHI (3DCHI): Touching and Designing 3D User Interfaces. In *CHI '12 Extended Abstracts on Human Factors in Computing Systems (CHI EA '12)*. ACM, New York, NY, USA, 2695–2698. https://doi.org/10.1145/2212776.2212698
[79] Frank Steinicke, Timo Ropinski, and Klaus Hinrichs. 2006. *OBJECT SELECTION IN VIRTUAL ENVIRONMENTS USING AN IMPROVED VIRTUAL POINTER METAPHOR*. Springer Netherlands, Dordrecht, 320–326. https://doi.org/10.1007/1-4020-4179-9_46
[80] STW Software. 2016. Gala 360. https://gala360site.wordpress.com/ (version 1.2.3).
[81] Zsolt Szalavári and Michael Gervautz. 1997. The Personal Interaction Panel – a Two-Handed Interface for Augmented Reality. *Computer Graphics Forum* 16, 3 (1997), C335–C346. https://doi.org/10.1111/1467-8659.00137
[82] Jonathan Taylor, Lucas Bordeaux, Thomas Cashman, Bob Corish, Cem Keskin, Toby Sharp, Eduardo Soto, David Sweeney, Julien Valentin, Benjamin Luff, Arran Topalian, Erroll Wood, Sameh Khamis, Pushmeet Kohli, Shahram Izadi, Richard Banks, Andrew Fitzgibbon, and Jamie Shotton. 2016. Efficient and Precise Interactive Hand Tracking Through Joint, Continuous Optimization of Pose and Correspondences. *ACM Trans. Graph.* 35, 4, Article 143 (July 2016), 12 pages. https://doi.org/10.1145/2897824.2925965
[83] The New York Times Company. 2017. NYT VR. http://www.nytimes.com/marketing/nytvr/ (version 2.2.0).
[84] John W. Tukey. 1949. Comparing Individual Means in the Analysis of Variance. *Biometrics* 5, 2 (1949), 99–114. http://www.jstor.org/stable/3001913
[85] University of Southern California E2C2 Team. 2018. Blueshark. http://e2c2.ict.usc.edu
[86] Dimitar Valkov, Frank Steinicke, Gerd Bruder, and Klaus Hinrichs. 2011. 2D Touching of 3D Stereoscopic Objects. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. ACM, New York, NY, USA, 1353–1362. https://doi.org/10.1145/1978942.1979142
[87] Jacob O. Wobbrock, Leah Findlater, Darren Gergle, and James J. Higgins. 2011. The Aligned Rank Transform for Nonparametric Factorial Analyses Using Only Anova Procedures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. ACM, New York, NY, USA, 143–146. https://doi.org/10.1145/1978942.1978963
[88] Po-Chen Wu, Robert Wang, Kenrick Kin, Christopher Twigg, Shangchen Han, Ming-Hsuan Yang, and Shao-Yi Chien. 2017. DodecaPen: Accurate 6DoF Tracking of a Passive Stylus. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology (UIST '17)*. ACM, New York, NY, USA, 365–374. https://doi.org/10.1145/3126594.3126664
[89] ByungIn Yoo, Jae-Joon Han, Changkyu Choi, Kwonju Yi, Sungjoo Suh, Dusik Park, and Changyeong Kim. 2010. 3D User Interface Combining Gaze and Hand Gestures for Large-scale Display. In *CHI '10 Extended Abstracts on Human Factors in Computing Systems (CHI EA '10)*. ACM, New York, NY, USA, 3709–3714. https://doi.org/10.1145/1753846.1754043
[90] Chun Yu, Yizheng Gu, Zhican Yang, Xin Yi, Hengliang Luo, and Yuanchun Shi. 2017. Tap, Dwell or Gesture?: Exploring Head-Based Text Entry Techniques for HMDs. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 4479–4488. https://doi.org/10.1145/3025453.3025964