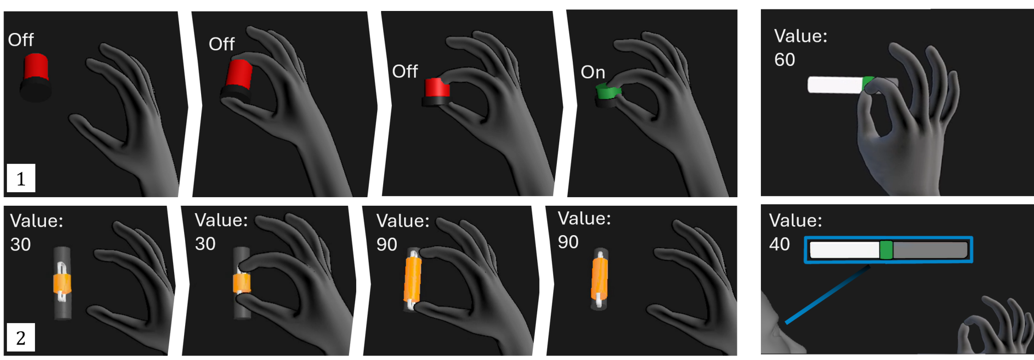
A Comparative Investigation of Continuous Thumb-Index Pinching for Hand-Based Input in Extended Reality

MARTIN SCHMITZ, University of Koblenz, Germany
SIMON SEIBERT, Saarland University, Saarland Informatics Campus, Germany
JÜRGEN STEIMLE, Saarland University, Saarland Informatics Campus, Germany
THOMAS KOSCH, HU Berlin, Germany



(a) Continuous thumb-index pinching
(b) Direct pinch and gaze+pinch

Fig. 1. We assess continuous thumb–index pinching for hand-based XR input by evaluating (A1) a pinch button, similar to a traditional pinch-to-click button, that activates on self-contact when pinching thumb and index together, and (A2) a pinchable slider controlled by grasping its two handles with thumb and index, pinching to extend or compress to the desired value, then releasing by moving off-axis. (B) We compare to established hand-based XR inputs (direct pinch/touch and gaze- or hand raycast) and derive guidelines.

Hand-based input is widely used for interaction in Extended Reality (XR). While the most established input techniques, direct-touch and pinch-to-click, have been extensively explored and combined with gaze- and hand-ray casting, previous research has proposed more continuous pinching gestures to set a greater range of values by adjusting the distance between the thumb and index finger. While promising, prior studies have not systematically assessed the factors contributing to these input techniques and their impact on usability and efficiency relative to established techniques. We address this gap by systematically combining and comparing selection (direct, by gaze, and by ray-casting) and manipulation methods (touch, pinch-to-click, and continuous pinching) to understand effective hand-based input techniques in XR. To this end, we transfer the concept of continuous thumb-index pinching to XR and present the results of a controlled experiment comparing it with established hand-based XR input.

CCS Concepts: • **Human-centered computing** → **Human computer interaction (HCI)**; **User studies**.

Additional Key Words and Phrases: Input; Pinching; Direct Touch; Extended Reality; Eye Gaze; User Studies

Authors' Contact Information: Martin Schmitz, University of Koblenz, Koblenz, Germany, martin@uni-koblenz.de; Simon Seibert, Saarland University, Saarland Informatics Campus, Saarbrücken, Germany, s8siseib@stud.uni-saarland.de; Jürgen Steimle, Saarland University, Saarland Informatics Campus, Saarbrücken, Germany, steimle@cs.uni-saarland.de; Thomas Kosch, thomas.kosch@hu-berlin.de, HU Berlin, Berlin, Germany.

ACM 2573-0142/2026/8-ARTMHCI4531
https://doi.org/10.1145/3821665

ACM Reference Format:
Martin Schmitz, Simon Seibert, Jürgen Steimle, and Thomas Kosch. 2026. A Comparative Investigation of Continuous Thumb-Index Pinching for Hand-Based Input in Extended Reality. *Proc. ACM Hum.-Comput. Interact.* 10, 5, Article MHCI4531 (August 2026), 26 pages. https://doi.org/10.1145/3821665

1 Introduction

In Extended Reality (XR), a body of work has explored hand-based input methods, such as direct touch [56], pinch-to-click [47], and gaze-supported interactions [58]. In addition to these established input methods, research proposed approaches that make increasing use of the dexterity of the human hand, ranging from on-finger [29] and thumb-to-finger interactions [14, 15, 29, 69] to micro- [8, 66] and two-handed gestures [43, 53]. Among these, recent work [61] has demonstrated that continuous thumb-index pinching, where the distance between the thumb and index finger serves as a continuous input, is a promising, general hand-based input method that enables millimeter-level granularity. Yet, they have only investigated thumb-index pinching in a distant setup with a linear scale on a display. Thus, it remains unclear how this promising method translates to XR input and how well it performs compared to established XR input methods.

Existing work on pinching and hand-based XR input addresses only parts of this problem space. Prior XR research has mainly treated pinch-to-click as a binary trigger for selection or manipulation, often in combination with direct touch, gaze, or raycasting, but without exploiting intermediate thumb-index states. Conversely, continuous thumb-index pinching has been studied on physical or on-screen sliders in non-XR settings, focusing on tactile feedback and feasibility rather than spatial interaction or comparison to established XR techniques. As a result, we lack (i) XR widgets that make continuous pinching usable in typical button and slider controls, and (ii) a systematic comparison that orthogonally varies selection and manipulation methods to understand when continuous pinching is preferable to standard XR input techniques.

This paper compares different methods for hand-based XR input. To this end, we first transfer the display-based, indirect thumb-index pinching from [61] to spatial XR controls. Compared to previous work, this allows thumb-index pinching to be used either directly in situ or indirectly via gaze selection or hand raycasting. We focus on the most common UI controls, buttons, and sliders (see Figure 1a): While pinchable buttons are mostly equivalent to a traditional pinch-to-click button, only serving for baseline comparison in our study, the pinchable sliders operate as follows: inspired by the metaphor of compressing or extending a spring, users control them by grasping two handles with their thumb and index, pinching to extend or compress to the desired value, then releasing by moving off-axis. We then systematically compare these controls with established hand-based XR input (see Figure 1b) by examining three *selection* methods – direct touch, gaze, and hand raycasting – in combination with three *manipulation* methods – standard pinch, extended index finger, and continuous thumb-index pinch.

We found that gaze-based selection paired with a standard pinch or a continuous thumb-index pinch is more comfortable, faster, and yields lower errors and a lower workload compared to hand raycast interactions, while direct finger touch excels for near buttons. In addition, hand raycasting, especially when combined with finger manipulations, underperforms compared to other methods in terms of time, workload, and comfort. Our findings offer practical guidance: (1) designers of hand-based XR interfaces should opt for a combination of gaze with a pinch manipulation for distant or precision-sensitive controls, (2) keep direct finger touch for proximal discrete actions, and (3) deprioritize hand raycast as a common selection method.

In summary, this paper makes three contributions:

(1) *XR-adapted continuous pinching input:* We introduce pinchable buttons and sliders that map continuous thumb-index distance to discrete and continuous XR controls and can be used both in situ and at a distance (via gaze or hand raycast).
(2) *A systematic* 3×3 *comparison of hand-based XR input techniques:* In a controlled experiment with 18 participants, we orthogonally vary selection (direct, gaze, hand raycast) and manipulation (finger, pinch, continuous pinch) for both buttons and sliders.
(3) *Design guidelines for hand-based XR input:* Based on accuracy, time, workload, usability, comfort, and occlusion, we derive guidelines that show when to prefer gaze-plus-pinch for distant or precision-sensitive controls, when to keep direct finger touch for near buttons, and when to deprioritize hand raycast as a selection method.

2 Related Work

Our work builds on previous research on hand-based gestures, direct touch, pinching interaction, and indirect interaction using raycasting and eye gaze.

2.1 Hand-Based Gestures

Hand-based gestures are a widely used interaction modality in XR, enabling intuitive, immersive user experiences without requiring participants to hold additional devices. Hence, they have been used for a wide range of settings, including surface computing [44, 78, 80], televisions [12, 83], and automotive interfaces [57]. Hand-based gestures have also been evaluated extensively as an interaction modality in XR [15, 53, 62, 70]. Specifically, mid-air hand-based gestures have been proposed to enable interaction with virtual elements in XR [73]. The design space of mid-air gestures is large, and many variations have been presented, ranging from shortcuts for tool selection and usage in a 3D interaction. [45, 70], and hand-based pointing selection techniques specifically targeting target selection in dense VR environments [67] to a condensed set of mid-air gestures [28] and investigations of their social implications [33, 41].

While previous work conducted further research into variations of mid-air gestures, including two-handed gestures [43, 53] and microgestures [8, 66], research has also started to investigate thumb-to-finger gestures [14, 15, 29, 69], utilizing the dexterity of the human hand.

2.2 Touch and Pinching

A simple method for manipulating UI elements in space is using a single finger, often called *direct touch*. This technique is widely employed on touch-enabled devices for interacting with interface elements [65], text entry [10, 19, 20], and performing gestures [6, 21]. Past research has also investigated ways to distinguish between different fingers [11] and finger elements [24]. This approach is utilized in XR for mid-air interactions [35, 76]. Pinching gestures refer to thumb-to-finger interactions in which the user varies the distance between the thumb and one or more fingers, typically the index finger. It is particularly well recognizable by common capturing technologies. Originating in early touch-based interfaces [2, 36, 78], pinching gestures, often associated with zooming in or out on a touchscreen, were popularized with the rise of multitouch devices following the introduction of the first iPhone. Since then, researchers used pinching in navigation [7, 23, 59, 60], item selection and manipulation [50, 76], fabric interactions [22, 32, 52] and in XR [47]. It has been demonstrated that pinching is an intuitive [2] and comfortable [16, 29] gesture to perform. In XR, pinching is frequently combined with hand movements to expand input possibilities, such as for sliding [76], scrolling [23], panning [60], text input [31], precise selection via magnification [85], or performing two-handed 3D manipulations [7]. Pinching gestures can also be paired with selection based on raycasting or gaze detection for interaction [5, 34, 55]. In today's XR systems, pinching is typically used as an interaction similar to a mouse click, where the system only recognizes whether the pinch is closed (i.e., thumb and index finger touching) or open (i.e., thumb and index finger apart) without accounting for intermediate states.

Recent work [61] demonstrated that continuous thumb-index pinching is a promising, generalized input technique, offering a fine-grained input space between thumb and index finger on a linear on-screen scale, and explores the role of tactile feedback for this modality. However, their study is conducted outside XR, uses a single eyes-free 1D task, and does not compare with other hand-based XR techniques. In contrast, we treat continuous thumb–index pinching as a hand-based XR input method that must compete and combine with direct touch, binary pinch-to-click, and gaze- or ray-based selection. Our work, therefore, focuses on (i) integrating continuous pinching into spatial XR widgets (buttons and sliders) and (ii) systematically quantifying its performance and user experience relative to established XR techniques.

2.3 Distant Input using Raycasting and Eye Gaze

Raycasting was first introduced by Mine et al. [42] as a method to overcome the challenge of interacting with objects beyond reach in large XR environments constrained by physical limits. Raycasting involves projecting a visual or invisible beam from the user's hand [4, 5, 42], finger [1], or controller [18, 39, 51, 77], which dynamically adapts to changes in orientation and position. Virtual objects that intersect with the ray cast are then marked for selection. Previous researchers confirmed raycasting as intuitive [47, 71], efficient [77], and demands minimal physical effort [4]. However, raycasting is affected by imprecision, mainly due to jitter [55] and the Heisenberg effect [5], which can occur when the same device or body part is used for both aiming the ray and confirming the selection. These issues become more pronounced as the distance to the target increases. Further work regarding confirming ray cast selection using button presses [18, 39] or gestures [1, 55] has been presented in the past to modify discrete (e.g., button press) and continuous values (e.g., slider values). Previous research also investigated the use of gaze input [9, 30, 42, 55], where eye gaze is employed to select targets and hand gestures such as pinching to confirm a selection to avoid the *Midas Touch Problem* [30, 55]. Gaze-based interaction can be quick, require little physical effort [55, 68, 75, 84], and can mitigate the *Gorilla Arm Effect* by allowing the arm to rest on the lap or hang naturally while standing. However, it is constrained by human visual capabilities. Accurately pointing at objects, especially small or distant ones, can be challenging when using eye tracking [55]. This difficulty can be addressed through target snapping techniques [54, 64] or gaze-hand alignment [74]. Compared to hand-based raycasting, gaze-based pointing methods are less effective in speed, ease, and controllability [51]. In summary, prior work has explored (a) direct mid-air gestures and touch-like input in XR, (b) binary pinch-to-click combined with gaze or hand raycasting, and (c) continuous thumb–index pinching on non-XR displays. However, these lines of work have not been brought together: existing studies do not systematically disentangle selection (direct, gaze, raycast) from manipulation (finger, binary pinch, continuous pinch) for common XR widgets, nor do they assess continuous thumb-index pinching as one option within that broader design space. This motivates our experimental comparison of nine hand-based input techniques and the resulting design guidelines for XR interfaces.

3 Transferring Continuous Thumb-Index Pinching to Extended Reality

This paper compares the performance of continuous thumb–index pinching relative to established hand-based XR input techniques. Because prior work examined this gesture only on a linear on-screen scale (cf. [61]), we first need to transfer this interaction technique to Extended Reality (XR). Specifically, we designed a spatially deployable visual UI control that enables principled comparison across input methods. Beyond serving as experimental apparatus, these controls are generic XR widgets: They require only commodity hand tracking, can be instantiated at arbitrary positions in 3D space, and support both in-situ and distant interaction. We therefore present the pinchable button and slider as reusable building blocks for XR interfaces that wish to exploit continuous thumb-index input without additional hardware.

We intentionally focus on buttons and sliders, which are traditional UI elements that have also been widely adopted in XR interface design. Their prevalence stems from their simplicity and directness: buttons enable discrete actions, while sliders allow continuous adjustments (e.g., volume, brightness). Given their foundational role in interaction design, we focus this paper on these UI primitives that can be integrated into more complex XR user interfaces. In the following, we explain the functionality of the pinchable button and the pinchable slider.

3.1 Pinchable Button

This button is mostly equivalent to conventional discrete buttons activated via a standard pinch-to-click. Yet, it necessitates a complete, more deliberate pinching motion between two fingers to activate (see Figure 1a1). This adds a more intentional physical action, reducing accidental activations and ensuring that users engage more consciously with the interface. We included the button to serve as a baseline condition to support comparative evaluation across different selection and manipulation techniques.

More specifically, the pinchable button operates as follows. As the user approaches the button in the spatial interface with their thumb and index finger pinched apart, the button snaps to the relative distance of their fingertips. The user can gradually decrease the distance between their fingers to move the button toward its "on" state. Once the pinching distance crosses a predefined threshold, the button is activated, indicated by a color change.

3.2 Pinchable Slider

This slider modifies a traditional 1D slider to utilize continuous thumb-index pinching (see Figure 1a2). We designed it as a cylindrical element with handles on either side, resembling the handles of a pair of scissors that people grab to pinch them together or move them apart.

The pinchable slider operates as follows: the user approaches the slider in the interface and grasps both handles with their fingers. The slider value is adjusted by pinching the handles closer together or pulling them apart, similar to compressing or stretching a spring. The distance between the thumb and index finger dynamically determines the value, making it a one-degree-of-freedom system. This means that both fingers were tracked independently, and the slider handles were dynamically anchored to their positions whenever the user approached them. Once the desired value is set, the user releases the slider perpendicular to its orientation.

Additionally, the pinchable slider supports intuitive manipulation beyond simple value adjustments, reducing the user's need to keep their hand stationary or in an awkward posture: The handles can be grabbed to reposition or rotate the slider, enabling users to adjust its orientation to a more comfortable hand position. For this, the entire slider is always aligned to the midpoint between the thumb and index, keeping it symmetric at all times, preventing either side from extending farther than the other, and avoiding direction-based asymmetries.

3.3 Extending Both to Distant Interactions

Both pinchable buttons and sliders are not confined to direct interactions in which the user engages with the UI element in place. It can be readily adapted for indirect, distant interaction styles, that, for instance, employ gaze- or hand-based ray casting. In these cases, pinchable buttons and sliders can still be controlled with the same thumb-index pinching gesture. To further assist the user, virtual representations of their fingers can be incorporated directly within the UI, providing real-time guidance during interaction. The following user study explicitly investigates how both controls can be used directly and indirectly, via gaze or hand raycast selection.

4 Methodology

The main contribution of this paper is to systematically evaluate the design space for hand-based XR input, offering insights between and beyond currently established techniques. We therefore performed a controlled experiment investigating the accuracy, efficiency, and usability of different hand-based XR input techniques that are composed of one means to (1) select and then (2) manipulate a UI control. We defined the following research questions:

RQ1 How does the SELECTION technique of a button or slider affect the accuracy, efficiency, and usability?

RQ2 How does the MANIPULATION technique of a button or slider affect the accuracy, efficiency, and usability?

RQ3 Are there interaction effects between these factors concerning accuracy, efficiency, and usability?

4.1 Task

We employ a visual target acquisition and selection task (similar to [17, 61]): After a countdown, either a button or a slider appeared in the participant's frontal view. The participant was then asked to either press a button or set a slider to a highlighted yellow target mark (see Figure 2b). Once the participants believed they had pressed the button or set the slider correctly to the target mark, they pressed an external clicker to finish the task.

4.2 Design

4.2.1 Independent Variables. To gain a comprehensive understanding of the factors that influence the accuracy, efficiency, and usability of hand-based input in XR, we varied two independent variables (see Figure 2a):

SELECTION. How users select a UI element in XR is a significant factor influencing the overall interaction. We therefore consider three representative selection methods with known strengths and weaknesses to accomplish a complete space. At the level *direct*, participants moved their hand directly to the spatial location of the UI control for interaction. Direct selection is precise and familiar from touch-based interaction and controller-free hand tracking in VR, but is negatively affected by arm fatigue and occlusion when used in mid-air or at greater distances [27, 40]. In the level *gaze*, participants distantly selected the UI element to manipulate via their gaze. Gaze-based selection has been shown to be fast and low-effort and can mitigate the Gorilla Arm effect by allowing the arm to rest [68, 72, 84], but its accuracy is limited for small or distant targets and often requires snapping or filtering [47, 51, 55]. In the *raycast* level, participants used a ray emanating from their hands, determined by the orientation of their palms, to point at a distant UI element for selection. Hand-based raycasting is intuitive, efficient, and widely adopted as a far-field selection technique in VR [3, 51, 72], yet it is affected by jitter and errors such as the Heisenberg effect when the same hand is used for aiming and confirming [81]. Together, they cover the main spectrum of near- and far-field hand-based selection in XR.

MANIPULATION. After selecting a UI control, manipulating its state is another influential factor for XR interaction. Again, we informed our levels by the most commonly used techniques from related work and tested them against our continuous pinching technique. In the level *finger*, participants

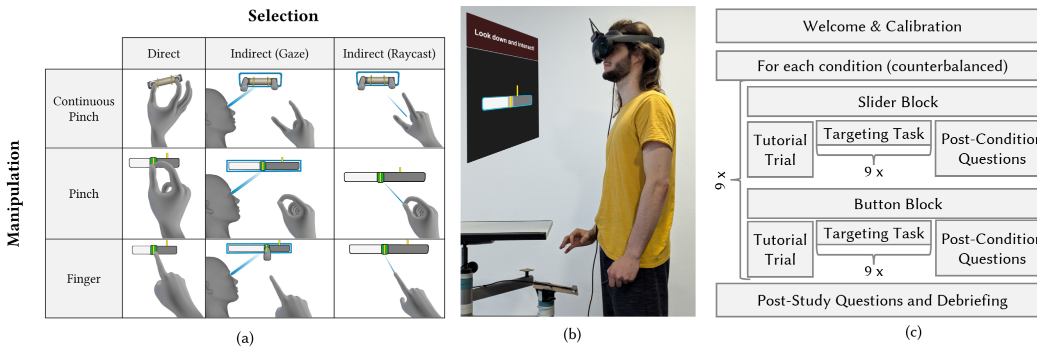


Fig. 2. (a) By varying the independent variables SELECTION and MANIPULATION with three levels each, we investigate a total of nine different hand-based input techniques in terms of accuracy, efficiency, and usability. (b) Users were tasked to press buttons or align a slider to a yellow target mark. Our setup combines two Leap Motion 2 controllers (at the HMD and below the hand) for precise hand tracking, also at the outer peripherals. (c) The study procedure shows the process of how participants performed slider and button tasks.

manipulated the UI control using their extended index finger. Direct finger manipulation is a well-established baseline on touch surfaces and in XR, providing high spatial precision and immediate feedback [40, 69], but it increases visual occlusion of the target and can be physically demanding when used in mid-air [27]. In the level *pinch*, participants pinched their thumb and index finger together to grab the UI control and moved their hand to manipulate it, adopting a mouse-style click and movement interaction (often referred to as pinch-to-click). Binary pinch is the dominant hand-tracking equivalent of a trigger in commodity XR systems and enables low-effort discrete control [3, 47, 55], yet it only provides on/off input and can be affected by unintentional activations or repeated pinching fatigue [48, 49]. Finally, the level *continuous-pinch* consists of techniques that employ continuous thumb-index pinching. Building on [61], it offers a fine-grained 1D input, but has so far only been explored on a linear on-screen scale outside XR without comparison.

Rationale. We chose the selection and manipulation techniques to be representative of common hand-based XR interfaces and to span known trade-offs in precision, effort, reach, and occlusion (as outlined above). With our comparative evaluation, we aim to clarify these trade-offs empirically and provide more informed design guidance for hand-based XR interfaces. We intentionally divided between selection and manipulation techniques to also cover common interaction styles in commercially available XR devices: The combination of DIRECT and FINGER defines the direct-touch interaction style, in which a finger is positioned at the spatial location of a UI control. Similarly, combining PINCH with either DIRECT or RAYCAST replicates the standard interaction techniques used by many XR interfaces (such as the Microsoft HoloLens and Meta Quest). Lately, the combination GAZE and PINCH, presented first by Pfeuffer et al. [55], received a lot of attention with the release of the Apple Vision Pro, as it promises to reduce the Gorilla Arm Syndrome. By combining RAYCAST with PINCH, our study design also incorporates an indirect approach, similar to that employed in Meta's HMDs, which uses ray-casting originating from the hand.

We combine selection and manipulation techniques within-subjects for a total of $3 \times 3 = 9$ conditions that were repeated nine times (see Figure 2c). We counterbalanced their order in a Balanced Latin Square ($3 \times 3 = 9$ levels, repeated twice for 18 participants). To avoid frequent switching between UI elements that could confuse participants, we presented repetitions in blocks (randomized across participants; 9 repetitions each), resulting in 81 button and 81 slider trials per participant (in total, 2916 trials). The average study duration was 78 minutes (SD 25 minutes).

4.2.2 Dependent Variables. We recorded the following dependent variables for each trial:

Accuracy. For sliders, we define the OFFSET (in mm) between the target mark and the current value upon completing a click as an accuracy measure, because it quantifies the spatial deviation from the intended target and reflects the user's fine motor control and precision. For buttons, we measure PRESS ACCURACY by counting whether the button was pressed at the time the click completes.

crossings. The number of overshooting (from below) and undershooting (from above) the target mark, where overshooting is defined as when the user's input exceeds the target mark beyond the upper boundary, and undershooting is defined as when the user's input falls below the lower boundary of the target mark. This variable captures the total deviation from the desired target mark, reflecting both the magnitude and frequency of errors.

task completion time. The time between the click to display the task and the click to complete. This variable reflects how quickly and effectively users can understand and accomplish the task, which is crucial for assessing overall usability and optimizing interface design.

NASA Task Load Index & System Usability Scale. To assess task load and usability, participants also completed two standardized questionnaires, the NASA Task Load Index (NASA-TLX) [25, 26] and the System Usability Scale (SUS) [37] after each trial. They provide a validated and multidimensional assessment of perceived task load and overall system usability, helping to capture subjective user experience and cognitive workload that are essential for understanding task difficulty and user satisfaction comprehensively, instead of relying on quantitative feedback.

Custom Questions. Perceived comfort and potential visual obstruction are additional factors that potentially influence the user experience. Therefore, we also asked participants custom questions about comfort during interaction and whether they felt that their hands occluded the visuals (both on a 5-point Likert scale).

4.3 Study Setup and Apparatus

The study setup (see Figure 2b) was as follows: We employ a Meta Quest Pro (with Unity 2021.03.26f1) connected to a computer with Meta Link. To provide a reliable estimate with low latency, we decided to track the hand using two faster, more accurate Leap Motion 2 controllers (LM2Cs) rather than the Meta Quest Pro's embedded hand-tracking. While the pinchable controls themselves require only commodity hand tracking and run on the headset's built-in sensors, we use external LM2Cs in this study to isolate the effects of input technique from tracking-quality artifacts. The experimental setup is therefore a controlled measurement choice and not a hardware requirement of the controls. One LM2C was mounted on the HMD, and the other faced upward from below to track hands when they were likely hanging down for GAZE and RAYCAST. We calibrated the coordinate systems of the two LM2Cs and the HMD using a Meta Quest Pro controller fixated at a constant distance from the lower LM2C. Participants held an external clicker in their non-dominant hand to signal the completion of a trial.

4.4 Procedure

After welcoming our participants, each session commenced with handing out an informed consent form signed by the participants (see Figure 2c). We then verbally explained the experimental procedure and addressed any questions.

Subsequently, the participant was asked to stand in front of a bar table equipped with the HMD and an external clicker in their non-dominant hand. Participants found themselves in a grey, monochromatic space with a floating panel positioned 30 cm in front of their heads. The panel displayed a small text, which the participants were instructed to read aloud to ensure the HMD was correctly positioned and their vision was not blurred. Below the text, an image showed a hand performing a pinching pose. Participants were asked to replicate this pose with the largest possible pinch span that still felt comfortable. This pinch span was then saved internally as the maximum comfortable pinch span. This information was used to transform all slider values into an absolute coordinate system for analysis. For slider and button trials in which MANIPULATION was either GAZE or RAYCAST, the panels were placed 200 cm apart so participants with typical arm reach (ca. 80 cm) and step length (ca. 80 cm) could not easily access the controls in a single motion. To avoid differences in occlusion, we appropriately scaled up to ensure the UI control appeared identically sized for direct and indirect interactions. Participants then proceeded with the nine conditions for sliders and buttons (order randomized between participants). Each was identical in structure but employed a different combination of MANIPULATION and SELECTION.

4.4.1 Slider Trials. Each block began with a tutorial trial, during which a panel featuring the relevant slider and button appeared in front of the participant for them to interact with. Adjacent to the panel was a text description explaining how to use the UI elements, accompanied by a looping tutorial video. Once the participants felt ready, they pressed the external clicker for the first time to initiate nine slider trials.

Participants were instructed to position their hand in a neutral position, indicated by a blue hand at the same location for all participants, to ensure consistent measurements. They then had to direct their gaze to a neutral position by looking at a panel displaying a short countdown above the interaction panel. This countdown was reset if they looked away too early, prompting them to do it again. When the countdown completed, the slider appeared with a yellow target mark indicating the value the participant needed to set. This marks the beginning of the trial (used for task completion time). A blue border around the UI control indicated to the participant that it was actively selected (especially relevant for GAZE and RAYCAST). Once the participants were confident they had set the slider correctly to the target, they pressed the external clicker for a second time to finish the trial. This marks the end of the trial (used for task completion time). In addition to using the external clicker to signal readiness at the beginning and end of each trial, participants interacted solely through the assigned combination of selection and manipulation methods.

For the repetitions, the target marks were randomly distributed as follows to avoid a skew in the number of crossings or the task completion times: Each of the nine slider trials in a block received a random value from one of three value zones (inclusive): low (20%-39%), middle (40%-59%), and high (60%-80%). Each value zone was used three times per block, ensuring an equal distribution of values across the entire range of the slider. Participants completed post-condition questionnaires virtually after all trials of a single condition.

4.4.2 Button Trials. Participants then continued with a set of nine button trials. These trials were structured similarly to the slider trials, starting with a tutorial and then requiring the participant to position their hand and gaze in a neutral position at the start of each trial. The button appeared after the same countdown, and the participant was instructed to press it. Regardless of whether the button was pressed, the trial concluded when the participant pressed the external clicker in their hand. After the final button trial was completed, the same virtual questionnaire appeared.

Between trials, participants were asked to relax their arms to reduce fatigue. After completing all nine conditions, participants answered a final post-study questionnaire regarding their professional background, vision impairments, VR experience, and any general comments.

4.5 Participants

In line with related work (cf. [46, 61]), we recruited 18 participants (10 male, 7 female, 1 identified as gender variant/non-conforming), aged between 20 and 31 ($M = 24.72$, $SD = 2.8$). 15 participants reported being right-handed, and 3 reported being left-handed. Six participants had no prior experience with virtual reality; eight had used it infrequently (1–5 times per year); one had moderate experience (6–10 times per year); and three were regular users (>10 times per year). All participants voluntarily participated in the study and received compensation.

4.6 Analysis

Unless otherwise indicated, we analyzed the recorded data using two-way repeated-measures ANOVAs with the factors SELECTION and MANIPULATION to identify significant effects. We tested the data for normality using Shapiro-Wilk's test. When indicating a severe deviation, we transformed the data (reciprocal transformation for task completion times, log transformation for offset). When Mauchly's test indicated a violation of the sphericity assumption, we corrected the tests using the Greenhouse-Geisser method. When the RM ANOVA indicated significant results, we used Holm-Bonferroni-corrected pairwise t-tests for post hoc analysis. We also report the generalized eta-squared η_G^2 as an estimate of the effect size when appropriate. Further, as an estimate of the mean response of the individual factors, we report, whenever applicable, the Estimated Marginal Means with 95% confidence intervals as proposed by Searle et al. [63] (denoted as M_E in the following). For the analysis of the questionnaires, we performed an Aligned Rank Transformation (ART) as proposed by Elkin et al. [13] and Wobbrock et al. [79], and report the means and standard deviation (denoted with M and SD).

5 Results

This section reports the results of the study. To enhance readability and focus on the most relevant findings, we report only statistically significant effects.

5.1 Accuracy

In terms of accuracy, we distinguish between sliders and buttons. For sliders, we define accuracy as the OFFSET between the target mark and the user value. For buttons, we measured PRESS ACCURACY as the number of times a button was actually pressed when the participant triggered the clicker.

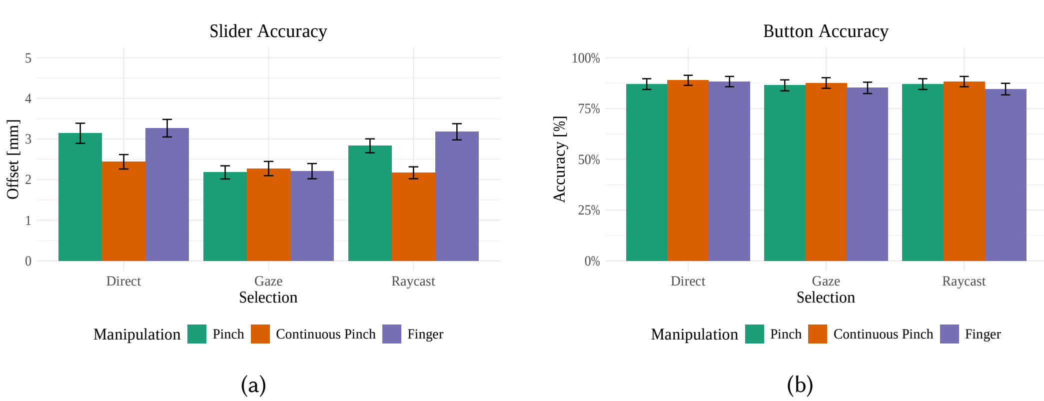


Fig. 3. The OFFSET when setting the target mark for all slider tasks (left) and the button accuracy (right) for the different MANIPULATION and SELECTION techniques.

Table 1. Main and interaction effects with significant post-hoc results for OFFSET (slider accuracy).

Effect	df	F	η^2_G		p
SELECTION	1.81, 30.77	6.11	0.05		<0.01
▶ *Post-hocs:*	GAZE (M_E = 1.35mm [1.09mm, 1.68mm]) < DIRECT (M_E = 1.79mm [1.36mm, 2.36mm])				<0.05
	GAZE (M_E = 1.35mm [1.09mm, 1.68mm]) < RAYCAST (M_E = 1.82mm [1.55mm, 2.13mm])				<0.01
Interaction	2.98, 50.61	3.55	0.03		<0.05
▶ *Post-hocs:*	For FINGER: GAZE (M_E = 1.23mm [0.84mm, 1.8mm]) < DIRECT (M_E = 2.23mm [1.64mm, 3.03mm])				<0.05
	For FINGER: GAZE (M_E = 1.23mm [0.84mm, 1.8mm]) < RAYCAST (M_E = 2.23mm [1.77mm, 2.82mm])				<0.05
	GAZE ⊕ PINCH (M_E = 1.45mm [1.11mm, 1.9mm]) < RAYCAST ⊕ FINGER (M_E = 2.23mm [1.77mm, 2.82mm])				<0.05

Slider Accuracy. As depicted in Figure 3a, we found values for OFFSET ranging from $M = 2.17mm$, $SD = 1.86mm$ (RAYCAST ⊕ CONTINUOUS-PINCH) to $M = 3.27mm$, $SD = 2.73mm$ (DIRECT ⊕ FINGER). We report all significant main and interaction effects in Table 1.

Button Accuracy. To predict this binomial data, we employ a logistic mixed model (estimated with BOBYQA optimizer) with the fixed effects SELECTION and MANIPULATION and the participant as a random effect. As depicted in Figure 3b, we measured press accuracies ranging from $M = 84.6\%$, $SD = 36.2\%$ (RAYCAST ⊕ FINGER) to $M = 88.9\%$, $SD = 31.5\%$ (DIRECT ⊕ CONTINUOUS-PINCH), but did not find any significant effect.

5.2 Slider Crossings

To better understand the challenge of hitting the target mark for sliders, we measured the number of crossings (defined as the instances of completely undershooting or overshooting the target mark). For statistical analysis, we fitted Poisson regression models (estimated with BOBYQA optimizer) with the fixed effects SELECTION and MANIPULATION and the participant as a random effect. We applied Type III Wald chi-square tests for significance testing and report the incidence rate ratios (IRRs) as an effect size for all significant contrasts.

As depicted in Figure 4, we found that the number of crossings ranges between conditions from $M = 1.74$, $SD = 1.70$ (RAYCAST ⊕ CONTINUOUS-PINCH) to $M = 3.91$, $SD = 3.21$ (GAZE ⊕ CONTINUOUS-PINCH). We report all significant main and interaction effects in Table 2.

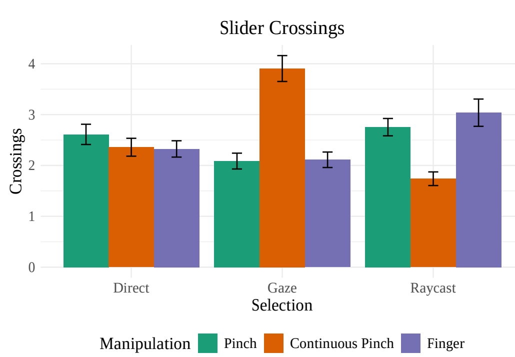


Fig. 4. The number of CROSSINGS (over- and undershooting the target mark) for all slider tasks for the different MANIPULATION and SELECTION techniques.

Table 2. Main and interaction effects with significant post-hoc results for slider CROSSINGS.

Effect	df	χ^2	p	IRR
SELECTION	2	16.036	<0.001	–
► *Post-hocs:*	*none*			
Interaction	4	179.136	<0.001	–
► *Post-hocs:*	For PINCH: DIRECT (M_E = 2.48 [2.08, 2.97]) > GAZE (M_E = 1.97 [1.65, 2.39])		<0.05	1.25
	For CONTINUOUS-PINCH: DIRECT (M_E = 2.27 [1.88, 2.72]) < GAZE (M_E = 3.74 [3.16, 4.44])		<0.001	0.61
	For CONTINUOUS-PINCH: DIRECT (M_E = 2.27 [1.88, 2.72]) > RAYCAST (M_E = 1.67 [1.36, 2.01])		<0.01	1.36
	For CONTINUOUS-PINCH: GAZE (M_E = 3.74 [3.16, 4.44]) > RAYCAST (M_E = 1.67 [1.36, 2.01])		<0.001	2.25
	For FINGER: DIRECT (M_E = 2.23 [1.86, 2.66]) < RAYCAST (M_E = 2.89 [2.44, 3.46])		<0.01	0.77
	For FINGER: GAZE (M_E = 2.01 [1.68, 2.41])< RAYCAST (M_E = 2.89 [2.44, 3.46])		<0.001	0.7
	GAZE ⊕ PINCH (M_E = 1.97 [1.65, 2.39]) < RAYCAST ⊕ PINCH (M_E = 2.61 [2.18, 3.13])		<0.01	0.76
	GAZE ⊕ PINCH (M_E = 1.97 [1.65, 2.39]) < GAZE ⊕ CONTINUOUS-PINCH (M_E = 3.74 [3.16, 4.44])		<0.001	0.53
	GAZE ⊕ PINCH (M_E = 1.97 [1.65, 2.39]) < RAYCAST ⊕ FINGER (M_E = 2.89 [2.44, 3.46])		<0.001	0.69
	GAZE ⊕ CONTINUOUS-PINCH (M_E = 3.74 [3.16, 4.44]) > DIRECT ⊕ FINGER (M_E = 2.23 [1.86, 2.66])		<0.001	1.68
	GAZE ⊕ CONTINUOUS-PINCH (M_E = 3.74 [3.16, 4.44]) > GAZE ⊕ FINGER (M_E = 2.01 [1.68, 2.41])		<0.001	1.85
	GAZE ⊕ CONTINUOUS-PINCH (M_E = 3.74 [3.16, 4.44]) > RAYCAST ⊕ FINGER (M_E = 2.89 [2.44, 3.46])		<0.001	1.29
	RAYCAST ⊕ PINCH (M_E = 2.61 [2.18, 3.13]) > RAYCAST ⊕ CONTINUOUS-PINCH (M_E = 1.67 [1.36, 2.01])		<0.001	0.7
	RAYCAST ⊕ PINCH (M_E = 2.61 [2.18, 3.13]) > GAZE ⊕ FINGER (M_E = 2.01 [1.68, 2.41])		<0.001	1.3
	RAYCAST ⊕ CONTINUOUS-PINCH (M_E = 1.67 [1.36, 2.01]) < DIRECT ⊕ FINGER (M_E = 2.23 [1.86, 2.66])		<0.01	0.75
	RAYCAST ⊕ CONTINUOUS-PINCH (M_E = 1.67 [1.36, 2.01]) < RAYCAST ⊕ FINGER (M_E = 2.89 [2.44, 3.46])		<0.01	0.57

5.3 Task Completion Time

To measure participants' efficiency, we recorded task completion time (TCT) as the time between clicking to start and confirming to end a trial.

Slider TCT. As shown in Figure 5, we found task completion times for the slider task that range from $M = 3.03s$, $SD = 1.40s$ (GAZE ⊕ CONTINUOUS-PINCH) to $M = 4.56s$, $SD = 2.57s$ (RAYCAST ⊕ FINGER). We report all significant main and interaction effects in Table 3 (top).

Button TCT. As shown in Figure 5, we found task completion times for the button task that range from $M = 1.84s$, $SD = 1.18s$ (DIRECT ⊕ FINGER) to $M = 3.44s$, $SD = 1.87s$ (RAYCAST ⊕ FINGER). We report all significant main and interaction effects in Table 3 (bottom).

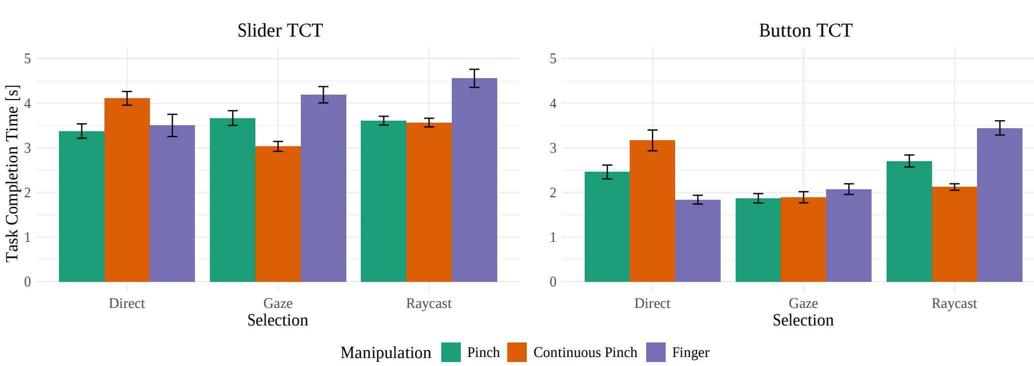


Fig. 5. The TASK COMPLETION TIME for both slider and button tasks for the different MANIPULATION and SELECTION techniques.

Table 3. Effects with significant post-hoc results for TASK COMPLETION TIME of slider (top) and button (bottom).

Effect	df	F	η^2_G	p
SELECTION	1.82, 30.97	5.14	0.05	<0.01
► *Post-hocs:*	RAYCAST (M_E = 3.38s [3.79s, 3.04s]) > DIRECT (M_E = 2.93s [3.42s, 2.56s])			<0.05
	RAYCAST (M_E = 3.38s [3.79s, 3.04s]) > GAZE (M_E = 2.99s [3.34s, 2.71s])			<0.05
Interaction	3, 51.02	10.2	0.1	<0.001
► *Post-hocs:*	For DIRECT: CONTINUOUS-PINCH (M_E = 3.49s [4.16s, 3.01s]) > PINCH (M_E = 2.78s [3.24s, 2.42s])			<0.01
	For DIRECT: CONTINUOUS-PINCH (M_E = 3.49s [4.16s, 3.01s]) > FINGER (M_E = 2.65s [3.19s, 2.27s])			<0.01
	For RAYCAST: FINGER (M_E = 3.58s [4.46s, 2.99s]) > PINCH (M_E = 2.78s [3.24s, 2.42s])			<0.05
	For RAYCAST: FINGER (M_E = 3.58s [4.46s, 2.99s]) > CONTINUOUS-PINCH (M_E = 2.65s [3.19s, 2.27s])			<0.01
	GAZE ⊕ CONTINUOUS-PINCH (M_E = 2.67s [2.99s, 2.4s]) < RAYCAST ⊕ PINCH (M_E = 3.3s [3.69s, 2.98s])			<0.05
SELECTION	1.99, 29.84	22.3	0.25	<0.001
► *Post-hocs:*	DIRECT (M_E = 1.83s [2.12s, 1.62s]) > GAZE (M_E = 1.53s [1.71s, 1.39s])			<0.05
	RAYCAST (M_E = 2.25s [2.5s, 2.05s]) > DIRECT (M_E = 1.83s [2.12s, 1.62s])			<0.05
	RAYCAST (M_E = 2.25s [2.5s, 2.05s]) > GAZE (M_E = 1.53s [1.71s, 1.39s])			<0.01
Interaction	3.26, 48.94	14.91	0.23	<0.001
► *Post-hocs:*	DIRECT ⊕ PINCH (M_E = 1.91s [2.18s, 1.7s]) > GAZE ⊕ CONTINUOUS-PINCH (M_E = 1.45s [1.7s, 1.27s])			<0.05
	DIRECT ⊕ PINCH (M_E = 1.91s [2.18s, 1.7s]) > DIRECT ⊕ FINGER (M_E = 1.44s [1.73s, 1.23s])			<0.05
	DIRECT ⊕ PINCH (M_E = 1.91s [2.18s, 1.7s]) < RAYCAST ⊕ FINGER (M_E = 2.74s [3.21s, 2.39s])			<0.05
	GAZE ⊕ PINCH (M_E = 1.48s [1.74s, 1.29s]) < RAYCAST ⊕ PINCH (M_E = 2.28s [2.57s, 2.04s])			<0.01
	GAZE ⊕ PINCH (M_E = 1.48s [1.74s, 1.29s]) < DIRECT ⊕ CONTINUOUS-PINCH (M_E = 2.38s [2.88s, 2.03s])			<0.01
	GAZE ⊕ PINCH (M_E = 1.48s [1.74s, 1.29s]) < RAYCAST ⊕ FINGER (M_E = 2.74s [3.21s, 2.39s])			<0.01
	GAZE ⊕ CONTINUOUS-PINCH (M_E = 1.45s [1.7s, 1.27s]) < DIRECT ⊕ CONTINUOUS-PINCH (M_E = 2.38s [2.88s, 2.03s])			<0.05
	GAZE ⊕ CONTINUOUS-PINCH (M_E = 1.45s [1.7s, 1.27s]) < RAYCAST ⊕ FINGER (M_E = 2.74s [3.21s, 2.39s])			<0.01
	DIRECT ⊕ CONTINUOUS-PINCH (M_E = 2.38s [2.88s, 2.03s]) > DIRECT ⊕ FINGER (M_E = 1.44s [1.73s, 1.23s])			<0.01
	DIRECT ⊕ CONTINUOUS-PINCH (M_E = 2.38s [2.88s, 2.03s]) > GAZE ⊕ FINGER (M_E = 1.69s [1.91s, 1.51s])			<0.01
	RAYCAST ⊕ PINCH (M_E = 2.28s [2.57s, 2.04s]) > DIRECT ⊕ FINGER (M_E = 1.44s [1.73s, 1.23s])			<0.05
	RAYCAST ⊕ PINCH (M_E = 2.28s [2.57s, 2.04s]) > GAZE ⊕ FINGER (M_E = 1.69s [1.91s, 1.51s])			<0.01
	RAYCAST ⊕ FINGER (M_E = 2.74s [3.21s, 2.39s]) > DIRECT ⊕ FINGER (M_E = 1.44s [1.73s, 1.23s])			<0.05
	RAYCAST ⊕ FINGER (M_E = 2.74s [3.21s, 2.39s]) > RAYCAST ⊕ CONTINUOUS-PINCH (M_E = 1.9s [2.19s, 1.68s])			<0.01
	RAYCAST ⊕ FINGER (M_E = 2.74s [3.21s, 2.39s]) > GAZE ⊕ FINGER (M_E = 1.69s [1.91s, 1.51s])			<0.01

5.4 NASA Task Load Index

As depicted in Figure 6, we found overall NASA-TLX scores ranging from $M = 21.4$, $SD = 18.2$ (GAZE ⊕ CONTINUOUS-PINCH) to $M = 48.2$, $SD = 28.9$ (RAYCAST ⊕ FINGER) to assess task load. We report the result of the statistical analysis in Table 4. Individual TLX items can be found in Table 7.

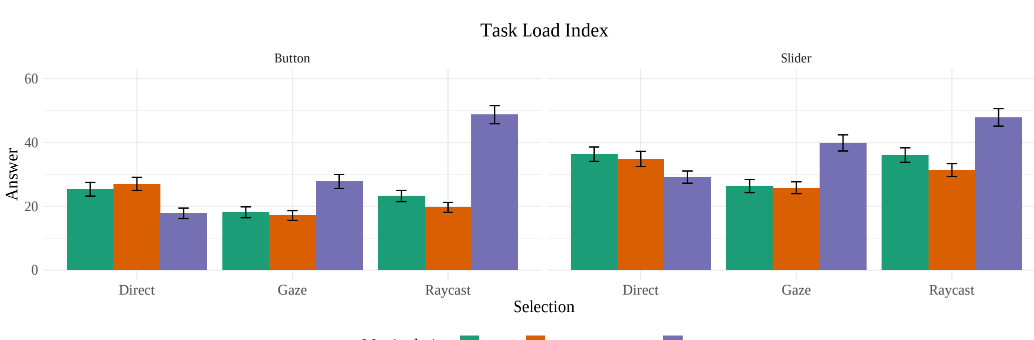


Fig. 6. The NASA-TLX scores for the different MANIPULATION and SELECTION techniques.

Table 4. Main and interaction effects with significant post-hoc comparisons for TLX SCORE (we report the mean and standard deviation in brackets).

Effect	df	F & η^2	p
SELECTION	2, 34	10.64 & 0.385	<0.001
► *Post-hocs:*	RAYCAST (M = 34.4, SD = 17.4) > GAZE (M = 25.8, SD = 14.6)		<<0.001
	RAYCAST (M = 34.4, SD = 17.4) > DIRECT (M = 28.4, SD = 13)		<<0.001
MANIPULATION	2, 34	27.91 & 0.621	<0.001
► *Post-hocs:*	FINGER (M = 35.2, SD = 18.5) > PINCH (M = 27.5, SD = 13.3)		<<0.001
	FINGER (M = 35.2, SD = 18.5) > CONTINUOUS-PINCH (M = 25.9, SD = 12.5)		<<0.001
Interaction	4, 68	16.25 & 0.489	<0.001
► *Post-hocs:*	DIRECT ⊕ PINCH (M = 30.8, SD = 14.4) > GAZE ⊕ CONTINUOUS-PINCH (M = 21.4, SD = 11.3)		<0.05
	DIRECT ⊕ CONTINUOUS-PINCH (M = 30.9, SD = 14.7) > GAZE ⊕ CONTINUOUS-PINCH (M = 21.4, SD = 11.3)		<0.05
	RAYCAST ⊕ FINGER (M = 48.2, SD = 19.4) > DIRECT ⊕ PINCH (M = 30.8, SD = 14.4)		<0.01
	DIRECT ⊕ FINGER (M = 23.4, SD = 8.13) < GAZE ⊕ FINGER (M = 33.8, SD = 17.1)		<0.05
	DIRECT ⊕ FINGER (M = 23.4, SD = 8.13) < RAYCAST ⊕ FINGER (M = 48.2, SD = 19.4)		<0.001
	GAZE ⊕ PINCH (M = 22.2, SD = 11.8) < GAZE ⊕ FINGER (M = 33.8, SD = 17.1)		<0.001
	GAZE ⊕ PINCH (M = 22.2, SD = 11.8) < RAYCAST ⊕ FINGER (M = 48.2, SD = 19.4)		<0.01
	GAZE ⊕ CONTINUOUS-PINCH (M = 21.4, SD = 11.3) < RAYCAST ⊕ PINCH (M = 29.6, SD = 12.6)		<0.05
	GAZE ⊕ CONTINUOUS-PINCH (M = 21.4, SD = 11.3) < GAZE ⊕ FINGER (M = 33.8, SD = 17.1)		<0.01
	GAZE ⊕ CONTINUOUS-PINCH (M = 21.4, SD = 11.3) < RAYCAST ⊕ FINGER (M = 48.2, SD = 19.4)		<0.01
	RAYCAST ⊕ FINGER (M = 48.2, SD = 19.4) > RAYCAST ⊕ CONTINUOUS-PINCH (M = 25.5, SD = 9.91)		<0.001
	RAYCAST ⊕ FINGER (M = 48.2, SD = 19.4) > RAYCAST ⊕ PINCH (M = 29.6, SD = 12.6)		<0.001
	RAYCAST ⊕ FINGER (M = 48.2, SD = 19.4) > GAZE ⊕ FINGER (M = 33.8, SD = 17.1)		<0.05

5.5 System Usability Scale

As depicted in Figure 7, we found SUS scores ranging from $M = 40.6$, $SD = 20.8$ (RAYCAST ⊕ FINGER) to $M = 80.8$, $SD = 14.2$ (GAZE ⊕ PINCH). We report all significant main and interaction effects in Table 5.

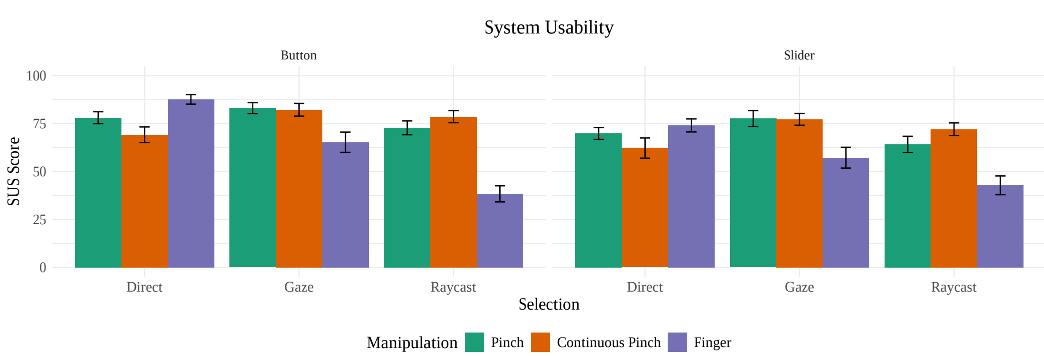


Fig. 7. The SUS scores for the different MANIPULATION and SELECTION techniques.

Table 5. Main and interaction effects with significant post-hoc comparisons for SUS SCORE (we report the mean and standard deviation in brackets).

Effect	df	F & η^2	p
SELECTION	2, 34	18.53 & 0.52	<0.001
► *Post-hocs:*	RAYCAST (M = 61.5, SD = 21) < DIRECT (M = 73.5, SD = 14.44)		<0.001
	RAYCAST (M = 61.5, SD = 21) < GAZE (M = 73.5, SD = 14.44)		<0.001
MANIPULATION	2, 34	18.62 & 0.52	<0.001
► *Post-hocs:*	FINGER (M = 60.9, SD = 23.3) < PINCH (M = 74.3, SD = 13.1)		<0.001
	FINGER (M = 60.9, SD = 23.3) < CONTINUOUS-PINCH (M = 73.6, SD = 15.1)		<0.001
Interaction	4, 68	20.84 & 0.55	<0.001
► *Post-hocs:*	DIRECT ⊕ FINGER (M = 80.8, SD = 10.3) > DIRECT ⊕ CONTINUOUS-PINCH (M = 65.7, SD = 17.8)		<0.01
	DIRECT ⊕ FINGER (M = 80.8, SD = 10.3) > GAZE ⊕ FINGER (M = 61.2, SD = 21)		<0.01
	DIRECT ⊕ FINGER (M = 80.8, SD = 10.3) > RAYCAST ⊕ PINCH (M = 68.5, SD = 15)		<0.001
	RAYCAST ⊕ FINGER (M = 40.6, SD = 16.9) < DIRECT ⊕ PINCH (M = 74, SD = 10.2)		<0.001
	RAYCAST ⊕ FINGER (M = 40.6, SD = 16.9) < DIRECT ⊕ CONTINUOUS-PINCH (M = 65.7, SD = 17.8)		<0.01
	GAZE ⊕ PINCH (M = 80.3, SD = 11.3) > RAYCAST ⊕ FINGER (M = 40.6, SD = 16.9)		<0.001
	GAZE ⊕ PINCH (M = 80.3, SD = 11.3) > GAZE ⊕ FINGER (M = 61.2, SD = 21)		<0.01
	GAZE ⊕ PINCH (M = 80.3, SD = 11.3) > DIRECT ⊕ CONTINUOUS-PINCH (M = 65.7, SD = 17.8)		<0.05
	GAZE ⊕ CONTINUOUS-PINCH (M = 79.7, SD = 12) > RAYCAST ⊕ FINGER (M = 40.6, SD = 16.9)		<0.001
	GAZE ⊕ CONTINUOUS-PINCH (M = 79.7, SD = 12) > GAZE ⊕ FINGER (M = 61.2, SD = 21)		<0.01
	GAZE ⊕ CONTINUOUS-PINCH (M = 79.7, SD = 12) > DIRECT ⊕ CONTINUOUS-PINCH (M = 65.7, SD = 17.8)		<0.05
	RAYCAST ⊕ FINGER (M = 40.6, SD = 16.9) < GAZE ⊕ FINGER (M = 61.2, SD = 21)		<0.05
	RAYCAST ⊕ FINGER (M = 40.6, SD = 16.9) < RAYCAST ⊕ PINCH (M = 68.5, SD = 15)		<0.001
	RAYCAST ⊕ FINGER (M = 40.6, SD = 16.9) < RAYCAST ⊕ CONTINUOUS-PINCH (M = 75.3, SD = 12)		<0.001

5.6 Custom Questions

Comfort. We were also specifically interested in participants' perceptions of comfort on a 5-point Likert scale (*"I felt that the interaction was comfortable"*; the higher the score, the greater the agreement). As shown in Figure 8 (left), we found values ranging from M = 2.42, SD = 1.16 (RAYCAST ⊕ FINGER) to M = 4.14, SD = 0.76 (GAZE ⊕ PINCH). We report all significant main and interaction effects in Table 6.

Occlusion. We were also specifically interested in whether the different techniques occluded the participant's view of the visuals on a 5-point Likert scale (*"I felt my hand occluded the visuals"*; the higher the score, the greater the agreement). As shown in Figure 8 (right), we found values ranging from M = 1.67, SD = 0.76 (GAZE ⊕ PINCH) to M = 2.72, SD = 1.39 (DIRECT ⊕ CONTINUOUS-PINCH). The analysis indicated a significant ($F_{2,34}$ = 10.97, $p < 0.001$) main effect of SELECTION on OCCLUSION. Post-hoc tests confirmed (both p<0.01) that the perceived OCCLUSION for DIRECT (M = 2.57, SD = 1.31) was significantly higher than for both GAZE (M = 1.82, SD = 0.91) and RAYCAST (M = 1.94, SD = 0.91).

6 Discussion and Guidelines

In the following, we discuss our findings and derive guidelines for hand-based XR input.

6.1 Gaze Selection and Pinch-Based Manipulation Complement Each Other

Our findings indicate a synergy between gaze for selection and pinch-based manipulation, with gaze yielding the lowest slider offsets on average and pinch or continuous-pinch providing strong efficiency (see Table 1 and Table 3), low workload (see Table 4), and high perceived usability (see Table 5) and comfort (see Table 6), especially at a distance where aiming to select and arm

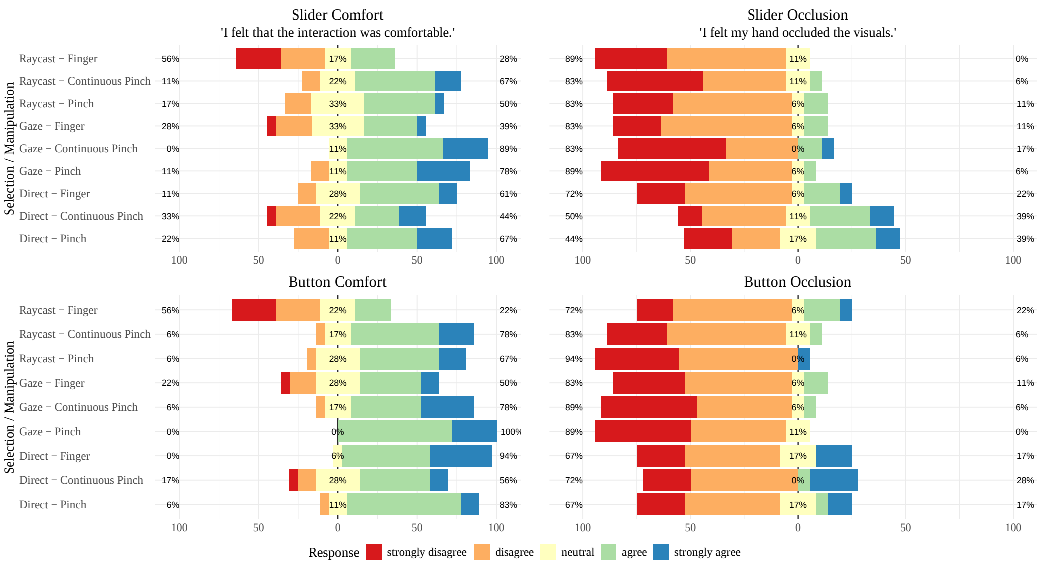


Fig. 8. Likert ratings for comfort and occlusion, subdivided by slider (top) and button tasks (bottom).

Table 6. Main and interaction effects with significant post-hoc comparisons for the custom question about COMFORT.

Effect	**df**	**F & η^2**		**p**
SELECTION	2, 34	9.49 & 0.36		<0.001
► *Post-hocs:*			RAYCAST ($M = 3.28$, $SD = 1.13$) < DIRECT ($M = 3.69$, $SD = 0.98$)	<0.01
			RAYCAST ($M = 3.28$, $SD = 1.13$) < GAZE ($M = 3.82$, $SD = 0.96$)	<0.01
MANIPULATION	2, 34	7.81 & 0.31		<0.01
► *Post-hocs:*			FINGER ($M = 3.2$, $SD = 1.19$) < CONTINUOUS-PINCH ($M = 3.76$, $SD = 0.97$)	<0.01
			FINGER ($M = 3.2$, $SD = 1.19$) < PINCH ($M = 3.83$, $SD = 0.86$)	<0.001
Interaction	4, 68	8.86 & 0.34		<0.001
► *Post-hocs:*			RAYCAST ⊕ FINGER ($M = 2.42$, $SD = 1$) < RAYCAST ⊕ PINCH ($M = 3.58$, $SD = 0.69$)	<0.01
			RAYCAST ⊕ FINGER ($M = 2.42$, $SD = 1$) < DIRECT ⊕ PINCH ($M = 3.78$, $SD = 0.75$)	<0.01
			GAZE ⊕ PINCH ($M = 4.14$, $SD = 0.59$) < RAYCAST ⊕ FINGER ($M = 2.42$, $SD = 1$)	<0.001
			DIRECT ⊕ CONTINUOUS-PINCH ($M = 3.33$, $SD = 1.04$) < GAZE ⊕ PINCH ($M = 4.14$, $SD = 0.59$)	<0.05
			DIRECT ⊕ FINGER ($M = 3.97$, $SD = 0.53$) > RAYCAST ⊕ FINGER ($M = 2.42$, $SD = 1$)	<0.01
			GAZE ⊕ PINCH ($M = 4.14$, $SD = 0.59$) > RAYCAST ⊕ FINGER ($M = 2.42$, $SD = 1$)	<0.01
			GAZE ⊕ PINCH ($M = 4.14$, $SD = 0.59$) > GAZE ⊕ FINGER ($M = 3.22$, $SD = 0.86$)	<0.01
			GAZE ⊕ CONTINUOUS-PINCH ($M = 4.11$, $SD = 0.61$) > GAZE ⊕ FINGER ($M = 3.22$, $SD = 0.86$)	<0.01
			GAZE ⊕ CONTINUOUS-PINCH ($M = 4.11$, $SD = 0.61$) > RAYCAST ⊕ FINGER ($M = 2.42$, $SD = 1$)	<0.01

fatigue can dominate the interaction. Specifically, gaze reduced the slider offset relative to both direct touch and hand raycast. The interaction analysis further showed that when manipulation was by finger, gaze delivered the smallest offsets among selection methods, which indicates that selection sets the ceiling for precision even before manipulation details come into play *(RQ1)*.

On the efficiency side, gaze consistently outpaced hand raycast for buttons, and the best button times after direct finger input appeared with gaze combined with either pinch or continuous-pinch

(RQ2). This aligns with low TLX scores, high SUS scores, and perceived comfort, suggesting that the combination of gaze-based selection and pinch-based manipulation minimizes cognitive overhead. Comfort and occlusion further strengthen this pairing, in line with prior work [54, 55, 64], where gaze with both pinch variants was among the most comfortable combinations, while gaze generally reduced perceived occlusion relative to direct, indicating that distant selection decouples the hand from the visual focus and reduces incidental blocking of the interface.

Continuous thumb-index pinching performed best when paired with gaze but is slow for direct interaction and can elevate perceived occlusion up close, suggesting that continuous pinching is most effective at a distance where the hand can remain relaxed and out of the way. However, the combination of gaze with continuous-pinch produced more crossings than the other conditions despite achieving fast task completion times, pointing to a speed–precision trade-off that likely stems from unregulated gain or insufficient endpoint stabilization in the continuous mapping *(RQ3)*. Our findings suggest that future work should further investigate adaptive velocity-to-gain mappings, damping near targets, or short "settle" windows that preserve speed during gross motion while reducing micro-oscillation during closing in on the target, thereby retaining the comfort and time benefits of continuous-pinch without incurring additional crossings.

As a consequence, selection matters more than manipulation (*RQ1* and *RQ2*): gaze selection reduced slider offset, improved times on both widgets, lowered TLX scores, and increased SUS scores relative to raycast, indicating that the stability and low visuomotor cost of gaze set a favorable stage for whichever manipulation is layered on top.

Moreover, manipulation is interaction-dependent *(RQ2)*. Pinch is a robust default interaction modality across distances, where continuous-pinch excels with gaze but slows down direct interaction and increases crossings for sliders. Fingers remain highly effective in direct, near-field button tasks but underperform in indirect or distant contexts, as indicated by workload and SUS scores.

6.2 Direct Touch Is Still Strong, but Has Occlusion Costs

Combining direct selection with finger manipulation remains a strong input method for near-field discrete actions, producing the fastest button completion times (see Table 3) and the highest SUS scores (see Table 5) among all combinations *(RQ3)*. This result validates direct touch when the control is close, targeting is trivial, and the action is discrete rather than continuous, echoing long-standing results on the efficiency of direct touch for discrete close-range interactions in mobile devices (cf. [6, 65]). In contrast, direct selection did not achieve the best slider precision and increased perceived occlusion relative to gaze and even raycast, indicating that the benefits of direct touch decrease when fine adjustment is required or when the hand blocks critical visual information near the control.

In addition, the combination of direct selection with continuous thumb-index pinch slowed tasks for both sliders and buttons, and it yielded higher perceived occlusion and higher TLX scores than gaze selection, suggesting that continuous modulation at arm's length can become visually intrusive and ergonomically demanding without adding precision gains to justify the cost *(RQ3)*. Our results show that direct finger interaction scored high in overall comfort, outperforming the combination of finger and hand raycast. Yet, combining continuous pinch with indirect gaze selection still surpassed its combination with direct hand selection in terms of comfort, which indicates that the same continuous manipulation becomes less taxing when selection keeps the hand away from the interface and possibly reduces posture strain.

Our results motivate clear guidelines: employ the combination of direct manipulation with the index finger for nearby, discrete interactions where its speed and familiarity shine, and occlusion is not critical. Yet for distant or precision-sensitive adjustments, prefer gaze with pinch or continuous-pinch to reduce occlusion, time, and workload. When continuous control is required in near-field contexts, designers should consider hybrid strategies, such as transient gain reduction in the last centimeters of motion or visuals that shift the locus of control out of the line of sight to mitigate occlusion and reduce the need for fine-motor corrections in front of the target.

6.3 Avoid Unassisted Hand Raycast for Selection

In our study, we used a raw, unassisted raycast as an empirical baseline that often lagged on both slider and button times, registered higher TLX scores, and received lower SUS scores than gaze and direct, indicating that it imposed stability and aiming overheads that were not compensated by the manipulation gesture's familiarity. This issue became most evident in the combination of hand raycast with finger manipulation, which clustered the worst experience attributes: slower times, higher workload, lower comfort, and the lowest SUS score. This might be due to jitter and continuous micro-corrections required to keep the ray aligned with the target during manipulation preparation, which is in line with prior work (cf. the Heisenberg effect in spatial interaction [81]). Yet, there are promising mitigation strategies for raycast, including increased damping and smoothing, target snapping or magnetic funnels at the endpoint, or alternative visualization that reduces visual clutter near the target.

The combination of continuous pinch with hand raycast, interestingly, achieved lower slider crossings than its combination with gaze. Yet, the broader pattern still disfavored unassisted hand raycast because the reductions in crossings did not translate into substantial time or workload advantages, indicating hidden coordination costs in maintaining the ray direction and timing the continuous control *(RQ3)*.

Our results suggest that using unassisted hand-based raycasting as the default selection technique for simple one-dimensional controls is not advisable when gaze input is available, particularly in contexts where speed and perceived usability are critical. If ray-based selection is needed, it should be strongly supported and is best paired with pinch gestures rather than finger-based input. This trade-off also applies when combining gaze with continuous pinching, as this configuration resulted in a high rate of crossing errors and may therefore benefit from mitigation strategies such as adaptive gain.

7 Limitations and Future Work

Our paper provides insights into various hand-based XR input techniques. However, this paper also has limitations that raise research questions for future work.

7.1 Validity

This paper advanced the understanding of different hand-based input techniques in XR. To achieve this, we conducted a controlled experiment to establish the fundamental properties of these techniques. However, due to the limited number of participants, the experiment, by definition, only captures a sample of actual human performance. Moreover, varying hand anthropometrics might especially impact comfort and occlusion. As a result, the findings may differ for other user groups, such as children or seniors.

7.2 Fatigue

Fatigue may be a potential limitation to the repeated use of the different hand-based input techniques. Participants did not report signs of fatigue. Also, the physical demand ratings from the TLX reported after each trial did not show a characteristic increase over the course of the study (likely due to relaxation phases between trials). While fatigue consequently did not appear to influence our study results regarding subjective physical demand, our study did not include objective measures of physical load (e.g., hand movement distance, fatigue-based models [38]). We acknowledge that incorporating such measures could provide additional insights into motor adaptation or compensatory behaviors during interaction. Therefore, fatigue remains a complex issue, especially for future interactive systems beyond study prototypes, which warrants further investigation. Future studies should investigate whether the hand becomes progressively fatigued with accumulated hand-based movements, particularly in continuous thumb-index pinching.

7.3 Hand Tracking

Our study employed two Leap Motion 2 sensors to obtain robust, low-latency hand tracking. Our setup offers better tracking quality compared to integrated hand-tracking solutions, particularly for extreme poses and specific occlusions. Yet, we were primarily interested in comparative measurements across the different interaction techniques and thus did not optimize the system for peak interaction performance. Therefore, the absolute performance values may differ when using built-in tracking, controllers, or more advanced hand motion tracking (e.g., recent work [82] already showed how hand motion tracking becomes increasingly accurate using only a single RGB camera).

7.4 Set of Selected Techniques

In this paper, we focus on the most commonly implemented hand-based input techniques for XR, including direct touch, pinch-to-click, and gaze-supported interactions. However, many other promising XR input techniques were not included in our evaluation. Future research should explore a broader range of interaction methods to determine how they compare to continuous thumb-index pinching, particularly in terms of usability, efficiency, and user comfort. Comparing a greater set of techniques will provide a more comprehensive understanding of the strengths and limitations of continuous thumb-index pinching.

Another important consideration relates to the Heisenberg effect in interaction [81]. This phenomenon suggests that certain errors might not only stem from the accuracy of the interaction technique itself, but also from the cognitive and motor demands imposed by the selection mechanism. Future work should further investigate this effect in the context of raycast-based thumb-index pinching.

7.5 Slider Design and Release

One potential confound in our study is the design of the pinchable slider, which differs from standard single-handle sliders used in XR. While this design was chosen to leverage pinching as an intuitive manipulation method, future work should investigate whether UI modifications (e.g., single-handle sliders, alternative snapping mechanisms) affect performance outcomes. Additionally, the high crossing rate for RAYCAST ⊕ CONTINUOUS-PINCH may be partly influenced by the slider design rather than solely by the interaction technique. Future studies could isolate UI factors to determine their impact on interaction effectiveness. Moreover, our evaluation prioritized assessing the influence of selection and manipulation and thus did not explicitly model the interaction's release phase. For this reason, we utilized an external clicker for all conditions so that performance differences reflect the techniques themselves, avoiding confounds from differing release methods (e.g., continuous pinching alone could support several release options without extra hardware like lifting the fingers off the slider, using the trapeziometacarpal joint to move the thumb out of plane, or using a brief rapid-pinch gesture). Consequently, this limits generalizability, as potential friction or errors inherent to this step in the interaction cycle may have been overlooked compared to those in established methods. Subsequent studies should more closely examine ecological settings with a range of release methods to fully characterize their impact on overall usability and efficiency.

8 Conclusion

In this work, we investigated continuous thumb-index pinching as a hand-based input technique for Extended Reality (XR), enabling pinchable buttons and sliders. To our knowledge, this is the first work to treat continuous thumb-index pinching as one option within a broader hand-based XR design space and to systematically compare it with direct touch, binary pinch-to-click, gaze-based, and raycast interactions for common UI controls. In a controlled experiment, we systematically compared them and found trade-offs among direct touch, pinch-to-click, gaze-supported, and raycasting-based interactions. Our findings show that continuous pinching is in some cases an alternative to the commonly known pinch interactions.

Our results suggest practical guidelines for hand-based XR interfaces: (1) designers should pair gaze with pinch-based manipulation when dealing with distant or precision-demanding controls, (2) use direct finger touch for nearby discrete actions, and (3) reduce the reliance on hand raycast. We hope these findings serve as a valuable foundation for future designs of hand-based XR interfaces.

Acknowledgments

We would like to thank Manuel Peters, all participants, and the reviewers for their valuable help. This work was supported by the German Research Foundation (DFG), CRC 1404: "FONDA: Foundations of Workflows for Large-Scale Scientific Data Analysis" (Project-ID 414984028).

References

[1] Nur Ameerah Binti Abdul Halim and Ajune Wanis Binti Ismail. 2021. Raycasting method using hand gesture for target selection on the occluded object in handheld Augmented Reality. In *2021 6th IEEE International Conference on Recent Advances and Innovations in Engineering (ICRAIE)*, Vol. 6. 1–6. https://doi.org/10.1109/ICRAIE52900.2021.9704035

[2] Jeff Avery, Mark Choi, Daniel Vogel, and Edward Lank. 2014. Pinch-to-zoom-plus: an enhanced pinch-to-zoom that reduces clutching and panning. In *Proceedings of the 27th annual ACM symposium on User interface software and technology (UIST '14)*. Association for Computing Machinery, New York, NY, USA, 595–604. https://doi.org/10.1145/2642918.2647352

[3] Hrvoje Benko and Andrew D. Wilson. 2010. Pinch-the-Sky Dome: Freehand Multi-Point Interactions with Immersive Omni-Directional Data. In *CHI '10 Extended Abstracts on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 3045–3050.

[4] 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 - SI3D '97*. ACM Press, Providence, Rhode Island, United States, 35–ff. https://doi.org/10.1145/253284.253301

[5] Doug A Bowman, Chadwick A Wingrave, Joshua M Campbell, and Vinh Q Ly. 2001. Using Pinch Gloves™ for both Natural and Abstract Interaction Techniques in Virtual Environments. *Proceedings of HCI International* (2001). https://eprints.cs.vt.edu/archive/00000547/01/PinchGlovePaper.pdf

[6] Andrew Bragdon, Eugene Nelson, Yang Li, and Ken Hinckley. 2011. Experimental analysis of touch-screen gesture designs in mobile environments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, Vancouver BC Canada, 403–412. https://doi.org/10.1145/1978942.1979000

[7] Nikolas Chaconas and Tobias Hollerer. 2018. An Evaluation of Bimanual Gestures on the Microsoft HoloLens. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, Reutlingen, 1–8. https://doi.org/10.1109/VR.2018.8446320

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

[9] Ishan Chatterjee, Robert Xiao, and Chris Harrison. 2015. Gaze+Gesture: Expressive, Precise and Targeted Free-Space Interactions. In *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction*. ACM, Seattle Washington USA, 131–138. https://doi.org/10.1145/2818346.2820752

[10] Mehrab Zaman Chowdhury, Mehedi Hassan, and Ahmed Iqbal Pritom. 2018. Easy Swipe Keyboard: A Touch based on Screen Text Entry Method for Smartphone. 05, 07 (2018).

[11] Ashley Colley and Jonna Häkkilä. 2014. Exploring finger specific touch screen interaction for mobile phone user interfaces. In *Proceedings of the 26th Australian Computer-Human Interaction Conference on Designing Futures: the Future of Design*. ACM, Sydney New South Wales Australia, 539–548. https://doi.org/10.1145/2686612.2686699

[12] Haiwei Dong, Ali Danesh, Nadia Figueroa, and Abdulmotaleb El Saddik. 2015. An Elicitation Study on Gesture Preferences and Memorability Toward a Practical Hand-Gesture Vocabulary for Smart Televisions. *IEEE Access* 3 (2015), 543–555. https://doi.org/10.1109/ACCESS.2015.2432679 Conference Name: IEEE Access.

[13] Lisa A. Elkin, Matthew Kay, James J. Higgins, and Jacob O. Wobbrock. 2021. An Aligned Rank Transform Procedure for Multifactor Contrast Tests. *Proceedings of the 34rd Annual ACM Symposium on User Interface Software and Technology* (2021). https://doi.org/10.1145/3472749.3474784

[14] Shariff Am Faleel, Michael Gammon, Kevin Fan, Da-Yuan Huang, Wei Li, and Pourang Irani. 2021. HPUI: Hand Proximate User Interfaces for One-Handed Interactions on Head Mounted Displays. *IEEE Transactions on Visualization and Computer Graphics* 27, 11 (Nov. 2021), 4215–4225. https://doi.org/10.1109/TVCG.2021.3106493

[15] Shariff A. M. Faleel, Michael Gammon, Yumiko Sakamoto, Carlo Menon, and Pourang Irani. 2020. User gesture elicitation of common smartphone tasks for hand proximate user interfaces. In *Proceedings of the 11th Augmented Human International Conference*. ACM, Winnipeg Manitoba Canada, 1–8. https://doi.org/10.1145/3396339.3396363

[16] Jacqui Fashimpaur, Kenrick Kin, and Matt Longest. 2020. PinchType: Text Entry for Virtual and Augmented Reality Using Comfortable Thumb to Fingertip Pinches. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–7. https://doi.org/10.1145/3334480.3382888

[17] Bruno Fruchard, Paul Strohmeier, Roland Bennewitz, and Jürgen Steimle. 2021. Squish This: Force Input on Soft Surfacesfor Visual Targeting Tasks. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. Association for Computing Machinery, New York, NY, USA, 1–9. https://doi.org/10.1145/3411764.3445623

[18] Jenny Gabel, Susanne Schmidt, Oscar Ariza, and Frank Steinicke. 2023. Redirecting Rays: Evaluation of Assistive Raycasting Techniques in Virtual Reality. In *29th ACM Symposium on Virtual Reality Software and Technology*. ACM, Christchurch New Zealand, 1–11. https://doi.org/10.1145/3611659.3615716

[19] Mayank Goel, Alex Jansen, Travis Mandel, Shwetak N. Patel, and Jacob O. Wobbrock. 2013. ContextType: using hand posture information to improve mobile touch screen text entry. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, Paris France, 2795–2798. https://doi.org/10.1145/2470654.2481386

[20] Aakar Gupta and Ravin Balakrishnan. 2016. DualKey: Miniature Screen Text Entry via Finger Identification. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, San Jose California USA, 59–70. https://doi.org/10.1145/2858036.2858052

[21] Kyohei Hakka, Toshiya Isomoto, and Buntarou Shizuki. 2019. One-Handed Interaction Technique for Single-Touch Gesture Input on Large Smartphones. In *Symposium on Spatial User Interaction*. ACM, New Orleans LA USA, 1–2. https://doi.org/10.1145/3357251.3358750

[22] Nur Al-huda Hamdan, Jeffrey R. Blum, Florian Heller, Ravi Kanth Kosuru, and Jan Borchers. 2016. Grabbing at an Angle: Menu Selection for Fabric Interfaces. In *Proceedings of the 2016 ACM International Symposium on Wearable Computers (ISWC '16)*. Association for Computing Machinery, New York, NY, USA, 1–7. https://doi.org/10.1145/2971763.2971786

[23] Teng Han, Jie Liu, Khalad Hasan, Mingming Fan, Junhyeok Kim, Jiannan Li, Xiangmin Fan, Feng Tian, Edward Lank, and Pourang Irani. 2019. PinchList: Leveraging Pinch Gestures for Hierarchical List Navigation on Smartphones. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, Glasgow Scotland Uk, 1–13. https://doi.org/10.1145/3290605.3300731

[24] Chris Harrison, Julia Schwarz, and Scott E. Hudson. 2011. TapSense: enhancing finger interaction on touch surfaces. In *Proceedings of the 24th annual ACM symposium on User interface software and technology*. ACM, Santa Barbara California USA, 627–636. https://doi.org/10.1145/2047196.2047279

[25] SG Hart. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Human mental workload/Elsevier* (1988).

[26] Sandra G. Hart. 2006. Nasa-Task Load Index (NASA-TLX); 20 Years Later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 50, 9 (2006), 904–908. https://doi.org/10.1177/154193120605000909

[27] Juan David Hincapié-Ramos, Xiang Guo, Paymahn Moghadasian, and Pourang Irani. 2014. Consumed endurance: a metric to quantify arm fatigue of mid-air interactions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, Toronto Ontario Canada, 1063–1072. https://doi.org/10.1145/2556288.2557130

[28] Masoumehsadat Hosseini, Tjado Ihmels, Ziqian Chen, Marion Koelle, Heiko Müller, and Susanne Boll. 2023. Towards a Consensus Gesture Set: A Survey of Mid-Air Gestures in HCI for Maximized Agreement Across Domains. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. ACM, Hamburg Germany, 1–24. https://doi.org/10.1145/3544548.3581420

[29] Da-Yuan Huang, Liwei Chan, Shuo Yang, Fan Wang, Rong-Hao Liang, De-Nian Yang, Yi-Ping Hung, and Bing-Yu Chen. 2016. DigitSpace: Designing Thumb-to-Fingers Touch Interfaces for One-Handed and Eyes-Free Interactions. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, San Jose California USA, 1526–1537. https://doi.org/10.1145/2858036.2858483

[30] Robert J. K. Jacob. 1990. What you look at is what you get: eye movement-based interaction techniques. In *Proceedings of the SIGCHI conference on Human factors in computing systems Empowering people - CHI '90*. ACM Press, Seattle,

Washington, United States, 11–18. https://doi.org/10.1145/97243.97246

[31] Haiyan Jiang, Dongdong Weng, Xiaonuo Dongye, and Yue Liu. 2024. PinchText: One-Handed Text Entry Technique Combining Pinch Gestures and Hand Positions for Head-Mounted Displays. *International Journal of Human–Computer Interaction* 40, 2 (2024), 278–294. https://doi.org/10.1080/10447318.2022.2115333 arXiv:https://doi.org/10.1080/10447318.2022.2115333

[32] Mengqi Jiang, Vijayakumar Nanjappan, Hai-Ning Liang, and Martijn Ten Bhömer. 2022. GesFabri: Exploring Affordances and Experience of Textile Interfaces for Gesture-based Interaction. *Proceedings of the ACM on Human-Computer Interaction* 6, EICS (June 2022), 1–23. https://doi.org/10.1145/3534522

[33] Anjali Khurana, Michael Glueck, and Parmit K. Chilana. 2024. Do I Just Tap My Headset? How Novice Users Discover Gestural Interactions with Consumer Augmented Reality Applications. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 7, 4, Article 165 (Jan. 2024), 28 pages. https://doi.org/10.1145/3631451

[34] Jinwook Kim, Sangmin Park, Qiushi Zhou, Mar Gonzalez-Franco, Jeongmi Lee, and Ken Pfeuffer. 2025. PinchCatcher: Enabling Multi-selection for Gaze+Pinch. In *CHI 2025 - Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, United States. https://doi.org/10.1145/3706598.3713530 Publisher Copyright: © 2025 Copyright held by the owner/author(s).; 2025 CHI Conference on Human Factors in Computing Systems, CHI 2025 ; Conference date: 26-04-2025 Through 01-05-2025.

[35] Woojoo Kim and Shuping Xiong. 2024. TouchView: Mid-Air Touch on Zoomable 2D View for Distant Freehand Selection on Virtual Reality User Interface. https://doi.org/10.20944/preprints202408.1478.v1

[36] Myron W. Krueger, Thomas Gionfriddo, and Katrin Hinrichsen. 1985. VIDEOPLACE—an artificial reality. *ACM SIGCHI Bulletin* 16, 4 (April 1985), 35–40. https://doi.org/10.1145/1165385.317463

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

[38] Yi Li, Benjamin Tag, Shaozhang Dai, Robert Crowther, Tim Dwyer, Pourang Irani, and Barrett Ens. 2024. NICER: A New and Improved Consumed Endurance and Recovery Metric to Quantify Muscle Fatigue of Mid-Air Interactions. *ACM Trans. Graph.* 43, 4, Article 102 (July 2024), 14 pages. https://doi.org/10.1145/3658230

[39] Yiqin Lu, Chun Yu, and Yuanchun Shi. 2020. Investigating Bubble Mechanism for Ray-Casting to Improve 3D Target Acquisition in Virtual Reality. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 35–43. https://doi.org/10.1109/VR46266.2020.00021 ISSN: 2642-5254.

[40] Alexander Masurovsky, Paul Chojecki, Detlef Runde, Mustafa Lafci, David Przewozny, and Michael Gaebler. 2020. Controller-Free Hand Tracking for Grab-and-Place Tasks in Immersive Virtual Reality: Design Elements and Their Empirical Study. *Multimodal Technologies and Interaction* 4, 4 (Dec. 2020), 91. https://doi.org/10.3390/mti4040091 Number: 4 Publisher: Multidisciplinary Digital Publishing Institute.

[41] Daniel Medeiros, Romane Dubus, Julie Williamson, Graham Wilson, Katharina Pöhlmann, and Mark McGill. 2023. Surveying the Social Comfort of Body, Device, and Environment-Based Augmented Reality Interactions in Confined Passenger Spaces Using Mixed Reality Composite Videos. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 7, 3, Article 113 (Sept. 2023), 25 pages. https://doi.org/10.1145/3610923

[42] Mark R Mine. 1995. Virtual Environment Interaction Techniques. (1995). https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=69ff1367d0221357a806d3c05df2b787ed90bdb7

[43] Pranav Mistry, Pattie Maes, and Liyan Chang. 2009. WUW - wear Ur world: a wearable gestural interface. In *CHI '09 Extended Abstracts on Human Factors in Computing Systems*. ACM, Boston MA USA, 4111–4116. https://doi.org/10.1145/1520340.1520626

[44] Meredith Ringel Morris, Jacob O. Wobbrock, and Andrew D. Wilson. 2010. Understanding Users' Preferences for Surface Gestures. 261–268. https://www.microsoft.com/en-us/research/publication/understanding-users-preferences-surface-gestures/

[45] Florian Müller, Niloofar Dezfuli, Max Mühlhäuser, Martin Schmitz, and Mohammadreza Khalilbeigi. 2015. Palm-based Interaction with Head-mounted Displays. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct* (Copenhagen, Denmark) *(MobileHCI '15)*. Association for Computing Machinery, New York, NY, USA, 963–965. https://doi.org/10.1145/2786567.2794314

[46] Florian Müller, Daniel Schmitt, Andrii Matviienko, Dominik Schön, Sebastian Günther, Thomas Kosch, and Martin Schmitz. 2023. TicTacToes: Assessing Toe Movements as an Input Modality. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 520, 17 pages. https://doi.org/10.1145/3544548.3580954

[47] Aunnoy K Mutasim, Anil Ufuk Batmaz, and Wolfgang Stuerzlinger. 2021. Pinch, Click, or Dwell: Comparing Different Selection Techniques for Eye-Gaze-Based Pointing in Virtual Reality. In *ACM Symposium on Eye Tracking Research and Applications*. ACM, Virtual Event Germany, 1–7. https://doi.org/10.1145/3448018.3457998

[48] Poh Kiat Ng, Meng Chauw Bee, Adi Saptari, and Nor Akramin Mohamad. 2014. A Review of Different Pinch Techniques. *Theoretical Issues in Ergonomics Science* 15, 5 (Sept. 2014), 517–533. https://doi.org/10.1080/1463922X.2013.796539

[49] Poh Kiat Ng, Adi Saptari, Qiao Boon, Meng Bee, Ka Chai, and Shiong Leh. 2014. Pinch Effort Variations with Torque, Shape, Size, Sensation and Technique. *Journal of Applied Sciences* 14 (Feb. 2014), 401–414. https://doi.org/10.3923/jas.2014.401.414

[50] Tao Ni, Doug A. Bowman, Chris North, and Ryan P. McMahan. 2011. Design and evaluation of freehand menu selection interfaces using tilt and pinch gestures. *International Journal of Human-Computer Studies* 69, 9 (Aug. 2011), 551–562. https://doi.org/10.1016/j.ijhcs.2011.05.001

[51] Tomi Nukarinen, Jari Kangas, Jussi Rantala, Olli Koskinen, and Roope Raisamo. 2018. Evaluating ray casting and two gaze-based pointing techniques for object selection in virtual reality. In *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology*. ACM, Tokyo Japan, 1–2. https://doi.org/10.1145/3281505.3283382

[52] Alex Olwal, Thad Starner, and Gowa Mainini. 2020. E-Textile Microinteractions: Augmenting Twist with Flick, Slide and Grasp Gestures for Soft Electronics. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–13. https://doi.org/10.1145/3313831.3376236

[53] Siyou Pei, Alexander Chen, Jaewook Lee, and Yang Zhang. 2022. Hand Interfaces: Using Hands to Imitate Objects in AR/VR for Expressive Interactions. In *CHI Conference on Human Factors in Computing Systems*. ACM, New Orleans LA USA, 1–16. https://doi.org/10.1145/3491102.3501898

[54] Ken Pfeuffer, Jason Alexander, Ming Ki Chong, and Hans Gellersen. 2014. Gaze-touch: combining gaze with multi-touch for interaction on the same surface. In *Proceedings of the 27th annual ACM symposium on User interface software and technology (UIST '14)*. Association for Computing Machinery, New York, NY, USA, 509–518. https://doi.org/10.1145/2642918.2647397

[55] Ken Pfeuffer, Benedikt Mayer, Diako Mardanbegi, and Hans Gellersen. 2017. Gaze + pinch interaction in virtual reality. In *Proceedings of the 5th Symposium on Spatial User Interaction*. ACM, Brighton United Kingdom, 99–108. https://doi.org/10.1145/3131277.3132180

[56] Ken Pfeuffer, Jan Obernolte, Felix Dietz, Ville Mäkelä, Ludwig Sidenmark, Pavel Manakhov, Minna Pakanen, and Florian Alt. 2023. PalmGazer: Unimanual Eye-hand Menus in Augmented Reality. In *Proceedings of the 2023 ACM Symposium on Spatial User Interaction*. ACM, Sydney NSW Australia, 1–12. https://doi.org/10.1145/3607822.3614523

[57] Bastian Pfleging, Stefan Schneegass, and Albrecht Schmidt. 2012. Multimodal interaction in the car: combining speech and gestures on the steering wheel. In *Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '12)*. Association for Computing Machinery, New York, NY, USA, 155–162. https://doi.org/10.1145/2390256.2390282

[58] Alexander Plopski, Teresa Hirzle, Nahal Norouzi, Long Qian, Gerd Bruder, and Tobias Langlotz. 2022. The Eye in Extended Reality: A Survey on Gaze Interaction and Eye Tracking in Head-worn Extended Reality. *ACM Comput. Surv.* 55, 3, Article 53 (mar 2022), 39 pages. https://doi.org/10.1145/3491207

[59] Vít Rusnák, Caroline Appert, Olivier Chapuis, and Emmanuel Pietriga. 2018. Designing Coherent Gesture Sets for Multi-scale Navigation on Tabletops. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, Montreal QC Canada, 1–12. https://doi.org/10.1145/3173574.3173716

[60] Kadek Ananta Satriadi, Barrett Ens, Maxime Cordeil, Bernhard Jenny, Tobias Czauderna, and Wesley Willett. 2019. Augmented Reality Map Navigation with Freehand Gestures. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 593–603. https://doi.org/10.1109/VR.2019.8798340 ISSN: 2642-5254.

[61] Martin Schmitz, Sebastian Günther, Dominik Schön, and Florian Müller. 2022. Squeezy-Feely: Investigating Lateral Thumb-Index Pinching as an Input Modality. In *CHI Conference on Human Factors in Computing Systems*. ACM, New Orleans LA USA, 1–15. https://doi.org/10.1145/3491102.3501981

[62] Dominik Schön, Thomas Kosch, Florian Müller, Martin Schmitz, Sebastian Günther, Lukas Bommhardt, and Max Mühlhäuser. 2023. Tailor Twist: Assessing Rotational Mid-Air Interactions for Augmented Reality. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 400, 14 pages. https://doi.org/10.1145/3544548.3581461

[63] Shayle R. Searle, F. Michael Speed, and George A. Milliken. 1980. Population Marginal Means in the Linear Model: An Alternative to Least Squares Means. *The American Statistician* 34, 4 (1980), 216–221.

[64] Baris Serim and Giulio Jacucci. 2016. Pointing while Looking Elsewhere: Designing for Varying Degrees of Visual Guidance during Manual Input. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. Association for Computing Machinery, New York, NY, USA, 5789–5800. https://doi.org/10.1145/2858036.2858480

[65] Mary E. Sesto, Curtis B. Irwin, Karen B. Chen, Amrish O. Chourasia, and Douglas A. Wiegmann. 2012. Effect of Touch Screen Button Size and Spacing on Touch Characteristics of Users With and Without Disabilities. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 54, 3 (June 2012), 425–436. https://doi.org/10.1177/0018720811433831

[66] Adwait Sharma, Joan Sol Roo, and Jürgen Steimle. 2019. Grasping Microgestures: Eliciting Single-hand Microgestures for Handheld Objects. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, Glasgow Scotland Uk, 1–13. https://doi.org/10.1145/3290605.3300632

[67] Rongkai Shi, Jialin Zhang, Yong Yue, Lingyun Yu, and Hai-Ning Liang. 2023. Exploration of Bare-Hand Mid-Air Pointing Selection Techniques for Dense Virtual Reality Environments. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*. ACM, Hamburg Germany, 1–7. https://doi.org/10.1145/3544549.3585615

[68] Linda E. Sibert and Robert J. K. Jacob. 2000. Evaluation of eye gaze interaction. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems (CHI '00)*. Association for Computing Machinery, New York, NY, USA, 281–288. https://doi.org/10.1145/332040.332445

[69] Mohamed Soliman, Franziska Mueller, Lena Hegemann, Joan Sol Roo, Christian Theobalt, and Jürgen Steimle. 2018. FingerInput: Capturing Expressive Single-Hand Thumb-to-Finger Microgestures. In *Proceedings of the 2018 ACM International Conference on Interactive Surfaces and Spaces*. ACM, Tokyo Japan, 177–187. https://doi.org/10.1145/3279778.3279799

[70] Zhaomou Song, John J. Dudley, and Per Ola Kristensson. 2023. HotGestures: Complementing Command Selection and Use with Delimiter-Free Gesture-Based Shortcuts in Virtual Reality. *IEEE Transactions on Visualization and Computer Graphics* 29, 11 (Nov. 2023), 4600–4610. https://doi.org/10.1109/TVCG.2023.3320257

[71] Hemant Bhaskar Surale, Fabrice Matulic, and Daniel Vogel. 2019. Experimental Analysis of Barehand Mid-air Mode-Switching Techniques in Virtual Reality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, Glasgow Scotland Uk, 1–14. https://doi.org/10.1145/3290605.3300426

[72] Eduardo Velloso, Jayson Turner, Jason Alexander, Andreas Bulling, and Hans Gellersen. 2015. An Empirical Investigation of Gaze Selection in Mid-Air Gestural 3D Manipulation. In *Human-Computer Interaction – INTERACT 2015 (Lecture Notes in Computer Science)*, Julio Abascal, Simone Barbosa, Mirko Fetter, Tom Gross, Philippe Palanque, and Marco Winckler (Eds.). Springer International Publishing, Cham, 315–330. https://doi.org/10.1007/978-3-319-22668-2_25

[73] Panagiotis Vogiatzidakis and Panayiotis Koutsabasis. 2018. Gesture Elicitation Studies for Mid-Air Interaction: A Review. *Multimodal Technologies and Interaction* 2, 4 (2018). https://doi.org/10.3390/mti2040065

[74] Uta Wagner, Mathias Lystbæk, Pavel Manakhov, Jens Emil Grønbæk, Ken Pfeuffer, and Hans Gellersen. 2023. A Fitts' Law Study of Gaze-Hand Alignment for Selection in 3D User Interfaces. In *2023 CHI Conference on Human Factors in Computing Systems*. ACM. https://doi.org/10.1145/3544548.3581423

[75] Colin Ware and Harutune H. Mikaelian. 1986. An evaluation of an eye tracker as a device for computer input2. In *Proceedings of the SIGCHI/GI Conference on Human Factors in Computing Systems and Graphics Interface (CHI '87)*. Association for Computing Machinery, New York, NY, USA, 183–188. https://doi.org/10.1145/29933.275627

[76] Kieran Waugh, Mark McGill, and Euan Freeman. 2022. Push or Pinch? Exploring Slider Control Gestures for Touchless User Interfaces. In *Nordic Human-Computer Interaction Conference*. ACM, Aarhus Denmark, 1–10. https://doi.org/10.1145/3546155.3546702

[77] Yannick Weiß, Daniel Hepperle, Andreas SieB, and Matthias Wolfel. 2018. What User Interface to Use for Virtual Reality? 2D, 3D or Speech–A User Study. In *2018 International Conference on Cyberworlds (CW)*. IEEE, Singapore, 50–57. https://doi.org/10.1109/CW.2018.00021

[78] Pierre Wellner. 1991. The DigitalDesk calculator: tangible manipulation on a desk top display. In *Proceedings of the 4th annual ACM symposium on User interface software and technology (UIST '91)*. Association for Computing Machinery, New York, NY, USA, 27–33. https://doi.org/10.1145/120782.120785

[79] 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*. Association for Computing Machinery, New York, NY, USA, 143–146.

[80] Jacob O. Wobbrock, Meredith Ringel Morris, and Andrew D. Wilson. 2009. User-defined gestures for surface computing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09)*. Association for Computing Machinery, New York, NY, USA, 1083–1092. https://doi.org/10.1145/1518701.1518866

[81] Dennis Wolf, Jan Gugenheimer, Marco Combosch, and Enrico Rukzio. 2020. Understanding the Heisenberg Effect of Spatial Interaction: A Selection Induced Error for Spatially Tracked Input Devices. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–10. https://doi.org/10.1145/3313831.3376876

[82] Zhengdi Yu, Stefanos Zafeiriou, and Tolga Birdal. 2025. Dyn-HaMR: Recovering 4D Interacting Hand Motion from a Dynamic Camera. arXiv:2412.12861 [cs.CV] https://arxiv.org/abs/2412.12861

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

[84] Shumin Zhai, Carlos Morimoto, and Steven Ihde. 1999. Manual and gaze input cascaded (MAGIC) pointing. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems (CHI '99)*. Association for Computing Machinery, New York, NY, USA, 246–253. https://doi.org/10.1145/302979.303053

[85] Fengyuan Zhu, Ludwig Sidenmark, Mauricio Sousa, and Tovi Grossman. 2023. PinchLens: Applying Spatial Magnification and Adaptive Control-Display Gain for Precise Selection in Virtual Reality. In *2023 IEEE International Symposium*

on Mixed and Augmented Reality (ISMAR). 1221–1230. https://doi.org/10.1109/ISMAR59233.2023.00139

A Appendix

Table 7. NASA-TLX means and SDs per condition for the slider (top) and button (bottom).

Selection ⊕ Manipulation	Mental Demand	Physical Demand	Temporal Demand	Performance	Effort	Frustration
direct ⊕ pinch	32.22 (19.34)	45.83 (27.67)	32.22 (26.25)	31.94 (17.92)	46.67 (23.33)	28.89 (19.97)
direct ⊕ continuous-pinch	32.78 (24.75)	42.78 (29.96)	28.33 (24.25)	25 (15.62)	50.28 (21.93)	29.72 (23.04)
direct ⊕ finger	25.28 (19.59)	29.44 (18.46)	31.39 (22.93)	23.61 (18.21)	42.22 (18.25)	22.78 (16.47)
gaze ⊕ pinch	30.83 (22.05)	23.89 (19.29)	23.89 (20.83)	18.89 (17.87)	38.61 (25.6)	21.67 (18.31)
gaze ⊕ continuous-pinch	23.06 (18.16)	28.61 (22.28)	24.44 (21.07)	23.33 (17.66)	34.72 (20.76)	20.56 (14.44)
gaze ⊕ finger	40.28 (29.03)	39.44 (26.56)	37.22 (24.69)	29.17 (24.21)	54.17 (23.53)	38.61 (26.34)
raycast ⊕ pinch	35.83 (23.9)	40.83 (26.42)	30.56 (21.07)	29.17 (20.16)	48.33 (25.15)	31.39 (20.35)
raycast ⊕ continuous-pinch	32.78 (22.18)	33.06 (23.4)	33.61 (20.71)	22.22 (16.02)	42.78 (20.02)	23.33 (18.86)
raycast ⊕ finger	50.83 (27.61)	46.67 (34.6)	34.72 (20.83)	39.44 (23.82)	65.83 (24.27)	49.44 (31.05)
direct ⊕ pinch	20.56 (21.27)	28.06 (21.5)	28.33 (27.06)	16.94 (16.99)	34.72 (24.76)	23.33 (19.78)
direct ⊕ continuous-pinch	23.33 (18.55)	31.67 (21.14)	25.56 (21.34)	16.94 (19.03)	38.61 (18.77)	25.83 (25.68)
direct ⊕ finger	12.78 (8.61)	16.94 (15.82)	25.83 (25.39)	11.67 (12.95)	26.39 (20.42)	13.06 (7.5)
gaze ⊕ pinch	17.78 (18.65)	13.06 (10.87)	24.17 (22.77)	10.83 (10.04)	25.56 (23.38)	17.22 (12.63)
gaze ⊕ continuous-pinch	15.28 (12.18)	15.56 (13.81)	23.61 (24.6)	12.78 (13.74)	22.22 (16.29)	13.06 (8.93)
gaze ⊕ finger	24.17 (17.84)	26.94 (22.7)	28.33 (24.73)	22.5 (24.57)	38.61 (25.19)	25.83 (18.81)
raycast ⊕ pinch	21.11 (15.77)	26.11 (19.82)	25.83 (23.09)	15.28 (11.69)	31.39 (21.06)	19.44 (14.44)
raycast ⊕ continuous-pinch	17.22 (11.4)	24.17 (17.43)	25.56 (24.67)	10.56 (10.27)	24.44 (12.71)	15.83 (10.33)
raycast ⊕ finger	49.17 (29.91)	51.94 (33.83)	38.89 (26.43)	40.83 (22.9)	64.72 (26.76)	46.39 (32.03)