PhoneInVR: An Evaluation of Spatial Anchoring and Interaction Techniques for Smartphone Usage in Virtual Reality

Fengyuan Zhu
University of Toronto
Toronto, Ontario, Canada
fyzhu@dgp.toronto.edu

Mauricio Sousa
University of Toronto
Toronto, Ontario, Canada
mauricio@dgp.toronto.edu

Ludwig Sidenmark
University of Toronto
Toronto, Ontario, Canada
lsidenmark@dgp.toronto.edu

Tovi Grossman
University of Toronto
Toronto, Ontario, Canada
tovi@dgp.toronto.edu

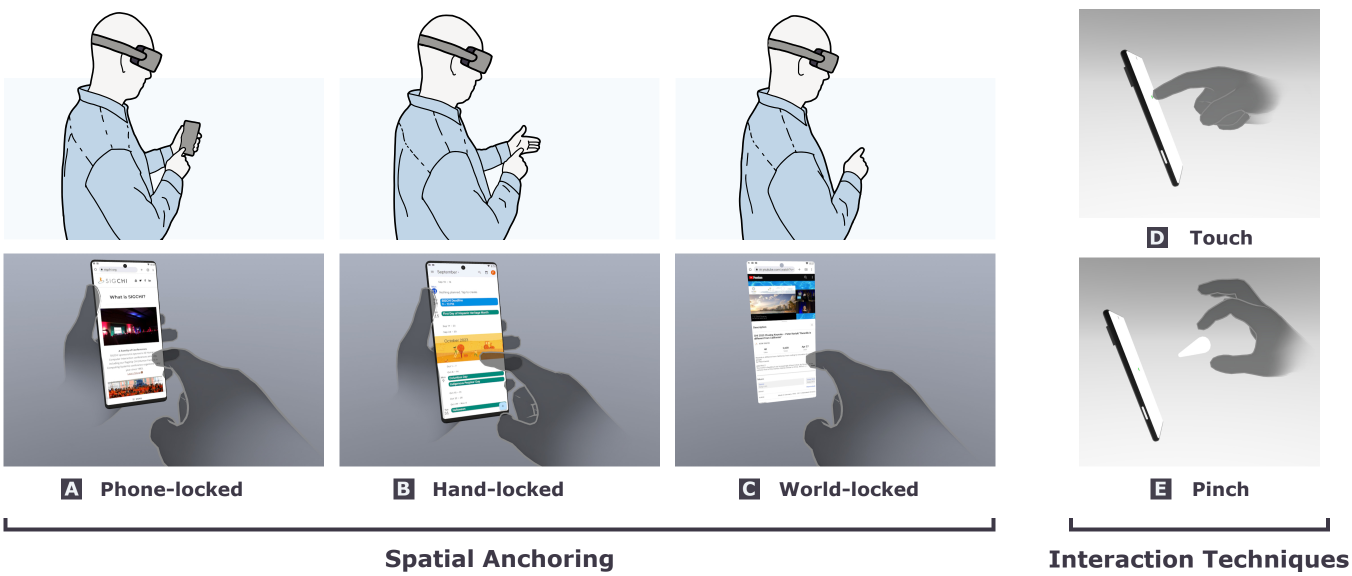

Figure 1: Overview of the spatial anchoring and interaction techniques investigated in our study. On the left, we present: A) the Phone-locked condition, where the virtual phone aligns perfectly with the physical phone; B) the Hand-locked condition, where the phone is anchored to the user's virtual hand without a corresponding physical entity; and C) the World-locked condition, where the phone is anchored at a fixed point spatially. On the right, the interaction techniques are depicted: D) Touch and E) Pinch.

ABSTRACT

When users wear a virtual reality (VR) headset, they lose access to their smartphone and accompanying apps. Past work has proposed smartphones as enhanced VR controllers, but little work has explored using existing smartphone apps and performing traditional smartphone interactions while in VR. In this paper, we consider three potential spatial anchorings for rendering smartphones in VR: On top of a tracked physical smartphone which the user holds (Phone-locked), on top of the user's empty hand, as if holding a virtual smartphone (Hand-locked), or in a static position in front of the user (World-locked). We conducted a comparative study of target acquisition, swiping, and scrolling tasks across these anchorings using direct Touch or above-the-surface Pinch. Our findings indicate that physically holding a smartphone with Touch improves accuracy and speed for all tasks, and Pinch performed better with virtual smartphones. These findings provide a valuable foundation to enable smartphones in VR.

CCS CONCEPTS

• **Human-centered computing** → **Interaction techniques**; **Virtual reality**; **User studies**.

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

ACM ISBN 979-8-4007-0330-0/24/05…$15.00
https://doi.org/10.1145/3613904.3642582

KEYWORDS

Virtual Reality, Smartphones, Touch Input

ACM Reference Format:
Fengyuan Zhu, Mauricio Sousa, Ludwig Sidenmark, and Tovi Grossman. 2024. PhoneInVR: An Evaluation of Spatial Anchoring and Interaction Techniques for Smartphone Usage in Virtual Reality. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24), May 11–16, 2024, Honolulu, HI, USA.* ACM, New York, NY, USA, 16 pages. https://doi.org/10.1145/3613904.3642582

1 INTRODUCTION

Modern smartphones give us immediate and convenient access to our contacts, communications, applications, and personal data. However, when using head-mounted displays, interacting with mobile devices and the information they hold is still cumbersome, and there is no consensus on how to perform these interactions. To glance at or interact with their personal devices, people either need to leave the immersion of the virtual environment by removing their headset, or use video pass-through techniques [40, 58]. Although video pass-through technologies are improving, they may still impede on interaction due to low resolution, latency, or distortion. Alternatively, users can interact with custom VR versions of smartphone apps, but this may require learning unfamiliar interface designs and interaction styles, and also introduces additional development costs for app developers. We argue that a promising alternative would be to allow people to interact directly with their smartphones in VR as a smartphone device.

Previous research has explored various approaches to integrating smartphones into virtual environments, such as enabling phone tracking in virtual space, using the phone as a 6-degree-of-freedom input device [48, 49, 52, 75], or expanding the smartphone view using a headset to enhance the experience [26, 54]. Furthermore, researchers have begun to investigate how to regain access to our smartphones in virtual environments [9, 23, 39]. These works enable phone usage in VR by replicating the interaction and display of smartphones as users would experience in the physical world, including tracking and touch functionalities.

While these approaches act as a good starting point, they raise further questions: 1) *Is there a need to strictly follow real-world smartphone usage conventions in a virtual space?*, 2) *Could simulating a virtual smartphone and allowing users to interact with the virtual instance be a viable solution to meet user needs?*, and 3) *Would mirroring the smartphone into a common spatially anchored user interface be more effective to enable access to the smartphone in VR?*

Considering these questions in light of existing research, we found a lack of comparative studies that address these specific issues. Most current research compares results in contexts such as target acquisition [48, 49, 61, 78], typing [9, 17, 26], or specialized tasks such as data visualization [22]. General operations, such as swipe gestures on a smartphone or scrolling through continuous content, have not been evaluated in the context of "using a smartphone as a smartphone in VR".

This paper addresses the existing gap in understanding smartphone interactions in VR. Specifically, we explore various spatial anchoring strategies to enable phone interfaces with different interaction techniques. We first classify possible smartphone interactions in VR into three main anchoring categories: *Phone-locked*, where the virtual screen is anchored directly above a physical smartphone; *Hand-locked*, where the virtual screen is positioned over a user's empty hand; and *World-locked*, where the virtual screen maintains a fixed position in the virtual world. In addition, we also considered two possible input modalities: direct *Touch*, mimicking typical phone input, and above-the-surface *Pinch*, a prominent input mechanism for immersive environments. In summary, we conducted a study to answer the following research questions:

RQ1: *How do different spatial anchoring strategies affect the performance of common smartphone interactions?*
RQ2: *How do different interaction techniques influence user experience and task efficiency?*
RQ3: *What are user preferences regarding the combination of spatial anchoring methods and input modalities?*

We compared our proposed spatial anchorings and interaction techniques in a study consisting of tasks that mirror everyday smartphone interactions (target acquisition, swiping, and content scrolling). Our results suggest that participants performed better in all tasks using direct touch interaction when holding a physical phone than in virtual representations, likely due to the haptic sensation and feedback it provides. However, if the user cannot access a physical phone for interaction, the indirect pinch interaction leads to higher accuracy for all tasks compared to Touch. In summary, our paper contributes to a deeper understanding of how smartphones can be used as smartphones within VR. Specifically, we contribute:

(1) An exploration of strategies to enable the use of smartphones in virtual reality for spatial anchoring and input modalities.
(2) Results of a user study that evaluates combinations of spatial anchoring techniques and input modalities across three everyday smartphone tasks.

2 RELATED WORK

Our research builds on previous work that adapts different strategies to anchoring spatial user interfaces (UIs), including fixed in space, attached to the body, or attached to physical objects. Additionally, we build on existing research focused on enabling phones within spatial environments.

2.1 Spatial User Interface Interactions and Anchoring

Spatial UIs allow for interactions between elements from physical, virtual, and immersive environments. This creates opportunities for these interactive systems to work as multi-device environments [70], and cross-reality systems [5, 79] that allow cross-device interactions [20], or simply using devices as controllers [9, 48, 77]. These opportunities create new challenges for interaction, as devices must be represented in multiple environments. While a smartphone is easily represented and interacted with in a physical environment, it is less obvious how to anchor it and interact with it in a virtual environment.

Inspired by desktop UIs, spatial user interfaces are often represented in VR as 3D menus that are spatially anchored to the surrounding virtual environment [12, 35, 41, 42, 56, 65–67]. Users commonly interact with these interfaces through ray-casting techniques enabled through controller or hand-tracking for interaction at a distance combined with a button click or hand gesture (e.g. pinching) as a trigger [2], or via *Direct Touch* techniques that aim to mimic the way users interact with physical buttons and other controls by "pressing" virtual widgets [2]. These *spatial UIs* lay the foundation of interaction with modern HMDs and are present in virtually all commercial VR devices (e.g., Meta Quest 2[1], PlayStation VR 2[2] or HTC Vive Pro[3]). Our work treats these "World-locked" interfaces as a baseline for smartphone interaction.

Although world-locked spatial UIs are efficient for system control, they can occlude other parts of the environments, and their anchoring to the world can be troublesome if users switch environments and lack tactile feedback as their physical counterparts. As such, researchers have proposed body-attached UIs that follow the user and enable haptic feedback by pressing on one's body [41]. These interfaces also benefit from our proprioception, as widgets are attached to specific parts of the body, helping the memorability of the interface and allowing interaction without looking [51, 71, 73]. Body-attached interfaces have been proposed for various easily accessible body parts such as fingers [18, 29], the palm [30, 68, 76], forearm [6, 14, 31, 32, 55, 60, 76], and different purposes such as cursor movements [69], text entry [30, 68, 76], and menu selections [18, 28, 30, 76]. Our work treats these "Hand-locked" interfaces as a suitable virtual representation of the smartphone interaction when no physical smartphone is available.

Previous work has shown that operating a spatial UI without haptics in VR can be a challenging task [46]. Chan et al. [21] found that users perform poorly in determining depth along the viewing axis of spatial targets without haptics. Thus, researchers have proposed physical surfaces to support touch interaction within virtual environments [4, 11, 72]. The most straightforward approach is to hold a physical object in hand, which provides real-world haptic feedback [44]. Previous work has explored how the spatially tracked tablet [3, 24, 63], smartphone [8, 9, 48, 78] could benefit the touch experience in the space environment. In our work, we treat these "Phone-locked" interfaces as directly reproduced smartphones in VR. However, this representation requires a physical smartphone to be available and tracked. Previous work has suggested that smartphone tracking errors may reduce usability due to the "fat finger" issue and virtual and physical touch timings [49, 78].

2.2 Enabling Phone Interaction in VR/AR

A significant benefit of using a physical smartphone in VR is to enable haptic feedback, often missing in VR [4, 15, 30, 71, 78]. As such, the smartphone is commonly introduced as a 6DOF controller to enable expressive and space-efficient input. Several works have introduced technical solutions to track the user's hands and the phone [9, 48, 52]. Phones have also been used as spatial controllers for specific tasks such as spatial content editing [49, 50, 52], text entry [9, 17, 26], proxy for content review [24, 26], or as input in space-constrained settings [38]. Typically, the smartphone screen is transformed into a custom UI tailored to the VR scenario rather than maintaining the traditional phone interface or phone applications.

A recent trend is to integrate daily smartphone usage into a VR environment by streaming the phone screen and incorporating hand tracking, thus emulating in VR the way users interact with their phones in the real world [9, 23]. While researchers have investigated the use of VR to enhance smartphones by varying screen size [54], accessibility testing tool [39], or to change its appearance to integrate with the virtual environment [47], few focus on performance evaluations, such as swipe gestures and content scrolling. Even design-based studies, such as those of Zhu and Grossman [77] and Zhang et al. [75], often emphasize the design space and user scenarios, while assuming that the virtual phone is always attached to a physical phone. This observation motivated us to investigate this assumption and compare smartphone anchorings and interaction techniques, offering a more comprehensive understanding of everyday tasks across various smartphone representations in VR.

3 PhoneInVR

When integrating a smartphone into VR, designers must consider how it is represented and how to interact with it. We address these considerations in the scope of the *spatial anchoring* and *interaction techniques* (Figure 2). We vary both factors based on their level of integration with a typical physical smartphone. The spatial anchorings vary from a direct relationship to a physical phone to a completely virtual smartphone, while the interaction techniques represent direct and indirect interaction with the smartphone.

3.1 Spatial Anchoring

We consider three spatial anchorings that each have different levels of representation of a physical smartphone in VR. It can be noted that two of these anchorings (phone-locked, hand-locked) are bimanual, and follow Guiard's Kinematic Chain model of bimanual control, where the non-dominant hand sets the reference frame for input and the dominant hand interacts using fine movements [27]. The third anchoring (World-locked) is unimanual.

3.1.1 Phone-Locked. A "Phone-locked" smartphone in VR is characterized by a user holding and interacting with a replicated phone from Reality in VR (Figure 1a). This approach is based on numerous previous studies that highlight the importance of tactile feedback for efficient interaction with smartphones in VR [9, 48, 49, 63]. Beyond replicating the physical experience, previous research has proposed different approaches to showing content on the VR smartphone. Some studies stream physical smartphone content directly into VR to allow users to use their smartphones while wearing HMDs [9, 48]. In contrast, other studies opt for a re-rendered VR interface distinct from the real-world phone screen to make the smartphone more flexible and to avoid potential latency issues arising from streaming the smartphone content into VR [22, 23, 36, 38, 63].

Our primary goal is to determine the effects of the phone's VR representation on typical phone-like interactions. Guided by this aim, our investigation will employ a re-rendered screen within the VR environment. This choice comes from our intention to avoid potential latency issues associated with streaming content to VR. It also addresses our concerns about the so-called "fat finger" problem in VR that could be aggravated by tracking inaccuracies.

[1]https://www.meta.com/us/quest/products/quest-2/
[2]https://www.playstation.com/en-us/ps-vr2/
[3]https://www.vive.com/us/product/#pro%20series

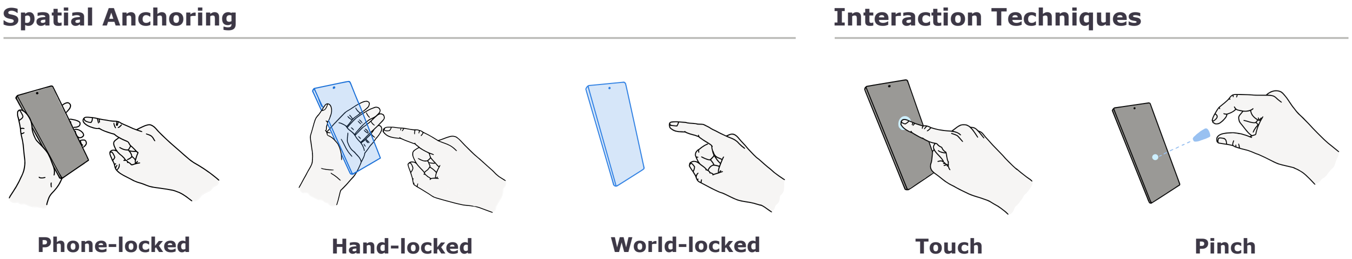

Figure 2: Spatial anchorings and interaction techniques considered in PhoneInVR.

3.1.2 Hand-Locked. A "Hand-locked" smartphone in VR aims to emulate the familiarity of a Phone-locked smartphone but without the presence of a physical smartphone that the user holds (Figure 1b). In this configuration, users will utilize their non-dominant hand to "hold" and interact with a virtual phone. The primary distinction from a Phone-locked smartphone is tangibility. No real device sits in the user's palm, transforming every touch into a "virtual" action without haptics.

One of our design considerations is whether we should apply skin-based haptic feedback to support interaction. Numerous hand-centric VR studies have stressed the importance of tangible haptic feedback and propose redirecting touch sensations to bodily areas such as the fingers [18, 29], the palm [13, 16, 28, 30], or the arm [43, 64, 76]. For studies employing hand-anchored UIs without haptic feedback, the common approach is to place the anchored UI near the hand, such as the area between both hands [7, 57]. This approach results in touch interactions against empty space rather than body regions like the palm.

Li et al. [43] posited that proprioception without haptic feedback could still enhance interaction. In our design, we mimic the interaction of real-world smartphones by having the virtual phone floating above the user's palm, directing touch actions towards it (as illustrated in Figure 2 Hand-Locked). We made this decision on two factors: the ubiquitous gesture of holding a smartphone and the fact that the typical phone's surface area will exceed the palm area. However, investigating how such proprioception without haptic feedback affects user performance is also an unexplored research avenue worth pursuing for future work.

3.1.3 World-Locked. A "World-locked" smartphone in VR is mirrored on a spatially anchored canvas (Figure 1c) [18]. Contrary to traditional VR canvases that are often rendered in larger dimensions (typically spanning meters), for the sake of consistency in our comparison with the Phone-locked and Hand-locked anchorings, we rendered the Spatial UI to match the screen size of a physical smartphone. This investigates whether a relatively compact UI can still accommodate the precise operations that users typically perform on their everyday mobile devices.

3.2 Interaction Techniques

We consider two techniques for interacting with a smartphone in VR: Touch and Pinch (Figure 2. Direct Touch is similar to the typical touch interactions with real-life smartphones, and the Pinch interaction allows indirect interaction without physical touch.

3.2.1 Touch. We use the term "touch" to denote a finger that interacts with a virtual surface, as defined by Hertel et al. [33]. While touch interactions may look visually the same regardless of whether a physical phone is present or not, the underlying techniques are slightly different. The distinctions are detailed as follows:

- **Touch without a physical phone.** Due to the lack of a physical smartphone, the touch event is purely virtual in the Hand-locked and World-locked anchorings. Therefore, we need to define the fingertip of the virtual hand in VR and calculate its collision with the virtual smartphone. We applied previous knowledge from *Touch the Droid* [78] and defined the fingertip in a contact point-based algorithm based on how the user's finger approaches the surface. A potential problem is smartphone penetration as there is no physical restriction [21]. We used a modified redirection process derived from the current Oculus Interaction SDK [4] to restrict the user's hand to stay over the surface when penetration occurs. For full details, please refer to the supplemental material.
- **Touch with a physical phone.** Integrating a physical smartphone into VR introduces challenges due to tracking inaccuracies. The touch point on the physical touch screen may not align with the collision point in VR. Additionally, an error in the z-depth direction could lead to a misalignment in the timing between the physical contact and the virtual contact between the fingertip and the screen. This can result in inconsistent visual feedback and haptic responses, which can cause problems such as the sensation of a floating touch or penetration in the virtual realm [78].

Drawing on previous research [21, 30, 76, 78], we adopted a hybrid approach. Our strategy utilizes the estimated visual point from VR to determine the touch location while using tangible contact with the phone to determine the touch timing. We integrated a target redirection technique to address potential inconsistencies, such as the virtual finger penetrating or hovering above the surface at the touch point. Specifically, if a physical touch has not occurred, the virtual hand's speed is reduced to prevent virtual contact before the physical contact takes place. Once the physical touch is registered, the virtual finger snaps onto the virtual screen. This design ensures the alignment of haptic feedback and visual cues. The details can be found in the supplemental material.

3.2.2 Pinch. The pinch gesture has become a standard interaction in VR [2]. The technique utilizes the moment when the tips of the

[4] https://developer.oculus.com/documentation/unity/unity-isdk-interaction-sdk-overview/

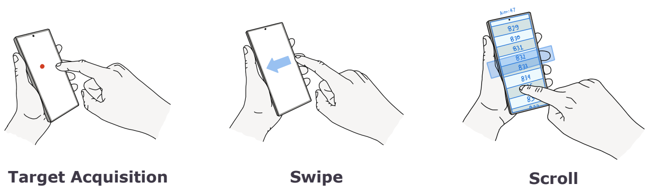

Figure 3: Evaluation tasks.

index finger and thumb come into contact with each other to cast a ray toward objects to determine an object for interaction. Previous research has proposed various methods to determine the pinch direction based on cues such as head orientation [45], hand position [62], and gaze direction [59]. These methods often assume that users interact with distant targets or expansive virtual canvases.

However, our research narrows its scope to interactions with interfaces the size of a smartphone positioned within the user's arm's reach. In such a close setting, it is challenging to directly implement the previously established methods. Therefore, we orient the ray direction perpendicular to the virtual screen surface, which we believe aligns more naturally with a user's natural inclination when performing a pinch gesture towards a phone in a VR environment.

4 EVALUATION

To answer our research questions (RQ1-RQ3), we implemented an evaluation setup that emulated the spatial anchorings and interaction techniques proposed in Figure 2. The evaluation design allowed participants to experience each combination of anchoring and interaction technique in three different smartphone tasks.

4.1 Tasks

We chose tasks that evaluate three of the most common smartphone operations: target acquisition, swiping, and scrolling (Figure 3).

4.1.1 Target Acquisition. In our target acquisition task, we aim to evaluate user performance with elements of common sizes in modern smartphone application design. Drawing on previous works [61, 78] and the smartphone design guidelines[5], we opted for target element sizes of 120*px* (~6mm) and 240*px* (~12mm) in width. These sizes represent the recommended keystroke dimension and the default icon size of the smartphone we used for our study. We designed the targets as rounded squares rather than circles, to mimic the shape of modern app icons. During the training sessions, participants were instructed to aim for the center of these targets. Additionally, to ensure a realistic spatial representation between targets, we settled on selection distances of 430*px* and 860*px*. These distances represent 1/8 and 1/4 of the screen's diagonal length.

For each combination of Technique, Anchoring, Size, and Distance, we perform nine repetitions. During the study, the subsequent target becomes visible upon the triggering of an acquisition event. This amounts to a total of 2 Techniques × 3 Anchorings × 2 Distances × 2 Sizes × 9 repetitions = 216 trials for each participant. For the Pinch technique, we also render a cursor (50px in width) to visualize the hover position before the final selection (Figure 4d). No cursor was used for the Touch technique (Figure 4a).

[5] https://developer.android.com/develop/ui/views/launch/icon_design_adaptive

4.1.2 Swipe. The swipe gesture has become integral to daily smartphone use [19, 53], and has replaced traditional buttons such as "back" and "home". Typically, directions such as Up, Down, Left, and Right are predefined and utilized to execute underlying commands, such as unlocking the device or switching content. For touch input, the swipe gesture is intuitive. However, when considering pinch input in VR, the definition and detection of a swipe gesture require further consideration. Drawing from previous research on gestures to simulate swipes in spatial UIs [1, 37], we directly projected the movement from the initiation of the pinch to its release on the screen to decide whether a swipe gesture is activated. As we project a point perpendicular to the virtual phone screen, the travel distance resembles a 1:1 movement between the pinch drag and the screen swipe. For our study, we introduced a set of filtering criteria based on two core assumptions for a successful swipe gesture:

- **Direction:** The swipe should not deviate substantially from the intended direction. Therefore, we define that the angle between the targeted direction and the actual direction executed by the user should not exceed 45 degrees.
- **Distance:** The travel distance for a valid swipe should not be short enough to confuse it with touch input. We established a minimum distance threshold of 240*px* (~1.2cm), representing the approximate size of a default Android icon.

For our study, we studied four directions: **Up**, **Down**, **Left**, **Right**. Each direction was depicted as an arrow extending 800*px* across the screen. Participants encountered a two-second countdown between each trial and were instructed to reset their right hand to its starting position after each gesture. We gathered data from 8 repetitions for each direction. Thus, we collected 2 Techniques × 3 Directions × 4 Directions × 8 repetitions = 192 swipes for every participant.

4.1.3 Scroll. In today's digital age, scrolling through content on social media platforms like Twitter[6], YouTube[7], or extensive webpages has become a routine activity for smartphone users. Unlike the traditional desktop experience [34, 74], evaluating smartphone scrolling performance requires us to consider specific features that represent the characteristics of typical smartphone applications. A defining characteristic of modern scrolling behavior is the perception of endless content, where new content is loaded as the user nears the end of what is currently visible. In our study, we designed a scrolling canvas that functions as an infinite loop to simulate this. We created endless lines of content using numbers from 0 to 1000 and set them to loop continuously during scrolling. Each line occupies a width equivalent to one-tenth of the screen's height(312*px*), ensuring that ten lines are visible at any given time. Participants were tasked with scrolling to locate target lines spaced at intervals of **5**, **10**, and **20 lines**, which correspond to half a screen, a full screen, and two screens' height, respectively.

Another feature of scrolling smartphone content is the inclusion of momentum feedback, which ensures that content continues to scroll for a short distance after the user finishes their scroll gesture. To mimic this experience, our study incorporated a default physics-based scroll function[8] to allow the content to continue scrolling

[6] https://twitter.com/
[7] https://www.youtube.com/
[8] https://docs.unity3d.com/Packages/com.unity.ugui@1.0/manual/script-ScrollRect.html

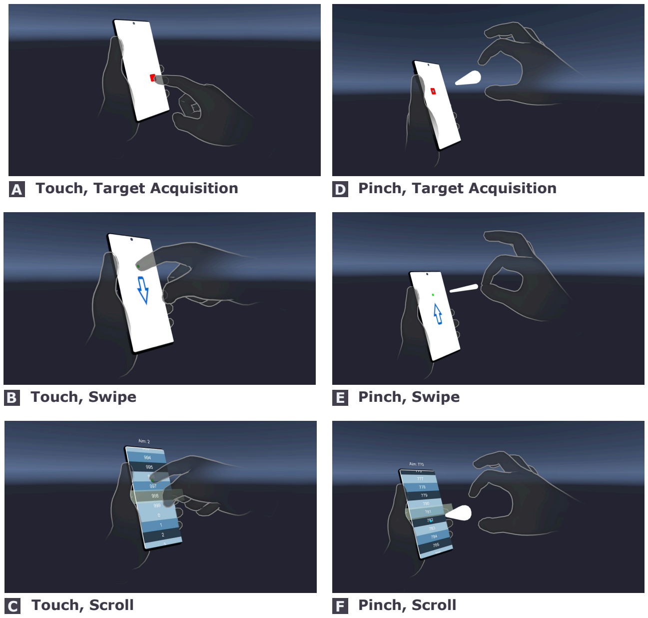

Figure 4: Virtual setup for each task and technique.

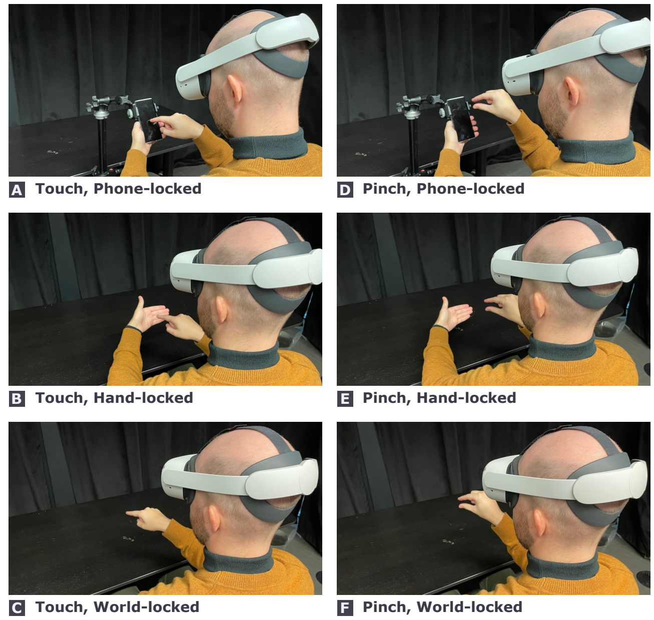

Figure 5: Physical setup for each technique and anchoring.

slightly after the user lets go of their touch or pinch gesture. Drawing from prior research [25, 74], it is essential to provide appropriate visual feedback to signify when a user has successfully identified the target. We placed a semitransparent area in the middle of the screen of 1.5 *times* the width of a content line to signify where the user has to scroll the intended target to (Figure 3). When the user scrolls the designated target within this area, the trial is set as successful, and the subsequent trial is started. The scroll canvas is first set to a random position for every trial. We then determine the target position based on the line distance, display the target line at the top of the screen, and instruct users to locate it. In summary, our scrolling study encompassed three distinct target line distances: **5**, **10**, and **20 lines**. The user performed eight repetitions for each condition. Thus, each participant completed a total of 2 Techniques × 3 Anchorings × 3 Distances × 8 repetitions = 144 scroll trials.

4.2 Setup and Apparatus

We used an Oculus Quest 2 and a Pixel 6 Pro for our study. Hand-tracking data were sourced from Oculus SDK. The Pixel 6 Pro phone is equipped with a 6.7-inch screen with a resolution of 1440 x 3120 pixels, leading to a pixel density (PPI) of 512. For conditions that required the physical phone, we deployed a background application on the smartphone. This application synchronizes the touch events from the phone to our study software through a UDP network. We developed our study software with Unity 2021.3.4f1. The software was run on a Lenovo R9000K laptop equipped with an RTX 3070 Graphics card and AMD-Ryzen 7 6800H CPU to guarantee smooth and uninterrupted performance during the study. Because spatially tracking a phone is difficult even with the current state-of-the-art techniques, and a user's hand would typically remain stationary during interaction, we locked the phone position during the study. This allows us to focus on the ideal input conditions and minimize any phone tracking errors.

4.2.1 Implementation of Anchoring. A critical aspect of our implementation was prioritizing user preference over a one-size-fits-all approach. We recognized that phone usage is a highly individual experience, and accommodating personal comfort was paramount. The following outlines our strategy for determining the optimal anchor point for different anchoring techniques.

- **Phone-Locked.** For the Phone-locked approach, the methodology is straightforward since the virtual phone is anchored to the 6-DOF tracked phone. To maintain consistency throughout our study, we instructed participants to identify a comfortable position to hold their phone. We then physically locked the phone using a tripod (Figure 5a).
- **Hand-Locked.** We placed a spatial phone within arm's reach and asked participants to find a position where they felt the phone was grasped correctly. Once they grasp the virtual phone comfortably, we establish that point in the user's hand as the anchor to attach the phone (Figure 5b). We instructed the participants to find a comfortable hand position and to maintain that posture during each task. As such, the virtual phone remained stationary.
- **World-Locked.** We began the World-locked procedure by first executing the Hand-locked approach. Users were instructed to operate the phone in a comfortable posture. Upon identifying this optimal position, we set that point in the world space as the anchor for the virtual phone and release the phone from the user's hand to the world (Figure 5c).

4.3 Participants

We recruited 24 participants aged 20 to 34 (12 female, M = 25.3 years, SD = 3.3). All but one participant reported previous experience with VR. Five reported being an expert, and five reported having intermediate experience. The remaining participants reported having limited or beginner-level experience. Our institutional ethics board approved the study and participants gave their informed consent before beginning the study.

4.4 Procedure

The first participants signed a consent form and answered a demographic questionnaire. Participants then put on the HMD, adjusted their headsets to sit comfortably, and adjusted the lenses to ensure a clear view of the virtual environment. The participants would then perform the tasks combining Technique and Representation. Participants would perform all tasks with a combination of Technique and Representation before continuing to the following combination. Participants always started with the target acquisition task, then the swiping task, and finally the scrolling task. Participants were instructed to have a one-minute rest between each task. Anchoring and Technique order was counterbalanced with a balanced Latin square. Before each task, the participants completed a training session to become familiar with the technique and task. After completing each set of tasks per condition, we asked participants to fill in a user preferences questionnaire. The questionnaire included statements scored on a 9-point Likert scale where a value of 1 meant that participants strongly disagreed with the statement and 9 indicated that they fully agreed with it. Table 1 shows the questions asked in the questionnaire and the results for each anchoring condition. The study took a total of around 90 minutes to complete. Participants received 30 CAD in compensation.

5 RESULTS

We performed our analysis for each task separately. As such, the dependent and independent variables varied per task. Unless otherwise stated, the analysis was performed with repeated measures ANOVA (RM-ANOVA) (α = .05). When the assumption of sphericity was violated, as tested with Mauchly's test, Greenhouse-Geisser corrected values were used in the analysis. Shapiro-Wilks tests were used to validate the assumption of normality. Bonferroni-corrected post-hoc tests were used when applicable. Effect sizes are reported as partial eta squared (η_p^2).

5.1 Target Acquisition Task

Our dependent variables were selection time and error distance for the target acquisition task (Figure 3 left). We performed a 4-way repeated measures ANOVA with Anchoring, Technique, Size, and Distance as independent variables.

5.1.1 Acquisition Time. We define "acquisition time" as the time between the appearance of the target and the initiation of a "touch" or "pinch". RM-ANOVA revealed a significant 4-way interaction for Anchoring × Technique × Size × Distance $F_{2,46}$=4.00, p=.025, η_p^2=.148, Figure 6). In terms of Size and Distance, post-hoc findings aligned with traditional Fitts' law results - longer distances required more time to complete (p<.001), while larger sizes resulted in shorter completion times (p<.001). Post-hoc tests further indicated that for all combinations of target size and distance, Touch combined with Phone-locked resulted in shorter touch times compared to Touch with World-locked and Hand-locked (all $p\leq$.033), showcasing the potential of haptic feedback for target acquisition. Regarding the difference between Pinch and Touch, Pinch proved consistently faster than Touch for World-locked (all $p\leq$.007). On the contrary, the Hand-locked and the Phone-locked did not show any significant differences between techniques.

We also found multiple main effects. An increased Distance ($F_{1,23}$=93.45, p<.001, η_p^2=.802), and a decreased Size ($F_{1,23}$=120.21, p<.001, η_p^2=.839) led to longer acquisition times. We also found that Anchoring had a significant main effect ($F_{2,46}$=9.52, p<.001, η_p^2=.293), where Phone-locked (1.03s) was significantly faster than Hand-locked (1.16s, p<.001) and World-locked (1.12s, p=.005).

5.1.2 Error Distance. We define the error distance as the Euclidean distance between the touch and target centers. We measured the error distance in the phone space and reported it in pixels (Figure 6). To satisfy the assumption of normality, we transformed the data using a base 10 logarithm transformation. Error distances included in figures or text represent untransformed data. RM-ANOVA of the transformed data revealed a significant four-way interaction effect for Anchoring × Technique × Size × Distance ($F_{2,46}$=4.62, p=.015, η_p^2=.167). Post-hoc tests revealed that for all combinations of Anchoring, Size, and Distance, Pinch consistently resulted in a smaller error distance compared to Touch (all $p\leq$.018). This observation is logical since the participants had access to a cursor for Pinch combined with Touch's "fat finger" issues. Regarding performance across Anchorings, we observed that the Phone-locked Touch outperformed the Hand-locked (all $p\leq$.006) but did not differ significantly from the World-locked Touch. This finding suggests that binding a virtual smartphone to the hand, as in Hand-locked, might amplify tracking jitters during touch, which increases error. Meanwhile, the similarity between the Phone-locked and World-locked can be attributed to visual hints being more dominant for target acquisition than physical cues.

We also found significant main effects for all independent variables. For Anchoring ($F_{2,46}$=38.15, p<.001, η_p^2=.624), post-hoc results showed that Hand-locked (55.3px) had a higher error distance than Phone-locked (82.0px, p<.001) and World-locked (65.4px, p<.001). World-locked also had a higher error distance than Phone-locked (p<.001). For Technique ($F_{1,23}$=107.01, p<.001, η_p^2=.823), post-hoc results showed that Pinch (55.5px) was significantly more accurate than Touch (79.6px). The results also showed that a longer target distance led to a smaller error distance ($F_{1,23}$=9.23, p=.006, η_p^2=.286). However, despite a generally low error rate, a larger size led to a higher error distance ($F_{1,23}$=106.49, p<.001, η_p^2=.822). This increase can be attributed to that determining the center of larger elements is more challenging, possibly leading to less precise touches.

5.1.3 Task Summary. The data from the Acquisition Time revealed a notable influence of haptic feedback on overall performance. Acquisition time was significantly reduced for Phone-locked conditions compared to Hand-locked and World-locked, which lack the same level of haptic feedback. Furthermore, selections with Touch on a haptic surface were faster than Pinch. However, without haptic

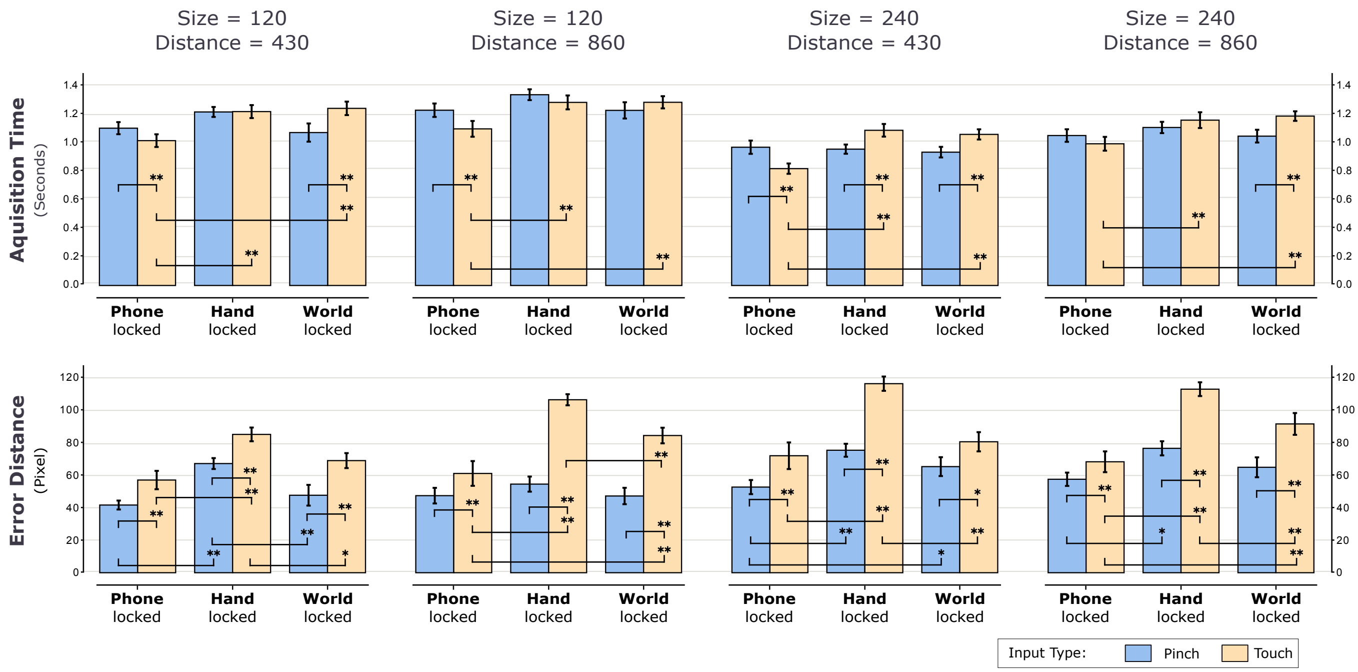

Figure 6: Results for the target acquisition task. Error bars represent the mean 95% confidence intervals. The symbol ∗ indicates $p < .05$ and ∗∗ indicates $p < .01$.

support, Pinch was faster than Touch. These results highlight the significant influence of haptic feedback on user interactions.

Regarding the error distance, pinch consistently resulted in a lower error distance across all conditions than touch. This can be attributed to the fact that pinch does not present the "fat finger" issue that touch might introduce. Furthermore, we found no significant differences between Phone-locked and World-locked conditions for Touch. This suggests that user accuracy is predominantly influenced by visual feedback rather than by the anchoring condition. The higher error distance from Touch on the Hand-locked condition than other anchorings might be due to the amplified tracking error.

5.2 Swipe Task

For the swipe task (Figure 3 middle), we measured swipe time, swipe distance, and swipe angle. We performed a 3-way repeated measures ANOVA with Anchoring, Technique, and Direction as independent variables.

5.2.1 Swipe Time. We define "swipe time" as the interval in seconds between the moment the press event is triggered and the moment the release is recorded (Figure 7). We found a significant Anchoring × Technique × Direction 3-way interaction ($F_{6,138}$=4.07, p<.001, η_p^2=.150). Post-hoc tests showed that there were no significant differences for Pinch. Phone-locked was significantly faster for the Touch technique than Hand-locked and World-locked for all conditions (all $p \leq .018$). For Pinch, the post hoc results showed no significant differences between the anchorings. For Touch, Phone-locked swipe times were consistently shorter than those in Hand-locked and World-locked conditions (all $p \leq .017$). Furthermore, for Phone-locked, all swipe times associated with the pinch gesture were significantly longer than those of the touch gesture (all $p \leq .002$). For World-locked, Pinch swipe times were significantly shorter than Touch for all directions (all $p \leq .009$) except upward (p=.853). Notably, no significant differences were observed between the pinch and touch gestures in the hand-locked condition.

We also found a significant main effect for Direction ($F_{3,33}$=10.94, p<.001, η_p^2=.322), where we found that the up direction was significantly faster than all other directions (all $p \leq .047$). Finally, we found a significant Anchoring main effect ($F_{2,46}$=14.88, p<.001, η_p^2=.393). Post-hoc results showed that Phone-locked (.35*s*) was significantly faster than Hand-locked (.44*s*, p<.001) and World-locked (.46*s*, p<.001). There was no significant difference between Hand-locked and World-locked.

5.2.2 Swipe Distance. We define the "swipe distance" as the distance between the swipe start and end points (Figure 7). RM-ANOVA showed a significant Anchoring × Technique 2-way interaction ($F_{2,46}$=29.70, p<.001, η_p^2=.566). The post-hoc tests showed that for Touch interaction, the Phone-locked Anchoring exhibited a significantly shorter swipe distance than Hand-locked (p<.001) and World-locked (p<.001). However, no significant differences were observed for Pinch. We also found that the Technique shows a significant difference for all reference frames, where a Pinch interaction leads to longer swipe distances than Touch in the Phone-locked (p<.001). However, Pinch led to significantly shorter swipe distances than Touch for Hand-locked (p<.001) and World-locked (p=.006).

We also found a significant Anchoring main effect ($F_{2,46}$=32.70, p<.001, η_p^2=.587). Post-hoc tests showed that Phone-locked (785*px*) had a significantly lower swipe distance than Hand-locked (1052*px*,

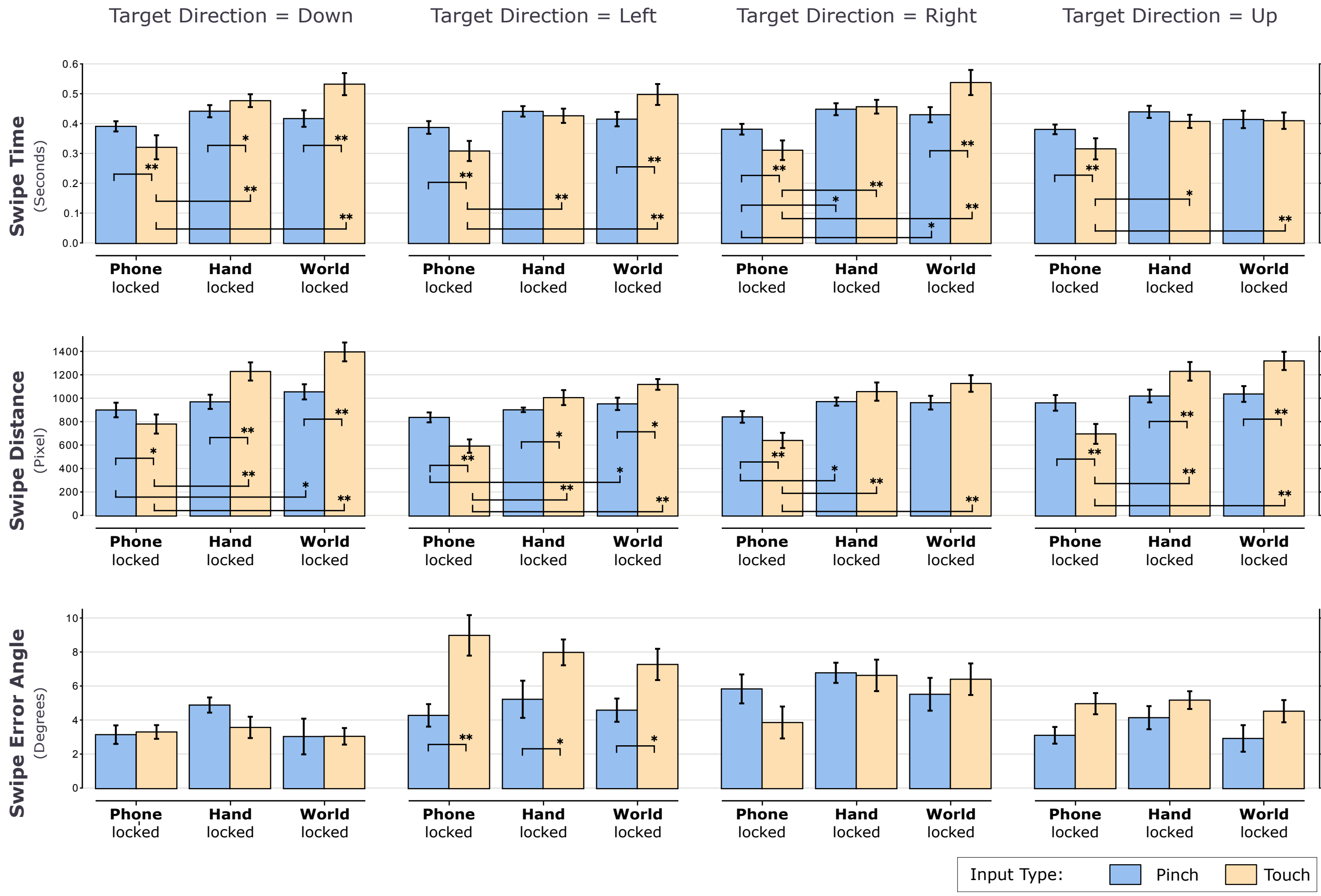

Figure 7: Swipe time, distance, and error angle results for the swipe task. Error bars represent the mean 95% confidence intervals. The symbol $*$ indicates $p < .05$ and $**$ indicates $p < .01$.

p<.001), and World-locked (1125px, p<.001). We also found a significant Direction main effect ($F_{3,69}$=13.36, p<.001, η_p^2=.367). Post-hoc tests showed that swipe distances were significantly longer for vertical directions than horizontal directions (all p≤.005).

5.2.3 Swipe Angle. The "swipe angle" is defined as the angular difference in degrees between the target direction and the vector representing the participant's swipe from start to end point (Figure 7). RM-ANOVA revealed a significant 2-way interaction between Technique × Direction ($F_{3,69}$=8.60, p<.001, η_p^2=.272). Post-hoc results showed, in general, smaller swipe angles for vertical directions. However, for all pairwise comparisons, potential differences were generally below 5°, indicating minimal practical impact.

Additionally, we found a main effect for Direction ($F_{3,33}$=15.69, p<.001, η_p^2=.588). Post-hoc results showed that the left (9.75°) and right (8.81°) directions had a significantly higher swipe angle than the down (6.27°, both p≤.031) and up (6.62°, both p≤.023) directions. We also found a significant Technique main effect ($F_{1,23}$=12.69, p=.002, η_p^2=.345. The results showed a small but statistically significant difference in which the Pinch swipe angle (4.49°) was smaller than the Touch swipe angle (5.51°).

5.2.4 Task Summary. Our analysis revealed distinct relationships between phone anchoring, input technique, and their effects on swipe-style tasks in different target directions. Specifically, the swipe time was notably influenced by both phone anchoring and input technique. When using Touch, the swipe times for both the World-locked and Hand-locked conditions were substantially longer compared to those executed on the physical screen. In comparison between Pinch and Touch techniques, Phone-locked Anchoring consistently showed shorter swipe times for Touch input. In contrast, the World-locked had shorter swipe times with Pinch. The Hand-locked condition showed no significant differences between the techniques. Our data indicated statistical significance under various conditions for the swipe distance and error angles. However, it is crucial to note that since swipe gestures predominantly manifest as cross-screen actions with a tolerance for variations in angle and distance, the differences we measured might not have significant implications for real-world applications.

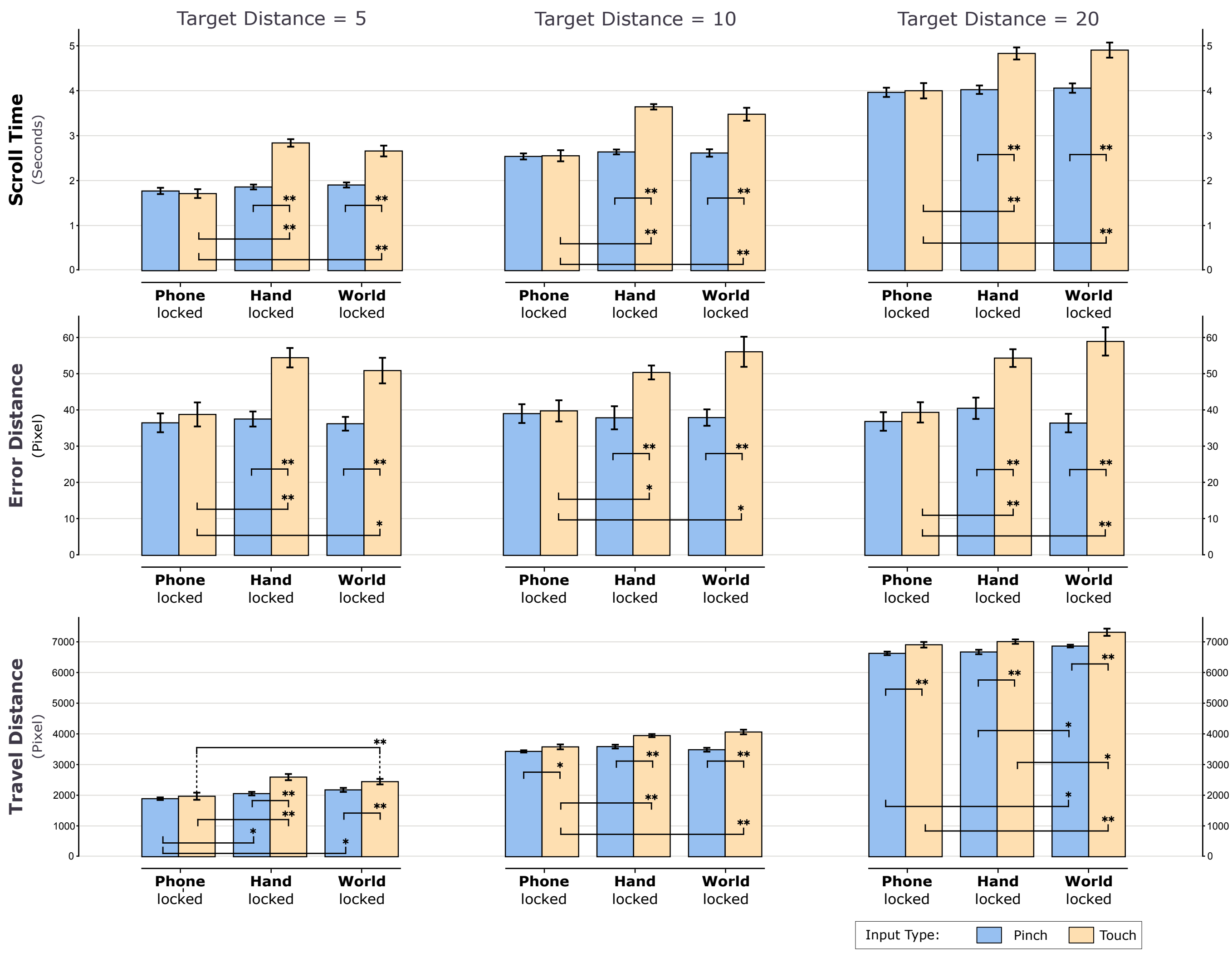

Figure 8: Time, error distance, and travel distance for the scroll task. Error bars represent the mean 95% confidence intervals. The symbol $*$ indicates $p < .05$ and $**$ indicates $p < .01$.

5.3 Scroll Task

For the scroll task (Figure 3 right), we measured scroll time, scroll distance, and error distance. We performed a 3-way RM-ANOVA with Anchoring, Technique, and Distance as independent variables.

5.3.1 Scroll Time. We defined "scroll time" as the duration from the onset of a user's trial to the moment they successfully locate the target line within the designated area (Figure 8). RM-ANOVA revealed a significant Anchoring × Technique 2-way interaction ($F_{2,46}$=35.60, p<.001, η_p^2=.614). For both Hand-locked (p<.001) and (p<.001) World-locked anchorings, we observed that the scroll time with Pinch was significantly shorter than with Touch. However, we found no difference between Touch and Pinch for Phone-locked. Furthermore, for the Touch technique, Phone-locked had a shorter scroll time than Hand-locked (p<.001) and World-locked (p<.001). These findings suggest that the haptic feedback in Phone-locked offers superior touch control.

All independent variables also showed significant main effects. For Anchoring ($F_{2,46}$=30.57, p<.001, η_p^2=.574), post-hoc tests showed that Phone-locked (2.76s) had a significantly shorter scroll time than Hand-locked (3.31s, p<.001) and World-locked (3.27s, p<.001). A significant Technique main effect showed that Touch was significantly faster than Pinch ($F_{1,23}$=36.66, p<.001, η_p^2=.810). Finally, Distance ($F_{2,46}$=174.28, p<.001, η_p^2=.970) showed a significant main effect where longer scroll distances led to longer scroll times.

5.3.2 Error Distance. We define "error distance" as the gap between the target and the screen center upon the completion of the scroll. Given our criteria that a successful scroll concludes only when the target is within 0.8cm (half the line width, 156 pixels) of the center, this inherently filters out certain data. The RM-ANOVA indicated a significant 2-way Anchoring × Technique interaction ($F_{2,46}$=19.03, p<.001, η_p^2=.463, Figure 8). For World-locked (p<.001) and Hand-locked (p=.006), the Pinch technique yielded a shorter error distance

Statements	Phone-locked	Hand-locked	World-locked
1. How mentally demanding was the task?	2 (2)*	3 (3)*†	2 (1.25)†
2. How physically demanding was the task?	3 (2)	4 (2.25)	3 (2)
3. How hurried or rushed was the pace of the task?	2 (1)	2 (1.25)	2.5 (1.25)
4. How successful were you in accomplishing the task?	8 (2)*	7 (2)*	7.5 (1.25)
5. How hard it was to accomplish your level of performance?	3 (2)*	4 (2.5)*	4 (2)
6. How insecure, discouraged, irritated, stressed, and annoyed were you?	1.5 (1)	2 (2)	1 (2)
7. How accurate were you during the task?	8 (1)	7 (.25)	7 (2)
8. How quick were you during the task?	8 (1.25)	7 (2)	7 (2.25)
9. (For touch only) I felt that my finger "touched" the smartphone.	9 (0)*§	5.5 (2.5)*	6 (3.5)§
10. (For pinch only) I felt the pinch reflected my intension to touch.	8 (.25)	8 (1)	7 (2.25)

Strongly Disagree 1 2 3 4 5 6 7 8 9 Strongly Agree

Table 1: User preferences results (median, interquartile range). ∗, † and § indicate statistically significant differences ($p < .01$).

than the Touch technique. Furthermore, with Touch, Phone-locked had a smaller scroll error than the Hand-locked (p<.001) and World-locked (p<.001). This aligns with our earlier swipe time findings, suggesting that the inherent haptic feedback of the Phone-locked Anchoring facilitates better touch control – an expected outcome considering the tactile feedback provided by physical interfaces.

We found significant main effects for Anchoring ($F_{2,46}$=15.08, p<.001, η_p^2=.396). Post-hoc tests showed that the Phone-locked (38.5px) had a significantly lower error distance than World-locked (46.0px, p<.001) and Hand-locked (46.8px, p<.001). A significant Technique effect ($F_{1,23}$=90.57, p<.001, η_p^2=.797) showed that Pinch (37.8px) had a significantly lower error distance than Touch (49.4px).

5.3.3 Scroll Travel Distance. We define "scroll travel distance" as the total distance (pixels) the canvas element travels before the user reaches the target. The scroll distance incorporates the user's direct scrolling action and the additional distance resulting from modern scrolling features, such as the "elastic bounce-back effect" and "momentum scrolling", where the content continues to scroll with inertia after the user's touch or drag action has ended. Results showed a significant 3-way Anchoring × Technique × Distance interaction ($F_{2.54,58.31}$=4.59, p=.009, η_p^2=.396, Figure 8). The post-hoc analysis showed significant differences between all distances for all combinations of technique and anchoring (all p<.001). Furthermore, there was a significant difference between Pinch and Touch for all combinations of Anchoring and Distance (all $p \leq$.001) except for the shortest distance in the Phone-locked anchoring (p=.204). Post hoc differences between Anchorings showed that for Touch, Phone-locked had significantly less travel distance than World-locked and Hand-locked in the shortest Distance (both $p \leq$.016), but only World-locked at the longest distance (p=.023). World-locked also had significantly longer travel distance than Hand-locked at the longest Distance (p=.015). There were no significant differences between anchorings at the middle distance. For Pinch, Phone-locked had significantly less travel distance than World-locked and Hand-locked in the shortest and middle Distances (all $p \leq$.003), while only being significantly shorter than World-locked for the longest Distance (p=.002). Meanwhile, Hand-locked had a significantly shorter travel distance than World-locked for the longest Distance (p=.013).

All independent variables showed significant main effects. For Anchoring ($F_{2,46}$=34.65, p<.001, η_p^2=.601), post-hoc tests revealed that Phone-locked (4068.2px) had a significantly shorter travel distance than Hand-locked (4311.7px, p<.001) and World-locked (4393.0px, p<.001). For Technique ($F_{1,23}$=71.84, p<.001, η_p^2=.757), the Pinch gesture (4088.8px) demonstrated a significantly shorter travel distance than the Touch gesture (4426.5px, p<.001). For Distance ($F_{2,46}$=9897.86, p<.001, η_p^2=.998), we observed that longer target distances naturally resulted in longer scroll distances, with values of 2190.1px, 3683.4px, and 6899.4px respectively (all p<.001).

5.3.4 Task Summary. For scroll tasks, the Phone-locked Anchoring showed superior performance compared to the World-locked and Hand-locked in shorter completion times and travel distances, and reduced error distances when using touch. However, when employing the Pinch, the difference was negligible in the scope of phone Anchoring. Without the support of a physical phone, as seen in World-locked and Hand-locked conditions, the Pinch outperformed the Touch, especially in the absence of haptic feedback, demonstrating advantages in time, travel distance, and accuracy. Notably, the inherent haptic feedback in the Phone-locked Anchoring provides enhanced touch control, underscoring its benefits.

5.4 User Preference Results

We report the questionnaire results in Table 1. A Friedman test suggests there is a significant difference in preferences between anchorings, $\chi^2(29) = 475.118$, $p < .001$). We conducted a post-hoc analysis with Wilcoxon signed-ranks tests with a Bonferroni correction. Results suggest that the Hand-locked condition is perceived as more mentally demanding than Phone-locked (Z = -3.055, p = .002) and World-locked (Z = -3.034, p = .002). Also, participants reported being less successful (Z = -2.974, p = .003) and with lower performance (Z = -2.886, p = .004) in the Hand-locked conditions when compared to the Phone-locked condition. Furthermore, the Wilcoxon signed-rank tests showed that the perception of feeling

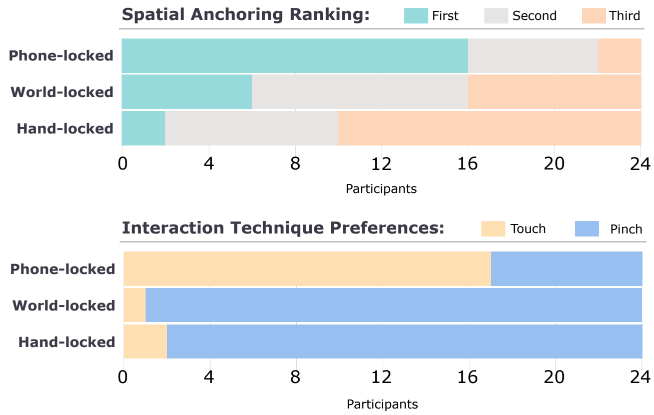

Figure 9: Participants' rankings of Anchoring (top), and technique preferences for each Anchoring (bottom).

a touch was more significant using the Phone-locked Anchoring when compared with the Hand-locked (Z = -4.118, $p < .001$) and the World-locked (Z = -4.119, $p < .001$).

After the evaluation session, we asked participants to rank the anchoring approaches and to pair the anchoring approach with their preferred interaction technique. Figure 9 shows the participants' rankings preferences. Regarding spatial anchoring, we observed that the Phone-locked and World-locked conditions were the most favored and balanced in popularity. We also found a significant influence of the physical phone's presence on the Interaction Technique preference. Most participants preferred Touch for the Phone-locked anchoring. However, Pinch was overwhelmingly preferred for both World-locked and Hand-locked spatial anchorings in scenarios where a physical phone was absent.

6 DISCUSSION

Our results demonstrate the differences between smartphone anchoring strategies and interaction techniques in VR, and their effects on multiple input tasks. We now reflect further on these results with respect to our research questions and then discuss limitations and future work.

6.1 Reflecting on Study Results

The main observations of our study can be summarized into the following considerations, which directly address the research questions we raised. Notably, the results suggest an interaction between RQ1 and RQ2: Holding a physical phone (Phone-locked) is preferable when using Touch compared to virtual phone representations (Hand-locked and World-locked) because of the haptic feedback. While in Hand-locked and World-locked conditions, the Pinch input performs better than the Touch. These results were consistent across the three tasks used in our study (target acquisition, swipe, and scroll). Additionally, the results addressing RQ3 reveal that the majority of participants preferred the Phone-locked spatial anchoring combined with the Touch interaction technique. In scenarios without a physical phone, participants predominantly favored the Pinch input over Touch for interaction without haptics.

The results suggest a clear advantage of haptic feeling from holding a physical phone (**RQ1**). For all tasks, interactions using Phone-locked anchoring consistently outperformed both Hand-locked and World-locked conditions regarding speed. Additionally, Phone-locked conditions exhibited fewer errors in both target acquisition, swipe, and scroll tasks, suggesting that, while tracking the smartphone might introduce additional costs, they are justified given the overall enhanced performance. Another possible benefit of Phone-locked anchoring may be the presence of the non-dominant hand holding the phone as a kinesthetic frame of reference [10], thus adhering to an important principle from Guiard's Kinematic Chain model of human bimanual action [27].

Without a physical phone, input using Pinch generally yields better accuracy for all tasks (**RQ2**). This improvement can be partially explained by the visual feedback provided when hovering over the virtual phone. This might have given the participants more visible space to accurately "aim for the target", mitigating the potential inaccuracies introduced by the "fat finger" occlusion issue inherent with touch interactions.

Beyond performance metrics, user preferences also played a significant role in our findings (**RQ3**). Participants had a strong preference for the Phone-locked condition combined with Touch interaction. However, participants strongly preferred using Pinch in both Hand-locked and World-locked conditions. Furthermore, participants preferred the World-locked phone when compared to Hand-locked. Therefore, for future endeavors that integrate a virtual phone in VR without tracking the actual smartphone, a recommended approach to facilitate interaction would be to use Pinch rather than virtual Touch and to favor a World-locked phone.

We also found that the Phone-locked condition that allows for the sense of touch generally yields shorter error distances (**RQ1 & RQ2**). But compared the touch performance with world-anchored conditions, whether a physical phone was present or absent did not significantly influence the overall error distance for target acquisition (**RQ1 & RQ2**). These results are reasonable; participants primarily relied on visual feedback to aim at targets. However, haptic feedback proved beneficial for swipe and scroll tasks. It assisted users in achieving more accurate directional inputs and enabled them to stop at the target with a minimized error distance while scrolling. These results indicate that for operations that require consistent movement, such as swiping and scrolling, where users traverse the surface—haptic feedback boosts user performance.

An interesting observation we made relates to the role of haptic feedback, particularly concerning the signal indicating the completion of a release action, rather than just the moment of touch activation. The lack of these haptic release signals led to an increase in errors when finalizing scroll and swipe tasks, affecting both accuracy and speed. This could explain why performance with touch on hand-locked and world-locked conditions typically underperforms compared to physically-locked conditions or in comparison to pinch gestures used in the same group (**RQ1 & RQ2**).

Furthermore, it is important to note that while we observed significant differences between anchorings and techniques in various tasks, the overall performance of all combinations of anchoring and technique yielded acceptable results for all tasks (**RQ1 & RQ2**). This suggests that while some combinations are more optimal in terms of performance, there is room for variety without severely

Figure 10: Navigating a web browser with a Phone-Locked and Touch interactive smartphone. Looking for vacation inspiration in Hawaii, A) the user navigates to the web browser via target acquisition, B) browses images through pictures with scrolling, and C) navigates pages through swipes.

hampering the user experience. For example, if a user’s phone is not easily accessible, they could choose to use the world-locked anchoring. Our implementation of PhoneInVR enables users to use their smartphone in VR as a smartphone with any of the spatial anchoring and interaction technique combinations, as Figure 10 shows. The implementation leverages screen streaming and input injection techniques, using a similar pipeline that has been demonstrated in previous work [9, 75]. As video pass-through techniques improve, this may soon be a viable alternative to enable a phone-locked anchoring with a true rendering of both the phone and hands.

6.2 Limitations and Future Work

We designed our study to focus primarily on performance and used abstract tasks. However, an essential aspect of using a phone in VR is the ability to read and digest content without obstructions. Evaluating the impact on readability across various techniques could be a promising direction for future research. Concerning readability, a common strategy is to render the virtual phone with a larger size. We opted not to pursue this approach because our starting point represented a physical phone. To maintain consistency, we did not explore the influence of size. Future studies of smartphones in VR could potentially investigate this aspect.

We also acknowledge that fixing the phone location (to simulate ideal tracking conditions) is a limitation of our study. The smartphone in the Phone-locked condition was aligned on a tripod to ensure stable tracking, and the smartphone in the Hand-locked condition was aligned to a stationary hand position. The performance of each technique and anchoring may change with in-the-wild tracking. This work provides a baseline for future research in more natural settings.

Finally, the overall approach of our work was based on the principle of replicating the form factor and user interface of contemporary smartphones in VR. This approach minimizes the need for extensive redesign and circumvents the need to relearn familiar interfaces, aligning with the concept of legacy bias in cross-device interactions [20]. Yet, applications could have different adapted versions of their UI specifically designed for VR, much like how Apple CarPlay and Android Auto offer apps adapted for use within the car. While the findings of our study should still apply to adapted app layouts, this does open up future research that compares the performance and preference between redesigned VR and traditional smartphone interfaces. Additionally, this raises broader questions regarding the effects of interacting with different versions or representations of the same application on user experience. This paper focused on maintaining content and interactions within the VR phone interface. Future work could investigate smartphone interfaces that change depending on current accessible devices or context, such as work on cross-reality systems or hybrid UIs [5, 48, 77].

7 CONCLUSION

In this paper, we contributed an investigation on how to present and interact with smartphones in VR. Our main objective was to understand what combination of spatial anchoring and interaction could lead to the best performance and preference to facilitate seamless integration of smartphone interactions within immersive virtual environments. We examined three spatial anchoring configurations, each representing varying degrees of smartphone embodiment: Phone-locked, Hand-locked, and World-locked. We paired these with two prevalent VR interaction methods: Touch and Pinch. Our research involved a comparative evaluation of these pairings across three fundamental tasks: target acquisition, swiping, and scrolling.

Our findings indicate that Touch interaction with a physical Phone-locked anchoring generally performed faster across different tasks and more accurately in swiping and scrolling. In contrast, the World-locked virtual phone with Pinch showcased its strengths in terms of accuracy and speed as well when a physical smartphone is not available. User preferences indicate that those combinations were also equally favored. The results of our study can guide and inform future research in integrating the usage of smartphones and other touch-enabled devices into VR environments.

ACKNOWLEDGMENTS

We would like to thank the anonymous reviewers for their feedback. This research was supported in part by the National Sciences and Engineering Research Council of Canada (NSERC) under Grant IRCPJ 545100-18.

REFERENCES

[1] Roland Aigner, Daniel Wigdor, Hrvoje Benko, Michael Haller, Alexandra Ion, and Shengdong Zhao. 2012. *Understanding Mid-Air Hand Gestures: A study of Human Preferences in Usage of Gesture Types for HCI.* Technical Report. Microsoft Research.

[2] Ferran Argelaguet and Carlos Andujar. 2013. A survey of 3D object selection techniques for virtual environments. *Computers & Graphics* 37, 3 (2013), 121–136. https://doi.org/10.1016/j.cag.2012.12.003

[3] Rahul Arora, Rubaiat Habib Kazi, Tovi Grossman, George Fitzmaurice, and Karan Singh. 2018. SymbiosisSketch: Combining 2D & 3D Sketching for Designing Detailed 3D Objects in Situ. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI ’18)*. Association for Computing Machinery, New York, NY, USA, 1–15. https://doi.org/10.1145/3173574.3173759

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

[5] Jonas Auda, Uwe Gruenefeld, Sarah Faltaous, Sven Mayer, and Stefan Schneegass. 2023. A Scoping Survey on Cross-Reality Systems. *ACM Comput. Surv.* 56, 4, Article 83 (oct 2023), 38 pages. https://doi.org/10.1145/3616536

[6] Takumi Azai, Shuhei Ogawa, Mai Otsuki, Fumihisa Shibata, and Asako Kimura. 2017. Selection and Manipulation Methods for a Menu Widget on the Human Forearm. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI EA ’17)*. Association for Computing Machinery, New York, NY, USA, 357–360. https://doi.org/10.1145/3027063.3052959

[7] Takumi Azai, Mai Otsuki, Fumihisa Shibata, and Asako Kimura. 2018. Open Palm Menu: A Virtual Menu Placed in Front of the Palm. In *Proceedings of the 9th Augmented Human International Conference* (Seoul, Republic of Korea) *(AH ’18)*. Association for Computing Machinery, New York, NY, USA, Article 17, 5 pages. https://doi.org/10.1145/3174910.3174929

[8] Teo Babic, Harald Reiterer, and Michael Haller. 2018. Pocket6: A 6DoF Controller Based On A Simple Smartphone Application. In *Proceedings of the 2018 ACM Symposium on Spatial User Interaction* (Berlin, Germany) *(SUI '18)*. Association for Computing Machinery, New York, NY, USA, 2–10. https://doi.org/10.1145/3267782.3267785

[9] Huidong Bai, Li Zhang, Jing Yang, and Mark Billinghurst. 2021. Bringing Full-Featured Mobile Phone Interaction Into Virtual Reality. *Comput. Graph* 97 (2021), 42–53. https://doi.org/10.1016/j.cag.2021.04.004

[10] Ravin Balakrishnan and Ken Hinckley. 1999. The Role of Kinesthetic Reference Frames in Two-Handed Input Performance. In *Proceedings of the 12th Annual ACM Symposium on User Interface Software and Technology* (Asheville, North Carolina, USA) *(UIST '99)*. Association for Computing Machinery, New York, NY, USA, 171–178. https://doi.org/10.1145/320719.322599

[11] Anil Ufuk Batmaz, Aunnoy K Mutasim, Morteza Malekmakan, Elham Sadr, and Wolfgang Stuerzlinger. 2020. Touch the Wall: Comparison of Virtual and Augmented Reality with Conventional 2D Screen Eye-Hand Coordination Training Systems. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 184–193. https://doi.org/10.1109/VR46266.2020.00037

[12] Klaus-Peter Beier. 2000. Web-Based Virtual Reality in Design and Manufacturing Applications. In *Proceedings of COMPIT*. 45–55.

[13] Joanna Bergström and Kasper Hornbæk. 2019. Human–Computer Interaction on the Skin. *ACM Comput. Surv.* 52, 4, Article 77 (aug 2019), 14 pages. https://doi.org/10.1145/3332166

[14] Joanna Bergstrom-Lehtovirta, Kasper Hornbæk, and Sebastian Boring. 2018. It's a Wrap: Mapping On-Skin Input to Off-Skin Displays. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–11. https://doi.org/10.1145/3173574.3174138

[15] Eric A. Bier, Maureen C. Stone, Ken Pier, William Buxton, and Tony D. DeRose. 1993. Toolglass and Magic Lenses: The See-through Interface. In *Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques* (Anaheim, CA) *(SIGGRAPH '93)*. Association for Computing Machinery, New York, NY, USA, 73–80. https://doi.org/10.1145/166117.166126

[16] Idil Bostan, Oğuz Turan Buruk, Mert Canat, Mustafa Ozan Tezcan, Celalettin Yurdakul, Tilbe Göksun, and Oğuzhan Özcan. 2017. Hands as a Controller: User Preferences for Hand Specific On-Skin Gestures. In *Proceedings of the 2017 Conference on Designing Interactive Systems* (Edinburgh, United Kingdom) *(DIS '17)*. Association for Computing Machinery, New York, NY, USA, 1123–1134. https://doi.org/10.1145/3064663.3064766

[17] Sabah Boustila, Thomas Guégan, Kazuki Takashima, and Yoshifumi Kitamura. 2019. Text Typing in VR Using Smartphones Touchscreen and HMD. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 860–861. https://doi.org/10.1109/VR.2019.8798238

[18] Doug A Bowman and Chadwick A Wingrave. 2001. Design and Evaluation of Menu Systems for Immersive Virtual Environments. In *Proceedings IEEE Virtual Reality 2001*. IEEE, 149–156. https://doi.org/10.1109/VR.2001.913781

[19] 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* (Vancouver, BC, Canada) *(CHI '11)*. Association for Computing Machinery, New York, NY, USA, 403–412. https://doi.org/10.1145/1978942.1979000

[20] Frederik Brudy, Christian Holz, Roman Rädle, Chi-Jui Wu, Steven Houben, Clemens Nylandsted Klokmose, and Nicolai Marquardt. 2019. Cross-Device Taxonomy: Survey, Opportunities and Challenges of Interactions Spanning Across Multiple Devices. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–28. https://doi.org/10.1145/3290605.3300792

[21] Li-Wei Chan, Hui-Shan Kao, Mike Y. Chen, Ming-Sui Lee, Jane Hsu, and Yi-Ping Hung. 2010. Touching the Void: Direct-Touch Interaction for Intangible Displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Atlanta, Georgia, USA) *(CHI '10)*. Association for Computing Machinery, New York, NY, USA, 2625–2634. https://doi.org/10.1145/1753326.1753725

[22] Neil Chulpongsatorn, Wesley Willett, and Ryo Suzuki. 2023. HoloTouch: Interacting with Mixed Reality Visualizations Through Smartphone Proxies. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI EA '23)*. Association for Computing Machinery, New York, NY, USA, Article 156, 8 pages. https://doi.org/10.1145/3544549.3585738

[23] Amit P Desai, Lourdes Peña-Castillo, and Oscar Meruvia-Pastor. 2017. A Window to Your Smartphone: Exploring Interaction and Communication in Immersive VR With Augmented Virtuality. In *2017 14th Conference on Computer and Robot Vision (CRV)*. 217–224. https://doi.org/10.1109/CRV.2017.16

[24] Tomás Dorta, Gokce Kinayoglu, and Michael Hoffmann. 2016. Hyve-3D and the 3D Cursor: Architectural co-design with freedom in Virtual Reality. *International Journal of Architectural Computing* 14, 2 (2016), 87–102. https://doi.org/10.1177/1478077116638921

[25] Jacqui Fashimpaur, Amy Karlson, Tanya R. Jonker, Hrvoje Benko, and Aakar Gupta. 2023. Investigating Wrist Deflection Scrolling Techniques for Extended 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 386, 16 pages. https://doi.org/10.1145/3544548.3580870

[26] Jens Grubert, Matthias Heinisch, Aaron Quigley, and Dieter Schmalstieg. 2015. MultiFi: Multi Fidelity Interaction with Displays On and Around the Body. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. Association for Computing Machinery, New York, NY, USA, 3933–3942. https://doi.org/10.1145/2702123.2702331

[27] Yves Guiard. 1987. Asymmetric Division of Labor in Human Skilled Bimanual Action. *Journal of Motor Behavior* 19, 4 (1987), 486–517. https://doi.org/10.1080/00222895.1987.10735426

[28] Sean G. Gustafson, Bernhard Rabe, and Patrick M. Baudisch. 2013. Understanding Palm-Based Imaginary Interfaces: The Role of Visual and Tactile Cues When Browsing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Paris, France) *(CHI '13)*. Association for Computing Machinery, New York, NY, USA, 889–898. https://doi.org/10.1145/2470654.2466114

[29] Taejin Ha, Steven Feiner, and Woontack Woo. 2014. WeARHand: Head-worn, RGB-D camera-based, bare-hand user interface with visually enhanced depth perception. In *2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 219–228. https://doi.org/10.1109/ISMAR.2014.6948431

[30] Chris Harrison, Hrvoje Benko, and Andrew D. Wilson. 2011. OmniTouch: Wearable Multitouch Interaction Everywhere. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology* (Santa Barbara, California, USA) *(UIST '11)*. Association for Computing Machinery, New York, NY, USA, 441–450. https://doi.org/10.1145/2047196.2047255

[31] Chris Harrison, Shilpa Ramamurthy, and Scott E. Hudson. 2012. On-Body Interaction: Armed and Dangerous. In *Proceedings of the Sixth International Conference on Tangible, Embedded and Embodied Interaction* (Kingston, Ontario, Canada) *(TEI '12)*. Association for Computing Machinery, New York, NY, USA, 69–76. https://doi.org/10.1145/2148131.2148148

[32] Chris Harrison, Desney Tan, and Dan Morris. 2010. Skinput: Appropriating the Body as an Input Surface. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Atlanta, Georgia, USA) *(CHI '10)*. Association for Computing Machinery, New York, NY, USA, 453–462. https://doi.org/10.1145/1753326.1753394

[33] Julia Hertel, Sukran Karaosmanoglu, Susanne Schmidt, Julia Bräker, Martin Semmann, and Frank Steinicke. 2021. A Taxonomy of Interaction Techniques for Immersive Augmented Reality based on an Iterative Literature Review. In *2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 431–440. https://doi.org/10.1109/ISMAR52148.2021.00060

[34] Ken Hinckley, Edward Cutrell, Steve Bathiche, and Tim Muss. 2002. Quantitative Analysis of Scrolling Techniques. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Minneapolis, Minnesota, USA) *(CHI '02)*. Association for Computing Machinery, New York, NY, USA, 65–72. https://doi.org/10.1145/503376.503389

[35] David Holman, Roel Vertegaal, Mark Altosaar, Nikolaus Troje, and Derek Johns. 2005. Paper Windows: Interaction Techniques for Digital Paper. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Portland, Oregon, USA) *(CHI '05)*. Association for Computing Machinery, New York, NY, USA, 591–599. https://doi.org/10.1145/1054972.1055054

[36] Sebastian Hubenschmid, Johannes Zagermann, Daniel Leicht, Harald Reiterer, and Tiare Feuchtner. 2023. ARound the Smartphone: Investigating the Effects of Virtually-Extended Display Size on Spatial Memory. 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, 1–15. https://doi.org/10.1145/3544548.3581438

[37] Pan Jing and Guan Ye-Peng. 2013. Human-Computer Interaction Using Pointing Gesture Based on An Adaptive Virtual Touch Screen. *International Journal of Signal Processing, Image Processing and Pattern Recognition* 6, 4 (2013), 81–91.

[38] Mohamed Kari and Christian Holz. 2023. HandyCast: Phone-Based Bimanual Input for Virtual Reality in Mobile and Space-Constrained Settings via Pose-and-Touch Transfer. 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 528, 15 pages. https://doi.org/10.1145/3544548.3580677

[39] Wonjung Kim, Kenny Tsu Wei Choo, Youngki Lee, Archan Misra, and Rajesh Krishna Balan. 2018. Empath-D: VR-Based Empathetic App Design for Accessibility. In *Proceedings of the 16th Annual International Conference on Mobile Systems, Applications, and Services*. 123–135. https://doi.org/10.1145/3210240.3211108

[40] Grace Kuo, Eric Penner, Seth Moczydlowski, Alexander Ching, Douglas Lanman, and Nathan Matsuda. 2023. Perspective-Correct VR Passthrough Without Reprojection. In *ACM SIGGRAPH 2023 Conference Proceedings* (Los Angeles, CA, USA) *(SIGGRAPH '23)*. Association for Computing Machinery, New York, NY, USA, Article 15, 9 pages. https://doi.org/10.1145/3588432.3591534

[41] Joseph J LaViola Jr, Ernst Kruijff, Ryan P McMahan, Doug Bowman, and Ivan P Poupyrev. 2017. *3D User Interfaces: Theory and Practice*. Addison-Wesley Professional.

[42] Zhen Li, Michelle Annett, Ken Hinckley, Karan Singh, and Daniel Wigdor. 2019. HoloDoc: Enabling Mixed Reality Workspaces That Harness Physical and Digital Content. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3290605.3300917

[43] Zhen Li, Joannes Chan, Joshua Walton, Hrvoje Benko, Daniel Wigdor, and Michael Glueck. 2021. Armstrong: An Empirical Examination of Pointing at Non-Dominant Arm-Anchored UIs in Virtual Reality. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 123, 14 pages. https://doi.org/10.1145/3411764.3445064

[44] R.W. Lindeman, J.L. Sibert, and J.K. Hahn. 1999. Hand-Held Windows: Towards Effective 2D Interaction in Immersive Virtual Environments. In *Proceedings IEEE Virtual Reality (Cat. No. 99CB36316)*. 205–212. https://doi.org/10.1109/VR.1999.756952

[45] Xinyi Liu, Xuanru Meng, Becky Spittle, Wenge Xu, BoYu Gao, and Hai-Ning Liang. 2023. Exploring Text Selection in Augmented Reality Systems. In *Proceedings of the 18th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry* (Guangzhou, China) *(VRCAI '22)*. Association for Computing Machinery, New York, NY, USA, Article 35, 8 pages. https://doi.org/10.1145/3574131.3574459

[46] Paul Lubos, Gerd Bruder, and Frank Steinicke. 2014. Analysis of Direct Selection in Head-Mounted Display Environments. In *2014 IEEE Symposium on 3D User Interfaces (3DUI)*. 11–18. https://doi.org/10.1109/3DUI.2014.6798834

[47] Akhmajon Makhsadov, Donald Degraen, André Zenner, Felix Kosmalla, Kamila Mushkina, and Antonio Krüger. 2022. VRySmart: A Framework for Embedding Smart Devices in Virtual Reality. In *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI EA '22)*. Association for Computing Machinery, New York, NY, USA, 1–8. https://doi.org/10.1145/3491101.3519717

[48] Fabrice Matulic, Aditya Ganeshan, Hiroshi Fujiwara, and Daniel Vogel. 2021. Phonetroller: Visual Representations of Fingers for Precise Touch Input with Mobile Phones in VR. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 129, 13 pages. https://doi.org/10.1145/3411764.3445583

[49] Fabrice Matulic, Taiga Kashima, Deniz Beker, Daichi Suzuo, Hiroshi Fujiwara, and Daniel Vogel. 2023. Above-Screen Fingertip Tracking With a Phone in Virtual Reality. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI EA '23)*. Association for Computing Machinery, New York, NY, USA, 1–7. https://doi.org/10.1145/3544549.3585728

[50] Alexandre Millette and Michael J McGuffin. 2016. DualCAD: Integrating Augmented Reality With a Desktop GUI and Smartphone Interaction. In *2016 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct)* (Merida, Yucatan, Mexico). IEEE. https://doi.org/10.1145/3411764.3445593

[51] Mark R. Mine, Frederick P. Brooks, and Carlo H. Sequin. 1997. Moving Objects in Space: Exploiting Proprioception in Virtual-Environment Interaction. In *Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '97)*. ACM Press/Addison-Wesley Publishing Co., USA, 19–26. https://doi.org/10.1145/258734.258747

[52] Peter Mohr, Markus Tatzgern, Tobias Langlotz, Andreas Lang, Dieter Schmalstieg, and Denis Kalkofen. 2019. TrackCap: Enabling Smartphones for 3D Interaction on Mobile Head-Mounted Displays. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–11. https://doi.org/10.1145/3290605.3300815

[53] Matei Negulescu, Jaime Ruiz, Yang Li, and Edward Lank. 2012. Tap, Swipe, or Move: Attentional Demands for Distracted Smartphone Input. In *Proceedings of the International Working Conference on Advanced Visual Interfaces* (Capri Island, Italy) *(AVI '12)*. Association for Computing Machinery, New York, NY, USA, 173–180. https://doi.org/10.1145/2254556.2254589

[54] Erwan Normand and Michael J. McGuffin. 2018. Enlarging a Smartphone with AR to Create a Handheld VESAD (Virtually Extended Screen-Aligned Display). In *2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 123–133. https://doi.org/10.1109/ISMAR.2018.00043

[55] Masa Ogata, Yuta Sugiura, Yasutoshi Makino, Masahiko Inami, and Michita Imai. 2013. SenSkin: Adapting Skin as a Soft Interface. In *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology* (St. Andrews, Scotland, United Kingdom) *(UIST '13)*. Association for Computing Machinery, New York, NY, USA, 539–544. https://doi.org/10.1145/2501988.2502039

[56] Tony Parisi. 2015. *Learning Virtual Reality: Developing Immersive Experiences and Applications for Desktop, Web, and Mobile.* O'Reilly Media, Inc.

[57] Siyou Pei, Alexander Chen, Jaewook Lee, and Yang Zhang. 2022. Hand Interfaces: Using Hands to Imitate Objects in AR/VR for Expressive Interactions. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 429, 16 pages. https://doi.org/10.1145/3491102.3501898

[58] Kevin Pfeil, Sina Masnadi, Jacob Belga, Jose-Valentin T Sera-Josef, and Joseph LaViola. 2021. Distance Perception with a Video See-Through Head-Mounted Display. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 528, 9 pages. https://doi.org/10.1145/3411764.3445223

[59] 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* (Brighton, United Kingdom) *(SUI '17)*. Association for Computing Machinery, New York, NY, USA, 99–108. https://doi.org/10.1145/3131277.3132180

[60] Stefan Schneegass and Alexandra Voit. 2016. GestureSleeve: Using Touch Sensitive Fabrics for Gestural Input on the Forearm for Controlling Smartwatches. In *Proceedings of the 2016 ACM International Symposium on Wearable Computers* (Heidelberg, Germany) *(ISWC '16)*. Association for Computing Machinery, New York, NY, USA, 108–115. https://doi.org/10.1145/2971763.2971797

[61] Daniel Schneider, Verena Biener, Alexander Otte, Travis Gesslein, Philipp Gagel, Cuauhtli Campos, Klen Čopič Pucihar, Matjazz Kljun, Eyal Ofek, Michel Pahud, Per Ola Kristensson, and Jens Grubert. 2021. Accuracy Evaluation of Touch Tasks in Commodity Virtual and Augmented Reality Head-Mounted Displays. In *Proceedings of the 2021 ACM Symposium on Spatial User Interaction* (Virtual Event, USA) *(SUI '21)*. Association for Computing Machinery, New York, NY, USA, Article 7, 11 pages. https://doi.org/10.1145/3485279.3485283

[62] 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* (Hamburg, Germany) *(CHI EA '23)*. Association for Computing Machinery, New York, NY, USA, Article 109, 7 pages. https://doi.org/10.1145/3544549.3585615

[63] Hemant Bhaskar Surale, Aakar Gupta, Mark Hancock, and Daniel Vogel. 2019. TabletInVR: Exploring the Design Space for Using a Multi-Touch Tablet in Virtual Reality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3290605.3300243

[64] Yutaro Suzuki, Kodai Sekimori, Buntarou Shizuki, and Shin Takahashi. 2019. Touch sensing on the forearm using the electrical impedance method. In *2019 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*. 255–260. https://doi.org/10.1109/PERCOMW.2019.8730739

[65] Balasaravanan Thoravi Kumaravel and Björn Hartmann. 2022. Interactive Mixed-Dimensional Media for Cross-Dimensional Collaboration in Mixed Reality Environments. *Frontiers in Virtual Reality* 3 (2022). https://doi.org/10.3389/frvir.2022.766336

[66] Balasaravanan Thoravi Kumaravel, Cuong Nguyen, Stephen DiVerdi, and Björn Hartmann. 2019. TutoriVR: A Video-Based Tutorial System for Design Applications in Virtual Reality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3290605.3300514

[67] Shuma Toyama, Mohammed Al Sada, and Tatsuo Nakajima. 2018. VRowser: A Virtual Reality Parallel Web Browser. In *Virtual, Augmented and Mixed Reality: Interaction, Navigation, Visualization, Embodiment, and Simulation*. Springer International Publishing, 230–244. https://doi.org/10.1007/978-3-319-91581-4_17

[68] Cheng-Yao Wang, Wei-Chen Chu, Po-Tsung Chiu, Min-Chieh Hsiu, Yih-Harn Chiang, and Mike Y. Chen. 2015. PalmType: Using Palms as Keyboards for Smart Glasses. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services* (Copenhagen, Denmark) *(MobileHCI '15)*. Association for Computing Machinery, New York, NY, USA, 153–160. https://doi.org/10.1145/2785830.2785886

[69] Cheng-Yao Wang, Min-Chieh Hsiu, Po-Tsung Chiu, Chiao-Hui Chang, Liwei Chan, Bing-Yu Chen, and Mike Y. Chen. 2015. PalmGesture: Using Palms as Gesture Interfaces for Eyes-Free Input. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services* (Copenhagen, Denmark) *(MobileHCI '15)*. Association for Computing Machinery, New York, NY, USA, 217–226. https://doi.org/10.1145/2785830.2785885

[70] Mark Weiser. 1999. The Computer for the 21st Century. *SIGMOBILE Mob. Comput. Commun. Rev.* 3, 3 (jul 1999), 3–11. https://doi.org/10.1145/329124.329126

[71] Katrin Wolf, Christian Müller-Tomfelde, Kelvin Cheng, and Ina Wechsung. 2012. Does Proprioception Guide Back-of-Device Pointing as Well as Vision?. In *CHI '12 Extended Abstracts on Human Factors in Computing Systems* (Austin, Texas, USA) *(CHI EA '12)*. Association for Computing Machinery, New York, NY, USA, 1739–1744. https://doi.org/10.1145/2212776.2223702

[72] Robert Xiao, Julia Schwarz, Nick Throm, Andrew D. Wilson, and Hrvoje Benko. 2018. MRTouch: Adding Touch Input to Head-Mounted Mixed Reality. *IEEE Transactions on Visualization and Computer Graphics* 24, 4 (2018), 1653–1660. https://doi.org/10.1109/TVCG.2018.2794222

[73] Yukang Yan, Chun Yu, Xiaojuan Ma, Shuai Huang, Hasan Iqbal, and Yuanchun Shi. 2018. Eyes-Free Target Acquisition in Interaction Space around the Body for Virtual Reality. In *Proceedings of the 2018 CHI Conference on Human Factors in*

Computing Systems (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3173574.3173616

[74] Shumin Zhai, Barton A Smith, and Ted Selker. 1997. Improving Browsing Performance: A Study of Four Input Devices for Scrolling and Pointing Tasks. In *Human-Computer Interaction INTERACT'97: IFIP TC13 International Conference on Human-Computer Interaction, 14th–18th July 1997, Sydney, Australia*. Springer, 286–293.

[75] Li Zhang, Weiping He, Huidong Bai, Qianyuan Zou, Shuxia Wang, and Mark Billinghurst. 2023. A Hybrid 2D–3D Tangible Interface Combining a Smartphone and Controller for Virtual Reality. *Virtual Reality* 27, 2 (2023), 1273–1291. https://doi.org/10.1007/s10055-022-00735-2

[76] Yang Zhang, Wolf Kienzle, Yanjun Ma, Shiu S. Ng, Hrvoje Benko, and Chris Harrison. 2019. ActiTouch: Robust Touch Detection for On-Skin AR/VR Interfaces. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology* (New Orleans, LA, USA) *(UIST '19)*. Association for Computing Machinery, New York, NY, USA, 1151–1159. https://doi.org/10.1145/3332165.3347869

[77] Fengyuan Zhu and Tovi Grossman. 2020. BISHARE: Exploring Bidirectional Interactions Between Smartphones and Head-Mounted Augmented Reality. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3313831.3376233

[78] Fengyuan Zhu, Zhuoyue Lyu, Mauricio Sousa, and Tovi Grossman. 2022. Touching The Droid: Understanding and Improving Touch Precision With Mobile Devices in Virtual Reality. In *2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, 807–816. https://doi.org/10.1109/ISMAR55827.2022.00099

[79] Cindy Ziker, Barbara Truman, and Heather Dodds. 2021. Cross Reality (XR): Challenges and Opportunities Across the Spectrum. *Innovative learning environments in STEM higher education: Opportunities, challenges, and looking forward* (2021), 55–77. https://doi.org/10.1007/978-3-030-58948-6_4