Armstrong: An Empirical Examination of Pointing at Non-Dominant Arm-Anchored UIs in Virtual Reality

Zhen Li
Chatham Labs & University of Toronto
Toronto, Canada
zhen@dgp.toronto.edu

Joannes Chan
Chatham Labs
Toronto, Canada
joannes@chathamlabs.com

Joshua Walton
Facebook Reality Labs
Redmond, United States
joshuawalton@fb.com

Hrvoje Benko
Facebook Reality Labs
Redmond, United States
benko@fb.com

Daniel Wigdor
Chatham Labs & University of Toronto
Toronto, Canada
daniel@dgp.toronto.edu

Michael Glueck
Chatham Labs
Toronto, Canada
mike@chathamlabs.com

ABSTRACT

In virtual reality (VR) environments, asymmetric bimanual interaction techniques can increase users' input bandwidth by complementing their perceptual and motor systems (e.g., using the dominant hand to select 3D UI controls anchored around the non-dominant arm). However, it is unclear how to optimize the layout of such 3D UI controls for near-body and mid-air interactions. We evaluate the performance and limitations of non-dominant arm-anchored 3D UIs in VR environments through a bimanual pointing study. Results demonstrated that targets appearing closer to the skin, located around the wrist, or placed on the medial side of the forearm could be selected more quickly than targets farther away from the skin, located around the elbow, or on the lateral side of the forearm. Based on these results, we developed Armstrong guidelines, demonstrated through a Unity plugin to enable designers to create performance-optimized arm-anchored 3D UI layouts.

CCS CONCEPTS

• **Human-centered computing** → **Virtual reality**; *Empirical studies in HCI*.

KEYWORDS

Arm-anchored UIs; Virtual reality; Bimanual interaction

ACM Reference Format:
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 *CHI Conference on Human Factors in Computing Systems (CHI '21), May 8–13, 2021, Yokohama, Japan.* ACM, New York, NY, USA, 14 pages. https://doi.org/10.1145/3411764.3445064

CHI '21, May 8–13, 2021, Yokohama, Japan

ACM ISBN 978-1-4503-8096-6/21/05…$15.00
https://doi.org/10.1145/3411764.3445064

1 INTRODUCTION

As commercial head-mounted displays for augmented reality (AR) and virtual reality (VR) have grown in prevalence, there has been a resurgence in egocentric 3D interaction techniques. Many of these techniques enable users to move throughout a 3D environment and directly manipulate objects in a virtual scene using hand-held controllers. For example, in Blocks, a double-sided palette is attached to a controller in the user's non-dominant hand [18], with various drawing tools on one side and selectable colors on the other side (Fig. 1a). Tilt Brush anchors a virtual prism with multiple layers of UI controls to the non-dominant controller [17], which the user can scroll through using the dominant controller (Fig. 1b).

A variety of techniques use this "palette and brush" metaphor [17, 18, 21, 49], which enables users to leverage their proprioceptive senses to locate their non-dominant hand in 3D space without needing to explicitly search for the menu [38, 44, 56, 65]. Proprioception thus creates a frame of reference that reduces the visual and mental effort needed to find the palette [26]. These techniques exploit the asymmetric nature of the hands by capitalizing on bimanual action feedback [8, 20] to minimize task flow interruptions, so that a user's attention is not diverted from her main task every time a palette selection is made.

As applications grow in complexity and body-tracking technologies mature, we envision that the design of UI control layouts will migrate beyond the space around the hands (hand-anchored UIs), to the space around the arms (arm-anchored UIs). A 3D palette that wraps around the non-dominant arm could provide a user with a larger working volume while leveraging the benefits of bimanual interactions and proprioception. This trend towards using the arm for interaction can be observed in on-skin interaction research, which has considered fingers, palms, and forearms as interactive surfaces [6, 14, 22, 24, 51, 61, 67]. We further evaluated the benefits of arm-anchored UI controls, through a prototype design exploration that inspired the current study. In a running scenario, a running app could show the user's current progress on one side of her arm and a music app on the other side could show the current playlist, enabling the user to quickly glance at an app by rotating her forearm (Fig. 1c). A user could also invoke functions using swipe gestures on her forearm (e.g., to pause music; Fig. 1d). These examples demonstrated the potential usefulness of capitalizing on a

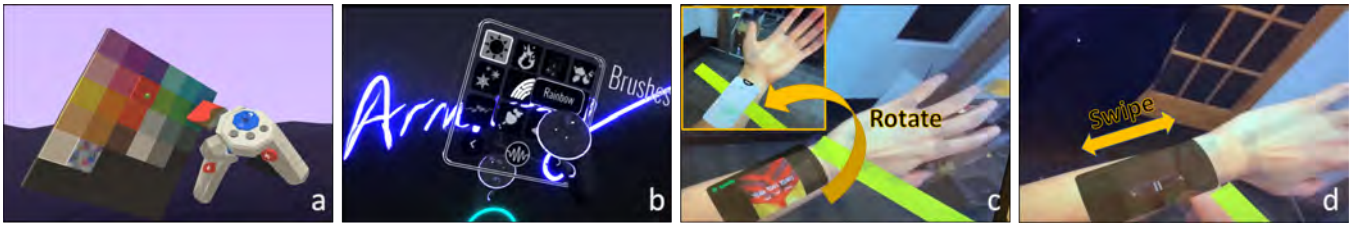

Figure 1: (a, b) Hand-anchored UIs from commercial applications. (a) A 2D color palette is attached to the user's non-dominant controller and a brush is attached to the dominant controller [18]. (b) A variety of brush types and UI controls are distributed on a triangular prism attached to the non-dominant controller and can be selected via ray-casting by the dominant controller [17]. (c, d) Arm-anchored UIs from our prototype: a music app and a running app are rendered on opposite sides of the user's arm and (c) the user can rotate her arm to switch between these apps or (d) swipe her arm to interact with the UI controls.

user's proprioceptive sensing of her arms to deploy arm-anchored UI controls in 3D environments.

Although commercial applications and prior research projects have explored attaching controls to user's hands, wrists, or arms, these applications have largely migrated UI design practices from 2D WIMP interfaces. In contrast, using the full 3D space around the user's limbs might enable opportunities to "*display more items in more varied layouts*" [5]. Limited research has focused on developing design guidelines for 3D UIs around a user's arms, leading to open research questions such as: What areas around a user's arm can be most quickly and accurately targeted? What is a performance-optimized layout of 3D UI controls around the arm? Can users perceive and leverage individual axes for interaction (e.g., along the longitude, latitude, or height dimensions around the arm)?

This work thus seeks to understand the appropriateness and usability of arm-anchored UIs within the context of AR and VR. Due to the limited field of view (FOV) of available AR devices, an experiment was conducted in VR to investigate user performance during pointing tasks, while targets were located in various positions around the arm. We measured the Fitts' *throughput* (TP) of user interactions [16, 30] to generate a heatmap and articulate the usefulness of different combinations of longitude, latitude, and height dimensions around the arm. We synthesized the quantitative and qualitative results to describe 72 regions in terms of visibility, reachability, and comfort.

The contribution of this work is thus twofold. First, we report analysis of quantitative and qualitative results from a pointing study in VR that identified regions with different levels of performance and subjective preference around the arm (e.g., targets closer to the skin, around the wrist, or on the medial side of the forearm can be selected quicker than targets in other locations) and unique movement strategies that resulted (e.g., moving the arm to the center of the FOV). Second, we present Armstrong design guidelines, and implemented a Unity plugin to demonstrate how to create performance-optimized arm-anchored 3D UI layouts.

2 RELATED WORK

This work was inspired and informed by previous research into theories of proprioception and peripersonal space, arm-anchored UI interaction techniques, and existing studies of pointing tasks around various body parts.

2.1 Proprioception and Peripersonal Space

As first identified by Sherrington in 1907 [53], *proprioception* enables humans to sense the position and movement of their limbs through their sensory neurons [35, 56]. Neuropsychological experiments have further identified that there is a *peripersonal* space that surrounds our bodies [12, 19, 47, 48]. This space acts as an intermediary between the visual space that we perceive through our eyes and the tactile, proprioceptive space we perceive on our body. Because this intermediary space is in such close proximity to our body, it enables us to form a unique connection with the objects within our reach. Prior studies have supported the existence of a "*hand-centered coordinate system*" within the peripersonal space [19, 33]. Makin et al., for example, identified brain areas that exhibited significantly stronger activation patterns to visual stimuli the closer that the stimuli were to the hand [34]. Participants were also found to detect targets near the hand with a faster response time than targets that appeared farther away from the hand, suggesting that "*the presence of the hand prioritized the space near the hand for attentional processing*" [46]. These findings suggest that there can be benefits to leveraging (i) the proprioceptive senses to determine the position and orientation of the arms and (ii) the peripersonal sensing of objects surrounding our arms for interaction in VR.

2.2 Arm-Anchored User Interfaces

Bimanual interactions induce asymmetric divisions of labor, where the motion of the non-dominant hand creates a *frame* into which the motion of the dominant hand inserts *content* [20, 32]. Guided by these findings, Bier et al. proposed Magic Lenses, where the non-dominant hand created a see-through interface as *context* and the dominant hand acted within that context [8]. The Worlds-in-Miniature (WIM) metaphor represented surfaces in a virtual environment held by the non-dominant hand, while the dominant hand held a "buttonball" for selection and manipulation [54]. The "hand-relative-to-hand" frame of reference has been demonstrated to provide perceptual cues that are independent of visual feedback [26], motivating the use of both hands for interaction in VR to increase the degree of manipulation. Commercial products have also explored attaching UI elements directly to the non-dominant hand (e.g., Hand Menu for HoloLens 2 [37] and Wearable Menu for Leap Motion [27]). This work focuses on 3D layouts of interfaces that are anchored to the non-dominant arm while interaction

is performed by the dominant hand, echoing the aforementioned benefits of asymmetric divisions of labor [8, 20, 32, 54].

Proprioception has been leveraged to help users locate UI elements projected directly onto the skin of their hands. Prototypes have utilized the skin of the palm or fingers as input or output devices, e.g., keyboards [22, 61, 67], trackpads [62], color palettes [22, 67], TV remotes [14], menu containers [10, 21], or displays for various applications [39]. As the palm and fingers have small surface areas, others have investigated the use of the entire arm for interaction. Azai et al. proposed a menu widget that would enable touching, dragging, sliding, and rotating operations on a user's forearm [4]. Researchers have also developed vision-based approaches and used bio-acoustic sensors to perceive touch input on the skin and render menu items on the forearm or palm for output [22–24, 67]. Other projects have used the forearm as a trackpad for 2D cursor movement or stroke commands [7, 43, 51, 67]. In our work, we hypothesize that arm-anchored 3D interfaces benefit from proprioception because the UI controls are always within reach and users can move their arms to minimize occlusions [38].

2.3 Empirical Studies on Pointing Tasks

A variety of empirical studies have evaluated user performance while acquiring targets in 3D environments. Users are able to store and recall a large number of applications on their hand and forearm if they use landmarks (e.g., fingers, scars, and tattoos) [6]. Dezfuli et al. evaluated the effectiveness of eyes-free targeting of nine landmarks on the user's palm and found that the center of the palm achieved the highest accuracy, whereas the pinky finger achieved the lowest accuracy [14]. Weigel et al. investigated six input locations on the upper limb and found that the forearm was most preferred for perceived ease and comfort, while the elbow and upper arm were least preferred [63]. Lin et al. found that users were able to precisely tap up to 6 distinct points between their wrists and elbows, and that haptic feedback could help to discriminate where the forearm was touched during eyes-free interaction [29]. Vechev et al. compared tapping speeds on six body parts and found that while cycling, the wrist was the fastest region, however, while running, the wrist was the slowest region because users may experience loss of balance while trying to tap on their wrist [59]. Wagner et al. defined 18 on-body target locations and found that upper body parts achieved higher efficiency compared to lower body parts, however, targets on the non-dominant arm were not included in the study [60]. Experiments by Lediaeva and LaViola found that spatial, hand, and waist menus were faster to choose from than arm menus when using ray-based pointing [28]. To our knowledge, our work is the first to systematically investigate direct selection of 3D targets positioned around a user's arm.

Furthermore, existing work exploring the 3D space around the arm also inspired our study. Xu et al. proposed the Hand Range Interface, which used wrist extension and flexion motions, illuminating the potential for body-centric interactions in VR [64]. Azai et al. implemented the Open Palm Menu, where virtual menu items were displayed around the palm of the non-dominant hand [5]. Dachselt and Hübner proposed a collection of classification criteria for 3D menus [13]. As a result of this work, opportunities exist to systematically investigate user's performance and preferences during pointing tasks in the 3D space around the arm. This work thus complements the literature's understanding of users' pointing behaviours. The corresponding Armstrong guidelines are proposed to support the future design of such interactions.

3 USER STUDY

To investigate the utility and feasibility of non-dominant arm-anchored UI, we conducted an empirical study. The primary goal was to evaluate user performance during pointing tasks at various locations around the arm. The secondary goal was to understand preferred target locations and qualitatively assess user arm movement strategies for different target awareness schemes (i.e., known or unknown target location). Thus, this study seeks to provide a complementary perspective to existing studies that have explored on-skin interfaces or compared the acquisition of targets on different body parts [6, 14, 59, 60, 63].

3.1 Participants

Twelve participants were recruited to participate in the study (8 females and 4 males; Mean=26 years old, range 19-52 years old). All participants were right-handed. Eight participants had used a VR headset less than once a month, whereas the remaining four had never used a VR headset before. The study took an average of 60 minutes to complete and participants were provided with a $20 honorarium as compensation for their time.

3.2 Apparatus

The study system consisted of a Windows 10 desktop computer, an Oculus Rift VR headset, a Vicon server, and 16 Vicon Vantage motion capture cameras that were mounted on the ceiling of the room where the study took place. The VR application was implemented in Unity 2018.2 and ran on a desktop computer.

To mimic free-hand interaction, a Vicon motion tracking system was used for hand and arm tracking. The user's left arm was tracked using an armband consisting of two 3D printed pads with retroreflective markers affixed to the top and bottom (Fig. 2a). The right arm was tracked using a similar pad placed on the dorsal side of the hand and another on the index finger. Position and orientation of the markers was streamed to a Unity application using UDP packets over a wired connection. Inverse kinematics were used to compute the parameters of the right index finger joints, based on the position and orientation of the right hand and index finger. The Unity application displayed virtual arms and hands to the participants that replicated their movements in the real world [41, 50].

All participants were right-handed, so targets were displayed around their left arm and they used their right index finger to select a target. A click to confirm selection was registered using the button of an Oculus remote [42] held in their right hand. All participants were able to freely press the button with their thumb while pointing with their index finger. Data recorded during trials included the position and rotation of the participant's left arm, right hand, and right index finger (25 fps), along with event-triggered logs whenever the button was pressed.

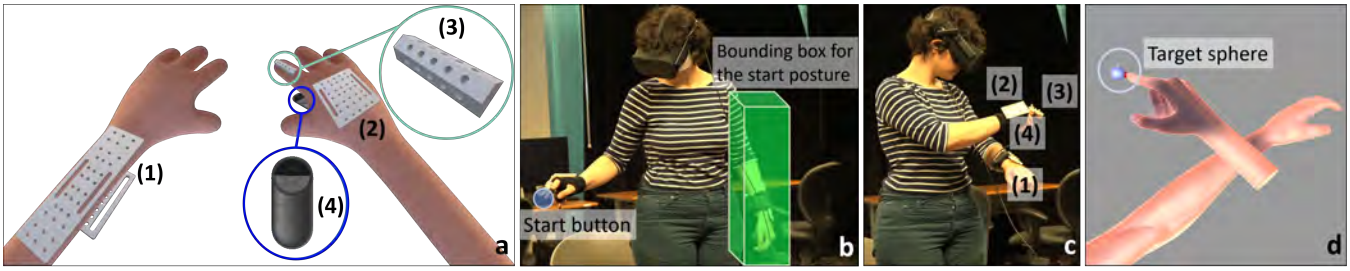

Figure 2: (a) 3D printed pads with retroreflective markers were used to track the user's (1) left forearm, (2) right hand, and (3) right index finger. (4) A hand-held Oculus remote was used to register a click. (b) The trial started when the participant performed the 'start posture' and clicked the start button. (c, d) The trial ended when the participant clicked the target sphere anchored to her left arm.

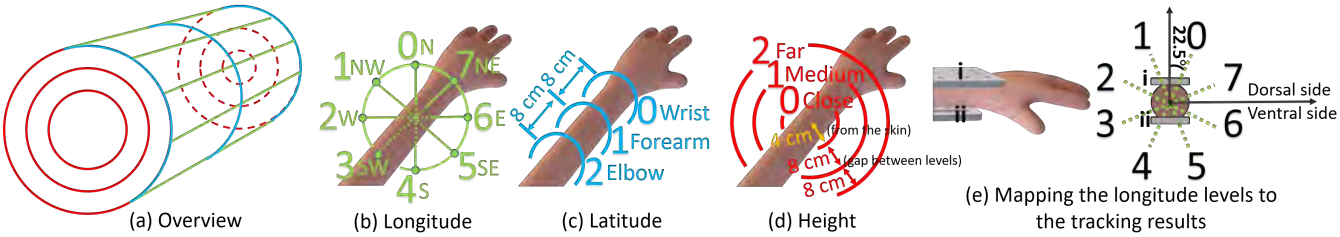

Figure 3: (a) An illustration of the 'conical frustum' grids and depictions of how the (b) longitude, (c) latitude, and (d) height target locations were situated around and along the arm. (e) An illustration of the mapping of the eight longitude levels to the tracking results of the left arm, when the arm was held forward with the 3D printed pads on the (i) dorsal side and (ii) ventral side parallel to each other. The longitude level 0 (N) has an angle of 22.5 degrees perpendicular to the two pads.

3.3 Study Design

During the study, participants performed a series of pointing tasks in VR. A start button was displayed on the right side of the participant's body and could only be pressed when the participant was in a "start posture", i.e., the participant was standing on a virtual footprint on the ground and held her left arm at her side, inside a virtual bounding box (Fig. 2b). After the start button was pressed, the timer started, and a colored sphere appeared around her left arm as the target. The participant was instructed to touch the target using her right index fingertip and press the Oculus remote button as quickly and accurately as possible (Fig. 2c, d). The timer stopped when the button was pressed. The target color changed as a visual cue when the fingertip entered the target and when the target was selected, and auditory feedback indicated whether the trial was successful. If the participant missed the target (i.e., pressing the remote without touching the target with the index fingertip), the trial was added to the end of the queue for the participant to repeat later. Regardless of the result, the participant could then press the start button to continue the study.

All possible target locations were distributed on virtual conical frustum grids centered along the participant's left arm, generating a 'multi-layer interaction space' ([36, 55]; Fig. 3a). Eight longitude levels were evenly distributed around the arm in a counterclockwise rotation. Cardinal and ordinal directions are used to annotate longitude levels, e.g., N for longitude 0 and NW for longitude 1. The longitude levels were positioned based on the tracking results of the 3D printed pads (Fig. 3e). Three latitude levels were placed on the arm moving from the Wrist (latitude 0) up to the Forearm (latitude 1) and Elbow (latitude 2), with an 8 cm gap between levels. Three height levels radiated away from the skin, with distances of Close (height 0, 4cm), Medium (height 1, 12cm), and Far (height 2, 20cm). The Close level was chosen to be close enough to the skin while avoiding collision with our arm-anchored tracking pads. The gap of 8cm between height levels was consistent with the gap between latitude levels, so targets were 'evenly' distributed. Given the differences in arm radii between participants, the base and top radius of each conical frustum was calibrated per participant.

Two target awareness schemes were used throughout the study, i.e., *unknown* or *known*. In the unknown scheme, the target location was not known before the start button was pressed. The participant would need to first identify the target before performing her selection. This scheme was used to simulate users who are not familiar with the layout of a UI. In the known scheme, the target location was visible prior to the trial start, and the participant was asked to locate the target before she pressed the start button. This scheme simulated an experienced user who knows the position of each UI element. By manipulating the target awareness scheme, it would become clear if target selection difficulty arose due to the time it took participants to find the target (i.e., search time) or the time it took them to move to the location where the target was located (i.e., acquisition time).

A pilot study was conducted to determine the size of the targets. In the pilot study (N=4), three sizes of target diameter were evaluated: small (1.0 cm), medium (2.2 cm), and large (4.5 cm). Smaller

targets took longer to find and had higher error rates than larger targets, but without interaction effects between the target size, awareness scheme, or target location. As a result, only medium size targets were used in the study. When rendered in the virtual scene, these targets had a diameter similar to the size of 'short-look icons' on an Apple Watch [1].

The experiment was organized into 10 groups of target selection trials. The first group contained 18 training trials to help participants practice the first target awareness scheme. The next four groups were all performed using the same target awareness scheme (i.e., group 2-5), with 72 trials per group. Following this, participants underwent another training group of 18 trials for the second target awareness scheme (i.e., group 6). They then underwent four additional groups of trials using the same target awareness scheme (i.e., group 7-10), with 72 trials per group. The target awareness scheme presentation order was counterbalanced across participants and the data from the training groups (i.e., first and sixth groups) were excluded from the analysis. Within each non-training group of trials, all 72 target locations (i.e., 8 longitude levels × 3 latitude levels × 3 height levels) appeared exactly once, and were randomly divided into four blocks of trials, with 18 trials per block. After the conclusion of each block of trials, participants were presented with their cumulative error rate. Participants were able to take breaks between blocks and groups of trials. In total, 576 valid trials (i.e., 18 trials per block × 4 blocks × 8 non-training groups) were collected per participant.

3.4 Study Procedure

A questionnaire was used to collect demographic information and prior VR experience. After the questionnaire, the participant donned the Vicon marker pads and Oculus headset. The participant then underwent a calibration procedure to ensure that the virtual arms she would see aligned with her actual limbs. The participant first held the Oculus controller and pressed a button (tracked by Oculus) with her right index fingertip (tracked by Vicon). This action recorded one data point pair between the two coordinates. Participants were then asked to repeat this process 12 times at various locations throughout the environment. The final calibration matrix M, i.e., a 4×4 matrix composed by a translation matrix T and a rotation matrix R, was calculated using a least-squares solution [3].

A tailor's tape measure was used to obtain the circumference of each participant's left wrist and elbow with their sleeves rolled up, so that the scale of the models would align to participants' actual limb measurements. The base and top radius of the conical frustum grids were then adjusted accordingly. All target locations were visible before the study, so participants could practice touch-and-select actions with their index fingertip and the Oculus remote.

The study was video-recorded and one of the researchers recorded field observations during the study. Semi-structured interviews were conducted with the participants after the study to probe their preferences and their self-assessment of the difficulties they had while interacting with various locations on their arm. They were also asked about the strategies they used to perform the pointing task and their reasons for using different arm postures for different targets (if applicable), based on the researcher's observations.

3.5 Metrics

During the study, a trial where the target was not selected would be added to the end of the block queue, until it was correctly completed. As a result, in each block, more than one data point may have been collected for a target if at least one incorrect target selection was made. For each target, the total errors made and the movement time (MT) for the last successful selection were computed. The error rate was calculated as the total errors made averaged over repetitions.

MT alone would not account for the varying distances across trials. Though the distance from the left forearm centroid to the right index fingertip was consistent due to the mandatory 'start posture' (Fig. 2b), the distance from each target to the fingertip varied due to the scale of the conical frustum grids, and MT varied based on participants' individual selection strategies. We considered using Speed as a measure, but this would not factor in the size of the target. To overcome the complexity of these variations, we adopt throughput (TP) as an 'index of performance' [15, 30, 31, 40, 45, 57] based on the actual starting distance (D), size of the target (W), and movement time (MT) in each trial. While this choice limits the granularity of our results, TP offers a consistent and comparable measure of the relative performance across targets, trials, and participants. Although our target width is fixed across trials, we felt using TP would enable the broader reproducibility of our results.

$$TP = \frac{ID}{MT}, \text{ where } ID = log_2\left(\frac{D}{W} + 1\right) \tag{1}$$

To complement the quantitative measures, a total of 12.2 hours of videos were also recorded. Post-study interviews were transcribed and linked to screen-captures from the videos. Themes were coded by two researchers using affinity diagramming [52], focusing on users' subjective preferences for different target locations and the movement strategies they used to acquire targets.

4 RESULTS

4.1 Quantitative Results

4.1.1 Fitts' Law Fitting of Bimanual Pointing Tasks to 3D Arm-Anchored Targets. We investigated whether Fitts' Law [16] holds for bimanual pointing tasks to 3D arm-anchored targets:

$$MT = a + b \cdot ID \tag{2}$$

In the known scheme, where expert behavior was simulated, participants knew the location of targets before they started to move. As a result, the total time that was recorded fulfills the definition of MT in the Fitts' Law model. Because ID was continuous, it needed to be partitioned into q-quantiles [11]. Averaged over 10-quantiles (i.e., 345 data points in each quantile), the model was $MT = 1.0 + 0.18 \cdot ID$ ($r^2 = .18$). In general, Fitts' Law does not hold for this bimanual 3D pointing task, however individual differences existed when the model was evaluated using data from individual participants. By averaging over 10-quantiles, P2's model was $MT = -3.2 + 0.96 \cdot ID$ ($r^2 = .90$) and P7's model was $MT = 1.6 + 0.097 \cdot ID$ ($r^2 = .024$). This distinction among participants could be explained by differing individual strategies that they used to find targets, which include moving the arms with or without rotation, and moving one arm

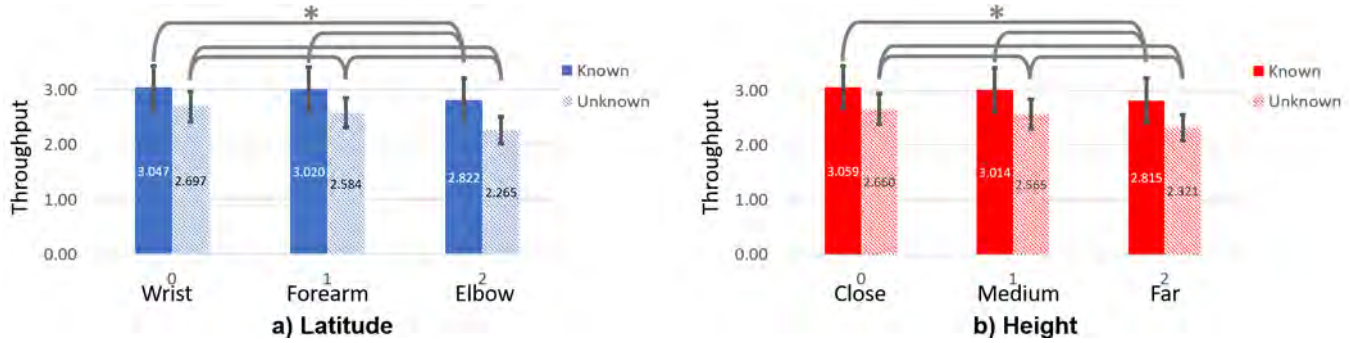

Figure 4: Mean throughput (TP) data segmented by (a) Awareness Scheme × Latitude and (b) Awareness Scheme × Height. Error bars show 95% CIs. Gray lines indicate selected significant pairwise comparisons within the same awareness scheme ($p < .0033$). Other comparisons across different awareness schemes are excluded for brevity.

only or moving both arms simultaneously. As a result, MT varied for targets at similar distances, but different rotation angles.

In addition, since all trials were started in the same posture (Fig. 2b), the variation of ID was only affected by the scale of the conical frustum and the variations allowed by the size of the bounding box ($range = 4.44 - 5.94, \mu = 5.30, \sigma = .23$). We believe that introducing diverse distances and target sizes in future work will increase the variation of ID and further validate the Fitts' Law fitting in such tasks.

4.1.2 Throughput Results. A repeated measures ANOVA was performed using the TP data on the within-subject factors, i.e., Awareness Scheme (2) × Longitude (8) × Latitude (3) × Height (3). Mauchly's tests indicated the assumption of sphericity had been violated for Longitude ($\chi^2_{27} = 57.20, p = .001$) and Awareness Scheme × Latitude ($\chi^2_2 = 17.33, p < .001$), and were corrected using Greenhouse-Geisser (G-G) estimates. Bonferroni corrections were adopted to adjust the significance level α. Mean TP results are presented for each factor (Fig. 6).

Awareness Scheme × Latitude. A two-way interaction was found between Awareness Scheme × Latitude ($F_{1.10,12.07} = 13.02, p = .003, \eta_p^2 = .54$). Paired t-tests ($\alpha = .0033$) revealed that the latitude location of the target influenced participants' TP (Fig. 4a). When the target location was known, TP was significantly higher when the target was located at the Wrist ($\mu = 3.047, \sigma = .625$) compared to the Elbow ($\mu = 2.822, \sigma = .628; t_{11} = 6.48, p < .0001$) and when the target was located at the Forearm ($\mu = 3.020, \sigma = .624$) compared to the Elbow ($t_{11} = 6.82, p < .0001$). Differences of TP comparing targets at the Wrist and Forearm yielded no significance under this condition ($t_{11} = 1.291, p = .2232$). When the target location was unknown, TP was significantly higher when the target was located at the Wrist ($\mu = 2.697, \sigma = .433$) compared to the Forearm ($\mu = 2.584, \sigma = .424; t_{11} = 5.90, p = .0001$) and when the target was located at the Forearm compared to the Elbow ($\mu = 2.265, \sigma = .388; t_{11} = 8.67, p < .0001$). Targets at the Wrist also had higher TP compared to the Elbow ($t_{11} = 9.35, p < .0001$). As for the impact of awareness scheme, higher TP was achieved when the target location is known compared to unknown, and targets at the Forearm ($t_{11} = 3.74, p = .0033$) and Elbow ($t_{11} = 5.00, p = .0004$) revealed significant differences. Targets at the Wrist also had higher TP in the known awareness scheme, though the difference was not significant ($t_{11} = 2.81, p = .0170$). In summary, TP was lower as the target was located farther from the Wrist, regardless of the awareness scheme, and the influence of various latitude locations was more prominent when the target location was unknown.

Awareness Scheme × Height. A two-way interaction was also found between Awareness Scheme × Height ($F_{2,22} = 6.09, p = .008, \eta_p^2 = .36$). Paired t-tests ($\alpha = .0033$) revealed that TP decreased as targets moved farther away from the arm for both known and unknown schemes (Fig. 4b). In the known scheme, TP decreased significantly when targets moved from the Close ($\mu = 3.059, \sigma = .615$) to Far distance ($\mu = 2.815, \sigma = .635; t_{11} = 10.28, p < .0001$), and from the Medium ($\mu = 3.014, \sigma = .625$) to Far distance ($t_{11} = 8.06, p < .0001$). Targets at the Close distance also had higher TP compared to the Medium distance, but the differences were not significant ($t_{11} = 2.60, p = .0246$). When the target location was unknown, TP decreased significantly as the targets moved from Close ($\mu = 2.660, \sigma = .438$) to Medium ($\mu = 2.565, \sigma = .418$) to Far ($\mu = 2.321, \sigma = .376$; i.e., Close vs. Medium: $t_{11} = 5.47, p = .0002$, Medium vs. Far: $t_{11} = 11.31, p < .0001$, and Close vs. Far: $t_{11} = 13.08, p < .0001$). In addition, TP increased when participants knew the target location, particularly for targets at the Medium ($t_{11} = 3.95, p = .0023$) and Far distances ($t_{11} = 4.07, p = .0019$). Differences between the two awareness schemes were not significant when targets were at the Close distance ($t_{11} = 3.52, p = .0048$). Similar to Awareness Scheme × Latitude, differences due to height were more pronounced when the target location was unknown. In general, TP was lower the farther the target was from the arm's surface.

Longitude × Latitude. A two-way interaction was found between Longitude × Latitude ($F_{14,154} = 6.66, p < .001, \eta_p^2 = .38$; Fig. 5). Paired t-tests ($\alpha = .00018$) revealed that TP was higher when the target was located around the Wrist or Forearm compared to the Elbow for most longitude levels: in particular, significant differences were found between the Wrist and Elbow when the target was in the W ($t_{11} = 7.22, p < .0001$), SW ($t_{11} = 10.13, p < .0001$), S ($t_{11} = 8.40, p < .0001$), and SE ($t_{11} = 7.50, p < .0001$) areas. Significant differences were also found between the Forearm and Elbow when the target was in the W ($t_{11} = 7.38, p < .0001$), SW ($t_{11} = 14.18, p < .0001$), and S ($t_{11} = 10.67, p < .0001$)

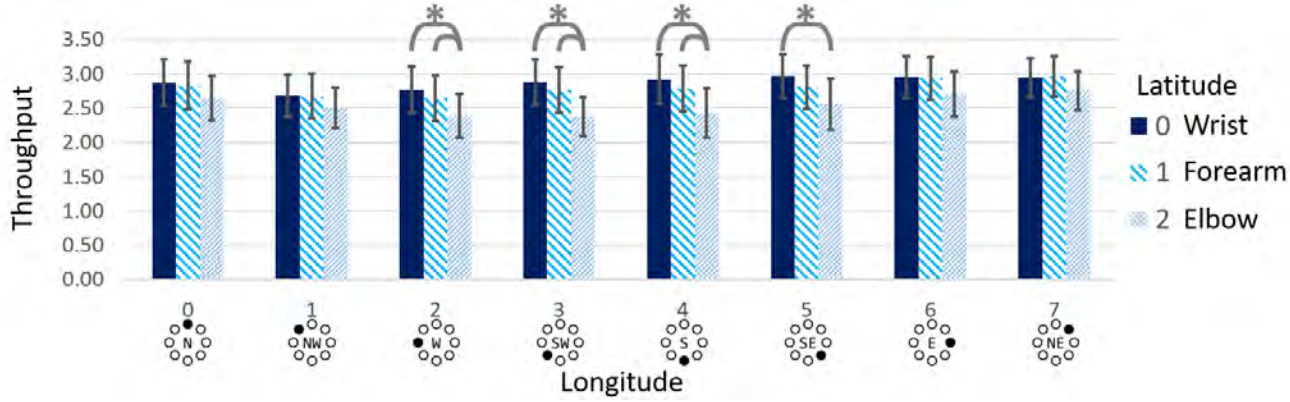

Figure 5: Mean throughput (TP) data segmented by Longitude × Latitude. Error bars show 95% CIs. Gray lines indicate selected significant pairwise comparisons within the same longitude level ($p < .00018$). Other comparisons are excluded for brevity.

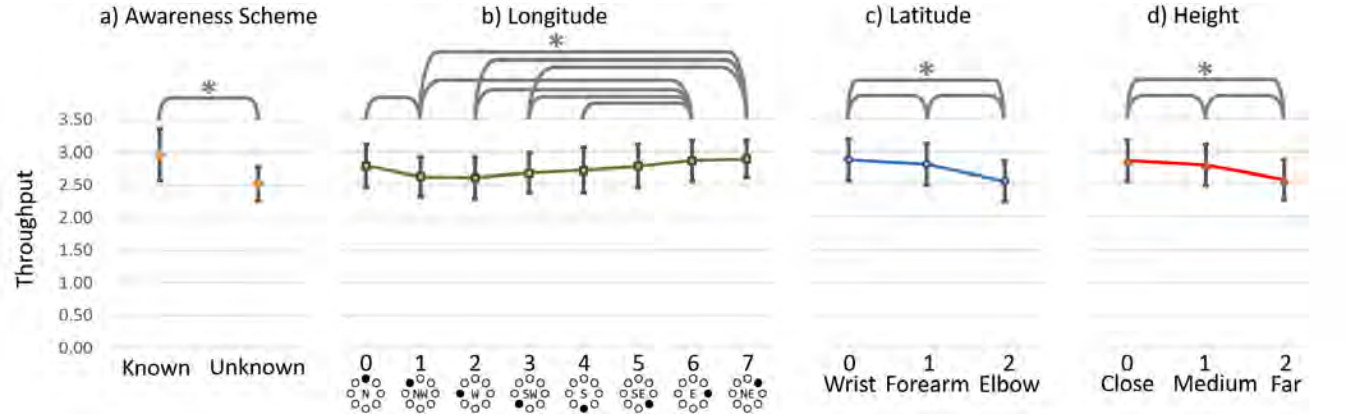

Figure 6: Mean throughput (TP) data segmented by the (a) Awareness Scheme, (b) Longitude, (c) Latitude, and (d) Height. Gray lines indicate significant pairwise comparisons ($\alpha = .05, .0018, .017$, and $.017$, respectively). Error bars show 95% CIs.

areas. No significant results were found while comparing the Wrist and Forearm for targets in the same longitude area. In summary, TP decreased when the target moved from the Wrist to Forearm to Elbow, and the differences between latitude levels were more prominent in the W, SW, S, and SE areas.

Paired t-tests also revealed different orders of longitude levels for each latitude level, though the differences were not significant after Bonferroni corrections. For targets located in the Wrist area, highest TP was achieved in the SE area ($\mu = 2.965, \sigma = .499$), and lowest TP was achieved in the NW area ($\mu = 2.677, \sigma = .482$). For the Forearm, targets in the NE area had the highest TP ($\mu = 2.964, \sigma = .460$), while targets in the W area had the lowest TP ($\mu = 2.649, \sigma = .520$). Finally, for the Elbow, targets in the NE area had the highest TP ($\mu = 2.752, \sigma = .451$), while targets in the SW area had the lowest TP ($\mu = 2.373, \sigma = .449$). In summary, targets placed on the *medial* side of the arm (the side toward the midline of the body, e.g., SE and NE areas) had higher TP, while targets placed on the *lateral* side (the side away from the midline of the body, e.g., NW, W, and SW areas) had lower TP.

Main Effects and Additional Interactions. A main effect of Awareness Scheme was found ($F_{1,11} = 15.10, p = .003, \eta_p^2 = .58$), with TP being significantly higher when the target location was known ($\mu = 2.963, \sigma = .623$) and lower when unknown ($\mu = 2.515, \sigma = .409$; see Fig. 6a). This difference meets our expectation, and is attributable to extra search time in the unknown scheme, which negatively impacted their TP.

A main effect was found for Longitude ($F_{3.11,34.23} = 9.75, p < .001, \eta_p^2 = .47$; Fig. 6b). Paired t-tests ($\alpha = .0018$; see Appendix A: Table 1) suggested that targets located in the NE area enabled participants to exhibit significantly higher TP than targets in the NW, W, and SW areas. Targets in the E area also supported significantly higher TP than targets in the NW, W, SW, and S areas. Those targets located in the N area enabled significantly higher TP than targets located in the NW area. None of the other pairwise comparisons revealed any significant differences. These results suggest that targets located on the medial side of the arm (N, NE, E, SE) are easier for users to interact with than those on the lateral side of the arm (NW,

W, SW, S), perhaps because they are not occluded by the arm and they do not require arm orientation adjustments.

A main effect of Latitude was found ($F_{2,22} = 93.86, p < .001, \eta_p^2 = .90$; Fig. 6c). Paired t-tests ($\alpha = .017$) revealed that TP decreased as targets moved from the Wrist ($\mu = 2.872, \sigma = .493$) to Forearm ($\mu = 2.802, \sigma = .494$; $t_{11} = 3.76, p = .003$) and from the Forearm to Elbow ($\mu = 2.543, \sigma = .485$; $t_{11} = 10.62, p < .001$). Significant difference was also found between the Wrist and Elbow ($t_{11} = 10.53, p < .001$). These results suggest that targets closer to the elbow or upper arm are harder to interact with, perhaps due to ergonomics impacting seeing and touching these targets.

A main effect was found for Height ($F_{2,22} = 156.05, p < .001, \eta_p^2 = .93$; Fig. 6d), with TP decreasing significantly as targets moved from the Close ($\mu = 2.859, \sigma = .497$) to Medium ($\mu = 2.790, \sigma = .493$; $t_{11} = 6.28, p < .001$) distance and from the Medium to Far ($\mu = 2.568, \sigma = .477$; $t_{11} = 12.53, p < .001$) distance. Significant difference was also found between the Close and Far distances ($t_{11} = 13.67, p < .001$). These results reveal that targets are more difficult to interact with as they move farther away from the arm surface, which may indicate a boundary of the peripersonal space around our arms.

In addition, one 2-way interaction between Longitude × Height ($F_{14,154} = 3.74, p < .001, \eta_p^2 = .25$) and one 3-way interaction between Awareness Scheme × Latitude × Height ($F_{4,44} = 2.85, p = .035, \eta_p^2 = .21$) were also found. Due to their small effect sizes, and thus decreased importance, a further examination of these factors was not conducted. No other interactions were found to be significant ($p > .05$).

4.1.3 Error Rate. The overall error rate of all trials was 5.0%. A repeated measures ANOVA was performed using the error rate data on the within-subject factors, i.e., Awareness Scheme (2) × Longitude (8) × Latitude (3) × Height (3). No significant main effects or interactions were found (all $p > .05$), except for a 3-way Awareness Scheme × Longitude × Latitude interaction that had a small effect size ($F_{14,154} = 1.99, p = .022, \eta_p^2 = .15$). All target sizes were held constant, which was a possible cause of the result that no significant differences were found for different target locations. Increasing the task difficulty (i.e., decreasing the target size) in future studies may influence the pointing task accuracy around the arm.

4.1.4 Summary. A TP heatmap is presented for each 2D grid of Longitude × Latitude, as if the 3D conical frustum was unwrapped at each Height level (Fig. 7). It qualitatively visualizes the diverse distribution of TP results of various locations and awareness schemes. In summary, we found that:

(1) TP was higher when participants knew where the target was (awareness scheme: known) and lower when they did not know where the target was (unknown)
(2) Targets on the medial side of the arm (i.e., N, NE, E, SE) had higher TP than those on the lateral side of the arm (i.e., NW, W, SW, S)
(3) TP decreased as targets moved up the arm: Wrist > Forearm > Elbow
(4) TP decreased as targets moved away from the skin: Close > Medium > Far

Among the entire TP heatmap, targets with the highest TP were located at S/SE/E/NE × Wrist/Forearm × Close/Medium areas, which should be the first choices for deploying UI elements around the arm. On the contrary, targets located at the Elbow or in the Far region often exhibited lower TP, so these regions should be avoided while placing UI elements. Some exceptions to these observations do, however, exist: e.g., targets in the E/NE × Elbow × Far areas had higher TP compared to the W × Forearm × Medium area when the target location was known. This suggests that as a user becomes more familiar with a UI layout, she may be able to use more regions. Developers and designers can use these results to position UI controls based on expectations of interaction frequency and user expertise.

4.2 Qualitative Results

The participants' subjective assessments of their performance in various task conditions provided a valuable lens on the quantitative results. Targets around the wrist or closer to the arm were found to be easier, and the interactions that appeared to exist between different target longitude and latitude locations were also supported by the interview results.

4.2.1 Subjective Assessment. Participants used their own words to describe the easiest and hardest regions to select targets. Ten of the twelve participants agreed that targets closer to the wrist were easier to select. Many of their explanations were coupled with longitude descriptions, i.e., "*(the easiest region is) around the top or right side of my wrist, where I can bring them to me*" (P5) and "*on top of my wrist because you just raise your hand and you see the targets*" (P12). Their preferences for targets around the wrist was not surprising, and aligned with the quantitative results demonstrating that the Wrist had the highest TP (Fig. 6c).

However, it was interesting that participants had diverse preferences for the longitude levels. Other than the aforementioned quotes from P5 and P12, P1 preferred targets on the 'right' side (medial side, e.g., E/NE) "*because you don't need to rotate the arm*", P8 found "*bottom side of my wrist is easier because you naturally raise your arm*", and P6 favored "*top or left or right side*". Referring to the TP of longitude and latitude results (Fig. 5), targets located around the Wrist had the highest TP regardless of the longitude level, whereas targets in the S/SE/E/NE areas (i.e., 'bottom/right side') had relatively higher TP. P10 also commented that "*the easiest (target) is around the wrist ... doesn't matter what other axis is*". As for the other latitudes (i.e., the Forearm or Elbow), the E/NE areas had the highest TP, followed by N/SE areas, echoing the participants' comments that targets on the right side are easier to select.

As for height, two participants felt that targets at lower heights were easier to select, as P6 commented, "*when it is closer to the skin I don't need to judge the distance or height, I just need to move closer to the skin.*" Other participants did not discuss height. This aligns with the quantitative results, where targets closer to the arm exhibited higher TP.

When asked about the most difficult task conditions, eight of twelve participants mentioned that targets near the elbow were hard to approach. P7 considered the region near the elbow to be an "*awkward position*" and P2 commented that "*when it gets too*

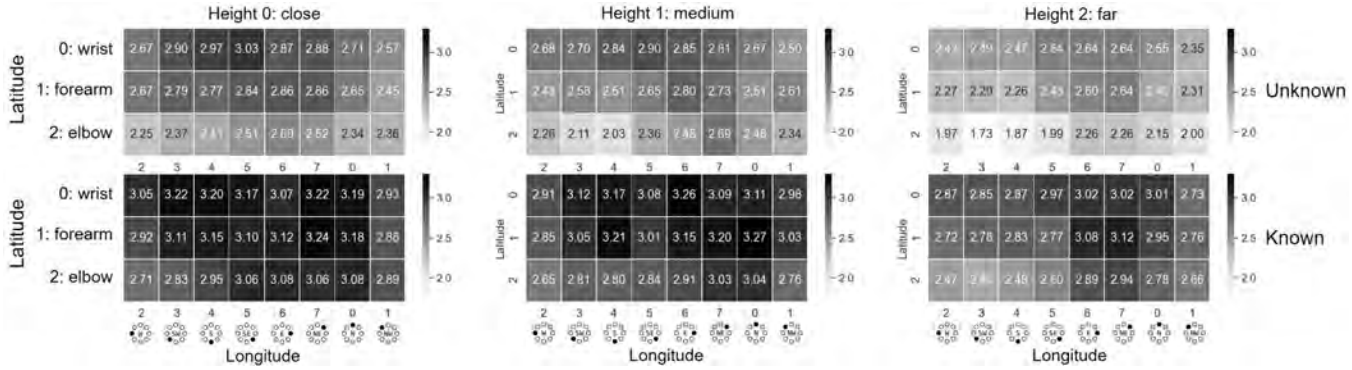

Figure 7: Throughput (TP) heatmap for targets at varying locations, for both unknown (1st row) and known (2nd row) awareness schemes. The x-axis is reordered to [2, 3, … 7, 0, 1] to better illustrate the clustered areas of higher and lower TP.

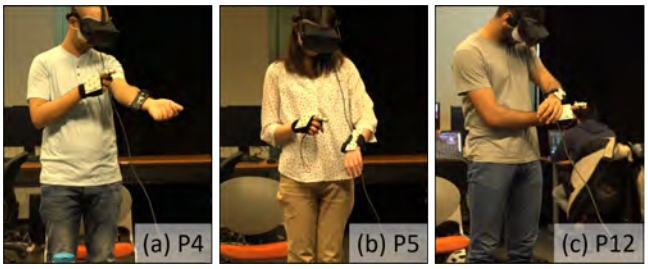

Figure 8: Photos from the user study, demonstrating how P4, P5, and P12 were trying to center the target in front of their body for convenience and better visibility by (a) rotating their arm, (b) lifting their arm, and (c) bending their arm.

close to me, especially when it is here close to the elbow and high enough … it is very close to my head and that is hard." It is interesting that another four participants also used the term "*close to my head*", or descriptors that referenced body parts in the head, to describe these difficult conditions, including "*close to my eyes*" (P4), "*close to me*" (P6), and "*under your eyes*" (P7). In such scenarios, participants might lose track of the target when they moved their arms or the target might be too close to their headset, due to the limits of the nearest rendering plane. As a result, targets in the NE/N/NW × Elbow × Far areas, i.e., "*close to head*" region, had lower TP, especially in the unknown scheme where the participants needed to search for the target during the trial (Fig. 7).

Four participants also found that the targets on the 'left' side (lateral side, e.g., W/SW) were hard to reach, especially when the targets were far away from the arm. Participants sometimes could not see targets in these regions easily (P5/P11), or they had trouble approaching them, as P9 stated, "*it is difficult when it's outside of the arm and I had to turn my arm … it's in very uncomfortable position.*" This could be explained by our TP heatmap: when targets were located in the Close × Forearm areas, SW area had a higher TP than the NW area, however, this contradicts the results that were found in the Medium/Far × Forearm areas (Fig. 7).

4.2.2 Strategies. Participants were also observed using different movement strategies in the known and unknown schemes. While performing the target selection task in the known scheme, four participants mentioned that they would raise their arm if the target was under their wrist (i.e., S/SE × Wrist areas). Three participants said they would twist their arm first if the target was under their elbow (i.e., S/SE × Elbow areas). While approaching targets on the lateral side (W/SW), P10 adopted a more unique strategy, "*I tried to limit the movement (of my left arm) … I would actually use my left arm to stabilize my right arm to make sure the depth was appropriate.*" While many participants were trying to speed up their movements by moving both arms in parallel, P10's strategy enabled them to focus on the accuracy and stability of their target selection instead of their speed.

Further, participants also developed other, more location-generic strategies to help them with their selections. For example, six participants tried to center the target in their field of view (FOV) by moving their left arm, i.e., "*I was trying to bring the targets to the same region, so I can move my right hand to the same region, sort of in front of me*" (P5; see Fig. 8b). Some participants applied this strategy to overcome occlusion issues, such as "*what I was doing is rotating my arm to bring the data points to my visibility*" (P4; Fig. 8a) and "*it's natural for me to rotate my arm to move the target and I can conveniently see the target*" (P11). Others adopted this strategy to ease the ergonomics of their target selections, such as "*if it is further than me I would like to turn my arm closer to me so it's easier to click*" (P9), and "*I try to bend my arm inside to bring the target easy to tap*" (P12; Fig. 8c). Thus, this single goal to bring the target into the center of one's FOV manifested itself via a variety of different user behaviors.

When the target location was unknown, four participants looked for the target by introducing visual motion cues within their FOV, i.e., "*lift my arm and twist at the same time … it gave me a view of everything*" (P8). P6 further explained that "*when I move my left arm, I can see the blur of the blue ball moving, and I knew where to accurately twist my arm.*" Interestingly, many users often apply a similar strategy when searching for their mouse cursor on desktop computers, i.e., quickly moving the cursor back and forth to make the cursor more distinct ("shake mouse pointer to locate" [2]).

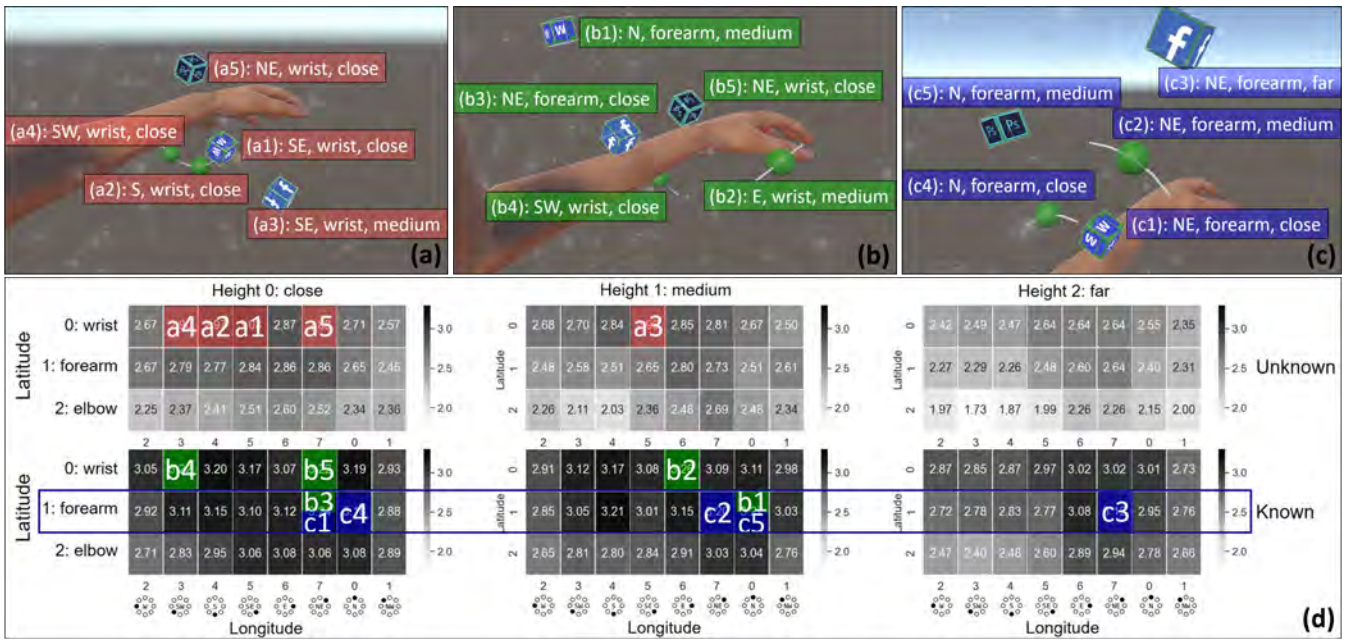

Figure 9: (a, b) Five UI controls are arranged using the heatmap to achieve the highest possible TP, for both unknown (a1–a5) and known (b1–b5) schemes. (c) A slice of the known scheme dataset (e.g., latitude preference: Forearm), could also be used to achieve the highest possible TP (c1–c5). (d) The heatmap annotated with corresponding labels for the UI elements in panels a, b, and c.

5 ARMSTRONG: DESIGN GUIDELINES

The results of the user study demonstrated the potential that the heatmap and the interview results could have to assist in the optimization of UI layouts. We next discuss the design of arm-anchored UIs for unknown, diverse user populations, and suggest how these findings could be integrated within a plugin for real world use.

5.1 Optimize Aggregated Throughput

The heatmap results could be used to optimize overall TP while arranging UI controls around the arm. For example, if a designer wishes to deploy a collection of icons on a smart wristband, the heatmap results for the Wrist × Close dimensions revealed that SW /S /SE /E /NE areas would be suitable choices (Fig. 7). Furthermore, if more icons need to be added, the heatmap suggests a number of possible axes along which to position icons while maintaining overall TP. Based on the goal of the design, the *unknown* or *known* heatmap data could be used to find the best arrangement plan (Fig. 9a, b). The designer could also make use of partial data from the heatmap. For example, if the designer wishes to have all icons along the Forearm, the corresponding row slice from the heatmap could be used to place the icons on the same surface along the Forearm and achieve the highest aggregated TP (Fig. 9c, d).

Differences in TP between the Wrist and Forearm areas were found to be smaller for the NW /N /NE /E areas (Fig. 5). Therefore, it may be possible that some directional UI controls, such as sliders, could be placed in these regions (e.g., from Wrist to Forearm in the N area). As a result, smooth interaction could be achieved when using controls that make use of this axis. The interview results revealed that participants preferred the targets around the wrist, which is not surprising given that users are already familiar with interacting with watches and wristbands. Interestingly, as the latitude level increases, targets get closer to the user's shoulder and head, and these locations were considered to be "*awkward*" (P7) and were not preferred by the participants. Designers should exercise caution when placing targets close to the elbow or upper arm or reserve these locations for controls where false activation is harmful.

The Close and Medium areas also had higher TP, suggesting that combining on-skin locations with in-air locations could result in increased TP. This result complements prior research that has utilized the surface of the skin as an I/O device [22, 24]. The interviews indicated many participants preferred to center the target to their field of view, which was easier when the target was closer to their arm.

It is worth noting that locations on the skin do not always ensure increased TP, for example, targets in the SE /E /NE × Medium areas in the heatmap had higher TP than targets in the NW /W × Close areas. This suggests that there may be a non-trivial distribution of "hot spots" around the arm that could be utilized. As a result, designers should consider the 3D space around the arm as a whole, rather than simply placing UI controls on the skin. In addition, because targets in the W /NW /N /NE /E areas had similar TP at the Close and Medium heights (Fig. 7), linear UI

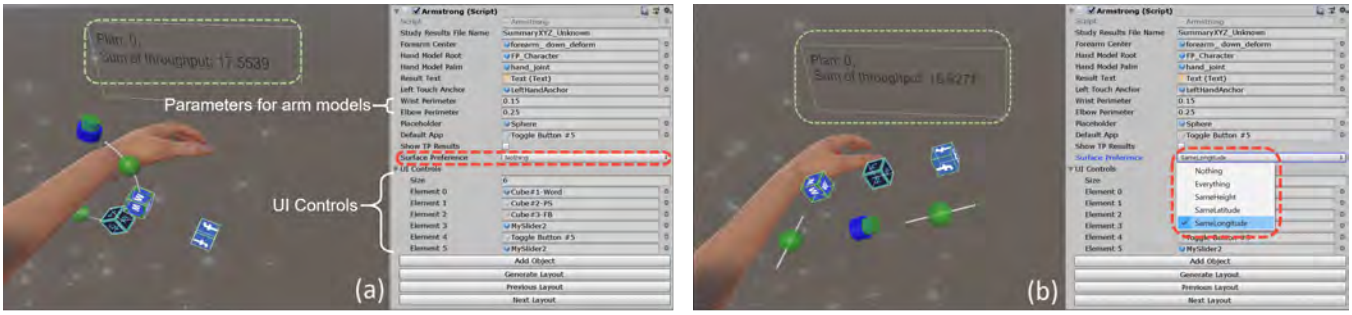

Figure 10: Screenshots of the Unity plugin for Armstrong guidelines. Possible layouts for UI controls are displayed to the user. (a) When no 'surface preference' is designated by default (red box), the layout is optimized to achieve the highest overall TP (green box). (b) When a preferred surface is selected, e.g., "same longitude" (red box), the distributions of targets with highest overall TP are limited to that surface. More than one preference can be selected at a time.

containers like the 'drop-down' menus could be positioned perpendicular to the arm in these regions (e.g., extending from the skin into the air in the N area).

Based on the results from our study, we summarize the following Armstrong design guidelines:

(1) When TP is of priority, the UI controls should be placed on the medial side of the arm (i.e., N, NE, E, and SE), around the wrist, and close to the skin. Alternately, if UI controls are expected to be used infrequently or avoiding accidental activation is critical, they could be deliberately placed in the 'low-TP' areas, such as the lateral side of the arm (i.e., S, SW, W, and NW), around the elbow, and far away from the skin.
(2) When arranging a number of UI controls around the arm, the TP heatmap could be referenced to optimize the layout. Partial data from the heatmap could also be leveraged if certain constraints exist, such as when arranging controls around a wristband, a row slice of the heatmap dataset at Wrist × Close could be used.
(3) For UI controls that span multiple cells of the heatmap (e.g., a slider or a drop-down menu), TP results of individual cells could help to optimize the arrangement of the sub-components: the most commonly used sub-component shall be placed at the position with the highest TP.
(4) UI controls close to the user's shoulder or head are 'awkward' to interact with, as they are more likely to overlap with the user's head when the user is raising or bending her arm (e.g., N × Elbow × Far is a risky area). Designers should be cautious when placing targets close to these areas.
(5) Users who are not familiar with the UI layout will need extra time to find the target UI controls. Visualization techniques could be used to create motion cues to help find UI controls, especially when they are placed in the low-TP areas.

5.2 Unity Plugin for Armstrong Guidelines

To illustrate how these guidelines could be used in practice, a Unity plugin was implemented. The TP heatmap generated from the study (Fig. 7) was used as input and potential locations of UI controls were distributed around the user's arm according to the conical frustum grids (Fig. 3). These locations were stored in a sorted queue based on their TP. Designers can create UI controls in Unity and drag them to this plugin. After loading the perimeters of the user's arm and setting surface preferences for the UI controls, the plugin will automatically position and orient desired UI controls around a virtual arm model optimizing the highest aggregated TP (Fig. 10). Designers can navigate through various layouts generated by the plugin and compare their TP values in real time. By default, the targets are arranged freely in space to optimize aggregated TP when no surface preference is designated (Fig. 10a). UI controls like sliders could also be optimized for orientation based on the aggregated TP results and their individual shape. If a designer prefers to have all UI controls along the same axis, e.g., the same longitude level, the corresponding surface preference checkbox could be selected to enable a customized layout (Fig. 10b). Then, all targets positioned in the E area, for example, could optimize the aggregated TP and the sliders would be re-orientated to follow these restrictions (i.e., either along the latitude or height directions). More than one option, including "same longitude", "same latitude", and "same height", are available to meet various designer's requirements.

6 DISCUSSION

The user study investigated user performance and preference for targets along the longitude, latitude, and height axes. The results revealed that participants tended to center targets in their FOV and that targets closer to the skin, located around the wrist, or placed on the medial side of the arm could be selected more quickly than targets in other locations. These results informed the Armstrong guidelines, based on the TP heatmap, which are demonstrated in a plugin to help with the design process of arm-anchored UIs. Herein, the implications and limitations of this work are discussed.

6.1 Design for Both Known and Unknown Schemes

While placing UI controls that need to be universal across applications, such as the window-frame controls, users will gradually learn the target locations as they use the system and thus prepare their arm motions before approaching possible target locations. As a result, the TP heatmap results and participants observations in

the known scheme could be leveraged for the placement, and the pre-selection information could be utilized within a recognition algorithm to predict target locations or provide hover-like feedback to the users about their selection before they make it.

When designing layouts for UI controls that are dynamic or adaptable across different applications, users who are not familiar with the interface will need to spend extra time searching for targets. Participants were observed quickly twisting their arms to generate visual cues to help them with their search. Inspired by this behavior and the "shake mouse pointer to locate" feature from Mac OS [2], such arm-anchored interfaces could enlarge UI controls if they detect that a user is beginning to twist her arm or utilize other motion-based cues to help find targets.

6.2 Granularity of Arm-Anchored Targeting

This study investigated the impact of the four variables, i.e., longitude, latitude, height, and awareness scheme, on pointing task performance. The quantitative results compared different discrete levels; however, the number of levels was not large enough to enable smooth interpolation between the levels, especially for the latitude and height variables. As the interviews demonstrated, selecting targets close to the head felt awkward and challenging, so it appears that there may be an upper bound on the latitude (upper arm or shoulder) and height (up to the head) for arm-anchored target selection. As we design for a larger space around the arm, ergonomics becomes another important factor: the maximum rotation angle and the comforts of the arm shall be considered to maintain the system's usability. Future work could increase the number of levels of each variable to increase the possible areas of interaction and also reduce the gap between levels to ensure that more fine-grained, 'continuous' data can be obtained and used to interpolate between the various levels.

6.3 Targeting in VR versus AR

The user study was conducted in VR due to the limited FOV and tracking capabilities available for AR devices today. The study results and design guidelines could help designers to create arm-anchored interfaces in VR, while further investigation is required to generalize the results to AR. Although the size of the VR arm models was calibrated for each participant using their own wrist and elbow measurements, the arm models inevitably looked different from their real arms, which may have caused them to behave and move differently than they would have if they were looking at their own actual arms. The surface texture details have been found to have a high impact on the visual perception in a study analyzing anatomical structures of the liver in VR [25]. On the other hand, arm-anchored UIs in AR may face unique challenges, for example, the user's own clothes may clash with the overlaid content or interfere with the rendered colors. As AR devices have been improving with larger FOV and better body-tracking functionality, future research should explore how these results apply to other hardware and modalities such as AR.

6.4 Quantitative Design Techniques for Various UI Controls

The Unity plugin for Armstrong guidelines adopted a simple objective function to maximize the sum of TP to achieve the best overall performance. As more UI controls are involved, the disambiguation between adjacent controls becomes another important factor, especially when future work explores an increased number of interface layers. Involving a 'density factor' in the objective function might be a promising direction [9]. The importance or usage frequency of each UI control could be another useful factor to adjust the weights in the function. Zhai et al.'s work also inspired us to consider the equivalent of 'digraph frequency' in the context of 3D UI [66]. A weighted combination of TP results and the aforementioned factors could be involved in a multi-objective function to optimize the UI layout [58].

This study investigated target selection performance using a colored sphere with status feedback that simulated a 3D button. This design decision enabled for a comparison of the performance between different locations without bias from the design of the actual UI control (e.g., size, shape, or function). The interaction between the size and the location of a target is worth investigating further because it appears that different target properties influence selection. For example, targets in the Elbow × Far areas could be enlarged to achieve higher TP, enabling for a smooth and consistent user experience over the entire space. In addition, future research could investigate compound UI controls (e.g., a number of targets in a drop-down list) and analyze how the individual grid cells of the TP heatmap could help in such scenarios.

6.5 Limitations

The results of this study are limited by the sparse distribution of potential locations and the upper bounds of latitude and height levels. In addition, though simulated skin textures were applied to participants' virtual arms and size adjustments were made to the virtual arm models, they inevitably looked different from the user's own arms. Further studies are thus necessary to validate the feasibility of transferring these results to an AR system. We demonstrated that the pointing strategies adopted by individual participants were inconsistently modeled by Fitts' Law. Future work to develop a model that accounts for the contributions of positional and rotational movements to MT will help to better understand the dynamics of these bimanual interactions. This work also focused more on the impact of various target locations around the arm, rather than on the implications of individual target properties (e.g., size or shape of the UI element). Opportunities exist to, for example, evaluate how targets might be enlarged in low-TP locations to compensate for the difficulty users have selecting them.

7 CONCLUSION

In summary, this work has explored the efficiency and accuracy of pointing tasks with targets situated around the non-dominant arm. The impact of the target's longitude, latitude, and height levels are investigated, contributing to the literature on proprioceptive sensing of arm-anchored UIs. Our study results demonstrated that targets that were closer to the skin, located around the wrist, or placed on the medial side of the arm were selected more quickly

than those far away from the skin, located around the elbow, or placed on the lateral side of the arm. Findings derived from participants' subjective comments and strategies were analyzed to supplement the quantitative results and found that participants tended to center the target in their field of view, and they would also introduce motion cues to assist their finding process by rotating their arm. The high-level findings on individual factors, as well as the 'best spots' revealed by the interaction effects of these factors were summarized into a 3D TP heatmap around the arm, and Armstrong design guidelines were proposed to be used by designers to position UI controls around the arm while optimizing for aggregated TP. A Unity plugin was also presented to help designers and developers to automate this process and design better arm-anchored UIs by enabling them to navigate through potential layouts for UI controls.

ACKNOWLEDGMENTS

The authors would like thank Tovi Grossman and Michelle Annett for their valuable comments, and Dan Clarke for recording voice-overs for our video. We also thank our participants for their involvement in our research and their presentation in our video.

REFERENCES

[1] Apple Inc. 2020. App Icon - Visual - WatchOS. https://developer.apple.com/design/human-interface-guidelines/watchos/visual/app-icon/

[2] Apple Inc. 2020. Change Cursor preferences for accessibility on Mac. https://support.apple.com/en-ca/guide/mac-help/mchl5bb12e1e/mac

[3] K. S. Arun, T. S. Huang, and S. D. Blostein. 1987. Least-Squares Fitting of Two 3-D Point Sets. *IEEE Transactions on Pattern Analysis and Machine Intelligence* PAMI-9, 5 (1987), 698–700.

[4] 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 (CHI '17 EA)*. 357–360. https://doi.org/10.1145/3027063.3052959

[5] 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 (AH '18)*. 1–5. https://doi.org/10.1145/3174910.3174929

[6] Joanna Bergstrom-Lehtovirta, Sebastian Boring, and Kasper Hornbæk. 2017. Placing and Recalling Virtual Items on the Skin. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '17)*. 1497–1507. https://doi.org/10.1145/3025453.3026030

[7] 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 SIGCHI Conference on Human Factors in Computing Systems (CHI '18)*. 1–11. https://doi.org/10.1145/3173574.3174138

[8] 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*. 73–80. https://doi.org/10.1145/166117.166126

[9] Renaud Blanch and Michael Ortega. 2011. Benchmarking pointing techniques with distractors. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. 1629–1638. https://doi.org/10.1145/1978942.1979180

[10] D.A. Bowman and C.A. Wingrave. 2001. Design and evaluation of menu systems for immersive virtual environments. In *Proceedings IEEE Virtual Reality 2001*. 149–156. https://doi.org/10.1109/vr.2001.913781

[11] Olivier Chapuis, Renaud Blanch, and Michel Beaudouin-lafon. 2007. *Fitts ' Law in the Wild : A Field Study of Aimed Movements*. Technical Report. https://hal.archives-ouvertes.fr/hal-00612026

[12] Justine Cléry, Olivier Guipponi, Claire Wardak, and Suliann Ben Hamed. 2015. Neuronal bases of peripersonal and extrapersonal spaces, their plasticity and their dynamics: Knowns and unknowns. *Neuropsychologia* 70 (2015), 313–326. https://doi.org/10.1016/j.neuropsychologia.2014.10.022

[13] Raimund Dachselt and Anett Hübner. 2006. A Survey and Taxonomy of 3D Menu Techniques. In *Proceedings of the 12th Eurographics Symposium on Virtual Environments (EGVE '06)*. 89–99.

[14] Niloofar Dezfuli, Mohammadreza Khalilbeigi, Jochen Huber, Florian Benjamin Müller, and Max Mühlhäuser. 2012. PalmRC: Imaginary Plam-based Remote Control for Eyes-free Television Interaction. In *Proceedings of the 10th European Conference on Interactive TV and Video (EuroITV '12)*. 27–34. https://doi.org/10.1080/0144929X.2013.810781

[15] Sarah A. Douglas, Arthur E. Kirkpatrick, and I. Scott MacKenzie. 1999. Testing pointing device performance and user assessment with the IS0 9241, Part 9 standard. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '99)*. 215–222. https://doi.org/10.1145/302979.303042

[16] Paul M. Fitts. 1954. The Information Capacity of the Human Motor System in Controlling the Amplitude of Movement. *Journal of Experimental Psychology* 47, 6 (1954), 381–391. http://www2.psychology.uiowa.edu/faculty/mordkoff/InfoProc/pdfs/Fitts1954.pdf

[17] Google LLC. 2016. Tilt Brush by Google. https://www.tiltbrush.com/

[18] Google LLC. 2017. Blocks - Create 3D Models in VR. https://arvr.google.com/blocks/

[19] Michael S.A. Graziano. 1999. Where is my arm? The relative role of vision and proprioception in the neuronal representation of limb position. *Proceedings of the National Academy of Sciences of the United States of America* 96, 18 (1999), 10418–10421. https://doi.org/10.1073/pnas.96.18.10418

[20] Yves Guiard. 1987. Asymmetric division of labor in human skilled bimanual action: The kinematic chain as a model. *Journal of Motor Behavior* 19, 4 (1987), 486–517. https://doi.org/10.1080/00222895.1987.10735426

[21] Sean Gustafson, Bernhard Rabe, and Patrick 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 (CHI '13)*. 889–898. https://doi.org/10.1145/2470654.2466114

[22] 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 (UIST '11)*. 441–450. https://doi.org/10.1145/2047196.2047255

[23] Chris Harrison, Shilpa Ramamurthy, and Scott E. Hudson. 2012. On-Body Interaction: Armed and Dangerous. In *In Proceedings of the Sixth International Conference on Tangible, Embedded and Embodied Interaction (TEI '12)*. 69–76. https://doi.org/10.1145/2148131.2148148

[24] 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 (CHI '10)*. 453–462. https://doi.org/10.1145/1753326.1753394

[25] Julian Hettig, Sandy Engelhardt, Christian Hansen, and Gabriel Mistelbauer. 2018. AR in VR: assessing surgical augmented reality visualizations in a steerable virtual reality environment. *International Journal of Computer Assisted Radiology and Surgery* 13, 11 (2018), 1717–1725. https://doi.org/10.1007/s11548-018-1825-4

[26] Ken Hinckley, Randy Pausch, and Dennis Proffitt. 1997. Attention and Visual Feedback: The Bimanual Frame of Reference. In *Symposium on Interactive 3D Graphics*. 121–126. https://doi.org/10.1145/253284.253318

[27] Leap Motion Inc. 2020. Documentation – Leap Motion Developer. https://developer.leapmotion.com/documentation

[28] Irina Lediaeva and Joseph J. LaViola. 2020. Evaluation of Body-Referenced Graphical Menus in Virtual Environments. In *Proceedings of the Graphics Interface Conference 2020 (GI '20)*. 1–9.

[29] Shu Yang Lin, Chao Huai Su, Kai Yin Cheng, Rong Hao Liang, Tzu Hao Kuo, and Bing Yu Chen. 2011. PUB - Point upon body: Exploring eyes-free interaction and methods on an arm. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology (UIST '11)*. 481–487. https://doi.org/10.1145/2047196.2047259

[30] I. Scott MacKenzie. 1992. Fitts' Law as a Research and Design Tool in Human-Computer Interaction. *Human–Computer Interaction* 7, 1 (1992), 91–139. https://doi.org/10.1207/s15327051hci0701_3

[31] I. Scott MacKenzie, Abigail Sellen, and William a S Buxton. 1991. A comparison of input devices in element pointing and dragging tasks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '91)*. 161–166. https://doi.org/10.1145/108844.108868

[32] P. F. MacNeilage, M. G. Studdert-Kennedy, and B. Lindblom. 1984. Functional precursors to language and its lateralization. *American Journal of Physiology - Regulatory Integrative and Comparative Physiology* 15, 6 (1984), 912–914. https://doi.org/10.1152/ajpregu.1984.246.6.r912

[33] Tamar R. Makin, Nicholas P. Holmes, and H. Henrik Ehrsson. 2008. On the other hand: Dummy hands and peripersonal space. *Behavioural Brain Research* 191, 1 (2008), 1–10. https://doi.org/10.1016/j.bbr.2008.02.041

[34] Tamar R. Makin, Nicholas P. Holmes, and Ehud Zohary. 2007. Is that near my hand? Multisensory representation of peripersonal space in human intraparietal sulcus. *Journal of Neuroscience* 27, 4 (2007), 731–740. https://doi.org/10.1523/JNEUROSCI.3653-06.2007

[35] D. I. McCloskey. 1978. Kinesthetic Sensibility. *Physiological Reviews* 58, 4 (1978), 763–811.

[36] Tim Menzner, Travis Gesslein, Alexander Otte, and Jens Grubert. 2020. Above surface interaction for multiscale navigation in mobile virtual reality. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces*. 372–381. https://doi.org/10.1109/vr46266.2020.00057

[37] Microsoft Corporation. 2019. Hand Menu – Microsoft Docs. https://docs.microsoft.com/en-us/windows/mixed-reality/hand-menu

[38] Mark R Mine, Frederick P Brooks Jr, The Problem, and Frederick P Brooks Jr. 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)*. 19–26. https://doi.org/10.1145/258734.258747

[39] 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 (MobileHCI '15)*. 963–965.

[40] Atsuo Murata and Hirokazu Iwase. 2001. Extending Fitts' law to a three-dimensional pointing task. *Human movement science* 20, 6 (2001), 791–805.

[41] NatureManufacture. 2017. VR Hands and FP Arms Pack. https://assetstore.unity.com/packages/3d/characters/humanoids/vr-hands-and-fp-arms-pack-77815{#}releases

[42] Oculus Inc. 2016. Oculus Rift remote. https://support.oculus.com/950065025058681/

[43] Masa Ogata, Yuta Sugiura, and Yasutoshi Makino. 2013. SenSkin: adapting skin as a soft interface. In *Proceedings of the ACM Symposium on User Interface Software and Technology (UIST '13)*. 539–543. https://doi.org/10.1145/2501988.2502039

[44] Ken Pfeuffer, Benedikt Mayer, Diako Mardanbegi, and Hans Gellersen. 2017. Gaze + Pinch interaction in virtual reality. In *Proceedings of the 2017 Symposium on Spatial User Interaction (SUI '17)*. 99–108. https://doi.org/10.1145/3131277.3132180

[45] Alexandros Pino, Evangelos Tzemis, Nikolaos Ioannou, and Georgios Kouroupetroglou. 2013. Using kinect for 2D and 3D pointing tasks: performance evaluation. In *International Conference on Human-Computer Interaction*. Springer, 358–367.

[46] Catherine L. Reed, Jefferson D. Grubb, and Cleophus Steele. 2006. Hands up: Attentional prioritization of space near the hand. *Journal of Experimental Psychology: Human Perception and Performance* 32, 1 (2006), 166–177. https://doi.org/10.1037/0096-1523.32.1.166

[47] Giacomo Rizzolatti, Luciano Fadiga, Leonardo Fogassi, and Vittorio Gallese. 1997. The Space Around Us. *Science* 277, 5323 (1997), 190–191. https://doi.org/10.1126/science.277.5323.190

[48] Giacomo Rizzolatti, Cristiana Scandolara, Massimo Matelli, and Maurizio Gentilucci. 1981. Afferent properties of periarcuate neurons in macaque monkeys. II. Visual responses. *Behavioural Brain Research* 2, 2 (1981), 147–163. https://doi.org/10.1016/0166-4328(81)90053-X

[49] George Robertson, Maarten van Dantzich, Daniel Robbins, Mary Czerwinski, Ken Hinckley, Kirsten Risden, David Thiel, and Vadim Gorokhovsky. 2000. The Task Gallery: A 3D Window Manager. In *Proceedings of the SIGCHI conference on Human factors in computing systems (CHI '00)*. 494–501. https://doi.org/10.1145/332040.332482

[50] RootMotion. 2020. Final IK. https://assetstore.unity.com/packages/tools/animation/final-ik-14290

[51] 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 (ISWC '16)*. 108–115. https://doi.org/10.1145/2971763.2971797

[52] Raymond Scupin. 1997. The KJ method: A technique for analyzing data derived from Japanese ethnology. *Human Organization* 56, 2 (1997), 233–237. https://doi.org/10.17730/humo.56.2.x335923511444655

[53] C. S. Sherrington. 1907. On the Proprio-ceptive System, especially in its Reflex Aspect. *Brain* 29, 4 (1907), 467–482.

[54] Richard Stoakley, Matthew J. Conway, and Randy Pausch. 1995. Virtual reality on a WIM: interactive worlds in miniature. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '95)*. 265–272.

[55] Sriram Subramanian, Dzimitry Aliakseyeu, and Andrés Lucero. 2006. Multi-layer interaction for digital tables. In *Proceedings of the 19th Annual ACM Symposium on User Interface Software and Technology (UIST '06)*. 269–272. https://doi.org/10.1145/1166253.1166295

[56] J. L. Taylor. 2009. *Proprioception*. Academic Press. 1143–1149 pages. https://doi.org/10.1016/B978-008045046-9.01907-0.

[57] Robert J Teather and Wolfgang Stuerzlinger. 2013. Pointing at 3d target projections with one-eyed and stereo cursors. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 159–168.

[58] Kashyap Todi, Daryl Weir, and Antti Oulasvirta. 2016. Sketchplore: Sketch and explore with a layout optimiser. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems (DIS '16)*. 543–555. https://doi.org/10.1145/2901790.2901817

[59] Velko Vechev, Alexandru Dancu, Simon T. Perrault, Quentin Roy, Morten Fjeld, and Shengdong Zhao. 2018. Movespace: on-body athletic interaction for running and cycling. In *Proceedings of the 2018 International Conference on Advanced Visual Interfaces (AVI '18)*. 1–9. https://doi.org/10.1145/3206505.3206527

[60] Julie Wagner, Mathieu Nancel, Sean G. Gustafson, Stephane Huot, and Wendy E. Mackay. 2013. Body-centric design space for multi-surface interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. 1299–1308. https://doi.org/10.1145/2470654.2466170

[61] 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 (MobileHCI '15)*. 153–160. https://doi.org/10.1145/2785830.2785886

[62] Cheng-Yao Wang, Min-Chieh Hsiu, Po-Tsung Chiu, Chiao-Hui Chang, Liwei Chan, Bing-Yu Chen, and Mike Y. Chen. 2015. Palm gesture: 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 (MobileHCI '15)*. 217–226. https://doi.org/10.1145/2785830.2785885

[63] Martin Weigel, Mehta Vikram, and Jürgen Steimle. 2014. More Than Touch: Understanding How People Use Skin as an Input Surface for Mobile Computing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. 179–188. https://doi.org/10.1145/2556288.2557239

[64] Xuhai Xu, Alexandru Dancu, Pattie Maes, and Suranga Nanayakkara. 2018. Hand range interface: information always at hand with a body-centric mid-air input surface. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '18)*. 1–12. https://doi.org/10.1145/3229434.3229449

[65] 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 SIGCHI Conference on Human Factors in Computing Systems (CHI '18)*. 1–13. https://doi.org/10.1145/3173574.3173616

[66] Shumin Zhai, Michael Hunter, and Barton Smith. 2002. Performance Optimization of Virtual Keyboards. *Human-Computer Interaction* 17, 2 (2002), 229–269. https://doi.org/10.1207/S15327051HCI172&3_4

[67] 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 ACM Symposium on User Interface Software & Technology (UIST '19)*. 1151–1159.

A TABLES

Table 1: Pairwise comparisons of throughput (TP) data between Longitude levels (α = .0018). For each comparison, the former longitude level (bold) has a larger TP than the latter one. Non-significant pairs were excluded for brevity.

Pairwise Comparison			$\mu\Delta$	$\sigma\Delta$	p
NE	vs.	SW	.213	.175	.0014
NE	vs.	W	.286	.206	.0005
NE	vs.	NW	.269	.137	<.0001
E	vs.	S	.153	.110	.0005
E	vs.	SW	.191	.132	.0004
E	vs.	W	.263	.162	.0002
E	vs.	NW	.246	.170	.0004
N	vs.	NW	.164	.114	.0004