Push, Tap, Dwell, and Pinch: Evaluation of Four Mid-air Selection Methods Augmented with Ultrasonic Haptic Feedback

TAFADZWA JOSEPH DUBE, University of California, Merced, United States
YUAN REN, University of California, Merced, United States
HANNAH LIMERICK, Ultraleap Ltd., United Kingdom
I. SCOTT MACKENZIE, York University, Canada
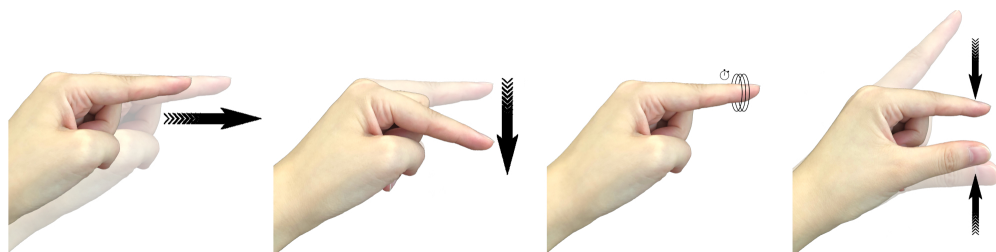AHMED SABBIR ARIF, University of California, Merced, United States

Fig. 1. The four mid-air selection methods explored in this work, with two types of ultrasonic haptic feedback. From left, *Push*, users move the index finger forward like pushing an elevator key, *Tap*, users flick the index finger downwards like tapping on a touchscreen, *Dwell*, users hold the current position of the index finger for 800 ms, and *Pinch*, users pinch using the thumb and index finger.

This work compares four mid-air target selection methods (*Push*, *Tap*, *Dwell*, *Pinch*) with two types of ultrasonic haptic feedback (*Select*, *Hover & Select*) in a Fitts' law experiment. Results revealed that *Tap* is the fastest, the most accurate, and one of the least physically and cognitively demanding selection methods. *Pinch* is relatively fast but error prone and physically and cognitively demanding. *Dwell* is slowest by design, yet the most accurate and the least physically and cognitively demanding. Both haptic feedback methods improve selection performance by increasing users' spatial awareness. Particularly, *Push* augmented with *Hover & Select* feedback is comparable to *Tap*. Besides, participants perceive the selection methods as faster, more accurate, and more physically and cognitively comfortable with the haptic feedback methods.

CCS Concepts: • **Human-centered computing → Gestural input**; **Pointing**; **Haptic devices**.

Additional Key Words and Phrases: Fitts' law, gestural input and interaction, mid-air, in-air, selecting, pointing

Authors' addresses: Tafadzwa Joseph Dube, Inclusive Interaction Lab, University of California, Merced, 5200 N. Lake Road, Merced, California, United States, 95343, tdube@ucmerced.edu; Yuan Ren, Inclusive Interaction Lab, University of California, Merced, 5200 N. Lake Road, Merced, California, United States, 95343, yren5@ucmerced.edu; Hannah Limerick, Ultraleap Ltd., Glass Wharf, Bristol, , United Kingdom, , hannah.limerick@gmail.com; I. Scott MacKenzie, Electrical Engineering and Computer Science, York University, 4700 Keele Street, Toronto, ON, Canada, M3J 1P3, mack@cse.yorku.ca; Ahmed Sabbir Arif, Inclusive Interaction Lab, University of California, Merced, 5200 N. Lake Road, Merced, California, United States, 95343, asarif@ucmerced.edu.

1 INTRODUCTION

Mid-air gestural interaction is more natural and intuitive than traditional interaction methods since it enables direct control of virtual objects using analogies from the real-world [8, 18, 23, 32, 67]. Due to the unreliability of early tracking systems and gesture recognition methods, early work in the area focused on improving gesture detection and recognition. There is also considerable work on eliciting mid-air gestures from users to increase their guessability [61]. Recently, the growing availability, affordability, and reliability of commercial gesture recognition products (e.g., Leap Motion Controller and Microsoft Kinect) there is increased use of three-dimensional (3D) mid-air gestures to interact with two-dimensional (2D) displays (e.g., interactive tabletops and walls, smart televisions, and desktop monitors) and content (e.g., menus and keyboards).

A survey of 80 recently published papers on mid-air gestures [33] revealed that 50% of the surveyed prototypes contained 2D displays and content, 29% contained 3D content, and the remaining 21% of the prototypes did not develop any digital content, but instead asked participants to point at an analog target. This growing interest in interacting with 2D displays and content with 3D gestures is presumably because 3D gestures are more natural [8, 23, 67] and do not necessarily require holding or wearing an extramural device for interaction, like a mouse or a controller. The spread of COVID-19 has also inspired interest in investigating mid-air gestures to enable contactless interaction with public devices, including ATMs and kiosks [17, 28]. However, little work has focused on comparing different mid-air selection methods for desktop and situated displays to identify the best performed gestures in terms of speed, accuracy, and user preference.

Mid-air gestures are difficult to perform due to the lack of spatial feedback. We experience real-world 3D reality by exploring spatial relationships between real-world objects, and performing gestures relative to these objects [23, 51]. The absence of a spatial reference makes mid-air gestures a consciously calculated activity rather a simple and effortless process [5], and this affects speed, accuracy, and ergonomics. To address this, various haptic feedback mechanisms have been proposed, including vibrotactile feedback through wearables, ultrasound, magnetic field repulsion, and air vortex. Most of this work, however, compares novel haptic feedback methods with traditional feedback methods (visual and auditory) rather than comparing the effects of different types of haptic feedback produced by the method (e.g., proximity and actions-based) on mid-air gestures.

In this work, we compare the four most commonly used mid-air selection methods (*Push*, *Tap*, *Dwell*, *Pinch*) [8] with two types of ultrasonic haptic feedback (*Select*, *Hover & Select*). We used a Fitts' law experiment to identify the best performed and most preferred mid-air gestures and haptic feedback methods for target selection. The remainder of the paper is organized as follows. We begin with a review of the relevant literature, then describe the standard Fitts' law experimental protocol. We then discuss the design and development of the investigated mid-air gestures and feedback methods. This is followed with the methodology of the Fitts' law experiment, then the results and design recommendations based on the findings. Finally, we conclude with reflection on potential future extensions of the work.

2 RELATED WORK

2.1 Mid-air Gestures

Performing mid-air gestures is considered a more natural and intuitive mode of interaction than traditional interaction methods as it enables direct control of virtual objects using analogies from the

Table 1. Performance of mid-air selection methods from the literature. Only the highest reported means are listed. The table does not include findings from studies involving virtual and augmented reality and whole-body or multi-modal interaction methods since they are unlikely to be applicable to this work. "Pro." signifies Fitts' law experimental protocol, "Leap" indicates Leap Motion Controller, "IR" signifies infrared cameras, "Bps" indicates throughput in bits/second, and "Bi." signifies bimanual.

Reference	Gesture	Pro.	Baseline	Tracker	Haptics	Bps
Bachmann et al. [4]	*Tap*	1D	*Mouse*	Leap	None	2.7
Foehrenbach et al. [15]	*Pinch + Vibration*	1D	*Pinch*	IR	Vibratory	3.0
Jones et al. [29]	*For-Back Push*	2D	*Mouse*	Leap	None	1.2
Jude et al. [30]	*Dwell (500 ms)*	2D	*Mouse, Touchpad*	Leap	None	2.6
Seixas et al. [50]	*Tap*	2D	*Mouse, Bi. Gesture*	Leap	None	2.3

real-world [8, 18, 23, 32, 67]. Yet, the most commonly used mid-air gestures are not well investigated for desktop and situated displays. There is a large body of work on eliciting and gathering intuitive mid-air gestures for desktop, situated, and large displays [63, 66, 68] and virtual and augmented reality [3]. Researchers have also investigated mid-air hand and whole-body gestures on various platforms, including desktop and situated displays [4, 10, 15, 29–31, 50], large public displays and spaces [1, 42, 45, 64], and augmented and virtual reality [6, 8, 14, 43, 64]. Some have also combined mid-air gestures with other interaction modalities, particularly touch [42], physical buttons [7], eye-gaze [10, 41, 46], and speech [21] to enable multi-modal interaction. Most of this work, however, focuses on comparing mid-air gestures with traditional interaction methods rather than comparing the most commonly used mid-air gestures with each other in terms of performance, user preference, and comfort [32].

Push, *Tap*, *Dwell*, and *Pinch* are the most commonly used mid-air gestures for target selection [4, 29, 30, 58]. Bachmann et al. [4] compared *Tap* with a mouse in a one-dimensional (1D) Fitts' law experiment, where the gesture yielded a 36% lower throughput (2.7 bps) than the mouse (4.2 bps). Foehrenbach et al. [15] conducted a 1D Fitts' experiment to compare the *Pinch* gesture with and without vibratory haptic feedback provided via a digital glove. The study failed to identify any significant difference between the two methods (both yielded about 2.5–3 bps throughputs). Jude et al. [30] compared *Dwell* (500 ms) with a mouse and a touchpad in a two-dimensional (2D) Fitts' law experiment. In the study, the gesture yielded 45% and 28% lower throughputs (2.6 bps) than the mouse (4.8 bps) and the touchpad (3.7 bps), respectively, with the dominant hand. In a similar study, Jones et al. [29] compared *Forward–Backward Push* with a mouse in a 2D Fitts' law experiment, where the gesture yielded a 71% lower throughput (1.2 bps) than the mouse (4.0 bps). The gesture investigated by Jones et al. is different than the one studied in this work. Jones et al. required users to make a forward then a backward push, while in our work users only had to make a forward push to select a target. Seixas et al. [50], in contrast, compared *Tap* with a mouse and a bimanual "grab" gesture. In the study, *Tap* yielded 54% and 15% lower throughputs (∼2.3 bps) compared to the mouse (∼5.0 bps) and the bimanual gesture (∼2.7 bps), respectively. In a different line of research, Cabreira and Hwang [8] showed that users find pointing with the index finger the most natural compared to other pointing approaches. Table 1 summarizes the findings of these works.

2.2 Mid-air Haptic Feedback

In the real-world, we experience 3D reality by experimenting with spatial relationships between tangible objects, and tend to perform gestures relative to these objects [23, 51]. Mid-air gestures are difficult to perform in 3D user interfaces due to the lack of this spatial reference. This increases the

physical and cognitive efforts needed to perform these gestures and compromises their performance by affecting both speed and accuracy. Many novel mid-air haptic feedback methods have been proposed to address these, including vibrotactile feedback [34, 35, 40, 49, 55, 69], magnetic field repulsion [65], and air vortex [19, 52, 53]. Augmenting mid-air gestures with a mid-air haptic feedback method has shown to improve user performance and the overall interaction experience [17, 26, 36, 48, 59]. Badler et al. [5] reported that providing users with spatial reference in 3D selection tasks reduces the effort needed to perform the tasks. Cornelio Martinez et al. [12] demonstrated that mid-air interaction accompanied by mid-air haptic feedback increases users' intentional binding. Yet, most of these methods require users to wear digital bands, rings, or gloves, or use extramural devices that are bulky, intrusive, and often impractical.

Ultrasonic haptic feedback, proposed in early 2000s [24, 25], is a non-intrusive solution that provides touch sensation by sending ultrasonic waves to a target (e.g., fingertip) at different wavelengths [9, 60]. The shear wave induced in the skin tissue triggers the mechanoreceptors within the skin to generate a haptic sensation that is somewhat comparable to a vibratory sensation. The mechanoreceptors response to vibrations between 0.4 to 500 Hz. For a comprehensive review of ultrasonic haptic feedback and its applications see a recent survey [47]. While this method has been compared in empirical studies with traditional feedback methods like auditory and visual [9, 36, 60], to the best of our knowledge, no prior work has investigated the effects of different types of ultrasonic haptic feedback on the performance of the most commonly used mid-air gestures for target selection.
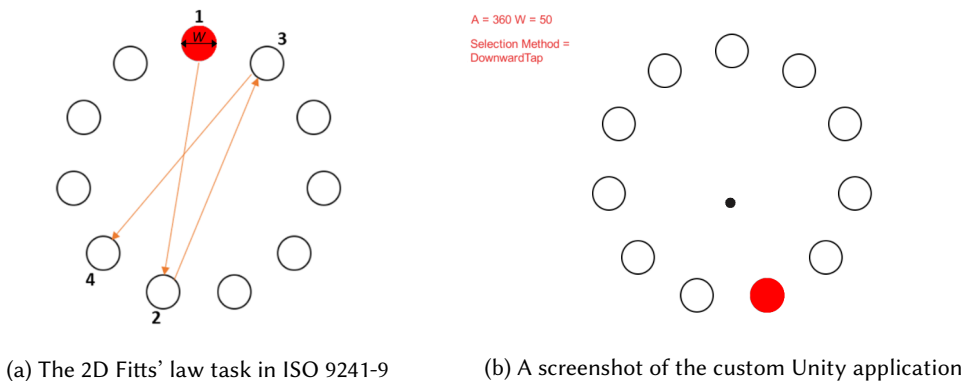
(a) The 2D Fitts' law task in ISO 9241-9   (b) A screenshot of the custom Unity application

Fig. 2. (a) The target to select is highlighted in red. The arrows and numbers demonstrate the selection sequence. (b) Example sequence of trials. The black dot is the cursor.

3 FITTS' LAW PROTOCOL

Fitts' law is a well-established method for evaluating target selection on computing systems [38]. In the 1990s, it was included in the ISO 9241-9 (revised: ISO 9241-411) standard for evaluating non-keyboard input devices by using Fitts' throughput as a dependent variable [54]. The most common multi-directional protocol evaluates target selection movements in different directions. The task is 2D with targets of width W equally spaced around the circumference of a circle (Fig. 2a). Participants select the targets in a sequence moving across and around the circle, starting and finishing at the top target. Each movement covers an amplitude A, which is the diameter of the layout circle. A *trial* is defined as one target selection task, whereas completing all tasks with a given amplitude is defined as a *sequence*. Throughput cannot be calculated on a single trial because

a sequence of trials is the smallest unit of action in ISO 9241-9. Traditionally, the difficulty of each trial is measured in bits using an index of difficulty (ID), calculated as follows:

$$ID = log_2(\frac{A}{W} + 1) \tag{1}$$

The movement time (MT) is measured in seconds for each trial, then averaged over the sequence of trials. It is then used to calculate the performance throughput (TP) in bits/second (bps) using the following equation:

$$TP = \frac{ID}{MT} \tag{2}$$

The revised ISO 9241-9 (9241-411) used in this work [27] measures throughput using an effective index of difficult ID_e, which is calculated from the effective amplitude A_e and the effective width W_e to make sure that the real distance traveled form one target to the next is measured. It also takes into account the spread of selections about the target center.

$$TP = \frac{ID_e}{MT} \tag{3}$$

$$ID_e = log_2(\frac{A_e}{W_e} + 1) \tag{4}$$

The effective amplitude is the real distance travelled by the participants, while the effective width is calculated as follows, where SD_x is the standard deviation of the selection coordinates projected on the x-axis for all trials in a sequence. This accounts for any targeting errors by the participants, assuming that participants were aiming at the center of the targets.

$$W_e = 4.133 \times SD_x \tag{5}$$
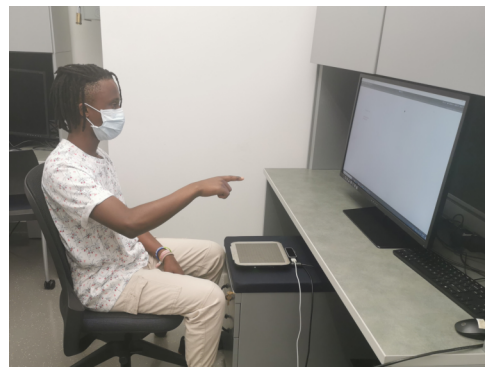
4 EXPERIMENTAL SYSTEM

We developed the experimental system with Unity3D 2019.4.8f1, Leap Motion Orion 4.0.0 SDK, Leap Motion Unity Core Assets 4.4.0, and Ultraleap Unity Core Assets 1.0.0 Beta 9. The system enables users to control a cursor on a computer display by moving the hand. A Leap Motion Controller [56] tracks hand movements 200 mm above the surface, which is the ideal distance recommended by the manufacturer [37], and translates its position into x-y coordinates of the cursor on a vertical display. The system uses the following four most commonly used mid-air gestures for target selection (Fig. 1) [4, 29, 30, 58].

- **Push.** With this method, users point at a target with the index finger then make a forward push to select it. Due to human physiology, this also moves the hand forward, which we exploited to detect *Push* gestures. Based on multiple trials, we used a threshold of 100 mm/s—when the forward velocity (i.e., along the z-axis) of the palm is over this threshold, a push is detected, otherwise, the system interprets it as movements to position the cursor.
- **Tap.** With this method, users point at a target with the index finger then flick the finger downwards to select it. The system detects a tap based on the angle of the index finger. When users point at a target, the finger is usually extended, where the angle between the joints is almost 0° (Fig. 1). When they tap, the angle between the joints changes (the finger becomes non-extended). The system uses the default Leap Motion SDK function to detect this change in the index finger to interpret it as a tap. Users naturally extend the finger after performing a tap, which makes this method reliable and easy to detect. We also considered using the downward velocity of the index finger to detect a tap. But in lab trials, this method was unreliable, resulting in many false positives due to the continuous movement of the hand when positioning the cursor.

- **Dwell.** With this method, users point at a target with the index finger then hold the current position for 800 ms to select the target. We picked the dwell time in a pilot study where 4 participants (2 female, 2 male, 30.3 years, SD = 3.1) selected six circular targets of 40 pixels in diameter, arranged in a circle of 200 pixels in diameter, using 4 dwell times (400, 600, 800, and 1000 ms) in a random order. In the pilot, 800 ms performed the best in terms of accuracy and user preference. This threshold falls within usable dwell times reported in the literature [44].
- **Pinch.** With this method, users point at a target with the index finger then pinch using the thumb and the index finger to select the target. It is detected based on the distance between the index finger and the thumb. When the distance is less than 0.05 mm, a pinch gesture is recognized. The 0.05 mm threshold was selected in lab trials, which revealed that the Leap Motion Controller usually returns values between 0.01 and 0.05 in pinching actions. Like *Tap*, users have to pinch and un-pinch to select a target. Continuous pinching actions are ignored by the system to reduce accidental selections.
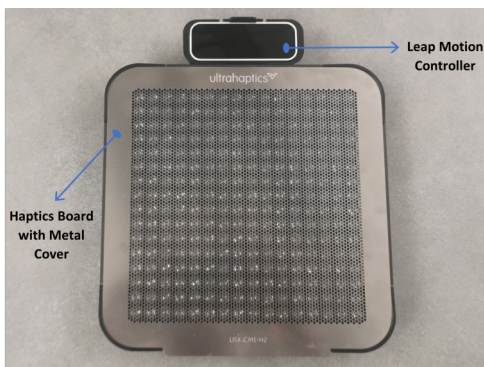
(a) Ultraleap Stratos Explore   (b) The complete experimental setup

Fig. 3. (a) The haptics device with the metal cover used in the study, (b) Participants sat about 700 mm away front of a display. The Ultraleap device was placed on a small table closer to the users for comfortable gesturing actions.

4.1 Ultrasonic Haptic Feedback

The system uses an Ultraleap Stratos Explore [57] haptics board (242 × 207 × 34 mm, 0.7 kg) to provide mid-air haptic feedback (Fig. 3a). The device is a phased array composed of 16 × 16 transducers that operate at a frequency of 40 kHz. The ultrasound waves produced by the transducers focus on a point within 600 mm above the device. When focused on the hand or a finger, the mechanoreceptors in human skin sense the waves as pressure or vibration [9]. The experimental system tracks the hand and the fingers using a Leap Motion Controller, then aims ultrasound waves at the tip of the index finger. Due to the tracking limitation of the controller, discussed earlier, it limits interactions between 200 to 600 mm above the haptics device. Its 700 × 700 mm haptic interaction zone [57] was mapped to a 812.8 mm display using the SDKs default linear function. The haptics board comes with two metal and three acoustic fabric frame-mounted cover materials. The system uses a metal cover (Fig. 3a), however, we were unable to identify any effect of the covers on user performance or preference in lab trials. The default Ultraleap SDK includes several ultrasonic sensations and enables developers to create new ones. We designed two different types of sensations to provide mid-air feedback, described below.

- **Select.** This method provides feedback on selection tasks by applying 30×15 mm sensation on the fingertip for 400 ms. It simulates a Lissajous curve with the default parameters ($a = 3, b = 2$) in the Ultraleap SDK. The sensation was drawn at a frequency of 40 Hz. The dimension of the sensation was picked based on the average human fingertip [13], while the duration was picked in lab trials (50–800 ms) as it provided the most comfortable and noticeable mid-air haptics feedback.
- **Hover & Select.** This method provides feedback on both hover (when the cursor is over a target) and selection tasks. It uses the same mechanism as the *Select* feedback (30×15 mm sensation via a Lissajous curve on the fingertip for 400 ms), but the hover sensation at 80% intensity and the select at 100% intensity (Fig. 4b).

5 METHOD

We conducted a Fitts' law experiment to investigate the performance of mid-air selection methods with and without ultrasonic haptic feedback.

5.1 Participants

Twelve participants took part in the experiment (M = 30.5 years, SD = 4.7). None participated in the pilot study or the lab trials described earlier. Four identified as female, eight as male. All had university-level education. None were experienced with ultrasonic or other mid-air haptic devices. However, two had used mid-air selection methods at least once in virtual reality. Ten self-identified as right-handed, one left-handed, and one ambidextrous. They received US $30 for participating in the study.

5.2 Apparatus

The system described in Section 4 was launched on an ASUS ROG GU501GM Gaming Laptop with Intel core i7 processor, 16 GB ram, NVIDIA GeForce GTX 1060 graphics card, running on Windows 10 operating system. It was connected to an external display, HP Omen 32" gaming monitor at 2569×1440 pixels, where 1 pixel equals to 0.3 mm. The Fitts' law experimental protocol described in Section 3 was developed with Unity3D 2019.4.8f1.

5.3 Operation Area

The operation area was a $400 \times 400 \times 400$ mm cubic area 200 mm above the haptic board (Fig. 4a). The system mapped finger movements in the x- and y-axes inside the area to x-y coordinates of the cursor on the computer display. Hence, the vertical operation plane was parallel to the display. When the cursor was over a target, the haptic board provided 30×15 mm sensation using a Lissajous curve on the fingertip for 400 ms. This fixed feedback area was selected based on multiple trials to provide comfortable feedback on the fingertip. The feedback area did not change based on the size of a target, instead the system dynamically changed feedback position based on movements in the x or y-axis, as appropriate. For example, when the user moved the finger along the the x- or y-axis but the cursor remained inside the target, the system adjusted feedback position to provide seamless feedback on the fingertip. Movements along the z-axis were ignored; that is, the cursor did not change position based on movements in the z-axis. However, when the velocity of the movement exceeded 100 mm/s, a *Push* gesture was registered. Movements in the z-axis were also used to adjust the feedback position. For example, when the user moved the finger along the z-axis as the cursor remained on a target, the system dynamically changed feedback position along the z-axis to provide seamless feedback on the fingertip.
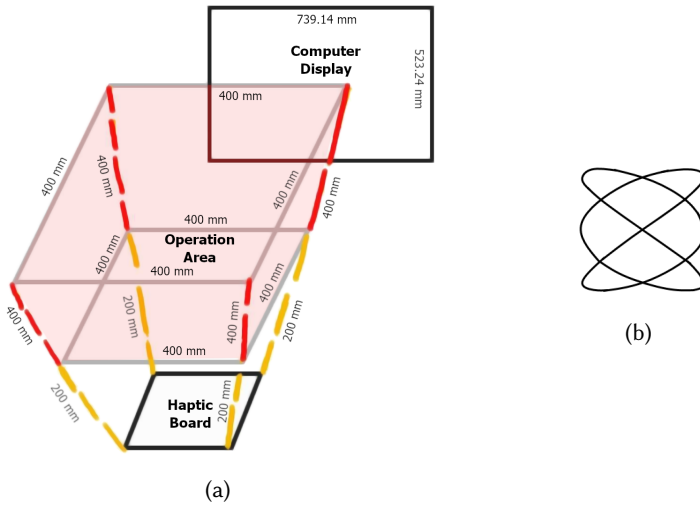
Fig. 4. (a) Operation area in the experiment setup (the red shaded area) and (b) Lissajous curve with parameters $a = 3$, $b = 2$.

5.4 Procedure

The study started with a researcher explaining the research and demonstrating the system to the participants. They then signed an informed consent form and completed a short demographics questionnaire. They sat about 700 mm from the display with the haptics board placed on a small table (Fig. 3b) to provide a comfortable and reliable gesturing position (i.e., 200 mm above the haptic board). At this distance, a target of 100 pixels presents a visual angle of 2.46°.

Participants were instructed to adjust the chair to a comfortable position, if needed. They then took part in a 10-minute training block, where they selected 11 circular targets of 40 pixels in diameter, arranged in a circle of 200 pixels in diameter, with the four mid-air gestures in a random order. The main study started after that, where participants selected targets using the four selection methods augmented with the three feedback methods in a counterbalanced order using a Latin square. They were instructed to select the targets as quickly and accurately as possible without compromising comfort. We enforced a 2-minute break after each three sequences and a 5-minute break after each condition, to avoid fatigue. Participants could also request breaks at any point or extend the duration of the mandatory breaks, when needed. After the completion of all conditions, participants completed the NASA-TLX questionnaire [20] to rate the perceived workload of *only* the four selection methods. The questionnaire was not used to rate all conditions to keep the duration of the study within 60–90 minutes (Section 6.5). Participants also completed a custom questionnaire to rate the examine haptic feedback methods' effects on their performance and preference.

5.4.1 Safety Measures for COVID-19. All researchers involved in this study were fully vaccinated for COVID-19. All participants were pre-screened for COVID-19 symptoms during the recruitment process by a researcher, and on the day of the experiment by the host institute. Both the researcher and the participants wore face coverings and sanitized their hands before a study session. The researcher also maintained a three-foot distance from the participants at all times. All study devices and furniture were disinfected before and after each study session. This protocol was reviewed and approved by the Institutional Review Board (IRB).

5.5 Design

The experiment was a $4 \times 3 \times 3 \times 3$ within-subjects design. The independent variables and levels were as follows:

- Selection method (*Push*, *Tap*, *Dwell*, *Pinch*) counterbalanced
- Haptic feedback (*None*, *Select*, *Hover & Select*) counterbalanced
- Amplitude (80, 360, 640 pixels)
- Width (25, 50, 75 pixels)

There were 11 trials per sequence. The three amplitudes were selected based on the capability of the haptic board and the motion sensor since they are not reliable with amplitudes outside the 80–640 pixels range. Likewise, the three widths were selected based on the smallest width recommended by the manufacturer (25 pixels) [37], while targets with widths over 75 pixels are unrealistic.

5.5.1 Performance Metrics. The dependent variables in the experiment were throughput (TP) and movement time (MT), as described in Section 3, as well as target re-entries (TRE) and error rate (ER). Target re-entries represent the total number of times the cursor re-entered the targets in a trial after having entered them once (count/trial). Error rate signifies the average percentage of incorrect target selections per trial (%), where users performed a selection gesture outside the target.

5.5.2 Graphical Feedback. The experimental software provided graphical feedback when the cursor was over a target by changing the color from red to blue. This feedback was included based on a pilot study, where some participants had difficulty selecting small targets in the no-haptic-feedback conditions as they could not always tell if the cursor is over the target or at the edge. Because this feedback was provided in all conditions, it is not a confounding variable in the study design. Instead, since "changes in object coloring" is the the most common type of feedback provided for target selection with mid-air gestures [33, 62], we argue that this decision increased the external validity of the work.

6 RESULTS

A complete study session took about two hours to complete, including demonstration, question-naires, and breaks. A Shapiro-Wilk test revealed that the response variable residuals were normally distributed. A Mauchly's test indicated that the variances of populations were equal. Hence, we used a repeated-measures ANOVA for all quantitative within-subjects factors (Section 5.5). We report effect size for all statistically significant results. Eta-squared uses the Cohen's [11] interpretation where $\eta^2 = 0.001$ constitutes a small, 0.5 constitutes a medium, and 0.1 constitutes a large effect. There were in total 1,296 observations, none were excluded from the analysis as outliers.

6.1 Throughput

The grand mean for throughput was 2.04 bps. The breakdown by selection method and haptic feedback is presented in Fig. 5a. By selection method, the highest throughput was 2.29 bps for *Tap*, followed by 2.21 bps (*Pinch*), 1.92 bps (*Push*), and 1.75 bps (*Dwell*). The differences were statistically significant ($F_{3,33} = 21.08, p < .0001, \eta^2 = .21$). By haptic feedback, the highest throughput was 2.09 bps for the *Select* and *Hover & Select*, followed by 1.96 bps for the *None*. The differences were statistically significant ($F_{2,22} = 5.80, p < .01, \eta^2 = .02$). *Pinch* with *Select* yielded the highest throughput (2.34 bps). However, the selection method × haptic feedback interaction effect was not statistically significant ($F_{6,66} = 1.64, p > .05$).

	Push	Tap	Dwell	Pinch
■ None	1.75	2.27	1.73	2.09
■ Select	1.94	2.30	1.77	2.34
■ Hover & Select	2.07	2.31	1.74	2.22

(a) Throughput (bps)

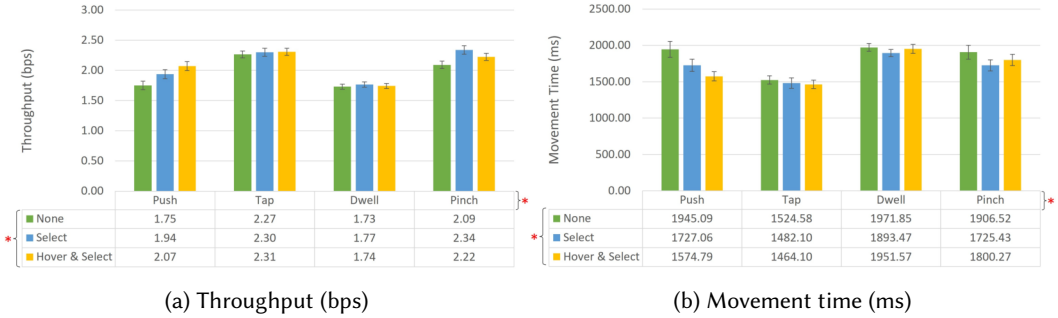|	Push	Tap	Dwell	Pinch
■ None	1945.09	1524.58	1971.85	1906.52
■ Select	1727.06	1482.10	1893.47	1725.43
■ Hover & Select	1574.79	1464.10	1951.57	1800.27

(b) Movement time (ms)

Fig. 5. Average throughput and movement time by selection method and haptic feedback. Error bars represent ±1 standard error. Significant main effects are highlighted with red asterisks.

The breakdown by amplitude and width is presented in Fig. 6. By amplitude, the highest throughput was 2.30 bps for 360 pixels, followed by 2.13 bps (640 pixels) and 1.71 bps (80 pixels). The differences were statistically significant ($F_{2,22} = 67.17, p < .0001, \eta^2 = .24$). By width, the highest throughput was 2.10 bps for 50 pixels, followed by 2.07 bps (75 pixels) and 1.96 bps (25 pixels). The differences were also statistically significant ($F_{2,22} = 5.31, p < .05, \eta^2 = .01$). There was also an amplitude × width interaction effect ($F_{4,44} = 6.09, p < .001$). 360 × 75 pixels yielded the highest throughput (2.38 bps).
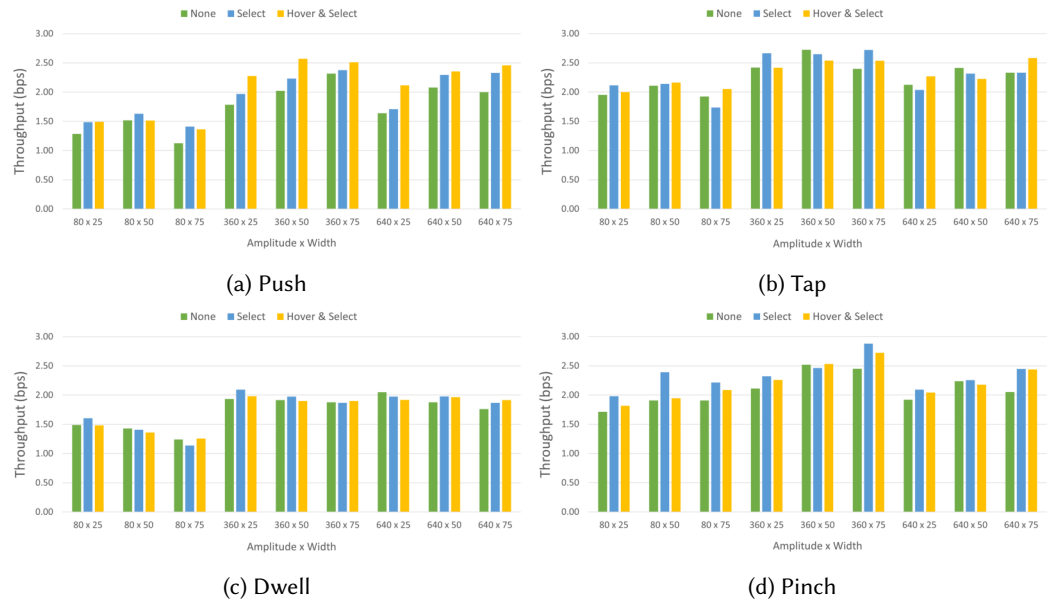

(a) Push

(b) Tap

(c) Dwell

(d) Pinch

Fig. 6. Average throughput (bps) for the four examined mid-air gestures by amplitude × width and haptic feedback.

6.2 Movement Time

The grand mean for movement time was 1747 ms. The breakdown by selection method and haptic feedback is presented in Fig. 5b. *Tap* was the fastest of all selection methods (1490 ms), followed by *Push* (1749 ms), *Pinch* (1810 ms), and *Dwell* (19396 ms). The differences were statistically significant ($F_{3,33} = 8.43, p < .0005, \eta^2 = .06$). By haptic feedback, *Hover & Select* was the fastest (1698 ms), followed by *Select* (1707 ms) and *None* (1837 ms). The differences were statistically significant ($F_{2,22} = 5.54, p < .05, \eta^2 = .01$). However, the selection method × haptic feedback interaction effect was not statistically significant ($F_{6,66} = 1.34, p > .05$). A Tukey-Kramer multiple-comparisons test revealed that *Push* and *Tap + Hover & Select* were significantly faster than the other methods (~20% faster).
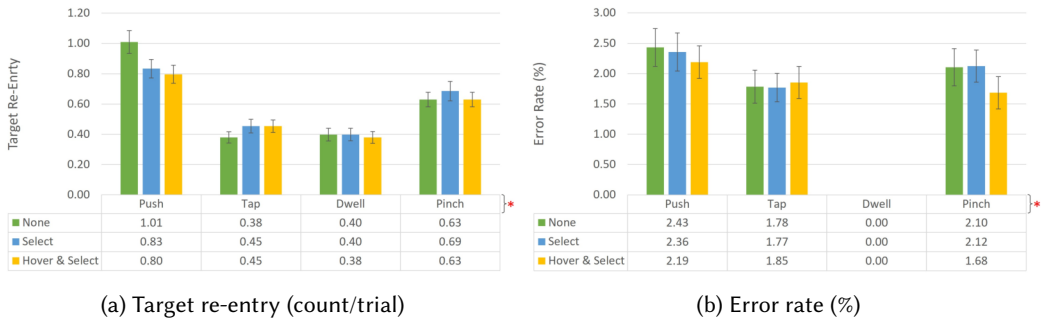

	Push	Tap	Dwell	Pinch
None	1.01	0.38	0.40	0.63
Select	0.83	0.45	0.40	0.69
Hover & Select	0.80	0.45	0.38	0.63

(a) Target re-entry (count/trial)

	Push	Tap	Dwell	Pinch
None	2.43	1.78	0.00	2.10
Select	2.36	1.77	0.00	2.12
Hover & Select	2.19	1.85	0.00	1.68

(b) Error rate (%)

Fig. 7. Average target re-entry and error rate by selection method and haptic feedback. Error bars represent ±1 standard error. Significant main effects are highlighted with red asterisks.

6.3 Target Re-entries

The grand mean for target re-entries was 0.59 count/trial. The breakdown by selection method and haptic feedback is presented in Fig. 7a. By selection method, *Dwell* required the least number of target re-entries (0.39 count/trial), followed by *Tap* (0.43 count/trial), *Pinch* (0.65 count/trial), and *Push* (0.88 count/trial). The differences were statistically significant ($F_{3,33} = 16.46, p < .0001, \eta^2 = .05$). By haptic feedback, *Hover & Select* required the least number of target re-entries (0.56 count/trial), followed by *Select* (0.59 count/trial) and *None* (0.60 count/trial). The differences were not statistically significant ($F_{2,22} = 0.30, ns$). There was also no significant effect of selection method × haptic feedback ($F_{6,66} = 1.18, p > .05$). However, a Tukey-Kramer multiple-comparison test revealed that *Tap* and *Dwell* caused significantly lower target re-entries that *Push* and *Pinch* (~50% lower).

6.4 Error Rate

The grand mean for error rate was 2.06%. The breakdown by selection method and haptic feedback is presented in Fig. 7b. By selection method, *Dwell* was the most accurate with a 0% error rate, followed by *Tap* (1.77%), *Pinch* (1.99%), and *Push* (2.32%). The differences were statistically significant ($F_{3,33} = 25.33, p < .0001, \eta^2 = .05$). By haptic feedback, *Hover & Select* was the most accurate (1.32%), followed by *Select* (1.56%) and *None* (1.57%). The differences were not statistically significant ($F_{2,22} = 1.28, p > .05$). There was also no significant effect of selection method × haptic feedback ($F_{6,66} = 1.11, p > .05$). A Tukey-Kramer multiple-comparison test identified *Push* to be significantly more error prone and *Dwell* to be significantly more accurate than the other methods. The performance of *Tap* and *Pinch* were comparable (1.8–1.9% error rate).

6.5 User Feedback

Participants completed two questionnaires upon the completion of conditions. A NASA-TLX questionnaire [20] to rate the perceived workload of the selection methods and a custom questionnaire to rate the perceived effects of the feedback methods on their performance (speed and accuracy) and physical and mental comfort on a 5-point Likert scale. We did not use the NASA-TLX questionnaire for all ($4 \times 3 = 12$) conditions to limit the duration of the study. Besides, we argue that the overall mental and physical workload of the selection methods and the perceived effects of the feedback methods on user performance is more relevant to this work than the perceived workload of the feedback methods. We used a Friedman test to compare user ratings of the examined selection and haptic feedback methods.

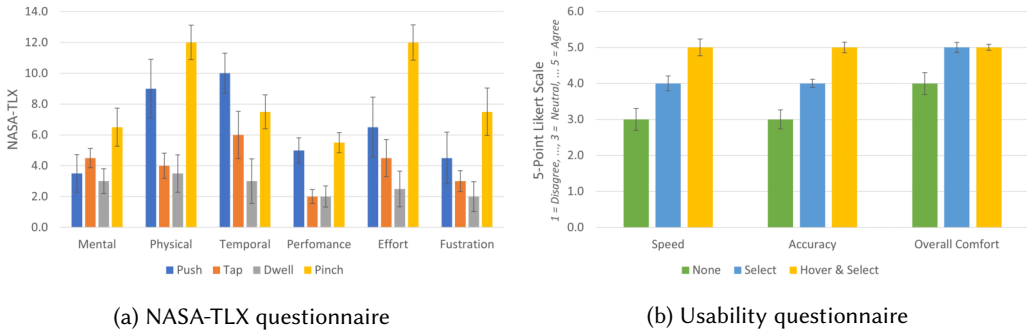
(a) NASA-TLX questionnaire

(b) Usability questionnaire

Fig. 8. Median perceived workload of the examined selection methods and perceived effects of the examined feedback methods on user performance and overall comfort (physical and cognitive). Error bars represent ±1 standard error.

6.5.1 Perceived Workload of the Selection Methods. A Friedman test identified a significant effect of selection method on mental demand ($\chi^2 = 10.32, df = 3, p < .05$), physical demand ($\chi^2 = 16.18, df = 3, p < .05$), performance ($\chi^2 = 11.24, df = 3, p < .05$), effort ($\chi^2 = 11.25, df = 3, p < .05$), and frustration ($\chi^2 = 15.32, df = 3, p < .005$). However, no significant effect was identified on temporal demand ($\chi^2 = 5.64, df = 3, p = .13$). Fig. 8a presents median user ratings of the four selection methods.

6.5.2 Perceived Effects of the Feedback Methods. A Friedman test identified a significant effect of feedback method on speed ($\chi^2 = 11.42, df = 2, p < .005$), accuracy ($\chi^2 = 19.46, df = 2, p < .0001$), and overall comfort ($\chi^2 = 8.67, df = 2, p < .05$). Fig. 8b presents median user ratings of the three feedback methods.

7 DISCUSSION

Tap and *Pinch* outperformed *Push* and *Dwell* in terms of throughput (~20% higher throughput, large effect size). A Tukey-Kramer multiple-comparison test identified these two groups to be significantly different. *Tap* was also significantly faster than the other methods (Fig. 5b). *Dwell* was the slowest of all methods, as expected, since users had to wait for an 800 ms timeout period to select a target. Target amplitude and width influenced the selection methods in accordance to the Fitts' law (large and small effect sizes, respectively, see Fig. 6).

Haptic feedback improved performance of all methods (small effect size). A Tukey-Kramer multiple-comparison test identified the methods with haptic feedback to be significantly faster

and more effective than without feedback. In particular, the performance of *Push + Hover & Select* and *Pinch + Select* elevated closer to *Tap* (Fig. 5a). A Tukey-Kramer multiple-comparison test identified this improvement to be statistically significant. It is important to note that haptic feedback improved performance of all methods despite all providing graphical feedback on collision to aid target selection (Section 5.5.2). This suggests that graphical feedback alone is not effective enough to facilitate mid-air gestural interaction.

We speculate, two factors contributed towards *Tap*'s superior performance. First, based on user responses, performing the gesture did not demand as much physical and cognitive effort as most other methods (Fig. 8a). Second, it did not require a high level of spatial awareness since there was no restriction on how much they could bend the finger, which reduced the total number of re-entries (Fig. 7a), improving its overall performance. A case in point, *Push* without feedback was significantly slower than *Tap* despite being a visually similar gesture (Fig. 5b). A Tukey-Kramer multiple-comparison test revealed that it resulted in significantly more target re-entries than *Tap*, which increased the physical and cognitive effort (Fig. 8a) and affected the overall performance (Fig. 5). Its 1.01 target re-entry rate suggests that participants frequently overshot the targets (Fig. 7a), presumably due to the lack of special reference. With *Push*, participants moved the index finger forward, like pressing a virtual button in the 3D space. Due to the human physiology, this also moved the hand. Without spatial references, it was difficult for the participants to estimate how far they should move the finger to select a target, often moving it too much, which the system interpreted as a pointing action rather than a selection action. This phenomenon has been observed in other 3D interfaces. Hinckley et al. [23] argued that *"to perform a [3D] task, the user's perceptual system needs something to refer to, something to experience"* and *"using a spatial reference [...] is one way to provide this perceptual experience"*. Consequently, target re-entries reduced by 18% and 21%, and throughput increased by 11% and 18%, when *Push* was augmented with *Select* and *Hover & Select* feedback methods, respectively, because the feedback provided the participants with a reference to which they can adjust the finger. Fig. 9 illustrates cursor traces from a random participant for *Push* with the three feedback conditions, where one can see that *Push* without haptic feedback caused multiple target re-entries but none when augmented with a haptic feedback method.
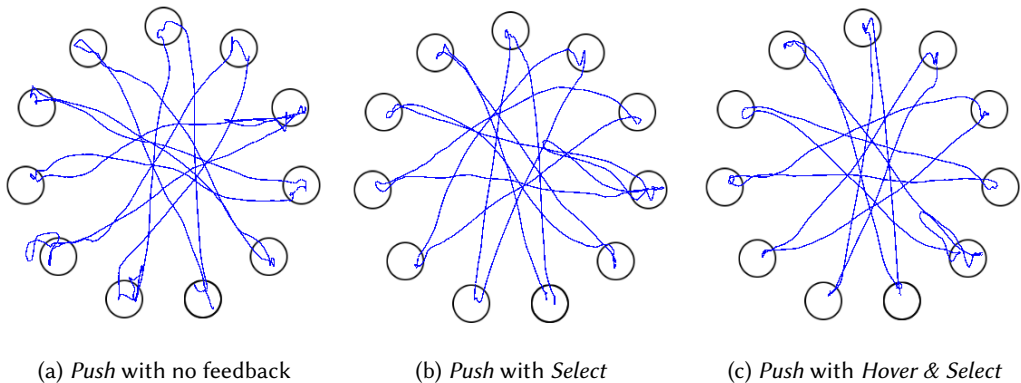

(a) *Push* with no feedback   (b) *Push* with *Select*   (c) *Push* with *Hover & Select*

Fig. 9. Cursor trace examples for *Push* ($A = 360$, $W = 50$ pixels) with the three feedback conditions.

Prior work reported that the performance of 3D interaction methods can improve substantially with practice when spatial references are provided. In an early work, Badler et al. [5] reported that providing users with spatial reference in 3D selection task can make a *"consciously calculated activity"* to a *"simple and effortless process"*. Hence, the performance of *Push* with a haptic feedback

can improve further over time. Relevantly, target re-entries for *Tap* and *Dwell*, which do not rely heavily on spatial awareness, are much lower than the other methods (Fig. 7a). A Tukey-Kramer multiple-comparison test revealed that *Dwell* was significantly more accurate (0% error rate) than the other methods, while *Push* was significantly less accurate than *Tap* and *Pinch*. This is not surprising since with *Dwell* the users did not have to perform any additional action but holding the current finger position for 800 ms. *Tap* was the second most accurate, presumably due to the reasons discussed earlier.

Participants found *Dwell* the least physically and cognitively demanding (Fig. 8a), regardless of it being significantly slower than the other methods. We speculate this is because *Dwell* did not require users to rely on their spatial awareness or perform a gesture that is different than the one used for moving the cursor. As a result, its performance did not improve much with haptic feedback (Fig. 5). *Tap* was the second least physically and cognitively demanding. Interestingly, participants found *Pinch* to be more physically demanding, effortful, and frustrating than the other gestures despite it being more effective than *Push* and *Dwell* in target selection. This could be either because *Pinch* was the only gesture that required the use of two fingers or since it was misrecognized a number of times during the study (about 1.5% of all instances). We discuss this further in Section 7.1.

All participants ($N = 12$) felt that haptic feedback improved their selection accuracy and the overall physical and cognitive comfort (Fig. 8b). Likewise, most participants ($N = 10$) felt that haptic feedback improved their selection speed, while the remaining participants ($N = 2$) were neutral about it.

7.1 Technical Issues

A few technical issues were recorded during the study. First, the Leap Motion Controller seldom stopped tracking the hand (0.01% of all cases). In such cases, we restarted the affected sequence. Second, in general, the system was able to recognize the mid-air gestures with about 100% accuracy, however, in a few occasions (about 1.5% of all cases), it was unable to recognize *Pinch*, in which case, participants performed the gesture again. Finally, the haptic feedback methods were not as effective when the hand was moving fast. However, our observation suggests that it did not affect performance since participants usually slowed down when the finger was closer to the target.

7.2 Generalizability in Different Postures

In the study, participants were in a seated position and selected targets at shoulder level with a bent or extended arm (Fig. 3b). One limitation of the work is it did not explore other possible positions (i.e., standing) and postures (i.e., interaction plane between the shoulder and the waist, at or below the waist, and with an bent arm). We speculate that the performance differences between the gestures will be comparable in different positions and postures in limited use. However, it is possible that the performance of some gestures will be affected more than the others in extended use due to increased "endurance", which is defined as *"the amount of time a muscle can maintain a given contraction level before needing rest"* [22]. Research showed that selecting targets at shoulder level with an extended arm consumes more endurance than targets between the shoulder and the waist [22]. The biomechanics of the upper limbs also suggest that selecting targets below the waist (like on a kiosk) is likely to consume the least endurance as it does not require extending the arm up, thus the arm remains closer to its resting position [16, 39]. Performing the gestures standing up, in contrast, can consume more endurance since users cannot rest their arms on the lap between the tasks. Further investigation is needed in this direction to fully understand the effects of different positions, postures, and gestures on endurance, and to find a definite answer to whether the findings of this work are generalizable to all positions or postures.

8 DESIGN RECOMMENDATIONS

Drawing on the findings of this work, we made recommendations for picking the most appropriate mid-air selection method based on the type of tasks, performance priorities, and technological limitations, summarized in Table 2. Although aiming for the top speed and accuracy in all interactive systems may appear desirable, it is neither necessary nor possible or cost effective in all scenarios. For instance, in a game where players score points by selecting big incoming targets as fast as they can (e.g., fruit slice or slashing games), aiming for a comfortable and fast gesture that supports repetitive performance is sufficient considering the target size and task frequency. Likewise, in scenarios where accuracy is most preferred than speed and the task is not repetitive (e.g., entering PIN on an ATM machine), a more accurate gesture is sufficient (since speed and comfort in non repetitive tasks are not vital). Repetitive tasks are performed repeatedly for a longer period, like the fruit slice game or in text entry. Non-repetitive tasks are performed occasionally, such as pressing a virtual button to exit a window, open a file, or to enter a few characters (e.g., PIN). Hence, the methods appropriate for repetitive actions could be used for non-repetitive actions as well, but not vice versa. We recommend using methods with high accuracy rates, especially for repetitive tasks, since users tend to get impatient and frustrated with error prone gesture-based methods and deem them unusable when error rate is over 3% [2]. However, all methods examined here yielded high accuracy (below 2.5% error rates). Note that the table reports comfort in limited use (within an hour) and does not account for fatigue in extended use.

Table 2. Recommendations for picking the most appropriate mid-air selection method based on the type of tasks (repetitive or not repetitive actions), performance priorities (top, moderate, low), and technological limitations (availability of haptic feedback). Comfort signify perceived workload. "Bps" indicates throughput in bits/second (only throughputs of the best performed haptic feedback are reported). The highlighted fields signify the best performed methods.

Haptic Feedback (Recommended)	Priority				Method	Bps
	Repetitions	Accuracy	Speed	Comfort		
Not Available	Low	Moderate	Top	Low	*Pinch*	2.09
	Low	Top	Low	Top	*Dwell*	1.73
	Moderate	Moderate	Moderate	Moderate	*Push*	1.75
	Top	Top	Top	Top	*Tap*	2.27
Available	Low	Moderate	Top	Low	*Pinch*	2.34
	Low	Top	Low	Top	*Dwell*	1.77
	Top	Moderate	Top	Top	*Push*	2.07
	Top	Top	Top	Top	*Tap*	2.31

9 CONCLUSION

We conducted a Fitts' law experiment to compare the performance of four mid-air selection methods: *Push*, *Tap*, *Dwell*, and *Pinch*, with and without two different types of ultrasonic haptic feedback: *Select*, *Hover & Select*. Results identified *Tap* as the fastest, the most accurate, and one of the least physically and cognitively demanding selection methods. *Pinch* performed well in terms of speed, but yielded a much higher error rate and perceived workload. *Dwell* was the slowest of all methods by design, but interestingly, the most accurate and the least physically and cognitively demanding. Both haptic feedback methods improved performance of the selection methods, presumably by increasing users' spatial awareness. Particularly, the performance of *Push*, which relies on users'

spatial awareness, improved substantially with haptic feedback, making it comparable to *Tap*. Besides, participants perceived the selection methods as faster, more accurate, and more physically and cognitively comfortable with the haptic feedback methods.

10 FUTURE WORK

In the future, we will compare the effects of graphical, auditory, ultrasonic, and hybrid feedback (combinations of graphical, auditory, ultrasonic feedback) on target selection performance. We will also replicate the work in virtual and augmented reality to investigate whether the findings of this work are pertinent to these settings. Further, we will explore the effects of ultrasonic haptic feedback on additional mid-air gestures.

REFERENCES

[1] Christopher Ackad, Andrew Clayphan, Martin Tomitsch, and Judy Kay. 2015. An in-the-Wild Study of Learning Mid-Air Gestures to Browse Hierarchical Information at a Large Interactive Public Display. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '15)*. Association for Computing Machinery, New York, NY, USA, 1227–1238. https://doi.org/10.1145/2750858.2807532

[2] Ahmed Sabbir Arif, Wolfgang Stuerzlinger, Euclides Jose de Mendonca Filho, and Alec Gordynski. 2014. Error Behaviours in an Unreliable in-Air Gesture Recognizer. In *CHI '14 Extended Abstracts on Human Factors in Computing Systems (CHI EA '14)*. Association for Computing Machinery, New York, NY, USA, 1603–1608. https://doi.org/10.1145/2559206.2581188

[3] Rahul Arora, Rubaiat Habib Kazi, Danny M. Kaufman, Wilmot Li, and Karan Singh. 2019. MagicalHands: Mid-Air Hand Gestures for Animating in VR. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology (UIST '19)*. Association for Computing Machinery, New York, NY, USA, 463–477. https://doi.org/10.1145/3332165.3347942

[4] Daniel Bachmann, Frank Weichert, and Gerhard Rinkenauer. 2015. Evaluation of the Leap Motion Controller as a New Contact-Free Pointing Device. *Sensors* 15, 1 (Jan. 2015), 214–233. https://doi.org/10.3390/s150100214

[5] Norman I. Badler, Kamran H. Manoochehri, and David Baraff. 1987. Multi-Dimensional Input Techniques and Articulated Figure Positioning by Multiple Constraints. In *Proceedings of the 1986 workshop on Interactive 3D graphics (I3D '86)*. Association for Computing Machinery, New York, NY, USA, 151–169. https://doi.org/10.1145/319120.319132

[6] 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, Washington, DC, USA, 184–193. https://doi.org/10.1109/VR46266.2020.00037 ISSN: 2642-5254.

[7] Michelle A. Brown and Wolfgang Stuerzlinger. 2016. Exploring the Throughput Potential of In-Air Pointing. In *Human-Computer Interaction. Interaction Platforms and Techniques (Lecture Notes in Computer Science)*, Masaaki Kurosu (Ed.). Springer International Publishing, Cham, 13–24. https://doi.org/10.1007/978-3-319-39516-6_2

[8] Arthur Theil Cabreira and Faustina Hwang. 2015. An Analysis of Mid-Air Gestures Used across Three Platforms. In *Proceedings of the 2015 British HCI Conference (British HCI '15)*. Association for Computing Machinery, New York, NY, USA, 257–258. https://doi.org/10.1145/2783446.2783599

[9] Tom Carter, Sue Ann Seah, Benjamin Long, Bruce Drinkwater, and Sriram Subramanian. 2013. UltraHaptics: Multi-Point Mid-Air Haptic Feedback for Touch Surfaces. In *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology (UIST '13)*. Association for Computing Machinery, New York, NY, USA, 505–514. https://doi.org/10.1145/2501988.2502018

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

[11] Jacob Cohen. 1988. *Statistical Power Analysis for the Behavioral Sciences* (2 ed.). Routledge, New York, NY, USA. https://doi.org/10.4324/9780203771587

[12] Patricia Ivette Cornelio Martinez, Silvana De Pirro, Chi Thanh Vi, and Sriram Subramanian. 2017. Agency in Mid-Air Interfaces. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 2426–2439.

[13] Kiran Dandekar, Balasundar I. Raju, and Mandayam A. Srinivasan. 2003. 3-D Finite-Element Models of Human and Monkey Fingertips to Investigate the Mechanics of Tactile Sense. *Journal of Biomechanical Engineering* 125, 5 (Oct. 2003), 682–691. https://doi.org/10.1115/1.1613673

[14] Katherine Fennedy, Jeremy Hartmann, Quentin Roy, Simon T. Perrault, and Daniel Vogel. 2021. OctoPocus in VR: Using a Dynamic Guide for 3D Mid-Air Gestures in Virtual Reality. *IEEE Transactions on Visualization and Computer Graphics* XX, X (2021), 1–1. https://doi.org/10.1109/TVCG.2021.3101854 Conference Name: IEEE Transactions on Visualization and Computer Graphics.

[15] Stephanie Foehrenbach, Werner A. König, Jens Gerken, and Harald Reiterer. 2009. Tactile Feedback Enhanced Hand Gesture Interaction at Large, High-Resolution Displays. *Journal of Visual Languages & Computing* 20, 5 (Oct. 2009), 341–351. https://doi.org/10.1016/j.jvlc.2009.07.005

[16] Andris Freivalds. 2011. *Biomechanics of the Upper Limbs: Mechanics, Modeling and Musculoskeletal Injuries, Second Edition.* CRC Press, Boca Raton, FL, USA. Google-Books-ID: fmk_F9ZWMoQC.

[17] Orestis Georgiou, Hannah Limerick, Loïc Corenthy, Mark Perry, Mykola Maksymenko, Sam Frish, Jörg Müller, Myroslav Bachynskyi, and Jin Ryong Kim. 2019. Mid-Air Haptic Interfaces for Interactive Digital Signage and Kiosks. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems.* Association for Computing Machinery, Glasgow Scotland Uk, 1–9. https://doi.org/10.1145/3290607.3299030

[18] Sukeshini A. Grandhi, Gina Joue, and Irene Mittelberg. 2011. Understanding Naturalness and Intuitiveness in Gesture Production: Insights for Touchless Gestural Interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* Association for Computing Machinery, New York, NY, USA, 821–824.

[19] Sidhant Gupta, Dan Morris, Shwetak N. Patel, and Desney Tan. 2013. AirWave: Non-Contact Haptic Feedback Using Air Vortex Rings. In *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '13).* Association for Computing Machinery, New York, NY, USA, 419–428. https://doi.org/10.1145/2493432.2493463

[20] Sandra G. Hart and Lowell E. Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Advances in Psychology.* Vol. 52. Elsevier, Amsterdam, The Netherlands, 139–183. https://doi.org/10.1016/S0166-4115(08)62386-9

[21] Benjamin Hatscher and Christian Hansen. 2018. Hand, Foot or Voice: Alternative Input Modalities for Touchless Interaction in the Medical Domain. In *Proceedings of the 20th ACM International Conference on Multimodal Interaction (ICMI '18).* Association for Computing Machinery, New York, NY, USA, 145–153. https://doi.org/10.1145/3242969.3242971

[22] Juan David Hincapié-Ramos, Xiang Guo, Paymahn Moghadasian, and Pourang Irani. 2014. Consumed Endurance: A Metric to Quantify Arm Fatigue of Mid-Air Interactions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14).* Association for Computing Machinery, New York, NY, USA, 1063–1072. https://doi.org/10.1145/2556288.2557130

[23] Ken Hinckley, Randy Pausch, John C. Goble, and Neal F. Kassell. 1994. A Survey of Design Issues in Spatial Input. In *Proceedings of the 7th annual ACM symposium on User interface software and technology (UIST '94).* Association for Computing Machinery, New York, NY, USA, 213–222. https://doi.org/10.1145/192426.192501

[24] Takayuki Hoshi, Takayuki Iwamoto, and Hiroyuki Shinoda. 2009. Non-Contact Tactile Sensation Synthesized by Ultrasound Transducers. In *World Haptics 2009 - Third Joint EuroHaptics conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems.* IEEE, Washington, D.C., USA, 256–260. https://doi.org/10.1109/WHC.2009.4810900

[25] Takayuki Hoshi, Masafumi Takahashi, Takayuki Iwamoto, and Hiroyuki Shinoda. 2010. Noncontact Tactile Display Based on Radiation Pressure of Airborne Ultrasound. *IEEE Transactions on Haptics* 3, 3 (July 2010), 155–165. https://doi.org/10.1109/TOH.2010.4 Conference Name: IEEE Transactions on Haptics.

[26] Seki Inoue, Yasutoshi Makino, and Hiroyuki Shinoda. 2015. Active Touch Perception Produced by Airborne Ultrasonic Haptic Hologram. In *2015 IEEE World Haptics Conference (WHC).* IEEE, Washington, DC, USA, 362–367. https://doi.org/10.1109/WHC.2015.7177739

[27] International Organization for Standardization. 2012. ISO/TS 9241-411:2012. https://www.iso.org/cms/render/live/en/sites/isoorg/contents/data/standard/05/41/54106.html

[28] Muhammad Zahid Iqbal and Abraham Campbell. 2020. The Emerging Need for Touchless Interaction Technologies. *Interactions* 27, 4 (July 2020), 51–52. https://doi.org/10.1145/3406100

[29] Keith S. Jones, Trevor J. McIntyre, and Dennis J. Harris. 2020. Leap Motion- and Mouse-Based Target Selection: Productivity, Perceived Comfort and Fatigue, User Preference, and Perceived Usability. *International Journal of Human–Computer Interaction* 36, 7 (April 2020), 621–630. https://doi.org/10.1080/10447318.2019.1666511

[30] Alvin Jude, G. Michael Poor, and Darren Guinness. 2014. An Evaluation of Touchless Hand Gestural Interaction for Pointing Tasks with Preferred and Non-Preferred Hands. In *Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational (NordiCHI '14).* Association for Computing Machinery, New York, NY, USA, 668–676. https://doi.org/10.1145/2639189.2641207

[31] Mohamed Khamis, Ludwig Trotter, Ville Mäkelä, Emanuel von Zezschwitz, Jens Le, Andreas Bulling, and Florian Alt. 2018. CueAuth: Comparing Touch, Mid-Air Gestures, and Gaze for Cue-based Authentication on Situated Displays.

Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 2, 4 (Dec. 2018), 174:1–174:22. https://doi.org/10.1145/3287052

[32] Panayiotis Koutsabasis and Panagiotis Vogiatzidakis. 2019. Empirical Research in Mid-Air Interaction: A Systematic Review. *International Journal of Human–Computer Interaction* 35, 18 (Nov. 2019), 1747–1768. https://doi.org/10.1080/10447318.2019.1572352

[33] Panayiotis Koutsabasis and Panagiotis Vogiatzidakis. 2019. Empirical Research in Mid-Air Interaction: A Systematic Review. *International Journal of Human–Computer Interaction* 35, 18 (Nov. 2019), 1747–1768. https://doi.org/10.1080/10447318.2019.1572352 Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/10447318.2019.1572352.

[34] Laurens R. Krol, Dzmitry Aliakseyeu, and Sriram Subramanian. 2009. Haptic Feedback in Remote Pointing. In *CHI '09 Extended Abstracts on Human Factors in Computing Systems*. ACM, Boston MA USA, 3763–3768. https://doi.org/10.1145/1520340.1520568

[35] Michael Julian Kronester, Andreas Riener, and Teo Babic. 2021. Potential of Wrist-Worn Vibrotactile Feedback to Enhance the Perception of Virtual Objects during Mid-Air Gestures. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, Yokohama Japan, 1–7. https://doi.org/10.1145/3411763.3451655

[36] David R. Large, Kyle Harrington, Gary Burnett, and Orestis Georgiou. 2019. Feel the Noise: Mid-Air Ultrasound Haptics as a Novel Human-Vehicle Interaction Paradigm. *Applied Ergonomics* 81 (Nov. 2019), 102909. https://doi.org/10.1016/j.apergo.2019.102909

[37] Ultraleap Ltd. 2020. Leap Motion Developer. https://developer.leapmotion.com

[38] I. Scott MacKenzie. 2018. Fitts' Law. In *The Wiley Handbook of Human Computer Interaction*. John Wiley & Sons, Ltd, Hoboken, NJ, USA, 347–370. https://doi.org/10.1002/9781118976005.ch17

[39] William S. Marras. 2006. Basic Biomechanics and Workstation Design. In *Handbook of Human Factors and Ergonomics*, Gavriel Salvendy (Ed.). Publisher: John Wiley & Sons Publishing Inc., New York, NY, USA, 340–370. Google-Books-ID: WxJVNLzvRVUC.

[40] Alex Mazursky, Shan-Yuan Teng, Romain Nith, and Pedro Lopes. 2021. Demonstrating Passive yet Interactive Soft Haptic Patches Anywhere Using MagnetIO. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, Yokohama Japan, 1–4. https://doi.org/10.1145/3411763.3451548

[41] Aunnoy K Mutasim, Anil Ufuk Batmaz, and Wolfgang Stuerzlinger. 2021. Pinch, Click, or Dwell: Comparing Different Selection Techniques for Eye-Gaze-Based Pointing in Virtual Reality. In *ACM Symposium on Eye Tracking Research and Applications*. Association for Computing Machinery, New York, NY, USA, Article 15, 1–7. https://doi.org/10.1145/3448018.3457998

[42] Jörg Müller, Gilles Bailly, Thor Bossuyt, and Niklas Hillgren. 2014. MirrorTouch: Combining Touch and Mid-Air Gestures for Public Displays. In *Proceedings of the 16th international conference on Human-computer interaction with mobile devices & services - MobileHCI '14*. ACM Press, Toronto, ON, Canada, 319–328. https://doi.org/10.1145/2628363.2628379

[43] Cisem Ozkul, David Geerts, and Isa Rutten. 2020. Combining Auditory and Mid-Air Haptic Feedback for a Light Switch Button. In *Proceedings of the 2020 International Conference on Multimodal Interaction (ICMI '20)*. Association for Computing Machinery, New York, NY, USA, 60–69. https://doi.org/10.1145/3382507.3418823

[44] Yesaya Tommy Paulus and Gerard Bastiaan Remijn. 2021. Usability of Various Dwell Times for Eye-Gaze-Based Object Selection with Eye Tracking. *Displays* 67 (April 2021), 101997. https://doi.org/10.1016/j.displa.2021.101997

[45] Giorgia Persichella, Calogero Luca Lomanto, Claudio Mattutino, Fabiana Vernero, and Cristina Gena. 2019. Experimenting Touchless Gestural Interaction for a University Public Web-based Display. In *Proceedings of the 13th Biannual Conference of the Italian SIGCHI Chapter: Designing the next interaction (CHItaly '19)*. Association for Computing Machinery, New York, NY, USA, 5.

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

[47] Ismo Rakkolainen, Euan Freeman, Antti Sand, Roope Raisamo, and Stephen Brewster. 2021. A Survey of Mid-Air Ultrasound Haptics and Its Applications. *IEEE Transactions on Haptics* 14, 1 (Jan. 2021), 2–19. https://doi.org/10.1109/TOH.2020.3018754

[48] Isa Rutten, William Frier, Lawrence Van den Bogaert, and David Geerts. 2019. Invisible Touch: How Identifiable Are Mid-Air Haptic Shapes?. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, Glasgow Scotland Uk, 1–6. https://doi.org/10.1145/3290607.3313004

[49] Samuel B. Schorr and Allison M. Okamura. 2017. Fingertip Tactile Devices for Virtual Object Manipulation and Exploration. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. Association for Computing Machinery, New York, NY, USA, 3115–3119. https://doi.org/10.1145/3025453.3025744

[50] Manuel César Bessa Seixas, Jorge C S Cardoso, and Maria Teresa Galvão Dias. 2015. One Hand or Two Hands? 2D Selection Tasks With the Leap Motion Device. In *Proceedings of the 8th International Conference on Advances in*

Computer-Human Interactions (ACHI '15). IARIA, Wilmington, DE, USA, 6.

[51] Roger N. Shepard and Jacqueline Metzler. 1971. Mental Rotation of Three-Dimensional Objects. *Science* 171, 3972 (1971), 701–703. https://www.jstor.org/stable/1731476 Publisher: American Association for the Advancement of Science.

[52] Ali Shtarbanov. 2018. *AirTap: A Multimodal Interactive Interface Platform with Free-Space Cutaneous Haptic Feedback Via Toroidal Air-Vortices*. Thesis. Massachusetts Institute of Technology. https://dspace.mit.edu/handle/1721.1/115718

[53] Rajinder Sodhi, Ivan Poupyrev, Matthew Glisson, and Ali Israr. 2013. AIREAL: Interactive Tactile Experiences in Free Air. *ACM Transactions on Graphics* 32, 4 (July 2013), 134:1–134:10. https://doi.org/10.1145/2461912.2462007

[54] R. William Soukoreff and I. Scott MacKenzie. 2004. Towards a Standard for Pointing Device Evaluation, Perspectives on 27 Years of Fitts' Law Research in HCI. *International Journal of Human-Computer Studies* 61, 6 (Dec. 2004), 751–789. https://doi.org/10.1016/j.ijhcs.2004.09.001

[55] Shan-Yuan Teng, Pengyu Li, Romain Nith, Joshua Fonseca, and Pedro Lopes. 2021. Demonstrating Touch&Fold: A Foldable Haptic Actuator for Rendering Touch in Mixed Reality. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, Article 203, 1–4. https://doi.org/10.1145/3411763.3451540

[56] Ultraleap. 2020. Leap Motion Controller. https://www.ultraleap.com/product/leap-motion-controller/

[57] Ultraleap. 2020. STRATOS Explore. https://www.ultraleap.com/product/stratos-explore/

[58] Florian van de Camp, Alexander Schick, and Rainer Stiefelhagen. 2013. How to Click in Mid-Air. In *Distributed, Ambient, and Pervasive Interactions (Lecture Notes in Computer Science)*, Norbert Streitz and Constantine Stephanidis (Eds.). Springer, Berlin, Heidelberg, 78–86. https://doi.org/10.1007/978-3-642-39351-8_9

[59] Chi Thanh Vi, Damien Ablart, Elia Gatti, Carlos Velasco, and Marianna Obrist. 2017. Not Just Seeing, but Also Feeling Art: Mid-Air Haptic Experiences Integrated in a Multisensory Art Exhibition. *International Journal of Human-Computer Studies* 108 (Dec. 2017), 1–14. https://doi.org/10.1016/j.ijhcs.2017.06.004

[60] Dong-Bach Vo and Stephen A. Brewster. 2015. Touching the Invisible: Localizing Ultrasonic Haptic Cues. In *2015 IEEE World Haptics Conference (WHC)*. IEEE, Washington, DC, USA, 368–373. https://doi.org/10.1109/WHC.2015.7177740

[61] Panagiotis Vogiatzidakis and Panayiotis Koutsabasis. 2018. Gesture Elicitation Studies for Mid-Air Interaction: A Review. *Multimodal Technologies and Interaction* 2, 4 (Dec. 2018), 65. https://doi.org/10.3390/mti2040065 Number: 4 Publisher: Multidisciplinary Digital Publishing Institute.

[62] Spyros Vosinakis and Panayiotis Koutsabasis. 2018. Evaluation of Visual Feedback Techniques for Virtual Grasping with Bare Hands Using Leap Motion and Oculus Rift. *Virtual Reality* 22, 1 (March 2018), 47–62. https://doi.org/10.1007/s10055-017-0313-4

[63] Robert Walter, Gilles Bailly, and Jörg Müller. 2013. StrikeAPose: Revealing Mid-Air Gestures on Public Displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 841–850. https://doi.org/10.1145/2470654.2470774

[64] Robert Walter, Gilles Bailly, Nina Valkanova, and Jörg Müller. 2014. Cuenesics: Using Mid-Air Gestures to Select Items on Interactive Public Displays. In *Proceedings of the 16th International Conference on Human-Computer Interaction with Mobile Devices & Services (MobileHCI '14)*. Association for Computing Machinery, New York, NY, USA, 299–308. https://doi.org/10.1145/2628363.2628368

[65] Malte Weiss, Chat Wacharamanotham, Simon Voelker, and Jan Borchers. 2011. FingerFlux: Near-Surface Haptic Feedback on Tabletops. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology (UIST '11)*. Association for Computing Machinery, New York, NY, USA, 615–620. https://doi.org/10.1145/2047196.2047277

[66] Daniel Wigdor, Hrvoje Benko, Michael Haller, David Lindlbauer, Ra Ion, Shengdong Zhao, Jeffrey Tzu, Kwan Valino Koh, Ars Electronica Futurelab, and Ars Electronica Gmbh. 2012. *Understanding Mid-Air Hand Gestures: A Study of Human Preferences in Usage of Gesture Types for Hci*. Technical Report. Microsoft Research. MSR–TR–2012–111 pages.

[67] Daniel Wigdor and Dennis Wixon. 2011. *Brave NUI World: Designing Natural User Interfaces for Touch and Gesture* (1st ed.). Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.

[68] Markus L. Wittorf and Mikkel R. Jakobsen. 2016. Eliciting Mid-Air Gestures for Wall-Display Interaction. In *Proceedings of the 9th Nordic Conference on Human-Computer Interaction (NordiCHI '16)*. Association for Computing Machinery, New York, NY, USA, 1–4. https://doi.org/10.1145/2971485.2971503

[69] John S. Zelek, Sam Bromley, Daniel Asmar, and David Thompson. 2003. A Haptic Glove as a Tactile-Vision Sensory Substitution for Wayfinding. *Journal of Visual Impairment & Blindness* 97, 10 (Oct. 2003), 621–632. https://doi.org/10.1177/0145482X0309701007