# Investigating Remote Tactile Feedback for Mid-Air Text-Entry in Virtual Reality

Aakar Gupta[*] Facebook Reality Labs
Majed Samad[†] Facebook Reality Labs
Kenrick Kin[‡] Facebook Reality Labs
Per Ola Kristensson[§] University of Cambridge
Hrvoje Benko[¶] Facebook Reality Labs

## Abstract

In this paper, we investigate the utility of remote tactile feedback for freehand text-entry on a mid-air Qwerty keyboard in VR. To that end, we use insights from prior work to design a virtual keyboard along with different forms of tactile feedback, both spatial and non-spatial, for fingers and for wrists. We report on a multi-session text-entry study with 24 participants where we investigated four vibrotactile feedback conditions: on-fingers, on-wrist spatialized, on-wrist non-spatialized, and audio-visual only. We use micro-metrics analyses and participant interviews to analyze the mechanisms underpinning the observed performance and user experience. The results show comparable performance across feedback types. However, participants overwhelmingly prefer the tactile feedback conditions and rate on-fingers feedback as significantly lower in mental demand, frustration, and effort. Results also show that spatialization of vibrotactile feedback on the wrist as a way to provide finger-specific feedback is comparable in performance and preference to a single vibration location. The micro-metrics analyses suggest that users compensated for the lack of tactile feedback with higher visual and cognitive attention, which ensured similar performance but higher user effort.

**Index Terms:** Human-centered computing—Human Computer Interaction—;——Human-centered computing—Keyboards—; Human-centered computing—Haptics——

## 1 Introduction

Commercial virtual reality (VR) headsets rely on a controller+ray-traced selection approach for text-input which has sub-optimal performance. As hand tracking in VR becomes reality (Leap Motion [36], Hololens [42], Oculus [61]), freehand text entry on a virtual floating Qwerty keyboard that resembles physical keyboard typing is increasingly being explored. Typing on physical keyboards is highly efficient since it uses small and chorded finger motion and enables novice-to-expert transition through motor memory. A fundamental obstacle in transplanting these positive traits to mid-air keyboard typing is the lack of tangibility.

This notion of tangibility can be broken down into two components: 1) Kinesthetic: the physical limit imposed by the keys and the keyboard surface on which the keys are mounted, and 2) Tactile: the finger-specific tactile feedback from every key. Existing work on freehand Qwerty typing in VR shows that participants perform significantly better on a flat table surface compared to typing in air [13]. The table surface provides both the kinesthetic and tactile components albeit with a much lower fidelity than physical keyboards. However, the availability of a dedicated surface or physical keyboard can only be assumed in very specific VR scenarios. While simulating the kinesthetic effect of a physically limiting surface in air is near-impossible without significantly encumbering the hands or instrumenting the space around them, providing tactile feedback in air is very much feasible.

Current approaches to providing mid-air haptic feedback can be divided into three parts - 1) handheld devices [3, 7, 9, 10, 26, 58, 59], 2) non-contact haptics (such as ultrasound, laser, and air vortexes) [6, 24, 51], 3) glove, ring, or wrist wearables [21, 22, 46, 47, 52, 68]. While handheld devices constrain 10-finger freehand interaction, non-contact haptics are specialized solutions that are not instantly portable and can be prohibitively expensive. In this paper, we focus on wrist or ring vibrotactile wearables which do not provide a collocated sensation (*remote*) on the finger-tips but are simple, inexpensive, and more practical. Prior work has shown that visual and tactile stimuli can vigorously interact even when they are not collocated but are close [53]. Further, smartwatches with vibrotactile motors are increasingly becoming popular and hold the potential to provide simple yet effective feedback for freehand VR interaction. There is surprisingly sparse research that investigates the effect of such remote tactile feedback on interaction performance in VR. Our work is therefore focused in this space and asks the following question: Does providing remote wrist or finger-level vibrotactile feedback have an effect on the user performance and experience for freehand text-entry on a mid-air Qwerty keyboard in VR?

To that end, we first designed a virtual keyboard with audio-visual feedback based on insights from existing literature and using high-fidelity hand tracking. We then designed four vibrotactile feedback conditions: *on-fingers*, *on-wrist spatial*, *on-wrist nonspatial*, and *audio-visual*. We conducted a text-input study with 24 participants across the four conditions and measured performance and preference metrics. We further used micro-metrics analyses and participant interviews to analyze the mechanisms underpinning the observed performance and user experience and discuss the wider implications if haptic feedback for text entry in VR.

The results showed comparable speeds and accuracies across feedback types; however, participants overwhelmingly preferred the tactile feedback conditions with 63% of the participants rating *audio-visual* as the least preferred condition among the four conditions. Participants further rated *on-fingers* feedback as significantly lower in mental demand, frustration, and effort. The micro-metric analyses and interviews suggest that participants compensated for the lack of tactile feedback with higher visual and cognitive attention, thus resulting in similar performance but increased mental load in conditions with lower or no tactile feedback.

## 2 Related Work

There have been numerous works on mid-air haptic feedback in VR that involve wrist wearables [47, 52], gloves (see [46] for a review), and controllers (see [54] for a review), there is no work to our

[*]e-mail: aakarg@fb.com
[†]e-mail: majed.samad@gmail.com (Now at Google Inc.)
[‡]e-mail: kenrickkin@fb.com
[§]e-mail: pok21@cam.ac.uk
[¶]e-mail: benko@fb.com

knowledge that investigates the effect of mid-air haptic feedback on text-input. There are existing works on finger-based haptics which could potentially be used for providing feedback for text-input such as Dexmo [19], Dextres [28], and WiredSwarm [60], but these are highly encumbering devices. We focus our discussion of related work on text-input in VR and divide it into the following three categories: encumbered, unencumbered, and freehand text-entry explorations in VR. *Encumbered* encompasses techniques where the user interacts with externally grounded devices, such as a physical keyboard, as well as techniques where the user's hands are significantly encumbered by controllers or other devices. *Unencumbered* includes techniques that minimally or not at all encumber the hands such as head motion, gaze, or rings. *Freehand* refers to unencumbered techniques that specifically focus on Qwerty keyboard typing using hand tracking.

### 2.1 Encumbered Text-entry in VR

Current commercial VR devices use controllers for text-entry where a ray cast from the controllers is used to select keys on a Qwerty keyboard [29,61]. Prior work has investigated dedicated handheld devices for VR typing such as Twiddler [5], 9-key keypads [15], smartphones [34], and bimanual touchpads with hover detection [55]. Speicher et al. [56] evaluated multiple techniques based on current commercial controllers including raycasted pointing, direct tapping, controller as gamepad, and found raycasted pointing as the fastest with 15.4 words-per-minute (WPM). Other studies have investigated non-Qwerty layouts with controllers including circular [15,67] and cubic [63] layouts. Multiple glove-based or optical tracking-based techniques have been proposed that map a keyboard layout on to the hand/fingers [15,31,45,49], reporting speeds in the range of 5–10 WPM.

The work on physical keyboards in VR can be classified into one of two categories: variations in visual feedback based on tracking hand-finger motion in the real world [4,18,33,38] and variations in hand representations [17,35,41]. Most studies in this category report typing speeds in the range of 25–45 WPM, which are much higher than other alternatives in VR. This lends credence to our approach of investigating 10-finger typing in VR and whether tactile feedback can help mitigate the problems arising from lack of tangibility in air. For a detailed review of physical keyboards in VR (and VR text-entry in general), we refer the reader to Dube et al. [12].

### 2.2 Unencumbered Text-entry in VR

Yu et al. [65] showed text-entry via headpointing leads to a speed of 10.6 WPM when using dwell for selection and 15.6 WPM when using a controller button press for selection. Rajanna et al. [50] showed that gaze typing results in speeds of 10 WPM in VR. An emerging thread of research is using wearable devices for typing in VR as well as for AR (augmented reality). This includes typing using a smartwatch [1], using a ring [22,32], and using the touchpad on the headset/glasses [16,20,66] with reported speeds in the range of 8–15 WPM.

### 2.3 Freehand Qwerty Text-entry in VR

In 2003, ARKB [37] used vision-based tracking of fingertips for multi-finger typing on a virtual Qwerty keyboard. Markussen et al. [40] showed that a single-finger mid-air vertical keyboard on a large display yields a speed of 13.2 WPM in its final session, after $\sim$ 75 mins of practice. VISAR [14] uses word-level decoding for a single-finger mid-air vertical keyboard in VR yielding 17.8 WPM after $\sim$ 90 minutes of practice. ATK [64] uses Leap Motion to implement a 10-finger mid-air horizontal qwerty keyboard supported by a word-level decoder and reports speeds of 29.2 WPM for a limited vocabulary phrase-set after $\sim$ 1 hour of practice. None of these works examined the influence of haptic feedback on typing performance. Dudley et al. [13] investigate the differences in on-surface vs. mid-air qwerty keyboard typing in VR and conclude that when using a Wizard of Oz decoder, the performance of 10-finger on-surface typing (51.6 WPM) is higher than mid-air (34.5 WPM). Wu et al. [62] propose a glove that provides vibration feedback on the fingertips for a mid-air Qwerty keyboard. However, they do not investigate typing performance.

Even though the performance reported in prior work is specific to their study design, physical keyboard style text-entry appears to be one of the most promising techniques for performant typing in VR. It is unencumbered, has potentially high speeds, and resembles physical keyboard typing. Our paper reports on the first ever investigation into the value of remote tactile feedback for Qwerty-based typing in mid-air.

## 3 Approach

Our approach towards this investigation consists of three broad steps. First, we build a keyboard prototype along with the varying levels of remote tactile feedback following a rigorous design process. The process includes the measurement of milisecond-level feedback latencies to ensure that tactile feedback latency is aligned with visual feedback latency, as well as a preliminary experiment to guide the spatial design of the feedback on the wrist. Second, we design, conduct, and report on a within-subjects experiment with 24 participants to compare the effects of the varying levels of tactile feedback. Third, we conduct analyses of text input micro-metrics and of our participant interviews to unravel the mechanisms that explain our results and discuss the implications of our findings for haptic feedback for text input in VR. In this section, we discuss our prototype design.

### 3.1 Tactile Feedback Design

We focus on vibrotactile feedback that could be provided via wrist or ring/glove wearables to ensure simple, inexpensive, and practical haptic feedback. We use Linear Resonant Actuators (LRAs) owing to their low-bulk instrumentation, ubiquity, low-cost, and low-power needs. We designed three levels of remote vibrotactile feedback for our study: *on-fingers*, *on-wrist spatial*, and *on-wrist nonspatial*. *On-Fingers* provided finger-specific feedback on the base of each of the 10 fingers. Having metallic motors at the finger-tip constrains the user from attending to external activities that routinely require fingers. The finger-base location is less inhibiting and can be enabled using open-finger gloves which would not need to be donned off and on as frequently as closed-finger gloves. *On-Wrist Spatial* provides vibrotactile feedback that is spatialized across the wrist using five vibrotactile actuators such that a different location is vibrated depending on the finger that is colliding or pressing. Here, the user still receives finger-specific feedback, though it is less discriminable due to the reduced acuity of the tactile sense at the wrist. We chose this condition to see if the user can use the spatialization at this remote location to inform which finger performed the key-press. *On-Wrist Nonspatial* provides vibrotactile feedback on a single location at the top of the wrist regardless of the finger. This provides non-finger-specific feedback, but requires only a single actuator, and is therefore suitable for current wrist wearables and smartwatches.

There are pros and cons of the different levels of haptic feedback we use. From on-fingers to on-wrist spatial to on-wrist nonspatial, the potential granularity of feedback decreases, but so do the actuation (and cost) requirements, physical inhibition, and power needs.

### 3.2 Text-Input Prototype Design

Since tactile feedback would require additional wearables whereas audio-visual (AV) feedback can be provided using only the headset, it is important to investigate the incremental benefits of tactile feedback for a well-designed AV keyboard. Demonstrating the utility of tactile feedback for a sub-optimal AV keyboard would be less informative, since further optimizations to audio and visual feedback may obviate the need for tactile feedback. We therefore build the AV keyboard carefully using insights from prior work on virtual keyboards.

#### 3.2.1 Hand Representation

Grubert et al. [18] showed that for typing on a physical keyboard in VR, providing fingertip-level feedback helped users retain 60% of their speed on a physical keyboard outside VR. Further, in a study comparing different hand representations for a physical keyboard in VR, Knierim et al. [35] found that realistic hands had the best aggregated NASA-TLX score and the highest score for Presence. We consequently used realistic hand silhouettes akin to the ones in Oculus Rift [44]. The virtual hands replicated the real hand movements down to the fingertip-level. For this, we used Han et al's passive markers+inverse kinematics approach [25] that generates a hand model using the data from a glove with fiducial markers. However, Grubert et al. [17] who also used a passive markers+inverse kinematics approach found that users rated realistic hands low in preference since they were occluding the keyboard too much. To solve occlusion, we displayed the hands at 50% transparency when they were $> 5$ cm away from the keyboard and at 95% transparency with opaque borders when they were $\leq 5$ cm away from the keyboard (Figure 1c). This offered a nice balance of having realistic hands while minimizing occlusion. In our pilot studies, participants anecdotally reported preferring this representation.

#### 3.2.2 Visual and Tactile Feedback

The visual and tactile feedback from the keyboard were designed taking inspiration from physical keyboards. In line with prior VR typing experiments [13, 14], we chose to keep the apparent key size the same for all users ($24 \times 24$ mm). Prior work [13, 64] on mid-air VR Qwerty keyboards register a tap when the finger collides with the key. However, in physical keyboards, the tap is registered when the key is depressed to a certain depth. We replicated this behavior such that the tap is registered when a key is depressed to a depth of 9 mm ("base depth") from its original position. What follows is the sequence of events that detail the visual feedback over the course of a key click: 1) (*"Hover"*) When the user's fingers are $< 5$ cm away from the keyboard, the user sees small purple spheres on the keyboard that indicate the locations over which the fingers are currently hovering; 2) (*"Collision"*) As soon as a finger touches (or collides with) a key, the key turns gray and starts depressing as per the location of the fingertip; 3) (*"Press"*) When the finger reaches the base depth (9 mm), the key turns yellow and stays that way until the finger leaves that depth; 4) (*"Release"*) When the finger leaves the *Press* state, the key turns back to its original color. The finger goes back from *Release* to *Hover* when it is no longer touching the key. Note that a key stays at the base depth even if a fingertip goes beyond it, to signal a hard stop just like a physical keyboard. The users also receive audio feedback in the form of a fixed duration click sound upon *Press* through the default speakers of the headset. These stages are depicted in Figure 1.

Analogous to physical key presses, users feel a subtle tactile actuation (0.1*maximum LRA amplitude) upon *Collision* and a stronger actuation (0.8*maximum LRA amplitude) upon *Press*. The subtle *Collision* actuation is played for the entire duration that the finger is in collision with a key. Similar to the fixed duration audio feedback, the strong *Press* actuation has a fixed duration (45 ms) to indicate to the user that a character has been entered.

#### 3.2.3 Position, Orientation, and Minimizing Coactivation

The keyboard was positioned parallel to the ground in a position where the user's hands reach the keyboard while keeping the shoulders relaxed. This minimizes arm fatigue since shoulder torque is the dominating cause for mid-air arm fatigue ("gorilla-arm effect") [27, 30]. Prior work [13, 64] shows that mid-air typing is very much an open problem primarily owing to its high coactivation errors—when hitting a key with a particular finger, the other fingers inadvertently hit the keyboard due to the lack of haptic feedback and the constrained individuation of fingers [57]. We conjecture the visual and haptic feedback upon *Collision* may help in avoiding such errors.

We conducted an initial pilot study with no tactile feedback that showed that inadvertent thumb presses was a specifically frustrating issue that dominated the user's experience. We sought to minimize this issue in our design and therefore restricted the thumb presses to register only on the space key and not on other keys. Further, inadvertent thumb presses on the space key were notably higher when the keyboard was angled towards the user. We consequently kept the keyboard orientation to be completely flat, parallel to the ground.

### 3.3 Measurement of Feedback Latencies

To ensure that the latency of tactile feedback is not too high in relation to the visual and audio feedback, we measured their end-to-end latencies using Di Luca et al.'s established method [11]. The virtual keyboard was aligned to the surface of a physical table and a microphone was positioned atop the table to pick up tapping sounds on the surface, which served as the baseline signal. The virtual scene inside a VR headset was a rectangle whose color changed from black to white upon any keypress from the virtual keyboard. A photo-diode was positioned on the lens of the VR headset to detect this luminance change. A microphone was also positioned adjacent to the HMD's headphones to pick up the audio generated from the virtual keyboard clicks. Finally, a contact microphone was attached to the vibrotactile actuator to detect the signal that was triggered by keypresses of the virtual keyboard. All four signals were attached to a multi-channel audio-card. When the experimenter tapped on the table, it triggered the baseline, visual, auditory, and tactile signals that were recorded on the same card and processed later to measure latency.

We performed 20 tap trials. The measured latencies (mean, sd) were as follows: Visual (75.8 ms, 11.5 ms), Tactile (64.8 ms, 10.8 ms), Audio (190 ms, 10.2 ms). The visual and tactile feedback latencies are comparable, but the audio is higher by $\sim 110$ ms. There are known latency issues for audio when it is transmitted over HDMI in VR headsets. However, the audio will remain the same across all conditions and in the interest of using the default headset behavior, we opted to not change the audio transmission channel.

## 4 3 VS 5 WRIST MOTORS

For the *Wrist Spatial* condition, five actuators on the wrist will not be as cutaneously distinct as they are on the fingers. Prior work [8] shows that users may only be able to localize four locations on the wrist. However, prior work [23] also shows that when users are asked to perform relative localization on the wrist (locate the current sensation relative to a previous one), they can be accurate for up to 8 locations. Therefore, it was unclear if five actuators would be useful to the user in our case or a lower number of actuators which may not have correspondence with all 10 fingers but ensure mutual discrimination. We therefore explored a three-actuator setup where the thumb and index finger presses correspond to individual actuators, but the middle, ring, and pinky finger presses all correspond to the third actuator. Due to the significantly low ring and pinky finger usage in 10-finger mid-air typing even in error-free Wizard of Oz scenarios [13], we chose to jointly represent their feedback with the middle finger reducing the number of actuators from 5 to 3 and increasing the space between actuators. This may enable higher spatial acuity while providing finger-specific feedback for the dominant fingers. We conducted a preliminary study to decide which on-wrist

(a) (b) (c) Hover (d) Collision (e) Press

Figure 1: a) Virtual keyboard in VR. b) Hands above the keyboard. c) *Hover:* Hands turn transparent closer to the keyboard. d) *Collision:* Finger colliding with but not pressing a key. e) *Press:* Finger presses a key.

spatial configuration was preferred. Five participants (2F, 3M, age: 22-30) wrote five phrases for each of the two conditions.

Figure 2 shows the apparatus for our preliminary and final studies. We implemented motion capture to track participants' head and hand movements using an Optitrack cage with 17 cameras (OptiTrack Prime 17W, 1664×1088 pixels, 70° FOV, 120 FPS, 2.8 ms latency) that precluded any marker occlusion issues. The virtual environment consisted of our virtual keyboard in the default Unity skybox scene. We used Oculus Rift as the VR headset which was tracked using fiducial markers on its surface. Participants wore a pair of ultra-thin and flexible power mesh gloves that were fitted with 19 fiducial markers each to enable high-res hand tracking following the approach outlined in [25]. We used three different-sized gloves to account for the variation in hand sizes. The vibration actuators were LRAs ML1040W* (Mplus, KR) with a resonant frequency at 170 Hz. For the five-actuator condition, double-sided velcro wristband was used with actuators placed 1.5 cm apart edge to edge. For the three-actuator condition, every alternate actuator was used skipping the 2nd and 4th actuators.

Three participants preferred the five-actuator condition, one had no preference, and one preferred three actuators. A major reason for the preference of the five-actuators was that even though participants did not use the ring and pinky fingers frequently, the fingers did have frequent inadvertent collisions with the keys, and the five-actuator condition provided finger-specific collision feedback in such cases. The three-actuator condition instead collapsed the inadvertent collisions of three fingers into a single wrist location due to which participants received no information on which finger(s) may be accidentally touching. Participants also mentioned that differentiating the five spatial locations was not significantly harder than three locations. We therefore chose five actuators for the on-wrist spatialized condition.

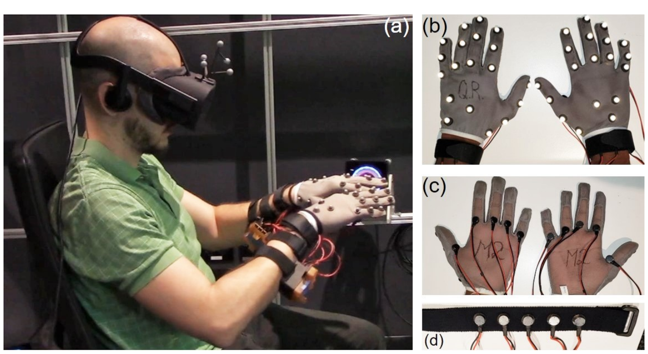


Figure 2: a) User with VR headset and wearing the hand tracking gloves as well as wristbands that provide tactile feedback. b) Fiducial markers. c) Finger-base actuators. d) Wristband actuators.

## 5 Effect of Tactile Feedback on Mid-air Typing

### 5.1 Participants

24 participants (8F, 16M, age range: 22–57, mean: 37, 2 left handed) did the study. Participants self-rated their typing proficiency on a physical Qwerty-keyboard on a 1–5 scale (increasing proficiency) yielding mean of 3.91 (sd: 0.65). Ten of them had a prior experience with VR, but none of them were habitual users. The same apparatus as earlier with the finger-base and five-actuator wrist set up was used. Participants wore the same set up across all conditions to keep the encumbrance constant.

### 5.2 Study Design

We adopted a within-subjects design with four conditions (FeedbackType): *audio-visual* (baseline), *on-fingers*, *on-wrist spatial*, and *on-wrist nonspatial*. The study consisted of four sessions. Each session consisted of all four feedback conditions. For each feedback condition within a session, participants transcribed five stimulus phrases. The order of feedback conditions within a session was counterbalanced using a Latin square across the 24 participants. For a particular participant, the ordering was kept constant for all four sessions. Each participant did 4 Sessions × 4 FeedbackTypes × 5 phrases = 80 phrases in total. The 80 stimulus phrases were randomly selected from the standard MacKenzie phrase-set [39] and then kept constant in their order of appearance across all participants in order to minimize confounds. As participants were evenly distributed across conditions, each condition was thus exposed to the same phrase set, which increases the internal validity.

The multi-session design allows us to test performance of the FeedbackType at four stages of proficiency with the mid-air keyboard starting with novice to increasingly proficient. The study was chosen to be within-subjects and not between-subjects to control for the variations between different participants in their immediate and over-time use of the mid-air keyboard.

### 5.3 Procedure

Upon arrival, participants were introduced to the task environment and apparatus. Participants sat on a chair and were asked to place their hands in a comfortable position as if to type on a horizontal keyboard in air while keeping their shoulders completely relaxed. Participants sat so that they do not get tired standing over the study duration. The keyboard was then placed under their fingers and adjusted according to their preference. The chair had lowered armrests; participants were allowed to use the armrests during breaks to minimize fatigue, but not during typing. Participants were then asked to complete two practice phrases on the keyboard without any tactile feedback. Participants were explicitly instructed on the keyboard design, the collision and press feedback, and that they were free to use as many fingers as they want to use. In keeping with prior unconstrained text-entry evaluations, participants were instructed to type as quickly and accurately as possible, and that they could correct errors (using *Backspace*) they noticed immediately, but could also choose to ignore errors which they notice after a few characters have been typed. Pressing *Enter* took them to the next phrase. To

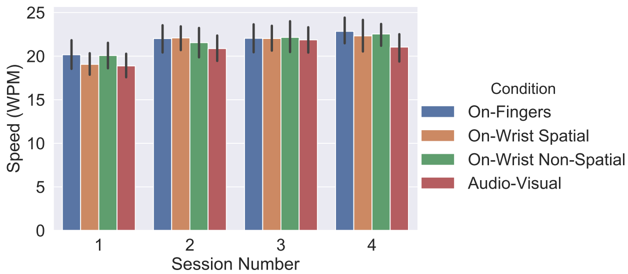


Figure 3: Average speed in words per minute as a function of condition (indicated by the color of the bars) and session number (indicated by the group of bars). Error bars represent the 95% CI.

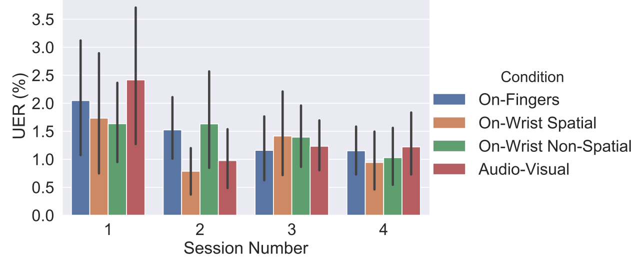


Figure 4: Average uncorrected error rate % as a function of condition (indicated by the color of the bars) and session number (indicated by the group of bars). Error bars represent the 95% CI.

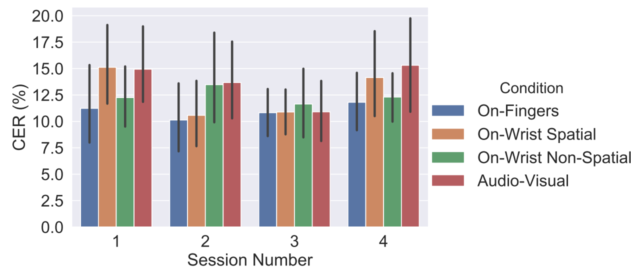


Figure 5: Average corrected error rate % as a function of condition (indicated by the color of the bars) and session number (indicated by the group of bars). Error bars represent the 95% CI.

prevent inadvertent *Enter* presses, the press was only registered if the minimum string edit distance (MSD) of the transcribed phrase from the stimulus phrase was $< 8$. On average, this means that the next phrase will not be displayed until a transcribed phrase has $< 30\%$ MSD from its stimulus. To minimize any fatigue, participants were given a 1 min break between different FEEDBACKTYPES within a session, and a 5 min break between conditions. Participants did a post-study preference and NASA-TLX questionnaire. The participant responses were not under observation during this time to minimize response bias. A semi-structured interview was conducted at the end. The entire experiment took 75–105 minutes to complete.

### 5.4 Measures

We measured *Speed*, *Uncorrected Error Rate (UER)*, and *Corrected Error Rate (CER)* using standard metrics [2]. Speed is measured in words-per-minute: $WPM = ((|T|-1)*60)/(S*5)$ where —T— is the transcribed phrase length and S is the time starting from the first key press until the last key press before *Enter* including time spent in correcting errors. $UER = MSD(P,T)*100/max(|P|,|T|)$ where P is the stimulus phrase. $CER = (C*100/|T|)$ where C is the number of corrections which translates to number of backspaced characters in our case.

Mid-air typing mechanics are not well understood. Prior work has therefore analyzed micro-metrics [13] such as press duration, finger travel, finger-key collisions, and finger usage for in-depth understanding of the input. We analyze such micro-metrics in our results to gain insights beyond the above measures.

## 6 RESULTS

We conducted our analyses as 2-way RM-ANOVAs with factors FEEDBACKTYPE and SESSION and dependent variables speed, UER, CER, and other micro-metrics and report them below.

### 6.1 Speed, Errors, and NASA-TLX

#### 6.1.1 Speed

We observed a main effect of SESSION on speed, $F(3,69) = 13.275, p < .001, \eta^2 = .366$. The effect of FEEDBACKTYPE is not significant, $F(3,69) = 2.688, p = .053$. There were no interaction effects. Pairwise comparisons show that there are significant differences in speed between session pairs 1–2, 1–3, and 1–4. Looking at mean values (Figure 3) further affirms that user speeds plateau after session 1.

#### 6.1.2 Uncorrected Error Rate (UER)

We observed a main effect of SESSION on UER, $F_{GG}(1.822, 41.910) = 3.609, p < .05, \eta^2 = .136$ (the subscript GG denotes the Greenhouse-Geisser correction for non-sphericity). Pairwise comparisons, however, did not show significant differences between any session pairs. Figure 4 shows that while the means from session 1 to session 2 have a clear decline, there is a large variance. The effect of FEEDBACKTYPE is not significant and there were no interaction effects.

#### 6.1.3 Corrected Error Rate (CER)

We observed a main effect of FEEDBACKTYPE on CER, $F(3,69) = 2.889$, $p < .05$, $\eta^2 = .112$. However, pairwise comparisons again did not show significant differences between any pairs. The *audio-visual* and *on-fingers* CER across all sessions ($\mu_{AV} = 13.7\%$, $\mu_{OF} = 11.0\%$, respectively) suggest that the *on-fingers* condition may have lower CER but that the variance in our sample was too large for it to be significant ($p = 0.07$). The effect of session is not significant and there were no interaction effects. Fig. 5 shows the CER.

#### 6.1.4 Preferences

Figure 6 (left panel) shows the preference choice counts for each of the four tactile conditions we tested. It can be seen that 75% (or 18 of the 24 participants) preferred the *on-fingers* condition. Participants preferred the *audio-visual* feedback type the least with 14 participants rating it as the least preferred technique. The two *on-wrist* feedback conditions were preferred at similar rates.

#### 6.1.5 NASA-TLX

Cronbach's alpha ($\alpha = 0.70$) showed the questionnaire to provide good internal consistency. We conducted a Friedman test on the NASA-TLX responses (see Fig. 6 right panel). Our analysis revealed that there were significant differences between FEEDBACKTYPES for Mental Demand ($\chi^2(3) = 8.46$, $p < .05$), Performance ($\chi^2(3) = 12.14$, $p < .01$), and Effort ($\chi^2(3) = 9.39$, $p < .05$). Pairwise Wilcoxon Signed-Rank tests showed that *on-fingers* had a significantly lower effort score than the other three conditions, a significantly lower mental demand than *audio-visual* and *on-wrist nonspatial*, and a significantly better performance score than *audio-visual*.

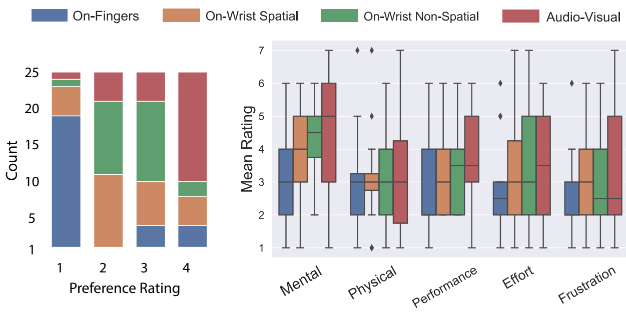


Figure 6: Subjective Results: (Left) Preference Ratings showing the number of participants who rated the various FEEDBACKTYPE conditions as their 1st, 2nd, 3rd or 4th choice. (Right) NASA-TLX Ratings. Lower is better.

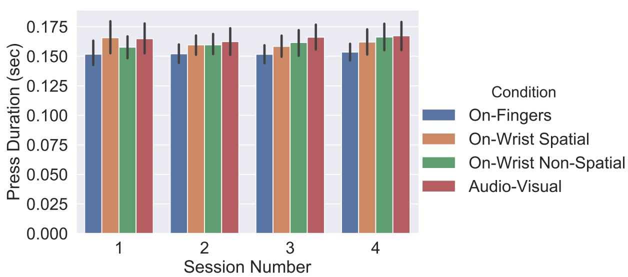


Figure 7: Average press duration in seconds as a function of condition (indicated by the color of the bars) and session number (indicated by the group of bars). Error bars represent the 95% CI

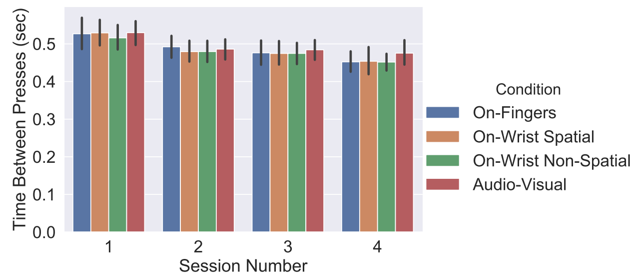


Figure 8: Average time between key presses in seconds as a function of condition (indicated by the color of the bars) and session number (indicated by the group of bars). Error bars represent the 95% CI.

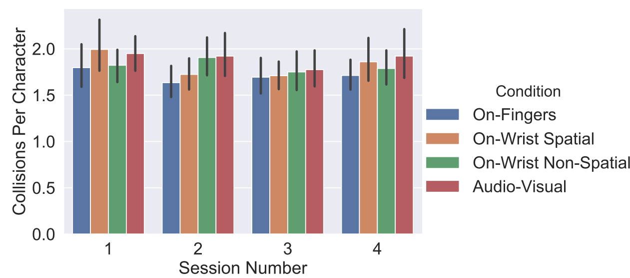


Figure 9: Number of collisions per character typed as a function of condition (indicated by the color of the bars) and session number (indicated by the group of bars). Error bars represent the 95% CI.

## 6.2 Micro-metrics

### 6.2.1 Press Duration

The press duration is defined as the time when a key starts getting pressed, until it is released ($keyrelease - keypress$) (Figure 7). We observe a main effect of FEEDBACKTYPE ($F_{GG}(1.964, 45.177) = 9.824$, $p < .001$, $\eta^2 = .299$). No other main or interaction effects were found. Pairwise comparisons showed that press duration in *on-fingers* is significantly lower than the other three ($p < .05$). *On-Finger* tactile feedback was found to be effective in lowering the press duration, which implies that users released the key quicker upon finger vibration. Looking at the means across all sessions ($\mu_{OF} = 152$ ms, $\mu_{WS} = 161$ ms, $\mu_{WN} = 161$ ms, $\mu_{AV} = 165$ ms), the difference was 13 ms between *on-fingers* and *audio-visual*. The lower press duration for *on-fingers* holds true across sessions, suggesting that on-fingers tactile feedback leads to lower press durations both for novice and longer-term users of the mid-air keyboard.

Participant comments also indicated that *on-fingers* not only provided an instant confirmation of their action, but also a confirmation that the intended finger led to the keypress. P10: *"On fingers lets me know which key was depressed. It gives me more confidence I hit the intended key. On wrist spatial does same thing, but it's not as clear. On wrist non-spatial just lets me know I hit something"*.

### 6.2.2 Press Depth (Finger Travel)

When pressing a key, press depth is the maximum depth that a participant's finger went down to relative to the key's original position. This is also known as "finger travel". In their comparison of a virtual keyboard in mid-air vs on-surface, Dudley et al. [13] found the on-surface keyboard to have a shorter press duration owing to shorter finger press depths. We analyzed finger press depth and found no significant effects. The mean finger-depth for all FEEDBACKTYPES were very similar (in the 2–3 mm range from base depth). This indicates that the lower press duration for *on-fingers* tactile feedback in our case is due to faster response time from the user.

### 6.2.3 Time between Presses

We observe a main effect of SESSION on the time between consecutive key presses ($F_{GG}(1.881, 43.267) = 19.097$, $p < .001$, $\eta^2 = .454$) (Figure 8). The main effect of FEEDBACKTYPE and the interaction effect were not significant. Pairwise comparisons show significant differences between all session pairs ($p < 0.05$ for all) except between sessions 2–3. Looking at the means in Figure 8, it implies that for Time between presses, Session1 $<$ Session2 $\sim$ Session3 $<$ Session4. Thus, participants did quicker presses as they became more familiar with the keyboard. We further analyzed inter-character time which is the *key press – previous key release* time. We observed a main effect of SESSION ($F_{GG}(1.649, 37.916) = 82.221$, $p < .05$, $\eta^2 = .781$) with the same pairwise differences as in time between presses.

Interestingly, lower press durations do not lead to a significant impact on speed. This could be because speed depends on the time between consecutive presses which may not be directly affected by a faster release of the previous key if a different finger is used to press the next key. This is supported by the fact that time between presses is not impacted by tactile feedback either.

### 6.2.4 Collisions Per Character

For every key press, we measured the number of collisions with other keys that did not get pressed, i.e., unintentional collisions with the keyboard (Figure 9). A main effect of FEEDBACKTYPE ($F(3, 69) = 5.773$, $p < .005$, $\eta^2 = .201$) was observed. Pairwise comparisons show that *on-fingers* has significantly fewer collisions than the other three and *audio-visual* has significantly more collisions than the other three. Thus, for collisions per character, *on-fingers* $<$ *on-wrist*

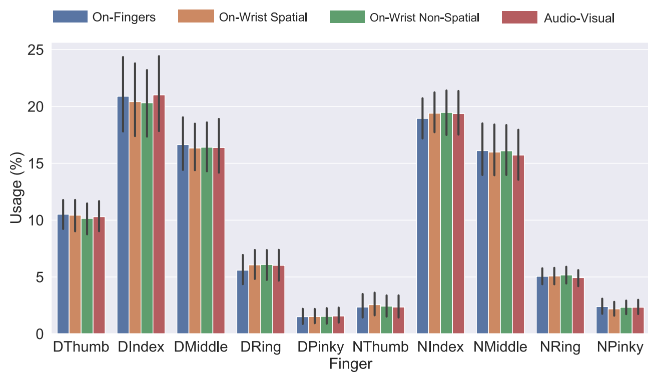


Figure 10: Finger usage across the four conditions for all 10 fingers. DThumb and NThumb refer to the thumbs of the user's Dominant and Non-dominant hand respectively.

*spatial* $\sim$ *on-wrist nonspatial* $<$ *audio-visual* ($\mu_{OF} = 1.71, \mu_{WS} = 1.82, \mu_{WN} = 1.82, \mu_{AV} = 1.89$).

The higher collisions in *audio-visual* reinforces our design choice of not using key collisions as presses directly. It will be useful to study the effect of tactile feedback on errors for a keyboard where collisions are regarded as presses. We speculate that the error rates would be higher for the non-tactile feedback conditions in that case.

#### 6.2.5 Typing Proficiency

Prior work [35] analyzes physical keyboard typing in VR by dividing users into 2 groups (<53 and >53 WPM) [23] and looks at them separately. However, user variations in physical keyboard typing proficiencies do not confound our results since our study is a repeated measures design where each user goes through all four FeedbackTypes. Further, since our investigation is on a mid-air keyboard, physical keyboard typing is not an objective baseline for it. We therefore did not collect physical keyboard typing baselines during our study.

We did however analyze normalized participant speeds relative to their speeds on the very first two phrases which is a more realistic mid-air typing proficiency baseline. The results mirrored the speed analysis above with a significant effect of SESSION on normalized speed and no effect of FEEDBACKTYPE.

#### 6.2.6 Finger Usage, Throughput

Figure 10 shows finger usage by SESSION and FEEDBACKTYPE. There seems to be no impact of FEEDBACKTYPE on finger usage. This aligns with existing work [13] that reported similar typing performance for two-finger and ten-finger mid-air typing.

We also analyzed Throughput, a recent metric proposed by Zhang et al. [43] that combines speed and error rate into a single metric. The analysis showed similar results to speed with a significant effect of SESSION, but not of FEEDBACKTYPE. This indicates that if the participants were to correct all errors, speed would follow a trend similar to our results. The average throughput across conditions and sessions was $\sim$7 bits/s.

## 7 DISCUSSION

We now discuss the results, the specific insights, and the directions for further work indicated by those insights.

### 7.1 Users compensate for lack of tactile feedback with higher visual and cognitive attention

In general, users had to expend pay higher visual and cognitive attention to their key-presses because the lack of tactile information of the keyboard layout made it hard to know the exact hand position and relative locations of the consecutive keys. Further, the in-air keyboard also lacks the traditional J and F key bumps act as navigational points. The fact that reduced collisions in the tactile feedback conditions did not result in a reduction in error rate suggests that the visual feedback upon key collision prior to the key being pressed in the *audio-visual* condition helped avoid unintended presses for at least some participants. Multiple participants reported varying eye gaze behavior depending on FEEDBACKTYPE. P21: *"With only the visual feedback, I had to look at the keys constantly. Vibrations felt freeing in that sense"*. Given that tactile feedback impacts intermediate metrics of press duration and collisions, but not the eventual speed and accuracies, it indicates a trend that participants compensated for the lack of tactile feedback with even higher visual and cognitive attention on the keyboard—P2: *"[In on-fingers], my eyes didn't need to look at the key to know that it was pressed. I felt that my eyes had to move more between the (phrase) display and my fingers when there were no vibrations"*.

The lower effort and mental demand scores reinforce this notion. Participants mentioned that the increased mental demand was mostly because they had to actively pay attention to avoid accidental key presses in the absence of tactile feedback. P1: *"Mental demand was mostly about making sure that the other fingers were not accidentally colliding with keyboard. The vibrations gave an early notice before clicking and gave me a chance to change my mind."* An investigation that uses eye-tracking to quantify visual attention with and without tactile feedback would be a useful follow-up.

### 7.2 Tactile feedback enables a more consistent maintenance of the hand position

The reduced accidental collisions in tactile feedback conditions shows that for the fingers that were not in use, participants were able to keep them away from the keys more consistently. Participant comments suggest that tactile feedback upon collisions helped them settle on a more consistent hand position over the keyboard. P5: *"The vibrations helped me to know how to position the hand to avoid the errors"*. The *on-fingers* feedback was again reported to be more useful—P1: *"On fingers gave me the best sense of what I was typing and where my fingers were on the keyboard. If I made a mistake that I needed to correct, for instance, I had a better idea of where I was accidentally resting my fingers and should not be."*

In their investigation of on-surface vs mid-air typing, Dudley et al. [13] noted the importance of a fixed reference plane yielded by the surface that enables the user to maintain a consistent hand position and therefore regulate their finger depths much more easily than in mid-air. While our tactile feedback helps maintain a more consistent hand position, participants mentioned the lack of a fixed reference plane. P13: *"When I type, I just use my fingers and my hands mostly stay fixed at the same place on the wrist restpad. I was trying to do that here, but it's so hard when you don't have any support"*. More continuous forms of haptic feedback may be able to provide better proxies of reference plane in air. For instance, squeezing feedback on the wrist is not annoying for continuous use [21] and is a good subject for future investigation. Another direction would be to explore how to replicate the feedforward behavior that is enabled by the tactile marks on the F and J keys on a physical keyboard.

### 7.3 Tangibility

Participants reported an overwhelming preference for tactile feedback. This was both due to the ease in performing the task as discussed above, and due to the physical feeling imparted by the tactile feedback. P4: *"When I am using a physical keyboard, I have something to feel, this gave me something to feel. The no vibrations was almost like I was just pushing buttons in the air, it had no substance. The vibrations made me feel like I was actually doing something."* At the same time, participants questioned if their performance improved in accordance with the better feeling—P9:

*"The feedback made it feel more like typing, but not sure if it made me type better."*

### 7.4 Encumbrance and feedback tradeoffs between the conditions

While a majority of the participants preferred finger vibrations, wrist vibrations, and no vibrations in that order, a few participants disliked finger vibrations and preferred other conditions more. For instance, P4 preferred *on-wrist nonspatial* the most: *"On fingers felt like too much buzzing. And the five different vibrations on wrist felt weird. The single one on the wrist was just enough to not be distracting and let me know if I was touching a key, so it was the best. No vibrations, I had to look to pay attention more, but it wasn't distracting."* P11 preferred *on-wrist spatial*: *"The under the finger location felt impeding. On-wrist spatial was the best since I had all the information and it felt freer."* P12 preferred *audio-visual*—*"With vibrations, it felt heavy, like a typewriter, whereas without vibrations it felt like a lighter keyboard. The sound was enough for me."*

Participant responses on the wrist conditions were split. While some mentioned that the feedback spatialization on wrist helped, others did not find any reasonable difference—P1: *"With Wrist Spatial I could tell if my pinkies are dropping onto the keyboard when they shouldn't be."* P24: *"The wrist vibrations almost blended together"*.

### 7.5 Prediction and Auto-correction

The time between presses metric can be characterized as the sum of the time the user waits to confirm their prior press ($T_c$), the time user spends in locating the next press ($T_l$), and the time to move from the current key to the next one ($T_m$). While $T_l$ and $T_m$ would be impacted most positively by the presence of a fixed reference plane, $T_c$ is the amount of uncertainty a user has about their input which could be reduced in alternative ways. The uncertainty can be broken down into two parts—whether their intended finger and that finger alone was the one that pressed the key, and whether the pressed key was the correct one or not. Participants reported that *on-fingers* feedback provided them certainty on the first part free of visual attention. One way to provide certainty on the second part is to use an accurate auto-correction decoder which increases user confidence in their input even if their input strays from their intended key. P2: *"I think if you include auto-correct here, that will help because even though vibrations sort of help with not looking at the keyboard all the time, I'm still always unsure if I pressed the exact key or not. With auto-correct I'll be more sure."* The effect of tactile feedback may thus be more evident on a keyboard with robust auto-correction and requires investigation.

### 7.6 Asymmetric Learning Effects

Even with perfect counterbalancing, within-subjects studies can be vulnerable to asymmetric learning effects [48] i.e. one condition may unduly influence another condition due to the presence of stronger feedback in one condition improving learning more for a lesser-feedback condition than vice versa. We therefore tested for order effects, i.e. whether the different orderings of the four feedback conditions had an asymmetric effect on the results. We conducted 3-way mixed ANOVAs on all reported measures with the order of feedback conditions as the between-subjects factor and SESSION and FEEDBACKTYPE as the within-subjects factors. We found no main or interaction effects of order on any of the measures.

We also analyzed the first five phrases for each participant using a between-subjects 1-way ANOVA to see the effect of FEEDBACKTYPE, but found no differences in speed and accuracy. However, participant comments suggested that certain tactile feedback may have been more useful for the participants in familiarizing themselves with the keyboard. According to P17: *"I tried to touch type initially, but that was impossible. With the finger vibrations I could kind of sense my fingers in space and know where they are. I think that really helped me to slowly start doing some touch-typing kind of thing."* A between-subjects study across several days would be needed to definitively answer if there are any asymmetric learning effects across different FEEDBACKTYPES over the long-term.

### 7.7 Transferability of Findings

Our work focuses on the specific question of tactile feedback for mid-air text-input in VR. However, our findings can inform investigations for other mid-air interaction modes in VR. For instance, 1) lower press durations with tactile feedback suggest a quicker user response time, which could be further investigated beyond key presses for virtual control/object manipulation for both discrete and continuous interactions. This could be especially useful in gaming and teleoperation scenarios where response time is crucial. 2) Reduced collisions indicate that tactile feedback helped users maintain a consistent hand position in air. This could be useful for other chording style virtual interactions where we want the user to interact with virtual objects using finger motion while keeping the hand position fixed. 3) Users reported lower visual and cognitive attention with tactile feedback. This could enable more relaxed, more eyes-free virtual object manipulations if similar trends are observed. One way to investigate this would be to use eye-tracking to quantify gaze behavior in different object manipulation tasks.

## 8 IMPLICATIONS

We now summarize the findings and their implications based on our results and discussion:

1) Tactile feedback resulted in fewer unwanted collisions per character thus indicating that users were more successful in maintaining careful hand and finger positions when tactile feedback was present. This also suggests that tactile feedback may lower the errors in a keyboard where key collisions are designed to result in key-presses.
2) Tactile feedback on the finger-base is better than spatial or nonspatial feedback on the wrist, which are in-turn better than only audio-visual feedback.
3) Tactile feedback on fingers resulted in a lower press duration than audio-visual suggesting a quicker response time. In some gaming scenarios every fraction of a second matters and tactile feedback could be useful there in speeding up freehand button clicks.
4) Participants overwhelmingly preferred tactile feedback over non-tactile feedback conditions, suggesting that users will value VR systems which integrate with existing wrist wearables with tactile feedback.
5) Tactile feedback on the finger-base was the most preferred and rated lower in effort, mental demand, and performance. This encourages the need for investigating the trade-offs between the benefits and constraints of fingertip placement.
6) Spatialized feedback on the wrist is comparable to a single-vibration motor on the wrist in almost all aspects, thus discounting the need for specific wrist hardware with multiple actuators as a way to provide finger-specific feedback.
7) The lower collisions and press durations were independent of keyboard familiarity suggesting that tactile feedback would continue to be effective and preferred over long-term use.
8) The introduction of the collision state in the mid-air keyboard as a distinct state from the press state appears to be useful for minimizing errors across all conditions, including the audio-visual only feedback.
9) Users compensate for lack of tactile feedback with higher visual and cognitive attention. This is an important implication in this context suggesting that more advanced forms of mid-air tactile feedback may be able to close the gap between mid-air typing and a physical keyboard.

## 9 CONCLUSION

Our work is the first investigation of the value of remote tactile feedback for mid-air text input in VR. Our results suggest that while tactile feedback does not result in significant improvements in user speed and accuracy, users indicated overwhelming preference for tactile feedback and scored it lower in terms of mental demand and effort. One potential reason for this trend is that in the absence of tactile feedback users use their visual and cognitive attention more, thus maintaining the same performance but expending more effort. This shows that the value of tactile feedback needs to be measured by going beyond traditional performance metrics and including evaluations that quantify user effort and mental load. We believe haptic feedback is a crucial component for text-input in VR and hope that our work serves as a guide for feedback design and as an impetus for future explorations.

## REFERENCES


[1] S. Ahn, S. Heo, and G. Lee. Typing on a Smartwatch for Smart Glasses. In *Proceedings of the Interactive Surfaces and Spaces - ISS '17*, pp. 201–209. ACM Press, New York, New York, USA, 2017. doi: 10.1145/3132272.3134136

[2] A. S. Arif and W. Stuerzlinger. Analysis of text entry performance metrics. In *2009 IEEE Toronto International Conference Science and Technology for Humanity (TIC-STH)*, pp. 100–105, Sep. 2009. doi: 10.1109/TIC-STH.2009.5444533

[3] H. Benko, C. Holz, M. Sinclair, and E. Ofek. Normaltouch and texturetouch: High-fidelity 3d haptic shape rendering on handheld virtual reality controllers. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, UIST '16, p. 717–728. Association for Computing Machinery, New York, NY, USA, 2016. doi: 10.1145/2984511.2984526

[4] S. Bovet, A. Kehoe, K. Crowley, N. Curran, M. Gutierrez, M. Meisser, D. O. Sullivan, and T. Rouvinez. Using Traditional Keyboards in VR: SteamVR Developer Kit and Pilot Game User Study. In *2018 IEEE Games, Entertainment, Media Conference (GEM)*, pp. 1–9. IEEE, aug 2018. doi: 10.1109/GEM.2018.8516449

[5] D. A. Bowman, C. J. Rhoton, and M. S. Pinho. Text Input Techniques for Immersive Virtual Environments: An Empirical Comparison. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 46(26):2154–2158, sep 2002. doi: 10.1177/154193120204602611

[6] T. Carter, S. A. Seah, B. Long, B. Drinkwater, and S. Subramanian. Ultrahaptics: multi-point mid-air haptic feedback for touch surfaces. In *Proceedings of the 26th annual ACM symposium on User interface software and technology*, pp. 505–514, 2013.

[7] D. K. Chen, J.-B. Chossat, and P. B. Shull. Haptivec: Presenting haptic feedback vectors in handheld controllers using embedded tactile pin arrays. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, p. 1–11. Association for Computing Machinery, New York, NY, USA, 2019. doi: 10.1145/3290605.3300401

[8] H.-Y. Chen, J. Santos, M. Graves, K. Kim, and H. Z. Tan. Tactor localization at the wrist. In M. Ferre, ed., *Haptics: Perception, Devices and Scenarios*, pp. 209–218. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008.

[9] I. Choi, H. Culbertson, M. R. Miller, A. Olwal, and S. Follmer. Grabity: A wearable haptic interface for simulating weight and grasping in virtual reality. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*, UIST '17, p. 119–130. Association for Computing Machinery, New York, NY, USA, 2017. doi: 10.1145/3126594.3126599

[10] I. Choi, E. Ofek, H. Benko, M. Sinclair, and C. Holz. Claw: A multifunctional handheld haptic controller for grasping, touching, and triggering in virtual reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, p. 1–13. Association for Computing Machinery, New York, NY, USA, 2018. doi: 10.1145/3173574.3174228

[11] M. Di Luca and A. Mahnan. *2019 IEEE World Haptics Conference*, 2019.

[12] T. J. Dube and A. S. Arif. Text Entry in Virtual Reality: A Comprehensive Review of the Literature. pp. 419–437. Springer, Cham, jul 2019. doi: 10.1007/978-3-030-22643-5_33

[13] J. Dudley, H. Benko, D. Wigdor, and P. O. Kristensson. Performance envelopes of virtual keyboard text input strategies in virtual reality. In *2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 289–300, Oct 2019. doi: 10.1109/ISMAR.2019.00027

[14] J. J. Dudley, K. Vertanen, and P. O. Kristensson. Fast and Precise Touch-Based Text Entry for Head-Mounted Augmented Reality with Variable Occlusion. *ACM Transactions on Computer-Human Interaction*, 25(6):1–40, dec 2018. doi: 10.1145/3232163

[15] G. González, J. P. Molina, A. S. García, D. Martínez, and P. González. Evaluation of Text Input Techniques in Immersive Virtual Environments. In *New Trends on Human-Computer Interaction*, pp. 1–10. Springer London, London, 2009. doi: 10.1007/978-1-84882-352-5_11

[16] T. Grossman, X. A. Chen, and G. Fitzmaurice. Typing on Glasses: Adapting Text Entry to Smart Eyewear. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services - MobileHCI '15*, pp. 144–152. ACM Press, New York, New York, USA, 2015. doi: 10.1145/2785830.2785867

[17] J. Grubert, L. Witzani, E. Ofek, M. Pahud, M. Kranz, and P. O. Kristensson. Effects of Hand Representations for Typing in Virtual Reality. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 151–158. IEEE, mar 2018. doi: 10.1109/VR.2018.8446250

[18] J. Grubert, L. Witzani, E. Ofek, M. Pahud, M. Kranz, and P. O. Kristensson. Text Entry in Immersive Head-Mounted Display-Based Virtual Reality Using Standard Keyboards. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 159–166. IEEE, mar 2018. doi: 10.1109/VR.2018.8446059

[19] X. Gu, Y. Zhang, W. Sun, Y. Bian, D. Zhou, and P. O. Kristensson. Dexmo: An inexpensive and lightweight mechanical exoskeleton for motion capture and force feedback in vr. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, p. 1991–1995. Association for Computing Machinery, New York, NY, USA, 2016. doi: 10.1145/2858036.2858487

[20] J. Gugenheimer, D. Dobbelstein, C. Winkler, G. Haas, and E. Rukzio. Facetouch: Enabling touch interaction in display fixed uis for mobile virtual reality. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, UIST '16, pp. 49–60. ACM, New York, NY, USA, 2016. doi: 10.1145/2984511.2984576

[21] A. Gupta, A. A. R. Irudayaraj, and R. Balakrishnan. Hapticclench: Investigating squeeze sensations using memory alloys. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*, UIST '17, p. 109–117. Association for Computing Machinery, New York, NY, USA, 2017. doi: 10.1145/3126594.3126598

[22] A. Gupta, C. Ji, H.-S. Yeo, A. Quigley, and D. Vogel. RotoSwype: Word-Gesture Typing using a Ring. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems - CHI '19*, pp. 1–12. ACM Press, New York, New York, USA, 2019. doi: 10.1145/3290605.3300244

[23] A. Gupta, T. Pietrzak, N. Roussel, and R. Balakrishnan. Direct manipulation in tactile displays. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, p. 3683–3693. Association for Computing Machinery, New York, NY, USA, 2016. doi: 10.1145/2858036.2858161

[24] S. Gupta, D. Morris, S. N. Patel, and D. Tan. 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, p. 419–428. Association for Computing Machinery, New York, NY, USA, 2013. doi: 10.1145/2493432.2493463

[25] S. Han, B. Liu, R. Wang, Y. Ye, C. D. Twigg, and K. Kin. Online optical marker-based hand tracking with deep labels. *ACM Transactions on Graphics*, 37(4):1–10, jul 2018. doi: 10.1145/3197517.3201399

[26] S. Heo, C. Chung, G. Lee, and D. Wigdor. Thor's hammer: An ungrounded force feedback device utilizing propeller-induced propulsive force. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, p. 1–11. Association for Computing Machinery, New York, NY, USA, 2018. doi: 10.1145/3173574.3174099

[27] J. D. Hincapié-Ramos, X. Guo, P. Moghadasian, and P. Irani. Con-

sumed endurance: A metric to quantify arm fatigue of mid-air interactions. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems*, CHI '14, pp. 1063–1072. ACM, New York, NY, USA, 2014. doi: 10.1145/2556288.2557130

[28] R. Hinchet, V. Vechev, H. Shea, and O. Hilliges. Dextres: Wearable haptic feedback for grasping in vr via a thin form-factor electrostatic brake. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology*, UIST '18, p. 901–912. Association for Computing Machinery, New York, NY, USA, 2018. doi: 10.1145/3242587.3242657

[29] HTC. Vive, 2019.

[30] S. Jang, W. Stuerzlinger, S. Ambike, and K. Ramani. Modeling cumulative arm fatigue in mid-air interaction based on perceived exertion and kinetics of arm motion. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, pp. 3328–3339. ACM, New York, NY, USA, 2017. doi: 10.1145/3025453.3025523

[31] H. Jiang, D. Weng, Z. Zhang, F. Chen, H. Jiang, D. Weng, Z. Zhang, and F. Chen. HiFinger: One-Handed Text Entry Technique for Virtual Environments Based on Touches between Fingers. *Sensors*, 19(14):3063, jul 2019. doi: 10.3390/s19143063

[32] J. Kim, W. Delamare, and P. Irani. ThumbText : Text Entry for Wearable Devices Using a Miniature Ring, 2018.

[33] S. Kim and G. J. Kim. Using keyboards with head mounted displays. In *Proceedings of the 2004 ACM SIGGRAPH international conference on Virtual Reality continuum and its applications in industry - VRCAI '04*, p. 336. ACM Press, New York, New York, USA, 2004. doi: 10.1145/1044588.1044662

[34] Y. R. Kim and G. J. Kim. HoVR-type: smartphone as a typing interface in VR using hovering. In *Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology - VRST '16*, pp. 333–334. ACM Press, New York, New York, USA, 2016. doi: 10.1145/2993369.2996330

[35] P. Knierim, V. Schwind, A. M. Feit, F. Nieuwenhuizen, and N. Henze. Physical Keyboards in Virtual Reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*, pp. 1–9. ACM Press, New York, New York, USA, 2018. doi: 10.1145/3173574.3173919

[36] Leap. Leap Motion, 2019.

[37] M. Lee and W. Woo. ARKB: 3D vision-based Augmented Reality Keyboard. In *International Conferece on Artificial Reality and Telexisitence*, pp. 54–57. Tokyo, Japan, 2003.

[38] J.-W. Lin, P.-H. Han, J.-Y. Lee, Y.-S. Chen, T.-W. Chang, K.-W. Chen, and Y.-P. Hung. Visualizing the keyboard in virtual reality for enhancing immersive experience. In *ACM SIGGRAPH 2017 Posters on - SIGGRAPH '17*, pp. 1–2. ACM Press, New York, New York, USA, 2017. doi: 10.1145/3102163.3102175

[39] I. S. MacKenzie and R. W. Soukoreff. Phrase sets for evaluating text entry techniques. In *CHI '03 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '03, pp. 754–755. ACM, New York, NY, USA, 2003. doi: 10.1145/765891.765971

[40] A. Markussen, M. R. Jakobsen, and K. Hornbæk. Selection-Based Mid-Air Text Entry on Large Displays. pp. 401–418. Springer, Berlin, Heidelberg, 2013. doi: 10.1007/978-3-642-40483-2_28

[41] M. McGill, D. Boland, R. Murray-Smith, and S. Brewster. A Dose of Reality: Overcoming Usability Challenges in VR Head-Mounted Displays. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15*, pp. 2143–2152. ACM Press, New York, New York, USA, 2015. doi: 10.1145/2702123.2702382

[42] Microsoft. HoloLens 2, 2019.

[43] J. O. W. Mingrui "Ray" Zhang, Shumin Zhai. Text Entry Throughput: Towards Unifying Speed and Accuracy in a Single Performance Metric. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 2019. doi: 10.1145/3290605.3300866

[44] Oculus. Oculus Rift, 2019.

[45] T. Ogitani, Y. Arahori, Y. Shinyama, and K. Gondow. Space Saving Text Input Method for Head Mounted Display with Virtual 12-key Keyboard. In *2018 IEEE 32nd International Conference on Advanced Information Networking and Applications (AINA)*, pp. 342–349. IEEE, may 2018. doi: 10.1109/AINA.2018.00059

[46] J. Perret and E. Vander Poorten. Touching virtual reality: A review of haptic gloves. In *ACTUATOR 2018; 16th International Conference on New Actuators*, pp. 1–5, June 2018.

[47] E. Pezent, A. Israr, M. Samad, S. Robinson, P. Agarwal, H. Benko, and N. Colonnese. Tasbi: Multisensory squeeze and vibrotactile wrist haptics for augmented and virtual reality. In *2019 IEEE World Haptics Conference (WHC)*, pp. 1–6, July 2019. doi: 10.1109/WHC.2019.8816098

[48] E. Poulton and P. Freeman. Unwanted asymmetrical transfer effects with balanced experimental designs. *Psychological Bulletin*, 66(1):1, 1966.

[49] M. Prätorius, D. Valkov, U. Burgbacher, and K. Hinrichs. DigiTap: an eyes-free VR/AR symbolic input device. In *Proceedings of the 20th ACM Symposium on Virtual Reality Software and Technology - VRST '14*, pp. 9–18. ACM Press, New York, New York, USA, 2014. doi: 10.1145/2671015.2671029

[50] V. Rajanna and J. P. Hansen. Gaze typing in virtual reality: Impact of keyboard design, selection method, and motion. In *Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications*, ETRA '18, pp. 15:1–15:10. ACM, New York, NY, USA, 2018. doi: 10.1145/3204493.3204541

[51] I. Rakkolainen, A. Sand, and R. Raisamo. A survey of mid-air ultrasonic tactile feedback. In *2019 IEEE International Symposium on Multimedia (ISM)*, pp. 94–944, 2019.

[52] M. Sakashita, S. Hashizume, and Y. Ochiai. Wrist-mounted haptic feedback for support of virtual reality in combination with electrical muscle stimulation and hanger reflex. In M. Kurosu, ed., *Human-Computer Interaction. Recognition and Interaction Technologies*, pp. 544–553. Springer International Publishing, Cham, 2019.

[53] M. Samad and L. Shams. Visual–somatotopic interactions in spatial perception. *Neuroreport*, 27(3):180–185, 2016.

[54] M. Sinclair, E. Ofek, M. Gonzalez-Franco, and C. Holz. Capstancrunch: A haptic vr controller with user-supplied force feedback. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*, UIST '19, p. 815–829. Association for Computing Machinery, New York, NY, USA, 2019. doi: 10.1145/3332165.3347891

[55] J. Son, S. Ahn, S. Kim, and G. Lee. Improving Two-Thumb Touchpad Typing in Virtual Reality. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems - CHI EA '19*, pp. 1–6. ACM Press, New York, New York, USA, 2019. doi: 10.1145/3290607.3312926

[56] M. Speicher, A. M. Feit, P. Ziegler, and A. Krüger. Selection-based Text Entry in Virtual Reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*, pp. 1–13. ACM Press, New York, New York, USA, 2018. doi: 10.1145/3173574.3174221

[57] S. Sridhar, A. M. Feit, C. Theobalt, and A. Oulasvirta. Investigating the Dexterity of Multi-Finger Input for Mid-Air Text Entry. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15*, pp. 3643–3652. ACM Press, New York, New York, USA, 2015. doi: 10.1145/2702123.2702136

[58] E. Strasnick, C. Holz, E. Ofek, M. Sinclair, and H. Benko. Haptic links: Bimanual haptics for virtual reality using variable stiffness actuation. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, p. 1–12. Association for Computing Machinery, New York, NY, USA, 2018. doi: 10.1145/3173574.3174218

[59] H. B. Surale, A. Gupta, M. Hancock, and D. Vogel. Tabletinvr: Exploring the design space for using a multi-touch tablet in virtual reality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, p. 1–13. Association for Computing Machinery, New York, NY, USA, 2019. doi: 10.1145/3290605.3300243

[60] E. Tsykunov and D. Tsetserukou. Wiredswarm: High resolution haptic feedback provided by a swarm of drones to the user's fingers for vr interaction. In *25th ACM Symposium on Virtual Reality Software and Technology*, VRST '19. Association for Computing Machinery, New York, NY, USA, 2019. doi: 10.1145/3359996.3364789

[61] UploadVR. Oculus Claims Breakthrough in Hand-tracking Accuracy, 2019.

[62] C.-M. Wu, C.-W. Hsu, T.-K. Lee, and S. Smith. A virtual reality keyboard with realistic haptic feedback in a fully immersive virtual environment. *Virtual Reality*, 21(1):19–29, Mar 2017. doi: 10.1007/s10055-016-0296-6

[63] N. Yanagihara and B. Shizuki. Cubic Keyboard for Virtual Reality. In *Proceedings of the Symposium on Spatial User Interaction - SUI '18*, pp. 170–170. ACM Press, New York, New York, USA, 2018. doi: 10.1145/3267782.3274687

[64] X. Yi, C. Yu, M. Zhang, S. Gao, K. Sun, and Y. Shi. ATK: Enabling Ten-Finger Freehand Typing in Air Based on 3D Hand Tracking Data. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology - UIST '15*, pp. 539–548. ACM Press, New York, New York, USA, 2015. doi: 10.1145/2807442.2807504

[65] C. Yu, Y. Gu, Z. Yang, X. Yi, H. Luo, and Y. Shi. Tap, Dwell or Gesture?: Exploring Head-Based Text Entry Techniques for HMDs. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17*, pp. 4479–4488. ACM Press, New York, New York, USA, 2017. doi: 10.1145/3025453.3025964

[66] C. Yu, K. Sun, M. Zhong, X. Li, P. Zhao, and Y. Shi. One-Dimensional Handwriting: Inputting Letters and Words on Smart Glasses. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16*, pp. 71–82. ACM Press, New York, New York, USA, 2016. doi: 10.1145/2858036.2858542

[67] D. Yu, K. Fan, H. Zhang, D. Monteiro, W. Xu, and H.-N. Liang. PizzaText: Text Entry for Virtual Reality Systems Using Dual Thumbsticks. *IEEE Transactions on Visualization and Computer Graphics*, 24(11):2927–2935, nov 2018. doi: 10.1109/TVCG.2018.2868581

[68] M. Zhu, A. H. Memar, A. Gupta, M. Samad, P. Agarwal, Y. Visell, S. J. Keller, and N. Colonnese. Pneusleeve: In-fabric multimodal actuation and sensing in a soft, compact, and expressive haptic sleeve. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, p. 1–12. Association for Computing Machinery, New York, NY, USA, 2020. doi: 10.1145/3313831.3376333