Exploring the effects of virtual hand appearance on midair typing efficiency

Dixuan Cui | Christos Mousas

Department of Computer Graphics Technology, Purdue University, West Lafayette, Indiana, USA

Correspondence
Christos Mousas, Department of Computer Graphics Technology, Purdue University, West Lafayette, IN 47907, USA. Email: cmousas@purdue.edu

Summary
Midair typing has gradually become the potential mainstream of virtual reality (VR) typing, as it eliminates the setup of additional hardware devices and provides convenience for portable head-mounted displays (HMDs). In midair typing, users type on a virtual keyboard while using their virtual hands. In addition to the lack of a dominant method for VR-based midair typing, we still do not know the impact that different appearances of virtual hands could have on text input in VR. Thus, we picked a widely used midair typing method (i.e., the VR-rendered QWERTY keyboard) and conducted a VR study to explore the effects of three hand appearances (i.e., abstract, mannequin, realistic hands) on typing efficiency in terms of typing accuracy, typing speed, and task load. Our within-group study revealed that (1) the mannequin hand caused significantly higher performance on typing speed than the abstract and realistic hands, (2) the abstract hand caused significantly lower performance (more typing errors) on typing accuracy compared to the mannequin and realistic hands, and (3) participants rated the task load of using the abstract hand significantly higher than the mannequin and realistic hands.

KEYWORDS
head-mounted display, midair typing, task load, text-input, typing accuracy, typing speed, virtual hand appearance, virtual reality

1 | INTRODUCTION

In most virtual reality (VR) applications, VR communication relies heavily on body gestures and voice chat. Collaborative VR applications allow users to interact with each other using gestures, body movements, and eye-gaze. A good example is VRChat, which allows users to interact with others and express themselves with avatar animation and audio. However, among the different methods, text-input methods were the least favored and considered, compared to voice chat, visual sharing, and avatars, in head-mounted display (HMD)-based communication. This is because text input needs to become more intuitive in virtual environments (VEs), as it takes more effort than other interaction methods. Currently, this situation is changing as there are demands on content production, such as creating annotations, labels, numeric values, and parameters on virtual objects in VR applications. Therefore, finding the optimal method for HMD-based text input is essential.

1https://hello.vrchat.com/.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2023 The Authors. Computer Animation and Virtual Worlds published by John Wiley & Sons Ltd.
Original methods, such as controller-based typing\(^8\) with ray-cast or drum-like keyboards,\(^9\) have been widely applied to customer-level devices. Researchers have also studied other methods, such as split keyboard\(^10\) and head-gaze typing,\(^8,11\) in coordination with hand-held controllers. In recent years, integration with a physical keyboard in VR\(^12-15\) has been a practical approach to reproduce the sense of realistic keyboard typing. Furthermore, researchers have made several attempts to develop midair typing in VR\(^6,7,16-18\) as it eliminates the need for a realistic keyboard or hand-held devices and provides convenience to portable devices. Although midair typing has the potential to become the mainstream VR text input method in the future, there is yet to be a dominant setup that is widely applied in VR applications. Moreover, in previous midair typing studies, there were no determined designs for effective virtual hand appearance, as researchers only designed hands with basic hand shapes and simple color textures.\(^19-21\) Thus, there is a lack of research on the effect of hand appearance in VR typing, as different body appearances can significantly affect task performance and embodiment in immersive environments.\(^21-26\)

To further understand the previously mentioned issues, we conducted a user study to explore the effect of three different virtual hand appearances (i.e., abstract, mannequin, and realistic hands) on typing speed (i.e., words per minute [WPM]), typing accuracy (i.e., total error rate [TER]), and task load (i.e., NASA’s task load index [TLX]). We asked the participants to perform three typing tasks using different phrase sets and virtual hand models. Based on the logged data collection and post-questionnaires, we aim to answer the following research questions:

1. **RQ1:** Do different hand appearances affect typing speed?
2. **RQ2:** Do different hand appearances affect typing accuracy?
3. **RQ3:** Do different hand appearances affect typing task load?

This paper is organized as follows: in Section 2, we present the related work of our study; in Section 3, we provide details on the methodology that we followed; in Section 4, we present our results; in Section 5, we discuss our findings and limitations; and finally, in Section 6, we conclude and discuss potential future directions.

## 2 | RELATED WORKS

### 2.1 | VR typing methodology

Text input has become essential in VR interactions, like other devices with operating systems. Researchers have taken different approaches to examine the optimized method for text input. The original method applied to customer-level HMDs was the controller-based technique. HMDs such as Oculus Quest\(^2\) applied ray casting onto a virtual keyboard using hand-held controllers.\(^8\) The handheld controller shoots a virtual ray pointing toward a specific key. Users can make the final confirmation by pressing a controller button. Another method was head-based orientation selection. Yu et al.\(^11\) studied different combinations of head-based text-entry methods with tapping, dwelling, and gestures. They found that users felt it easy to learn all three methods. Rajanna and Hansen\(^27\) applied gaze tracking in their VR bike simulation. They asserted that gaze typing worked best when the entire keyboard was in the field of view. The third method involving a drum-like keyboard using controllers. Google Daydream Labs\(^9\) first proposed the drum-like VR keyboard method. Users can use the controllers as sticks to strike the keys with downward movements. Finally, Whitmire et al.\(^10\) developed the split keyboard method to allow users to type using both hands with controllers. Users can select the keys using the touch-sensitive trackpad of the Vive controller and make the final confirmation by pressing the trackpad button. While these methods can be applied to VR applications, users often find that these types of typing need more efficiency because they are used to computer keyboard typing.\(^28\)

Other researchers have focused on integrating physical desktop keyboards within immersive environments to improve the original VR input methods. Users can now type in VR using a physical keyboard, either with\(^12-14\) or without\(^15,29\) a keyboard view. Study results indicated that providing a keyboard view in VR can increase typing speed and typing accuracy.\(^15,29\) Additionally, researchers have started to explore free-hand midair typing using virtual keyboards. Bowman et al.\(^6\) examined four immersive text input methods (i.e., pinch keyboard, pen and tablet, one-hand chord, and speech). They found that no single technology exhibited a high level of user satisfaction. Kim and Xiong\(^16\) examined three approaches (i.e., standard keyboard, pseudo-haptic keyboard, and pinch keyboard) with midair typing. They found that while

---

\(^8\)https://www.oculus.com/experiences/quest/.

---
standard and haptic keyboards generated stronger embodiment, the pinch keyboard was advantageous in terms of accuracy and satisfaction.

### 2.2 Virtual hand appearance

For VR text input, hand control is an essential aspect of optimized typing. Virtual hand illusion (VHI), which refers to how much participants perceive their virtual hands as their own hands, can significantly affect interactive task efficiency.21-23 Most studies have been conducted on the realism of virtual hand appearance to seek the effect on VHI. Past studies have focused on appearance realism, the size of the hand, and skin color's effect on hand illusion. Lin and Jörg30 studied how different hand appearances affect virtual hand illusions. They used six appearances (i.e., realistic, toony, very toony, zombie, robot, and wooden block) of virtual hand models. Although VHI was present for all participants, such an effect was perceived as weakest for the block model and strongest for the realistic human-hand model. In addition to changing different 3D hand models, Kilteni et al.31 performed a drumming study using different skin colors (i.e., white, casual dark-skinned, and formal light-skinned) on the self-avatar. In their study, participants experienced a full-body immersive experience with a mirror in the environment and were asked to play the drum following a virtual character in the scene. Results indicated that participants embodied in casual dark skin felt stronger body ownership and showed stronger movement patterns. Other studies21,32,33 followed similar approaches using different hand appearances as conditions and conducted studies with either mere hand observation or object manipulation with obstacles using the hands.

Studies were also conducted to explore the relationship between virtual hand appearance and task performance. Kocur et al.24 explored the effect of an avatar’s muscular appearance on physical performance. Their results indicated that being embodied in a muscular avatar can decrease the perception of effort and enhance physical performance. Singh et al.25 found a relationship between hand appearance and prediction error rate. Participants with realistic hands made more prediction errors regarding the distance of object manipulation in VR. Bailenson and Yee26 also explored the effect of avatar appearance and claimed that increasing the similarity between virtual and actual appearance can improve task performance. Based on previous findings, virtual hand appearance could be considered a factor that impacts interactive task performance and thus can potentially affect VR typing performance.

### 3 METHODOLOGY

We present the methodology we followed in this study in the following subsections.

#### 3.1 Participants

We conducted an *a priori* power analysis to determine the sample size using G*Power v.3.10 software.34 Based on a medium-effect size of .30,35 one group with three repeated measures, and a non-sphericity correction $\epsilon = .75$, to achieve an 80% power (1 − $\beta$ error probability), the analysis recommended a minimum of 24 participants. For our within-group study, we recruited 26 participants (age: $M = 24.92, SD = 3.45$) from our university. Fifteen were male, and 11 were female. All participants volunteered for this study and completed the experiment by responding to the questionnaires provided.

#### 3.2 VR application

We developed a VR application for our study in Unreal Engine 4 and deployed it on an Oculus Quest 2 HMD. We followed previous midair typing studies,16,19 so we used the VR-rendered QWERTY keyboard in the application. Results of earlier studies showed that the QWERTY keyboard fits the users’ preference for typing, as it can best replicate typical typing experiences.16,19 The application consisted of an empty virtual room with a black virtual keyboard in front of the virtual character floating in midair (see Figure 1). We used the base of a VR keyboard application from the UE4 marketplace.3

FIGURE 1  The midair keyboard for the experiment. Participants can type on the keyboard using their index fingers. Above the keyboard, we see the task information.

FIGURE 2  The hands used in this experiment. From left to right: abstract, mannequin, realistic male, and realistic female hand models.

modified the collision and functionality (i.e., backspace, space, and enter) to fit our study purposes. On top of the virtual keyboard were two lines of instructional text. The first line shows the experiment’s task instruction; the second line shows the context for the typing task. Participants in the application were embodied in a pair of virtual hands depending on the study condition and gender (see Figure 2). We used the free-hand tracking system from Oculus Quest 2 in our application to detect actual hand location and apply animation to the virtual hand.

We developed three conditions for this study resembling three models: abstract, mannequin, and realistic hand models. We followed previous VR keyboard studies and considered avoiding using multiple fingers for free-hand typing, as it can generate collision and detection issues. 16,19 Thus, we designed our typing task so that participants could only type on the keyboard using their index fingers on both hands. We used Mackenzie and Soukoreff’s 36 phrase set for each study condition and applied three phrases for the participants to type in. Mackenzie and Soukoreff evaluated 500 data sets via the internet and compiled phrases suitable for typing tasks. Researchers who conducted VR midair typing used the same sets in their studies and found them helpful in experiment design. 11,16,20,28 We provide our phrase sets in Table 1. For each condition, a random group of phrases was selected.

3.3  Measurements

In our study, we collected typing efficiency data regarding text-entry rate and typing accuracy to evaluate the three hand appearances. We computed words per minute (WPM) for the text-entry rate. WPM was the most widely reported
TABLE 1  The phrase sets used in the experiments.

<table>
<thead>
<tr>
<th>Set 1</th>
<th>Set 2</th>
<th>Set 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase 1</td>
<td>Phase 1</td>
<td>Phase 1</td>
</tr>
<tr>
<td>My preferred treat is</td>
<td>Physics and chemistry are hard</td>
<td>Wear a crown with</td>
</tr>
<tr>
<td>chocolate</td>
<td></td>
<td>many jewels</td>
</tr>
<tr>
<td>Phase 2</td>
<td>Phase 2</td>
<td>Phase 2</td>
</tr>
<tr>
<td>Question that must be</td>
<td>We are subjects and must obey</td>
<td>My bank account is</td>
</tr>
<tr>
<td>answered</td>
<td></td>
<td>overdrawn</td>
</tr>
<tr>
<td>Phase 3</td>
<td>Phase 3</td>
<td>Phase 3</td>
</tr>
<tr>
<td>There will be some fog</td>
<td>Great disturbance in the force</td>
<td>Movie about a nutty</td>
</tr>
<tr>
<td>tonight</td>
<td></td>
<td>professor</td>
</tr>
</tbody>
</table>

Note: For each condition, our application randomly picked a phrase.

measurement for text-entry performance.\textsuperscript{37,38} We should note that in typing experiment tasks, “word” was considered a five-character word, including spaces.\textsuperscript{39} The only considerations in the process were the length of the final string and the completion time.\textsuperscript{37} The equation, according to Wobbrock,\textsuperscript{37} for calculating WPM is as follows:

\[
WPM = \frac{|T| - 1}{S} \times 60 \times \frac{1}{5},
\]

where \(T\) refers to the final string length entered by the participants, which counts any inputs besides backspaces. \(S\) represents the time in units of seconds. Considering the time started when the first character was entered, we followed the same process as in previous studies and never timed the first character; thus, the \(-1\) was included in the string length;\textsuperscript{37,38} 60 refers to seconds per minute, and \(\frac{1}{5}\) refers to words per character.

Our second measurement was the total error rate (TER). We followed the Mackenzie and Soukoreff\textsuperscript{36} method to collect in-game variables regarding correct and incorrect characters and performed measurements, as shown in the following equation:

\[
TER = \frac{INF + IF}{C + INF + IF} \times 100\%,
\]

where \(C\) refers to correct keystrokes, \(INF\) refers to incorrect and not fixed keystrokes, which are unfixed errors in the final phrase, and \(IF\) refers to incorrect but fixed keystrokes, which are the number of times fixing errors during the input process.\textsuperscript{36}

We also evaluated the task load for each condition using NASA’s task load index (TLX) scale. The scale included measurement in six different subjective subscales: mental demand, physical demand, temporal demand, overall performance, effort, and frustration level. We used the paper-and-pencil version of the questionnaire and converted it into a computer-based survey using the Qualtrics online tool. We followed the weighting procedure from the official NASA TLX website\textsuperscript{4} and performed manual calculations using the same weight for each subscale. We eliminated the comparison process to use the raw TLX for our results. We then received our results as a workload score from 0 to 100.

### 3.4 Procedure

Upon the participants’ arrival at the lab, we provided them with a digital consent form approved by our university’s institutional review board. We then asked the participants to review and sign the consent form to indicate their participation

\textsuperscript{4}http://humansystems.arc.nasa.gov/groups/TLX/.
agreement. After the participants signed their consent forms, we asked them to provide basic demographic information in our Qualtrics survey. Later, we assisted our participants in setting up the HMD and adjusted the size of the headbands to fit their heads. After briefly introducing the study, we started the application and asked the participants to complete all the necessary tasks in all conditions (see Figure 3).

4 | RESULTS

We performed all analyses using IBM SPSS v.28.0 statistical analysis software. We conducted a one-way repeated measures analysis of variance (ANOVA) for WPM, TER, and TLX using different hand models as factors. We evaluated the normality assumption of logged data and self-reported ratings using the Shapiro-Wilk tests at the 5% level and graphically using Q-Q plots of the residual. For all statistical tests, $p < .05$ was deemed statistically significant. We provide boxplots of our results in Figure 4.

4.1 | Words per minute

The statistical analysis revealed a significant main effect of hand models on the words per minute in midair typing (Wilks’ $\Lambda = .293, F[2, 24] = 28.967, p = .000, \eta^2_p = .707$). Pairwise comparisons showed that participants typed significantly faster using the mannequin hand ($M = 11.65, SD = 3.41$) than using the realistic hand ($M = 9.27, SD = 3.25$) at $p = .001$ and

![Figure 4](image)

**Figure 4** Box plots of our data results: (a) Words Per Minute, (b) Error Rate, and (c) Task Load. Boxes enclose the middle 50% of the data. A thick horizontal line denotes the median. We denote significant levels as: * for $p < .05$, ** for $p < .01$, *** for $p < .005$, and **** for $p < .001$. 
abstract hand \((M = 7.32, SD = 2.69)\) at \(p < .001\). We also found that participants typed significantly faster when using the realistic hand than the abstract hand at \(p = .036\).

4.2 | Total error rate

The statistical analysis revealed a significant effect of different hand models on the total error rate of midair typing (Wilks’ \(\Lambda = .674, F[2, 24] = 5.809, p = .009, \eta^2_p = .326\)). Pairwise comparisons revealed that participants using the abstract hands \((M = 37.99, SD = 10.75)\) made significantly more errors compared to when they were using the mannequin hands \((M = 31.24, SD = 11.79)\) at \(p = .01\) and the realistic hands \((M = 31.86, SD = 11.62)\) at \(p = .024\).

4.3 | Task load index

Based on the statistical analysis, we found a significant effect of hand models on the task load of mid-air typing (Wilks’ \(\Lambda = .535, F[2, 24] = 10.449, p = .001, \eta^2_p = .465\)). Pairwise comparisons showed that participants using the abstract hand \((M = 71.08, SD = 15.50)\) rated higher their task load compared to when using the mannequin hand \((M = 60.46, SD = 14.59)\) at \(p = .003\) and realistic hand \((M = 59.31, SD = 13.89)\) at \(p = .001\). We did not find a significant difference between the mannequin and realistic hands.

5 | DISCUSSION AND LIMITATIONS

In this study, we explored how different virtual hand models affect typing speed, typing error rate, and task load in VR midair typing. We discuss our findings and our study’s limitations in the following subsections.

5.1 | Effect of hand appearance

For RQ1, we found that when we assigned our participants to the mannequin hand, they completed the typing task faster than when they used abstract and realistic hands. Moreover, when we assigned participants to realistic hands, they completed the typing task faster than when they used abstract hands. In previous VR midair typing studies, researchers suggested that struggling to guide a virtual finger on keystroke tapping would result in worse typing efficiency.\(^{19,40}\) We think we found such a result because of the virtual hand illusion (VHI). Two dimensions of the VHI (i.e., body ownership and agency) modulate proprioceptive drift and thus can affect the position perception of the virtual hand.\(^{41}\) While the realistic hand provided a realistic external appearance, mannequin hands were able to generate the strongest hand ownership in VR.\(^{42}\) Therefore, participants should have made more effort to position their mannequin hands for typing. In addition, the realistic hand and abstract hand generated a sense of style clash compared to the virtual keyboard and background environment, which decreased the sense of ownership.\(^{43}\)

Another possible reason for this result could be the keyboard’s size perception. Moskvin et al.\(^{44}\) stated that the larger the keyboard size, the faster participants could type. Based on previous studies, the results suggested that virtual hand fidelity significantly affected object size perception.\(^{45,46}\) Participants with less realistic hands would perceive objects as smaller. As for our study, we think the abstract hand provided smaller size recognition over the keyboard and slowed the typing process.

Regarding typing accuracy (RQ2), we found that when our participants used the abstract hand, they performed significantly more errors than when they used the other two hands. One possible reason is the discordance between virtual hand animation and actual hand movement. Grubert et al.\(^{17}\) mentioned in their previous study that matching motion between virtual and real hands can decrease the error rate. Furthermore, in virtual hand appearance studies, Cui and Mousas\(^{42,47}\) confirmed that virtual hands with less detail could cause confusion regarding the implicit perception (i.e., location of hand landmarks, joints, and fingers) of the virtual hands. The abstract hand had basic joints and connections as skeletons, while the other two hands had full mesh and could perform richer animations to match real hand actions. In addition, participants provided feedback indicating that the abstract hand tracking felt different compared to
the actual hand movements and caused issues in positioning their fingers. Therefore, during the typing process, participants sensed a discoordination between the abstract hand and their actual hand, and, as a result, this negatively impacted their performance compared to when using the other virtual hands.

Finally, for RQ3, we found that when the participants used the abstract hand, they rated their task load higher. Our results aligned with previous studies, which found that the abstract hand caused lower presence and a higher workload. Based on the feedback received from the participants, they commented that the abstract hand was “hard to recognize” and “unfamiliar.” In VR, Schwind et al. reported a similar finding: VR participants highly focus on virtual hands for interaction tasks. Therefore, in our study, the abstract representation did not appear clearly and caused an additional workload in the typing tasks.

One of the main issues in typing is occlusion with the keyboard. Thus, although the abstract hand provided a better view of the keyboard, since fingers were not attached to the wrist, it may lead to confusion when hands were close to each other. However, we can also assume that the abstract hand caused a poor experience because it was “too abstract” and had fewer depth cues. Specifically, both the mannequin and realistic hands were shaded and had thicker fingers. Thus, it would have performed better if the abstract hand consisted of rigid bodies, such as cylinders or cuboids. Compared to the other models, the mannequin’s hand has slightly pointed fingers, which we think has played a role in our participants’ performance.

5.2 Limitations

There are several limitations worth mentioning for this study. The first limitation was the virtual keyboard design. We tried to generate a realistic typing experience for our participants. However, we received feedback regarding the size, angle, and distance between different keyboard keys. Furthermore, based on our observation, “fat-fingering” onto other keys occasionally happened during the typing process, which can cause additional mental demand and error rate in the experiment. We think adjustments to the keyboard model’s design and collision volume are necessary for future studies.

The second limitation was the human body position for midair typing. We developed the application to suit the sitting participants. Nevertheless, the participants constantly had to lift their arms midair during typing. This potential risk could affect typing efficiency after a certain amount of time due to exhaustion. We would need to adjust the position of the keyboard and apply the armrest in the real world to eliminate fatigue factors.

Our last limitation was the hand tracking itself. Although Oculus Quest 2 free hand tracking was sufficient in designing the application, a more precise hand-tracking method would be preferred. However, we needed access to higher-quality hand trackers. Thus, researchers should apply fingertip trackers or VR hand gloves to refine the midair typing experiment with richer hand animation and a more accurate hand position.

6 CONCLUSION AND FUTURE WORK

We developed a VR typing application to conduct a within-group study to explore the effects of hand appearance on midair typing. Our results showed that when participants used the abstract hand, they made the most mistakes and rated the highest task load. Moreover, we found that when participants used the mannequin hand, they completed the typing task faster than when using the other two hand models.

In addition to the previously mentioned limitations, a few more directions are worth studying in the future. Researchers could investigate the effect of tactile feedback by applying a vibrotactile stimulus on the fingertip during a keystroke. This will give the research community more insight into realistic typing experiences in VR. Moreover, a comparison between different virtual keyboard styles is necessary. While we developed a keyboard with critical animation sinking when stroking, other studies used a static flat keyboard and found it to be effective in typing tasks. Study results about the effect of virtual keyboard design can also provide design considerations for VR typing. Furthermore, researchers could also study the size of the virtual keyboard to find the optimal setup for VR typing. Past studies have found that different virtual hand sizes affect object size perception. Due to the various sizes of human hands, it is essential to understand whether we need to dependently adjust the size of the virtual keyboard. Lastly, precision with different hand models could also impact typing performance. Study designs with Fitts’ law and throughput measurement can bring interesting results regarding the effect of pointing and typing precision.
REFERENCES


47. Cui D, Mousas C. Evaluating virtual hand illusion through realistic appearance and tactile feedback. Multimodal Technol Interact. 2022;6(9):76.


AUTHOR BIOGRAPHIES

Dixuan Cui holds a BS and MS in Computer Graphics Technology from Purdue University (West Lafayette, IN). Dixuan is currently a PhD student in the Department of Computer Graphics Technology at Purdue University. His research interests include virtual reality, tactile feedback, computer graphics, games, and human-computer interaction. Dixuan has received a Best Paper Award from ACM SIGGRAPH VRCAI 2022.

Christos Mousas is an Assistant Professor at Purdue University (West Lafayette, IN). His research revolves around virtual reality, virtual humans, computer graphics & animation, and human-computer interaction. He has previously been affiliated with the Department of Computer Science at Dartmouth College (Hanover, NH) as Postdoctoral Researcher, and with the Department of Cultural Technology and Communication of the University of the Aegean (Mytilene, Greece) as a Research Assistant. He holds a PhD in Informatics and an MSc in Multimedia Applications and Virtual Environments, both from the School of Engineering and Informatics of the University of Sussex (Brighton, UK), and an integrated Master’s degree in Audiovisual Science and Art from Ionian University (Corfu, Greece). He is a member of ACM and IEEE, and has been a member of the organizing and
program committees of many conferences in the virtual reality, computer graphics/animation, and human-computer interaction fields. Dr Mousas has received a Best Paper Award from ACM SIGGRAPH VRCAI 2022, and Honorable Mention Awards (top 5% of all accepted papers) from ACM CHI PLAY 2021 and ACM CHI 2022.

**How to cite this article:** Cui D, Mousas C. Exploring the effects of virtual hand appearance on midair typing efficiency. Comput Anim Virtual Worlds. 2023;e2189. [https://doi.org/10.1002/cav.2189](https://doi.org/10.1002/cav.2189)