Physical Keyboards in Virtual Reality: Analysis of Typing Performance and Effects of Avatar Hands

Pascal Knierim[1], Valentin Schwind[2], Anna Maria Feit[3], Florian Nieuwenhuizen[2], Niels Henze[2]
[1]LMU Munich, Munich, Germany, pascal.knierim@ifi.lmu.de
[2]University of Stuttgart, Stuttgart, Germany, valentin.schwind@vis.uni-stuttgart.de,
niels.henze@vis.uni-stuttgart.de, florian.nieuwenhuizen@gmail.com
[3]Aalto University, Helsinki, Finland, anna.feit@aalto.fi

ABSTRACT
Entering text is one of the most common tasks when interacting with computing systems. Virtual Reality (VR) presents a challenge as neither the user's hands nor the physical input devices are directly visible. Hence, conventional desktop peripherals are very slow, imprecise, and cumbersome. We developed a apparatus that tracks the user's hands, and a physical keyboard, and visualize them in VR. In a text input study with 32 participants, we investigated the achievable text entry speed and the effect of hand representations and transparency on typing performance, workload, and presence. With our apparatus, experienced typists benefited from seeing their hands, and reach almost outside-VR performance. Inexperienced typists profited from semi-transparent hands, which enabled them to type just 5.6 WPM slower than with a regular desktop setup. We conclude that optimizing the visualization of hands in VR is important, especially for inexperienced typists, to enable a high typing performance.

ACM Classification Keywords
H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems-Artificial, Augmented, and Virtual Realities

Author Keywords
Virtual reality, Text entry, Physical keyboard, Hands

INTRODUCTION
Head-mounted displays (HMDs) for Virtual Reality (VR) are finally available for the consumer market. Today, consumers mainly use VR for entertainment applications including 3D movies and games [16]. A large field of view (FOV), high visual fidelity as well as the visual and auditory encapsulation can create truly immersive experiences with almost unlimited opportunities. Gamepads and novel tracked input devices are used to interact with the application or game.

CHI 2018, April 21–26, 2018, Montreal, QC, Canada

ISBN 978-1-4503-5620-6/18/04...$15.00
DOI: https://doi.org/10.1145/3173574.3173919

While consumers currently use VR for entertainment, a wide range of serious applications have been proposed in the past and are currently explored by industry and academia [11, 17]. VR systems also offer great potential to create pleasant working or study environments for office workers or students. Commuters in trains and cars [11], or employees working from home could wear an HMD to sit virtually in their familiar working environment or attend business meetings far away. External visual and auditory distractions can be blocked completely, which would aid productive and focused work. Furthermore, VR allows the creation of entirely new environments with vast three-dimensional display space in any direction. VR user interfaces are no longer bound to rectangular two-dimensional displays limited by the size of our desks. Future VR systems could supersede current interaction paradigms and enable improved work performance.

Present HMDs deliver rich and high immersion due to the latest technology advances. They deliberately limit the connection to the real world to create a high level of immersion and strong sensation of presence. Unfortunately, visual immersion not only substitutes real world distractions but also makes it impossible to see the physical keyboard and mouse which are essential for high-bandwidth general purpose interaction [10]. Typewriting is the most used generic text input method for desktop computing. To enable users to work as efficiently in a virtual environment as in a real office they require high performance input devices. Especially for users who are not fluent touch typists, text input quickly becomes tiresome if they cannot see their hands or the keyboard.

Different solutions have been proposed for text input while immersed in VR. They embrace point and click solutions with tracked controllers, handwriting with a pen on a tablet or speech. Others overlay the virtual environment with a cropped video stream of the real world. None of these solutions can facilitate high-performance text input known from real world typing. Of course the user can take off the HMD every time a text input needs to be made, however this quickly becomes inconvenient and destroys the immersion.

To enable performant typing in VR, we developed a apparatus that visually represents the user's hands and the physical keyboard of a desktop workspace in VR. Keyboard and fingers are tracked and visualized in real time to support the user visually to interact with the peripheral. In a text input study, with particular emphasis on the hand visualization, we evaluated typing speed and accuracy in contrast to real life typing.

This work paves the way for efficient work in VR due to new techniques for generic text input in VR. We present design recommendations for the most suitable hand visualization depending on the typing proficiency. We contribute with our findings on the effect of avatar hands on typing performance in VR using a physical keyboard. With our apparatus, we regulate the workload while typing in VR and eliminate the frustration of taking off the HMD and destroying the sense of presence.

RELATED WORK

The first Virtual Reality HMD system was created in 1968 by computer scientist Ivan Sutherland [15]. With new technology advances companies like HTC, Oculus and Sony recently catapulted virtual reality into the living rooms. However, the potential which lies within education and business and beyond gaming is largely under-explored. We review text input solutions tailored for VR and works exploring the effect of avatar hands on VR environments.

Text input

Current game-controllers or gesture interfaces are suitable for micro inputs and games. Unfortunately, they fail to support high bandwidth generic input. Researchers, developers, and stakeholders have proposed a wide variety of different text input solutions for VR. A comparison of early works using pen, keyboard and gesture-based techniques showed that the achieved text input rates of less than 14 words per minute (WPM) stay far behind real world typing speeds on a physical keyboard [2]. Gonzales et al. [4] confirmed these results analyzing a different set of input devices developed for text input in VR.

All commercially available controllers enclosed with the HMD of HTC, Oculus and Sony support text input on a virtual keyboard. Users point a virtual ray with the controller onto a character and confirm the selection with a button press. Alternatively, a built-in touch pad can be used to move the pointer around. Similarly, R. Kim and J. Kim [7] use the touch screen with hovering capabilities of a smartphone for selecting characters on a virtual keyboard. With their input technique they achieve up to 9.2 WPM.

The Microsoft HoloLens supports text input in augmented reality applications through a holographic keyboard and a pointer which is controlled using head rotation. Yu et al. [21] also studied head-based text entry for VR and combined the concept with gesture-word recognition whereby experienced users perform up to 25 WPM. The VR system FaceTouch [5] leverages a touch-sensitive front cover of the HMD and the sense of proprioception to enable text input with up to ten WPM on a virtual keyboard.

None of these approaches can keep up high input speed and usability known from typing on a physical keyboard. Recently, researchers focused on augmenting VR by incorporating a video stream of reality into the virtual environment to compensate typing performance decrease [10]. Lin et al. [8] extend this approach by utilizing a depth camera to display a point cloud of a user's hands beside a rendered virtual representation of the physical keyboard. To compensate for the increased error rate introduced while typing in VR, Walker et al. propose to use decoders known from text entry on touchscreens to correct errors [18, 19]. Overall, it remains an open challenge how to build a VR system that supports accurate and fast text input that can compete with typing on a regular desktop setup.

Avatar Hands

The effectiveness of virtual environments has been linked to the subjective experience of being and acting at one place while physically situated at another [20]. New sensors can easily determine the hands pose and position to render them in VR accordingly. Displaying them increases the immersion and presence and further enables natural user interaction within the virtual environment [1]. Schwind et al. [13] investigated the effect of different hand renderings on presence. Results highlight the importance of users' diversity when designing virtual reality experiences.

We take the current body of related work and investigate how hand representation regarding model, texture, and transparency affect typing performance, workload and felt presence. We restrict the physical environment to a seated setup while the user feeds text into the system via a physical keyboard.

REALIZING TYPING IN VIRTUAL REALITY

For any physical keyboard based text input users execute, they need to localize and reach out to the keyboard in a first step. Localizing could either happen visually or haptically using the surface features of the keyboard. VR HMDs prevent the user from visually localizing any physical peripherals. A system realizing effortless typing in virtual realities should support the user with an easy to understand representation of the keyboard's location in relation to their fingers. According to Feit et al. [3], non-touch typists' gaze switches up to 1.2 times between the display and the keyboard within a sentence. They spend up to 41% of their attention looking at the keyboard. Hence, an accurate representation of the keyboard and hands seems necessary particularly for this group of typists.

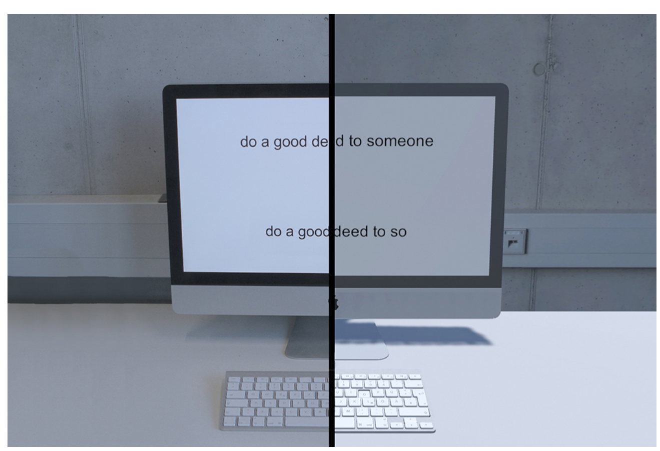

Figure 1: Side by side illustration of the real environment (left) and the virtual reality replica (right).

IMPLEMENTATION

To investigate the different aspects of typing in a virtual environment, we implemented our VR apparatus using an Oculus Rift Consumer Version 1 (CV 1). The Oculus camera tracks the headsets position. We incorporated a motion tracking system comprising eight OptiTrack 13W cameras and the Motive 1.10 motion capture software for very accurate finger and keyboard tracking. Twenty-three 4 mm retroreflective markers are affixed to anatomical landmarks of each hand ensure precise tracking of each joint and bone of the hand. During the application startup, markers are seamlessly analyzed and automatically mapped to the virtual skeleton. In case of losing track of a marker during typing due to occlusion, our software automatically reassigns it, when it reappears, to untracked joints following a nearest neighbor approach. The layout of the markers is depicted in Figure 2.

A second generation Apple wireless keyboard is used for text input. Four retroreflective markers are attached to the top of the keyboard to enable repositioning of the keyboard during runtime to allow comfortable typing. The precise and interactive virtual replica of the keyboard is rendered according to physical position and keypresses in the virtual environment.

Our apparatus uses the OptiTrack NetNat SDK for streaming position data of bones, joints, and keyboard in real time. Our application and the virtual environment are implemented using the Unity game engine 5.4.0.

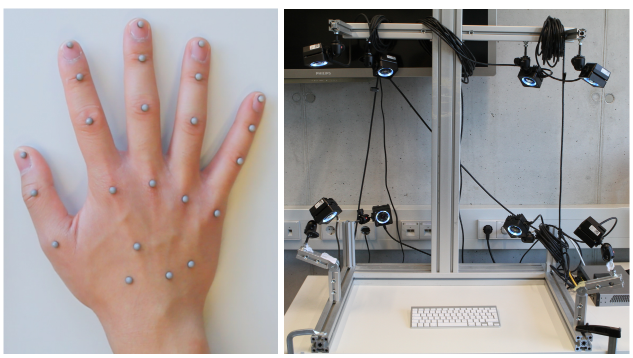

Figure 2: Hand with 23 retroreflective markers (left) and the hardware setup for finger and keyboard tracking (right).

METHOD

Our apparatus enables users to see the virtual representation of a physical keyboard and their own hands. The goal of this study is to evaluate the effect of virtual hand representation and hand transparency on typing performance of experienced and inexperienced typists in VR. Further, we investigate the overall typing experience by measuring the task load and sense of presence. We used a mixed nested factorial design with the nested within-subject variable Hand and Transparency and the between-subject variable Typing Experience. For Hand we had three different levels. All hands were presented with 0% and 50% Transparency. In addition, we use 100% Transparency resulting in no hand visualization and the real world scenario. An overview of all eight conditions is shown in Figure 3. Typing performance was measured while participants typed outside of VR on the real world apparatus, or inside of VR seeing different hands with varying transparency levels.

Subjects

In a first step, we asked 80 (5 female) participants to conduct a simple online typing test.[1] Based on their results ($M = 53.3$ WPM, $SD = 18.8$), we invited a random sample of 16 participants with more and 16 participants with less than 53.3 WPM to shape groups of inexperienced and experienced typists. The 32 participants (three female) were aged from 18 to 27 ($M = 21.9$, $SD = 2.3$). Thirteen participants had previous experience with VR. Fourteen of them were wearing corrective lenses during the study. Participants received a small gratuity and either 10 EUR or course credits as compensation for their participation.

Apparatus

The apparatus for this study comprised two individual setups. One facilitated the real world typing task, the other allowed users to type on a physical keyboard while immersed in VR.

Real World Apparatus

The real world setup served as a baseline and consisted of a sixth generation 27 inch Apple iMac with Intel Core i5 and a second generation Apple wireless keyboard. The computer was running a full-screen typing application showing one stimulus after another at the display. It was developed in Unity game engine 5.4.0.

Virtual Reality Apparatus

For the virtual reality setup, we used our developed VR apparatus. We designed an alike looking virtual environment representation of our laboratory including the real world study apparatus comprising the iMac. The real world apparatus next to the virtual replica is shown in Figure 1.

Our experiment was running on a Windows PC with an Intel i7-6700, 16GB RAM, and a Nvidia GTX980. The target frame rate was set to 90 frames per second (FPS) to match the refresh rate of the Oculus Rift CV 1. Of course, there is a latency between a user's finger movement and photons hitting the user's retina. We used the provided performance toolboxes to monitor the latency. The summed up calculated latency caused by motion tracking, rendering pipeline, and HMD never exceeded 30 ms during the study.

Task

In this study participants had to accomplish a simple text input task on a physical keyboard. Participants were asked to place their hands left and right next to the keyboard to mimic aperiodic typing. Being in this resting pose, a 3-second countdown, displayed on the (virtual) iMac, started. After it elapsed, a random phrase from the MacKenzie and Soukoreff [9] phrase set was displayed. Participants were asked to enter the phrase as accurately and fast as possible. Phrases were presented at the top of the (virtual) display while participants' input was shown underneath. Participants were allowed to correct errors but also to confirm inaccurate or incomplete phrases. Pressing

[1] https://10fastfingers.com

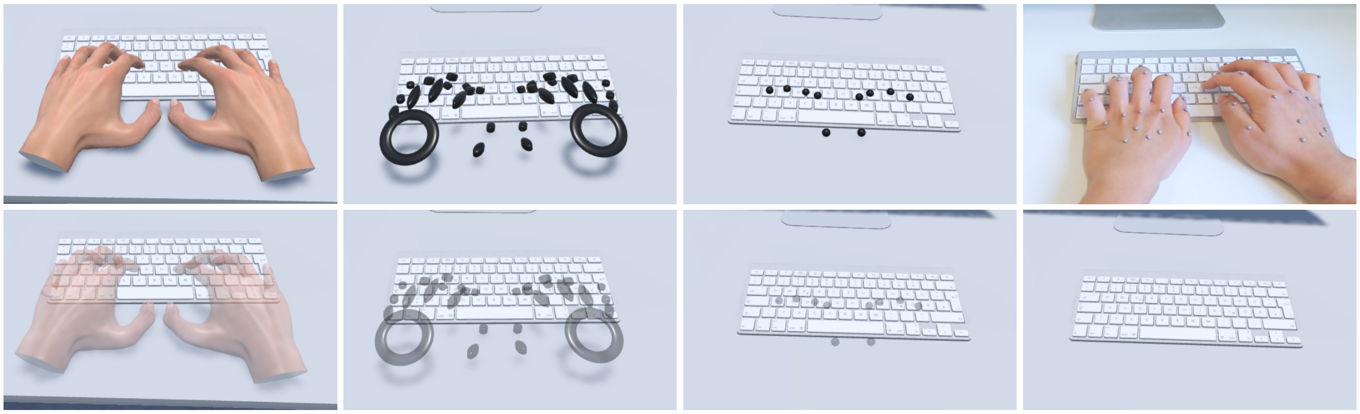

Figure 3: Pictures of the eight hand visualizations used in the study. Realistic, abstract, fingertips with no transparency and real hands (1st row) as well as 50% transparency and no hands (2nd row).

the enter key confirmed the input and the next phrase was displayed. For each condition, participants performed three sets of ten phrases. In between each set participants had to place their hands in the resting position again and wait for the countdown to elapse. The task was the same for all conditions inside and outside of the VR.

Procedure

After welcoming the participants, we asked them to sign the consent form and take a seat next to the apparatus. While attaching the 23 self-adhesive markers to each hand, we explained all devices and the course of the study to the participants. Afterward, the participant placed his hands within the tracking volume, and we defined the four markers at the dorsum of the hand as rigid bodies. In the last preparation step, we adjusted the HMD to the participant's head and calibrated it to the participant's inter pupil distance for best visual results. Then participants started with the typing task. After each task (three sets of 10 phrases), they had to fill out the NASA-TLX [6] and presence questionnaire (PQ) [20]. Subsequently, they repeated the procedure using the next hand representation. The first set of ten phrases at the start of each condition was a practice set to familiarize the participant with the different appearances. We did not include this set in our analysis. For the baseline outside of VIRTUAL REALITY, participants had to take off the HMD and move to the real setup to continue with the text input task. HANDS and VIRTUAL REALITY were presented in a counterbalanced order using a full latin square to prevent sequence effects. Throughout the study, we logged every keystroke including the timestamp for offline analyses. After all eight iterations, we asked for comments about their experience, typing performance, and which hand representation they finally preferred. Including debriefing and detaching the self-adhesive markers, participants completed the study in 70 to 110 minutes.

RESULTS

We conducted multiple four-way repeated measure analyses of variance (RM-ANOVA) with the within-subjects variables VIRTUAL REALITY, HAND, TRANSPARENCY, and the between-subjects variable TYPING EXPERIENCE. As previously mentioned, the within-subjects factor HAND is a nested factor of the VIRTUAL REALITY condition. TRANSPARENCY is nested into HANDS, which means that conditions of a nested factor cannot be compared with levels of factors above (*e.g.*, there is no transparency in the *Real World* condition). All significance levels are at $\alpha = .05$.

Objective Measures

One participant was removed from the analysis of the objective measures due to missing correct inputs (error rate: 100%) in multiple conditions. Hence, we invited one more participant from the same group of typists to compensate for the deficit. In total participants wrote 7680 phrases and we analyzed 5120 phrases since the first ten phrases of each condition were assigned for training. The results of the objective measures are shown in Figure 4. The mean values of all metrics are listed in Table 1.

Words per Minute (WPM)

The average typing performance is calculated in WPM where one word is defined to be five characters long [14]. Based on the logged keystrokes, we divided the length of the final input by the time the participant took to enter the phrase. We measured the time from the first to the confirm keypress to calculate the WPM.

We found a significant effect of VIRTUAL REALITY, $F(1,30) = 22.97, p < .001$, and an interaction effect of VIRTUAL REALITY × TYPING EXPERIENCE, $F(1,30) = 22.97, p < .001$. Furthermore, we found a significant effect of HAND, $F(3,90) = 8.336, p < .001$, but no interaction effect of HAND × TYPING EXPERIENCE, $F(3,90) = .439, p < .726$. And we found no significant effects of TRANSPARENCY, $F(3,90) = 1.596, p = .196$, and no interaction of TRANSPARENCY × TYPING EXPERIENCE, $F(3,90) = 1.022, p = .387$.

Post-hoc analysis was performed using Bonferroni corrected pairwise t-tests to determine statistically significant differences between the conditions. Due to the significant effects of TYPING EXPERIENCE, we compared the measures between

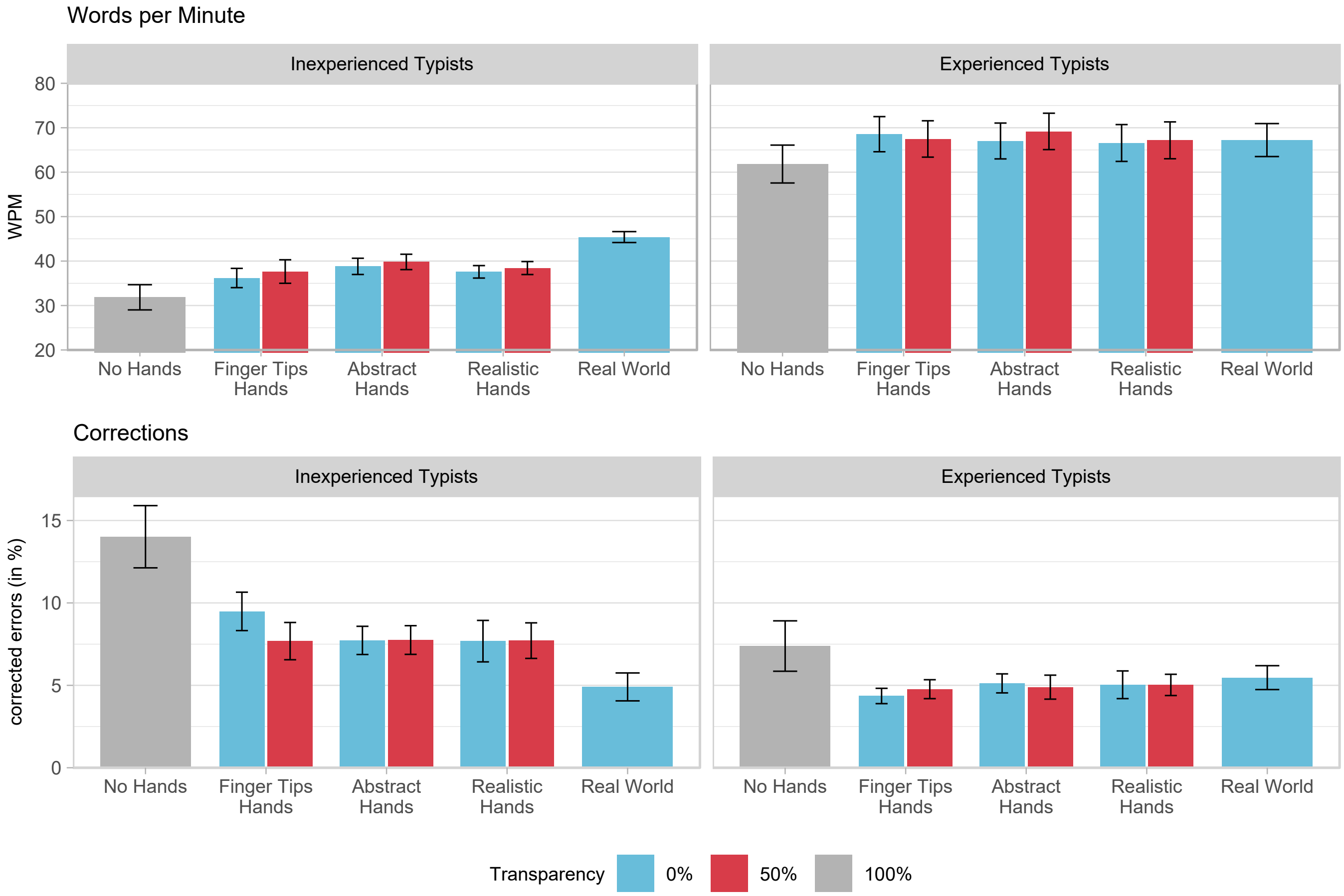

Figure 4: Mean values of words per minute and corrected error rate for each condition. Error bars show standard error of the mean (SE). Exact values are also listed in Table 1.

experienced and *inexperienced* users separately. Due to no statistically significant effects of TRANSPARENCY, the data were aggregated across the transparency levels. For *inexperienced* users and the HAND factor we found significant differences between *No Hands* and the *Real World* condition ($p < .001$), *No Hands* and *Abstract Hands* ($p = .024$), between *Finger Tips* and *Real World* ($p = .006$), between *Abstract Hands* and the *Real World* condition ($p < .001$), and between *Real World* and the *Realistic Hands* ($p = .041$). No significant differences were found by comparing the other hand pairs (all with $p > .05$). Furthermore, we found no significant differences between the hand conditions only considering *experienced* typing users in VR (all with $p = 1$).

We summarize that the rendering of hands in VR has a significant effect on the typing performance measured using the WPM for *inexperienced* users in VR. The actual appearance of hands had no significant effect on the WPM measure of *experienced* users in typing.

Error Rate

One measure as an indicator of the users' typing performance alongside the WPM is the number of errors in the transcribed string. The *Error Rate* is given by the minimum string distance (*MSD*) between the transcribed string (T) and the presented phrase (P). The *Error Rate* in percent is: $ErrorRate = \frac{MSD(P,T)}{max(|P|,|T|)} \times 100$. It captures the minimum number of insertions, deletions, or substitutions we have to perform to change one phrase into another [14].

We found a significant effect of VIRTUAL REALITY, $F(1,30) = 6.463, p = .016$, but no interaction effect of VIRTUAL REALITY × TYPING EXPERIENCE, $F(1,30) = 3.086, p = .089$ on the correction measure. There was no significant effect of HAND, $F(3,90) = 2.389, p < .073$ and no significant interaction of HAND × TYPING EXPERIENCE, $F(3,90) = 1.034, p = .381$. Both TRANSPARENCY, $F(3,90) = .158, p = .924$, as well as the interaction of TRANSPARENCY × TYPING EXPERIENCE, $F(3,90) = .337, p = .799$, were not significant. Pairwise post-hoc comparisons of the corrections showed no differences between the conditions of experienced and inexperienced users (all with $p > .05$)

Corrections

Neither WPM nor *Error Rate* captures the number of corrections and edits made during text input. The *Corrected Error Rate* [14] represents the effort put into correcting errors. We calculated the *Corrected Error Rate* by offline analysis of the keystroke log file. Therefore, we analyzed the log file and sought characters appearing in the keystroke log file, but not in the final transcribed text.

	VIRTUAL REALITY														REAL WORLD	
HAND	**No Hand**		**Finger Tips**				**Abstract**				**Realistic**					
TRANSPARENCY	**100%**		**0%**		**50%**		**0%**		**50%**		**0%**		**50%**		**0%**	
INEXPERIENCED TYPIST	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
WPM	31.848	11.338	36.189	8.669	37.648	1.576	38.809	7.307	39.808	6.919	37.581	5.611	38.430	5.839	45.398	4.909
Error Rate (in %)	.740	.538	1.244	1.528	1.304	1.127	.973	.707	1.121	.842	1.140	1.041	1.001	1.097	.713	.635
Corrected Error Rate (in %)	14.015	7.549	9.486	4.660	7.683	4.521	7.726	3.422	7.749	3.480	7.681	5.031	7.712	4.313	4.904	3.388
1^{st} correct Keypress (in s)	4.386	2.813	2.200	1.236	1.971	.632	1.986	.793	2.129	.842	2.456	1.290	1.864	.528	1.769	1.054
EXPERIENCED TYPIST	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
WPM	61.830	17.047	68.547	15.810	37.648	1.576	67.018	16.134	69.172	16.370	66.566	16.569	67.165	16.589	67.223	14.837
Error Rate (in %)	.540	.505	.757	.922	.846	.974	1.003	1.135	.687	.613	.449	.415	.362	.316	.597	.528
Corrected Error Rate (in %)	7.383	6.116	4.354	1.858	4.766	2.297	5.118	2.307	4.889	2.909	5.034	3.361	5.025	2.578	5.467	2.899
1^{st} correct Keypress (in s)	3.638	2.026	2.108	1.332	1.831	.670	1.791	.568	1.821	.561	1.953	.754	1.980	.692	1.370	.423

Table 1: Means and Standard Deviations (SD) of the typing performance indices of inexperienced and experienced typists: words per minute (WPM), error rate, corrected characters per phrase, and the time for the 1^{st} correct keypress.

We found a significant effect of VIRTUAL REALITY, $F(1,30) = 14.4, p < .001$, and an interaction effect of VIRTUAL REALITY × TYPING EXPERIENCE, $F(1,30) = 18.4, p < .001$ on the corrected error rate. There was a significant effect of HAND, $F(3,90) = 9.933, p < .001$, however, not interaction of HAND × TYPING EXPERIENCE, $F(3,90) = 2.03, p = .115$. Both TRANSPARENCY, $F(3,90) = 1.006, p = .393$, as well as the interaction of TRANSPARENCY × TYPING EXPERIENCE, $F(3,90) = 2.527, p = .062$, were not significant.

Pairwise post-hoc comparisons of the ratio between corrected and overall inputs considering *inexperienced* users in typing showed significant differences between all hands and the *No Hands* condition (all with $p < .05$). Further pairwise comparisons considering other pairs and pairwise comparisons of experienced typists were not significant (all with $p > .05$).

Response Time Until the 1^{st} Correct Keypress
For several applications, the time to react on a specific event using keyboard input is a critical measure of typing performance. After the expiration of the countdown, we recorded the time (in s) a user needed for the first correct keyboard input.

VIRTUAL REALITY had a significant effect on the reaction time, $F(1,30) = 22.85, p < .001$, however, there was no interaction of VIRTUAL REALITY × TYPING EXPERIENCE, $F(1,30) = .19, p = .666$. We found a significant effect of HAND, $F(3,90) = 17.947, p < .001$, however, not on HAND × TYPING EXPERIENCE, $F(3,90) = .374, p = .772$. There were no effects of TRANSPARENCY, $F(3,87) = 1.324, p = .271$, or TRANSPARENCY × TYPING EXPERIENCE, $F(3,90) = .872, p = .459$).

Pairwise post-hoc comparisons of the average response times until the first correct keyboard input revealed significant differences between all hands and the *No Hands* condition for inexperienced as well as experienced users in typing (all with $p < .001$). Other pairwise comparisons of the reaction time measure were not significant (all with $p > .05$). Thus, particularly to have *No Hands* in VR affected the initial response time for the first keyboard event negatively for both *inexperienced* and *experienced* users in typing.

Subjective Measures
Further analyses were conducted to assess how the participants subjectively perceived the virtual hands. We asked for perceived work load and presence. All measures are shown in Figure 5.

Task Load Index (NASA-TLX)
To assess the users' perceived task load of each hand we used the TLX score of the NASA-TLX questionnaire. We found significant main effects of VIRTUAL REALITY, $F(1,30) = 17.514, p < .001$, and HAND, $F(3,90) = 13.735, p < .001$, but no effect of TRANSPARENCY, $F(3,90) = 0.676, p = .569$. There were no interaction effects and none of the TLX measures was significantly affected by TYPING EXPERIENCE (all with $p > .05$).

Pairwise post-hoc comparisons of typing accuracy between the conditions considering the aggregated TLX measures across TRANSPARENCY and TYPING EXPERIENCE show statistically significant differences between *No Hands* ($M = 9.906, SD = 3.583$) and *Finger Tips Hands* ($M = 7.698, SD = 3.463$, with $p = .025$), between *No Hands* and *Abstract Hands* ($M = 7.021, SD = 3.348$, with $p < .001$), between *No Hands* and *Real Hands* ($M = 5.542, SD = 3.500$, with $p < .001$), between *No Hands* and *Realistic Hands* ($M = 6.625, SD = 2.954$, with $p < .001$), and between *Finger Tips Hands* and *Real Hands* ($p = .030$).

We summarize that having *No Hands* caused a significantly higher workload than the other conditions for both *experienced* as well as *inexperienced* users in typing. The lowest TLX score within the conditions of VIRTUAL REALITY was achieved by using *Realistic Hands*.

Presence
The presence questionnaire (PQ) was primarily designed to compare experiences in VR [20]. For the sake of completeness and to avoid potential biases, we asked for presence in the *Real World* condition as well. The overall score was averaged. Subscales are not considered in the following analysis. We found a significant effect of VIRTUAL REALITY, $F(1,30) = 99.62, p < .001$, and HAND, $F(3,90) = 13.269, p < .001$, but no effect of TRANSPARENCY, $F(3,90) = .549, p = .650$.

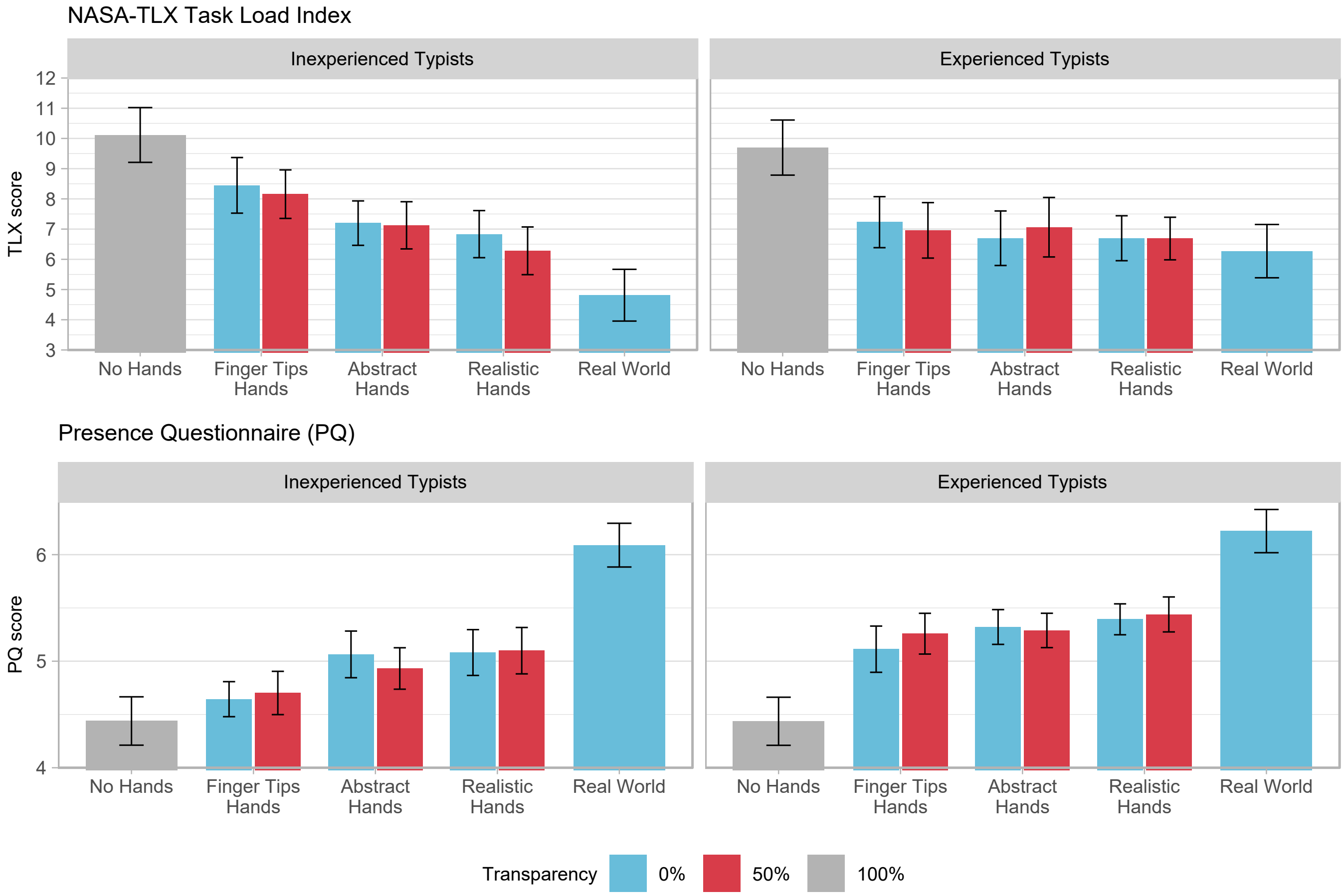

Figure 5: Subjective assessments of task load and presence. Error bars show standard error of the mean (SE).

There were no interaction effects and none of the PQ measures was significantly affected by TYPING EXPERIENCE (all with $p > .05$).

Pairwise post-hoc comparisons of the measures between the conditions considering the aggregated PQ scores across TRANSPARENCY and TYPING EXPERIENCE show statistically significant differences between *No Hands* ($M = 4.438, SD = .890$) and *Abstract Hands* ($M = 5.151, SD = .744$, with $p = .002$), between *No Hands* and *Real Hands* ($M = 6.155, SD = .806$, with $p = .001$), between *No Hands* and *Finger Tips Hands* ($M = 4.929, SD = .806$, with $p < .010$), between *No Hands* and *Realistic Hands* ($M = 5.253, SD = .753$, with $p < .001$), between *Finger Tips Hands* ($M = 9.906, SD = 3.583$) and *Real Hands* ($M = 7.698, SD = 3.463$, with $p < .001$), between *Abstract* and *Real Hands* (with $p < .001$), and between *Real World* and *Realistic Hands* ($p < .001$). Other pairwise comparisons (*Finger Tips* and *Abstract Hands*, *Finger Tips* and *Realistic Hands*, *Abstract* and *Realistic Hands*) were not significant (all with $p = 1.000$).

We summarize that the perceived presence was significantly affected by VIRTUAL REALITY and HANDS. The highest presence score was achieved using *Realistic Hands* while *No Hands* and *Finger Tips Hands* received the lowest presence scores.

DISCUSSION

Our results show that the typing performance of mainly inexperienced users using a physical keyboard in VR was significantly decreased compared to real world text input. This is confirmed by several works evaluating typing in VR [8, 10, 19]. Experienced typists' text input performances were not significantly affected by missing hands or the different hand visualizations. However, rendering virtual avatar hands significantly increases the typing performance, response time, and typing accuracy of inexperienced users. Renderings of each virtual hand pair brought their typing performance back to a level that did not significantly differ from measurements in the real world.

Our results neither confirm an effect of appearance nor of transparency. Related to the degree of realism or human likeness of a virtual avatar, previous work suggests an effect of the Uncanny Valley. As we found no effects between abstract and very realistic hands, we cannot confirm an effect of the Uncanny Valley on the typing performance in VR.

Since the mental workload increases while typing in the virtual world, we assume that users are rather focused on the typing task than on the appearance of their hands. This finding is supported by two studies by Schwind et al. [12, 13] which reported that participants were highly focused while performing a typing task using virtual hands and non-physical keyboards in VR. In the present study, we confirm these observations even using a physical keyboard in VR.

Our results show that the workload is statistically higher for all typists when no hands are visible. However, experienced typists' workload is not affected by typing in VR as long as hands are rendered. This leads to the assumption that hand rendering has less impact on typing performance since experienced typists do not rely as much on the visual cues. Further, Realistic Hands caused the lowest workload for all, while maintaining the highest presence scores for typing in VR. Abstract or the absence of hands causes lower presences and a higher workload. We assume that the possible negative effect of latency, tracking errors as well as limited headset resolution and field of view contribute to the increased workload for inexperienced typists since they rely on seeing the own hands while typing [3]. Video see-through solutions [10] could minimize some of this factors like tracking errors or latency, though at the expense of full control over the hand and keyboard rendering as well as higher levels of immersion.

Setting typing performance, workload, and measured presence into contrast, our results suggest a correlation in particular for inexperienced typists, who seem to struggle more with abstract hand representations. We assume they need more visual guidance and abstract hands look less familiar to them. For future systems that enable typing in VR, our findings imply rendering realistic looking hands for best typing performance as well as high presence.

Limitations and Future Work

To achieve precise tracking and visual accuracy, our apparatus relays on a high-quality motion capturing system. Hence, our setup is not mobile and self-adhesive retroreflective markers need to be attached to each hand. Large occlusion of markers or palm up-facing hand poses cause the tracking to fail. We evaluated the Leap Motion[2], a small sensor specific for hand tracking, to build a mobile version of our apparatus. Positional tracking is almost accurate enough, however, cannot match the precision of a professional motion capturing system. For the future, we envision a small, mobile sensor with high accuracy to build a truly mobile setup.

With our apparatus, experienced typists can perform text input as fast as in the real world. Inexperienced typists are on average only 5.6 WPM slower. We will further investigate different visualizations and layout to further optimize their performance or even outperform real world typing.

CONCLUSION

Consumer virtual reality is still in the fledgling stages and mainly targets entertainment use cases. We have shown the potential of VR for a wide variety of use cases by enabling natural generic text input on a physical keyboard while being immersed in a virtual environment. Our apparatus comprises a calibration free, low latency, and accurate finger tracking with a state of the art head mounted display. Thus we can create virtual environments allowing for effortless typing in VR. We conducted a text input study with a total of 32 typists of different skill levels and tested their typing performance based on various hand representations and transparency in a virtual environment. In the study, we found no significant difference in typing speed for the experienced typists using avatar hands. Inexperienced typists require hand visualizations to orient themselves in VR while transparency has no effects. They were just 5.6 WPM slower in VR using transparent realistic hands. Further, results show that all typists benefit from seeing a representation of their hands during non-contiguous typing. Results show that experienced typists are less affected by different hand rendering. For all typists, realistic hands rendering are in favor to generates the highest presence with the lowest workload.

A new combination of mobile HMDs, advanced finger trackers and a keyboard could allow us to have a truly mobile office. For the future, we envision working in well-known but highly flexible virtual environments completely independent of where we are physically located.

[2]https://www.leapmotion.com/

ACKNOWLEDGMENTS

This work was supported by the German Federal Ministry of Education and Research as part of the project Be-greifen (Grant No. 16SV7527) and the German Research Foundation (DFG) within project C04 of SFB/Transregio 161.

REFERENCES

1. Mahdi Azmandian, Mark Hancock, Hrvoje Benko, Eyal Ofek, and Andrew D. Wilson. 2016. Haptic Retargeting: Dynamic Repurposing of Passive Haptics for Enhanced Virtual Reality Experiences. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 1968–1979. DOI: http://dx.doi.org/10.1145/2858036.2858226

2. Doug A. Bowman, Christopher J. Rhoton, and Marcio S. Pinho. 2002. Text Input Techniques for Immersive Virtual Environments: An Empirical Comparison. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 46, 26 (2002), 2154–2158. DOI: http://dx.doi.org/10.1177/154193120204602611

3. Anna Maria Feit, Daryl Weir, and Antti Oulasvirta. 2016. How We Type: Movement Strategies and Performance in Everyday Typing. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 4262–4273. DOI: http://dx.doi.org/10.1145/2858036.2858233

4. Gabriel González, José P. Molina, Arturo S. García, Diego Martínez, and Pascual González. 2009. *Evaluation of Text Input Techniques in Immersive Virtual Environments*. Springer London, London, 109–118. DOI: http://dx.doi.org/10.1007/978-1-84882-352-5_11

5. Jan Gugenheimer, David Dobbelstein, Christian Winkler, Gabriel Haas, and Enrico Rukzio. 2016. 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). ACM, New York, NY, USA, 49–60. DOI: http://dx.doi.org/10.1145/2984511.2984576

6. Sandra G Hart. 2006. NASA-task load index (NASA-TLX); 20 years later. In *Proceedings of the human factors and ergonomics society annual meeting*, Vol. 50. Sage Publications, 904–908. DOI: http://dx.doi.org/10.1177/154193120605000909
7. Youngwon R. Kim and Gerard J. Kim. 2016. 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)*. ACM, New York, NY, USA, 333–334. DOI: http://dx.doi.org/10.1145/2993369.2996330
8. Jia-Wei Lin, Ping-Hsuan Han, Jiun-Yu Lee, Yang-Sheng Chen, Ting-Wei Chang, Kuan-Wen Chen, and Yi-Ping Hung. 2017. Visualizing the Keyboard in Virtual Reality for Enhancing Immersive Experience. In *ACM SIGGRAPH 2017 Posters (SIGGRAPH '17)*. ACM, New York, NY, USA, Article 35, 2 pages. DOI: http://dx.doi.org/10.1145/3102163.3102175
9. I. Scott MacKenzie and R. William Soukoreff. 2003. Phrase Sets for Evaluating Text Entry Techniques. In *CHI '03 Extended Abstracts on Human Factors in Computing Systems (CHI EA '03)*. ACM, New York, NY, USA, 754–755. DOI:http://dx.doi.org/10.1145/765891.765971
10. Mark McGill, Daniel Boland, Roderick Murray-Smith, and Stephen Brewster. 2015. 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)*. ACM, New York, NY, USA, 2143–2152. DOI: http://dx.doi.org/10.1145/2702123.2702382
11. Mark McGill, Alexander Ng, and Stephen Brewster. 2017. I Am The Passenger: How Visual Motion Cues Can Influence Sickness For In-Car VR. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 5655–5668. DOI: http://dx.doi.org/10.1145/3025453.3026046
12. Valentin Schwind, Pascal Knierim, Lewis Chuang, and Niels Henze. 2017a. "Where's Pinky?": The Effects of a Reduced Number of Fingers in Virtual Reality. In *Proceedings of the 2017 Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts (CHI PLAY Companion '17)*. ACM, New York, NY, USA. DOI: http://dx.doi.org/10.1145/3116595.3116596
13. Valentin Schwind, Pascal Knierim, Cagri Tasci, Patrick Franczak, Nico Haas, and Niels Henze. 2017b. "These Are Not My Hands!": Effect of Gender on the Perception of Avatar Hands in Virtual Reality. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 1577–1582. DOI: http://dx.doi.org/10.1145/3025453.3025602
14. R. William Soukoreff and I. Scott MacKenzie. 2003. Metrics for Text Entry Research: An Evaluation of MSD and KSPC, and a New Unified Error Metric. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '03)*. ACM, New York, NY, USA, 113–120. DOI: http://dx.doi.org/10.1145/642611.642632
15. Ivan E. Sutherland. 1968. A Head-mounted Three Dimensional Display. In *Proceedings of the December 9-11, 1968, Fall Joint Computer Conference, Part I (AFIPS '68 (Fall, part I))*. ACM, New York, NY, USA, 757–764. DOI: http://dx.doi.org/10.1145/1476589.1476686
16. Bruce H. Thomas. 2012. A Survey of Visual, Mixed, and Augmented Reality Gaming. *Comput. Entertain.* 10, 1, Article 3 (Dec. 2012), 33 pages. DOI: http://dx.doi.org/10.1145/2381876.2381879
17. Etienne van Wyk and Ruth de Villiers. 2009. Virtual Reality Training Applications for the Mining Industry. In *Proceedings of the 6th International Conference on Computer Graphics, Virtual Reality, Visualisation and Interaction in Africa (AFRIGRAPH '09)*. ACM, New York, NY, USA, 53–63. DOI: http://dx.doi.org/10.1145/1503454.1503465
18. J Walker, S Kuhl, and K Vertanen. 2016. Decoder-assisted typing using an HMD and a physical keyboard. In *Extended Abstracts of the the ACM Conference on Human Factors in Computing Systems, CHI*, Vol. 16.
19. James Walker, Bochao Li, Keith Vertanen, and Scott Kuhl. 2017. Efficient Typing on a Visually Occluded Physical Keyboard. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 5457–5461. DOI: http://dx.doi.org/10.1145/3025453.3025783
20. Bob G. Witmer and Michael J. Singer. 1998. Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence: Teleoper. Virtual Environ.* 7, 3 (June 1998), 225–240. DOI: http://dx.doi.org/10.1162/105474698565686
21. Chun Yu, Yizheng Gu, Zhican Yang, Xin Yi, Hengliang Luo, and Yuanchun Shi. 2017. 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)*. ACM, New York, NY, USA, 4479–4488. DOI: http://dx.doi.org/10.1145/3025453.3025964