Observing Virtual Avatars: The Impact of Avatars' Fidelity on Identifying Interactions

Florian Mathis
florian.mathis@glasgow.ac.uk
University of Glasgow
Glasgow, United Kingdom

Kami Vaniea
kvaniea@inf.ed.ac.uk
University of Edinburgh
Edinburgh, United Kingdom

Mohamed Khamis
mohamed.khamis@glasgow.ac.uk
University of Glasgow
Glasgow, United Kingdom

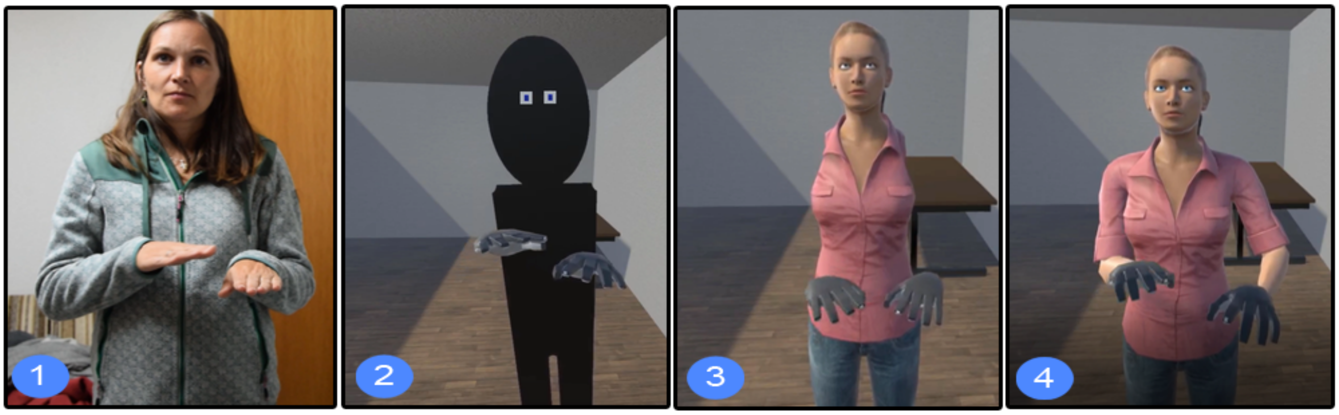

Figure 1: We investigate the impact of an avatar's fidelity on a bystander's performance when identifying the avatar's touch, mid-air, and eye gaze gestures. We had an abstract avatar (❷) and two more realistic avatars (❸, ❹), which are provided by Microsoft Research [24] and modified based on our investigation. We use touch, mid-air, and eye gaze gestures performed by a human in the real world (❶) as our baseline.

ABSTRACT

There are many cases where observing the interactions of a virtual avatar can be useful. However, it is unclear to what extent the avatar fidelity affects bystanders' performance when observing gestures performed by avatars. We, therefore, conducted an online study (N=28) with different avatars performing touch, mid-air, and eye gaze gestures. Our avatars range from an abstract avatar to two more realistic ones. Our study shows that an abstract avatar provides bystanders with the same interaction information as more realistic avatars. This implies that it is sufficient to use abstract avatar designs when observing interaction behaviour in virtual environments and researchers do not necessarily need to go through time-consuming and expensive implementations of highly realistic avatars to answer their research questions, making Virtual Reality (VR) studies more accessible to the broader research community. Finally, we discuss studying VR security systems as a potential research application for which abstract avatars can be leveraged.

Mindtrek '21, June 1–3, 2021, Tampere/Virtual, Finland

The definitive Version of Record was published in *Academic Mindtrek 2021 (Mindtrek '21), June 1–3, 2021, Tampere/Virtual, Finland*, https://doi.org/10.1145/3464327.3464329.

CCS CONCEPTS

• **Human-centered computing** → **HCI design and evaluation methods**; **Mixed / augmented reality**.

KEYWORDS

Virtual Reality, Virtual Environments, Avatars, Input Methods

ACM Reference Format:
Florian Mathis, Kami Vaniea, and Mohamed Khamis. 2021. Observing Virtual Avatars: The Impact of Avatars' Fidelity on Identifying Interactions. In *Academic Mindtrek 2021 (Mindtrek '21), June 1–3, 2021, Tampere/Virtual, Finland.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3464327.3464329

1 INTRODUCTION

Virtual avatars are used extensively in different research fields such as gaming [9, 37], education [11, 54, 81], or social virtual environments (VE) [15, 16, 29, 36, 88]. There is a large body of work that investigated the perception of virtual avatars (e.g., [62]) or studied people's body ownership of a virtual avatar (e.g., [25]). At the same time, surprisingly little is known about bystanders'[1] performance and perception when observing virtual avatars providing input with touch, mid-air, and eye gaze. For example, are bystanders able to distinguish between different mid-air gestures performed

---

[1] Note that in contrast to other works (e.g., [27]) that used the term *bystander* for Non-HMD users, we use the term *bystander* when we refer to other users in the same environment. This can be a person next to a human (in the real world, our baseline) or a user next to an avatar (in a virtual environment).

by an abstract avatar? Or do we need highly realistic avatars that better match with the interactions performed by a human in the real world? Finding answers to these questions can be particularly important. The recent restrictions in conducting face-to-face user studies due the COVID-19 pandemic [52] forced many researchers to find alternative ways to conduct human-centred research. There is already work that looked into using VR as a test-bed to, for example, evaluate user behaviour in front of public displays [39]. However, it remains unclear how avatar designs impact bystanders' performance and perception when observing corresponding human gestural movements. If abstract avatars can provide the same information to other co-located users as highly realistic avatars or even humans in the real world, then the additional effort required to set up more realistic avatars can be eliminated; making human-centred studies in virtual environments (also called "VR studies" [39, 41, 46, 63]) more accessible to the broader HCI community. VR studies can also be particularly helpful in times where face-to-face user studies are challenging to conduct (e.g., not having access to specific equipment and/or physical locations) [39] or even prohibited (e.g., due to COVID-19 restrictions) [44].

Our work discusses the results of an online study (N=28) where participants (in the role of bystanders) watched gestures performed by a human in the real world (baseline) and by different avatars in a virtual environment (Figure 1). We expected gestures performed by more realistic avatars to be easier to identify than those performed by an abstract avatar; however the difference was found to be negligible. Our work suggests that easier to create/less resource-based avatars can already provide researchers with sufficiently accurate findings and, depending on the research context, lead to similar research findings as more realistic avatars, thereby reducing the required effort to leverage VR for human-centred research.

1.1 Contribution Statement

The contribution of our work is threefold:

1. We explore the impact of avatar fidelity[2] on bystanders' interaction identification performance when performing touch, mid-air, and eye gaze gestures.
2. We show through a user study that an abstract avatar design can already provide bystanders with the same amount of information as observing a human in the real world (baseline).
3. We discuss the implications of our findings in the context of social VR and propose and discuss studying VR security systems as a potential application area for which abstract avatars can be leveraged (see Section 5.2).

2 RELATED WORK

To guide readers of this work from more general VR research to more specific avatar-focused research, we first discuss recent works which leveraged *VR studies* for human-centred research. We then review works which studied avatar designs in virtual environments and users' perception of different avatars. Finally, we discuss prior works which investigated the visual identification of human gestural movements.

2.1 Virtual Reality Studies ("VR studies")

Prior works on virtual reality research, including the works about avatar design that we discuss in Section 2.2, and also, for example, the different locomotion techniques (see [22] for an overview) or novel VR interaction techniques (e.g., [3, 7, 74]), provided the ground work to apply *VR studies* for more general human-centred research. For example, there is already research that applied VR to evaluate user behaviour in front of public displays [39] or studied a real-world authentication system's usability and security through VR replication studies [41]. In the work by Saffo et al. [63] the researchers even proposed remote VR experiments using social VR platforms to produce practically and ethically valid research results. Similar to the transition of lab to online studies through surveys or online platforms, VR studies can further inspire and support human-centred research in times where face-to-face user studies are not applicable or challenging to conduct (e.g., having access to specific physical locations [39]). There is a significant larger body of research that used VR as a platform for human-centred evaluations (e.g., [64, 86, 87]) and discussed the strengths and potential pitfalls of such (remote) VR studies (e.g., [44, 60]).

2.2 Research on Virtual Avatars

2.2.1 Avatars in Virtual Environments. There is a large body of work on avatars that aimed to understand body ownership and further effects on psychological theories. Work by Slater et al. [70] examined the role of the virtual body in immersive virtual environments and it became clear that, similar to our body in reality, the virtual body plays a primary role in virtual environments. Particularly, the virtual body is the representation of self and lays the foundation of an interaction model for body-centred interactions. Follow up work by Slater et al. [68] showed that there seems to be a relationship between body movement and presence, highlighting the importance of semantically appropriate body gestures (e.g., head movements) that are considered to be "natural" and can likely increase users' presence. In a similar vein, more recent works by, for example, Piumsomboon et al. [56] showed that the presence of an adaptive avatar significantly improved social presence and the overall experience of collaborations in virtual environments.

Merola et al. [45] even argued that "using an avatar is in many ways like donning a halloween costume" and that avatars can change users' behaviour. Saffo et al. [63] also emphasised the importance of the avatar selection in their guidelines for social VR studies. The findings by Heidicker et al. [29] even suggested that more abstract avatars (e.g., head and hands only) already produce an increased feeling of co-presence and behavioural interdependence, which implies that complete avatar bodies are not necessarily required in social VR [29] (e.g., see also Mozilla Hubs [47]). That being said, Gonzalez-Franco et al. [23] have recently argued that rigged avatars will be key to the future of VR and its wide adoption. They also open-sourced an avatar library for research and academic purposes [24], which can further help researchers to adopt VR for their research. All the previous works emphasise the importance of avatars in virtual environments.

2.2.2 Perception of Others' Avatars in Virtual Environments. Work by Slater et al. [69] explored computer generated audiences of avatars and how they are perceived when training individuals to

---

[2] Note that with the term "fidelity" we refer to the extent to which the different avatars convey the look of a human in the real world (similar to [83]).

give speeches. The proposed audience avatars were capable of moving their heads and came with different facial expressions. Slater et al. [69] concluded that human subjects respond appropriately to a negative or positive virtual audience. It has also been argued that avatars are more effective when perceived as humans. Ströfer et al. [75] showed that participants' electrodermal activity differences between deceivers and truth tellers were only significant when participants believed they interacted with a human-operated avatar, rather than a computer-operated avatar. Other work by Nowak and Rauh [48] suggested that more anthropomorphic avatars are perceived as more credible than less anthropomorphic avatars. Work by Khan et al. [33] showed that physically attractive avatars are rated higher on social competence, social adjustment, and intellectual competence, and Nowak and Rauh [49] showed that masculinity or femininity of an avatar significantly influenced the perception of avatars (e.g., masculine avatars were perceived as less attractive than feminine avatars [49]). Other work showed that a cartoon-like avatar created stronger co-presence but was less trustworthy than a realistic avatar [31]. Casanueva and Blake [8] found significant differences between co-presence generated by realistic human-like avatars, cartoon-like avatars, and unrealistic avatars, with the human-like ones achieving the highest sense of co-presence. The works mentioned above emphasise the importance of understanding users' perceptions of different virtual avatars.

2.3 Visual Identification of Human Gestural Movements

Despite the promising computer vision approaches that aim to analyse human movements (e.g., see [17] for an overview), there is also research on user-centred approaches where the visual identification task is performed by a user rather than an algorithm. For example, the work by Thornton et al. [79] investigated the distance at which an actor could be placed so that an observer would still be able to interpret human actions. In their work, users' action perception remained functional over a large range of distances (within a distance of 200 m there seems to be little impact of the distance on users' observation performance) [79]. Fademrecht et al. [13] showed that users' recognition performance of human gestures are high (> 66%) up to 30° of eccentricity and Lapenta et al. [35] even argued that human action perception is so powerful that movement identification is possible even in the absence of pictorial information (e.g., texture, facial expressions). Other work by Gleeson et al. [21] showed that different hand gestures performed by a robot can be interpreted with high accuracy by observers. There are some additional works where the visual identification of human gestural movements plays an important role (e.g., when observing user interactions to assess a system's security [1, 43]). Abdrabou et al. [1], for example, studied to what extent hand gestures and multimodal approaches (e.g., eye gaze + hand gestures) can be leveraged for user authentication and concluded that user interactions based on hand gestures only are easier to observe than multimodal approaches.

2.4 Summary and Research Gap

From previous research we learned that avatars play an important role in virtual environments (e.g., see also [84] for a review of studies using virtual characters). We also learned that Human-computer Interaction (HCI) researchers started to leverage VR for their non-VR related research (e.g., to study human behaviour in front of public displays [39]). Although it has been argued that VR could be a feasible research tool [39, 41, 86], it can require a significant amount of additional expertise and effort [39]. For example, many VR studies require complex setups and additional hardware (e.g., OptiTrack systems for finger tracking [34] or Microsoft Kinects for body tracking [73]). Furthermore, there are many ways to represent users in VR (e.g., [24, 29, 53]), and it is often not clear which one best imitates users. It is common to embody avatars of various appearances ranging from abstract to realistic [50]. Although there are human-centred works which studied visual identification of human gestural movements (e.g., [79]), the impact of an avatar's fidelity on users' gesture identification performance seems to be relatively unexplored.

As a result, we explore in this work the impact of different avatar fidelities on bystanders' interaction identification performance. In other words, we compare how well different avatars can be observed when performing touch, mid-air, and eye gaze gestures.

3 PREAMBLE

To find answers to bystanders' interaction identification performance, we covered three different avatars (ranging from an abstract avatar to a highly realistic avatar by Microsoft Research [24]) and used a human in the real world as our baseline. We evaluated the different avatars in front of a public display that allows input with three input methods: touch, mid-air, and eye gaze. All input methods and corresponding gestures are summarised in Table 1. We cover a wide range of different input methods and gestures, many of which are frequently used in virtual environments (e.g., touch: [42, 76], mid-air: [72, 77], eye gaze: [42, 55]). We decided to pre-record the interactions of both the human in the real world and the avatars in a virtual environment. We borrowed this approach from researchers who used video recordings to assess a system's resistance against observations (e.g., [4, 10]) or studied human action perception (e.g., [35]). Recording the interactions (vs live observations) also enabled us to conduct the entire study online and provide all participants with the exact same set of material (i.e., in live observations the experimenter's gestures likely vary between participants).

3.1 Apparatus and Implementation

We used Unity 3D (C# as programming language) to set up a virtual room and the different avatars. The experimenter performed all interactions outlined in Table 1 in (a) the real world and (b) in the virtual world. To support the experimenter performing the interactions, we replicated the interface of [32]'s public display to indicate the gesture direction for touch and mid-air (see also Figure 2). For eye gaze, moving targets were used as stimulus that move along the trajectories as outlined in Table 1. This approach is equivalent to prior work which used moving targets to enable gaze-based interaction (e.g., [32, 82]). All avatars used in our study are tracked through the VR headset's position and rotation in the space. Furthermore, we use a Leap Motion Controller for the finger tracking and the integrated Tobii eyetracker (together with the Tobii XR SDK [80]) for the eye gaze movement of the avatars. For the more realistic avatars, we used an avatar provided by Microsoft Research [24]

Input method	Gesture	Number of different Gestures	Description
touch	left/right	2	Touch gesture to the left/right side.
	up/down	2	Up- and downwards touch gesture.
	tap	1	Single tap gesture on the screen surface.
mid-air	left/right	2	Mid-air gesture to the left/right.
	up/down	2	Up- and downwards mid-air gesture.
	front	1	Mid-air gesture to the front.
eye gaze	linear diagonal	4	Diagonal eye movements (all four directions).
	circular CW/CCW	2	Clockwise and counter-clockwise circular eye movements.
	zigzag	2	Vertical/horizontal zigzag eye movements.

Table 1: Our set of gestures is based on the work by Khamis et al. [32]. We cover different touch and mid-air gestures and eye gaze movements (overall 18 unique gestures). We studied 5 gestures for touch and mid-air, and 8 variants for eye gaze gestures.

that we slightly modified based on the purpose of our research. We calculated the joint angles of the arms using Unity's Animation Rigging package and Inverse Kinematics (IK) [30] to increase realism and link the avatar's hands to the avatar's body. While this means we still do not rely on any additional high-end tracking systems (except two additional HTC VIVE trackers), implementing the highly realistic VR avatar requires significantly more expertise and effort compared to the implementation of the abstract avatar that does not require any inverse kinematics calculations nor additional HTC VIVE trackers and is based on simple 3D shapes (i.e., a cube for the head and three cuboids that are merged together for the body). We used the same hand asset (i.e., low polygon hands) for all avatars. We aimed to compare how additional body elements (e.g., arms) impact identifying users' interaction, rather than investigating the impact of different hand assets (and their mesh) on observations. Furthermore, at the point where we introduce different hand assets, we also introduce an additional confounding variable. It is also worth to note that the required implementation expertise and hardware for the finger tracking remains the same for all avatar designs, independent of the used hand asset.

For the VR headset and to record the interactions in the virtual environment, we used the Tobii HTC VIVE [78], which we connected to a VR-ready laptop *(Razer Blade 15, NVIDIA GeForce RTX 2080*), and OBS [71]. The interactions in the real world were recorded with a NIKON D5300 single-lens reflex camera.

4 USER STUDY

4.1 Methodology and Study Design

We used Qualtrics [59], an online survey tool that can be accessed via web browsers, and Prolific [57] to deploy the study online and evaluate the impact of the avatar representation on bystanders' interaction identification performance. Prolific is an established platform for online subject recruitment for scientific purposes and regularly used for HCI research (e.g., [2, 40]).

Our study was a within-subject experiment with two factors. The first was the input method, which had three levels: touch, mid-air, and eye gaze. The second was the avatar that represents the user, which had four levels: (1) human in the real world (baseline); (2) an avatar that shows eyes and hands only, similar to avatars used in social virtual rooms (e.g., Mozilla Hub [47]); (3) an avatar which included a realistic virtual body and head because previous work found that gaze is closely associated to head movements under natural conditions [5, 6, 67] and could thus affect bystanders when observing interactions; and (4) a full-body avatar where virtual eyes and hands are attached to a highly realistic avatar.

4.2 Procedure and Task

After obtaining informed consent, we collected demographics (e.g., age, gender, prior experience with virtual environments) and followed with explainer videos of all three input methods and all 18 gestures to introduce participants to the different input methods and their task. Figure 2 shows excerpts of the explainer videos for all three conditions. We used the same camera position and angle throughout the study. We added attention check questions to make sure all participants understood (1) the input methods and (2) their task. We then showed participants pre-recorded interactions performed by a human in the real world or one of the three avatars in the virtual environment. Conditions were counterbalanced using a Latin Square. Participants watched the gestures (e.g., left/right/up/down/tap in *touch*) once each before providing a guess. For each observation, we asked participants which gesture they observed. After each avatar, participants had to fill in the NASA-TLX questionnaire [28] to indicate their perceived workload when observing the interactions. The same was repeated for the other input methods. We concluded the study with 5-point Likert scales and a ranking of the avatars in terms of participants' preference when observing them. We also asked participants to justify their ranking and participants had the chance to provide additional feedback. Each participant received £7.5 for their participation. The study has been approved by the University of Glasgow College of Science & Engineering Ethics Committee.

4.3 Results

We recruited 28 participants of which all passed the attention check questions. We had to remove the data of eight participants due to several reasons. Some of them explicitly mentioned that they faced some issues with the video playback, while others provided low-quality feedback throughout the study; indicating that they did not meaningfully participate. The importance of cleaning data and removing low-quality responses to increase the ecological validity has already been discussed in previous work (e.g., [14, 46, 61]). Our

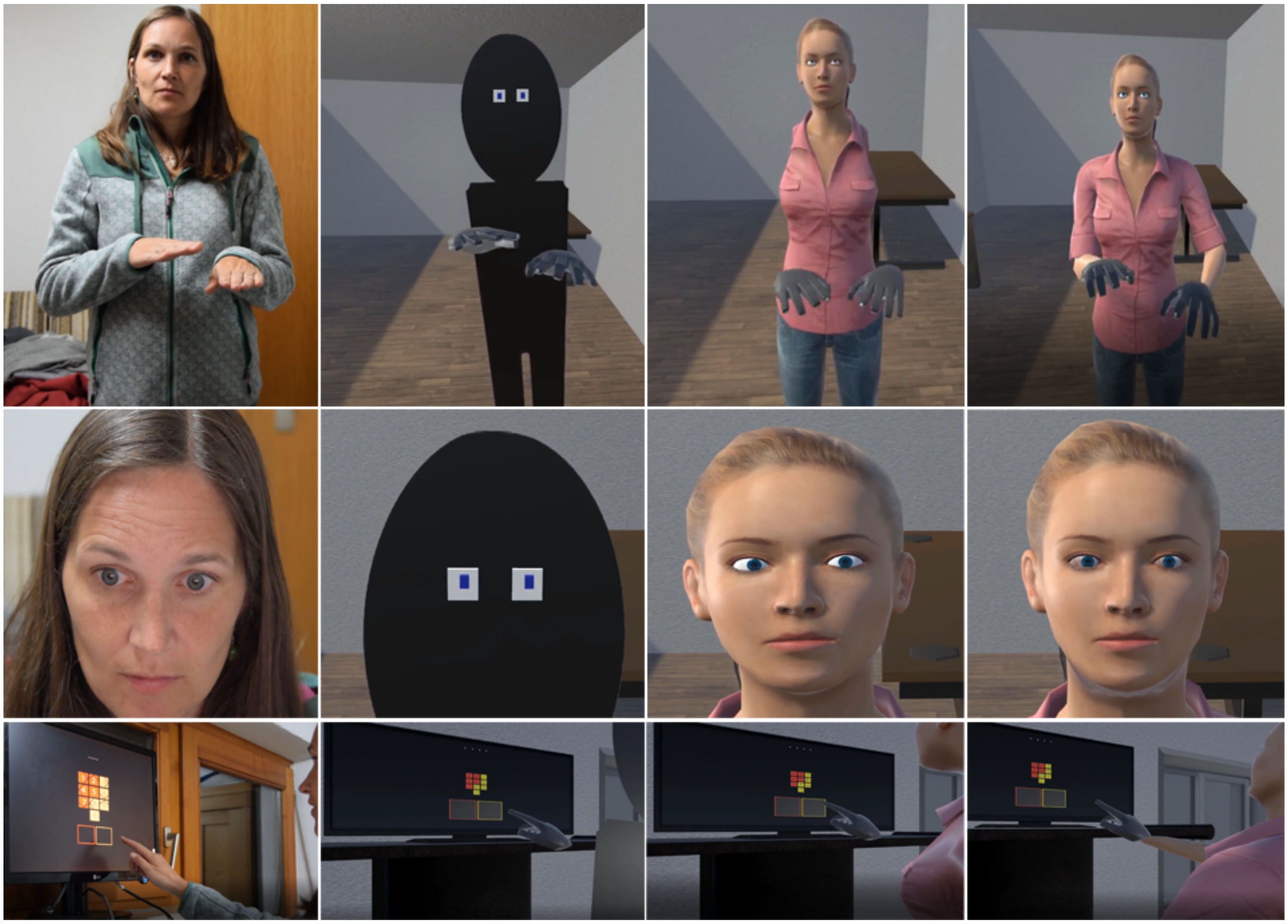

Figure 2: Participants watched pre-recorded videos where a human in the real world or avatars in a virtual environment performed (a) mid-air gestures, (b) eye gaze movements, and (c) touch gestures. In all interaction identification tasks participants had access to the same view as depicted in the figures above.

analysis is therefore based on 20 participants (11 male, 9 female, self-reported) aged between 18 and 54 (M=31.79, SD=10.51). Out of the 20 participants, 19 (95%) mentioned that they have heard about the term "Virtual Reality" before and 9 participants (45.0%) voiced that they had hands-on experience with VR. In the following, we report on the results of participants' (a) interaction identification performance, the number of correctly identified gestures for touch, mid-air, and eye gaze; (b) perceived workload when observing the interactions (NASA-TLX questionnaire [28]); (c) preference of the avatars; and (d) perceived realism of the avatars. Our analysis is based on (5 touch gestures + 5 mid-air gestures + 8 eye gaze gestures) × 20 participants = 360 observations.

4.3.1 Successful Interaction Identifications. A two-way repeated-measures ANOVA was conducted to examine the effect of avatar and input method on participants' number of successful interaction identifications. In other words, the number of gestures participants could successfully guess. Mauchly's test indicated that the assumption for sphericity had been violated for avatar × input method, $\chi^2(20) = 51.244$, $p<0.05$, and input method, $\chi^2(2) = 23.529$, $p<0.05$. Degrees of freedom were corrected using Greenhouse-Geisser correction. There was no statistically significant two-way interaction between avatar and input method on the number of successful interaction identifications, $F_{(3.529,67.043)}=2.068$, $p=0.103$, and no main effect of avatar, $F_{(3,57)}=0.285$, $p=0.836$, and input method, $F_{(1.156,21.973)}=1.255$, $p=0.297$, on the number of successful interaction identifications. The values are M=14.5 (SD=2.56) for observations on the real-world user, M=14.6 (SD=2.80) for the abstract avatar, and M=14.25 (SD=2.97) and M=14.45 (SD=2.65) for the two more realistic avatars. There is no evidence that observations on one avatar were more accurate than on the others and that observations on a human in the real-world were easier/more challenging than on the avatars. Figure 3 summarises the number of successful interaction identifications for each avatar and input method.

4.3.2 Perceived Workload. We were particularly interested in participants' perceived workload (NASA-TLX [28]) when observing the different avatars and the human in the real world. We ran a one-way repeated-measures ANOVA on participants' mean perceived workload when observing the different avatars and the human in the

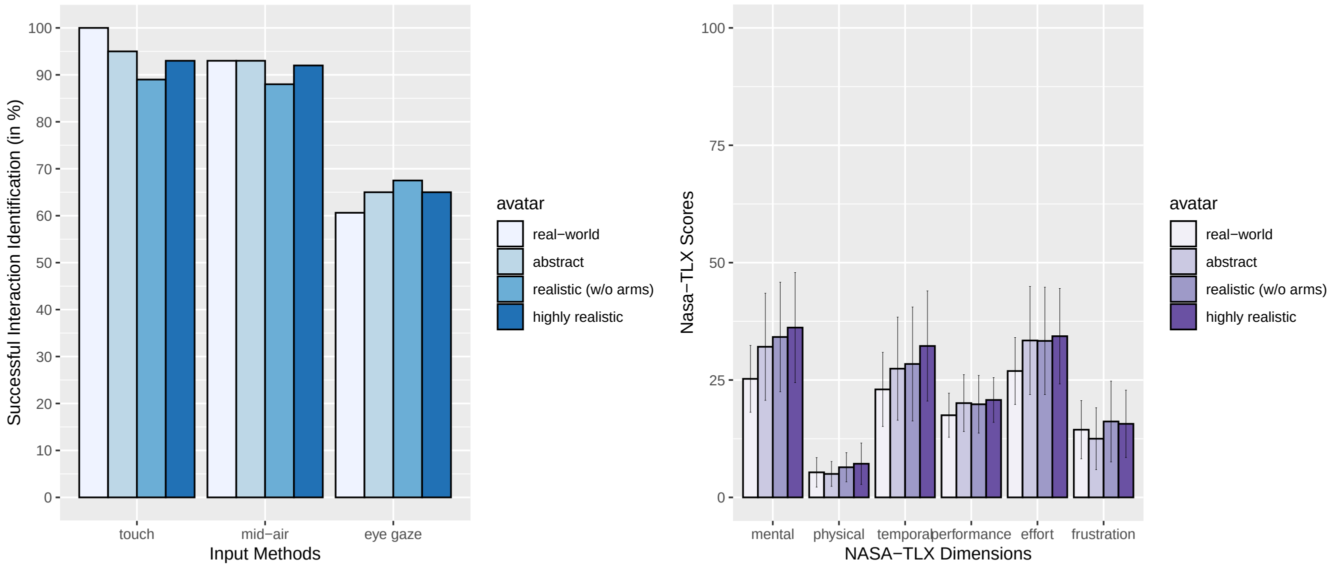

Figure 3: We measured participants' interaction identification performance and their perceived workload (using NASA-TLX [28]) when observing virtual avatars performing input with touch, mid-air, and eye gaze. Our analysis does not indicate that interactions performed by an abstract avatar are easier or more difficult to observe than more realistic avatars and a human in the real world (baseline). The same was found for users' perceived workload when observing the different avatars.

real world. Mauchly's test indicated that the assumption of sphericity had been violated, $\chi^2(5) = 11.712$, $p<0.05$, therefore degrees of freedom were corrected using Greenhouse-Geisser correction. There was no statistically significant effect of avatar on participants' perceived workload, $F_{(2.138,\ 40.626)}=2.922$, $p=0.062$. The mean raw NASA-TLX values were low for all four conditions: M=18.74 (SD=8.97) for observing the human in the real world, M=21.75 (SD=13.73) for the abstract avatar, and M=23.06 (SD=14.64) and M=24.39 (SD=14.12) for the more realistic avatars. There is no evidence that observations on one avatar were more demanding than on the others; however, participants' perceived workload is slightly lower in the real-world condition. Figure 3 shows the mean values of each NASA-TLX dimension and avatar.

4.3.3 Perceived Avatar Realism. We aimed to understand how the different avatars are perceived in terms of realism. We did this through a 5-point Likert scale (i.e., *"The avatar's behaviour matched real-world human movements."*). We ran a Friedman test on the 5-point Likert scale data to investigate participants' perceived avatar realism compared to the human in the real world. The analysis is performed on the level of each input method. For eye gaze, there was a statistically significant difference in perceived avatar realism depending on the type of the user representation, $\chi^2(2) = 9.435$, $p<0.05$. Post hoc analysis with Wilcoxon signed-rank tests was conducted with a Bonferroni correction applied. Both the more realistic avatars were perceived as more realistic than the abstract avatar ($p<0.05$). The median values are 2.5 for the abstract avatar and 4.0 for both the more realistic avatars. For mid-air, there was no statistically significant difference in perceived avatar realism depending on the type of the user representation, $\chi^2(2) = 1.660$, $p=0.436$. The median values are 4.0 for all three avatars. The same was found for touch, $\chi^2(2) = 1.714$, $p=0.424$, with median values of 4.0 for all three avatars. This suggests that the eyes in the more realistic avatars were perceived as more realistic compared to the abstract avatar, but there is no evidence of a significant difference between the avatars when providing input with mid-air and touch.

4.3.4 Avatar Preference. We asked participants to rank the avatars based on their preference of observing them (i.e., *"Which user representation made it easier to observe the different interactions (1=best;4=worst)?"*. Raw scores were multiplied by their weight factor: ×4 for rank 1, ×3 for rank 2, ×2 for rank 3, ×1 for rank 4, and summed up to compute weighted scores. We then calculated the average ranking score for each avatar (i.e. real-world user, abstract avatar, and the two more realistic avatars). The results suggest that the real-world user achieved the highest score (avg ranking score=68.33) next to the highly realistic full-body avatar with an average ranking score of 52. The other more realistic avatar (w/o arms) and the abstract avatar were ranked similarly with an average ranking score of 40.67 and 43.33. Participants' avatar preference is also reflected in the qualitative feedback that we present below.

4.3.5 Qualitative Feedback. We applied a high-level qualitative analysis to spot the main themes in users' qualitative feedback. The data set was fairly simplistic and small (1-2 sentences for each question), so one author did all the qualitative data analysis. We use participant numbers *(P1 to P20)* to ensure anonymity while presenting rich raw data from our participants.

When observing the mid-air interactions, there was a general consensus that the human in the real world was easiest to observe. Some participants mentioned that they needed arms for clarity (e.g., *"the real-world was easiest for me followed by the [most realistic* *one] which is closest to real-world. I think this is because of the fully developed arm extension"*, P20). Interestingly, the need for arms was mentioned in combination with the more realistic avatar, but not with the abstract avatar. This suggests that participants did not necessarily expect to see arms in the abstract avatar condition, but in the more realistic avatar designs. The same was mentioned in touch. Some participants mentioned that all avatars were perceived as easy to observe (e.g., *"they were all easy to identify so any would be fine"*, P1), while one participant, P3, mentioned that *"the lack of arms in the mid-fidelity avatar was off putting"*. Contrary to P3's comment, one participant, P6, voiced that the *"low-fidelity avatar was more obvious and less distracting as it's so basic"*. However, the general consensus was that the attached arms helped our participants identifying the different gestures. One participant mentioned that the movements in the real world were smoother and that the virtual hands of the avatars did not come with visual physical feedback: *"the physical feedback of the human hand touching the screen helped me see exactly what was happening. if the 3D finger bent back a little as it touched the screen (even if it was in an unrealistic uncanny way) then maybe that would help", P9.* In eye gaze, participants mentioned that the human in the real world slightly moved her head when performing the eye gaze gestures, whereas this seemed to be less present in the avatars. Other participants mentioned that observing the user's real world eye movements was easier compared to the avatars. One participant explained this around the fact that he is used to seeing eye movements from real humans rather than from avatars: *"the real world was the clearest to me. I think because I am most used to seeing [eye] movements from real humans."*, P20. Another participant, P15, mentioned that the real-world user blinked, which made it harder for him to observe eye movements in the real-world condition.

From the overall qualitative feedback we notice that the comparisons mostly happened between the real-world user and the avatars. Surprisingly little differences between the different avatars were mentioned. However, one point many participants addressed was the lack of arms in one of the more realistic avatars.

5 DISCUSSION

Our user study suggests that even an abstract avatar's interactions are distinguishable by bystanders to the same extent as interactions performed by a highly realistic avatar and a human in the real world (Section 4.3.1 and Figure 3). There is also no evidence that observing one of the avatars leads to higher perceived workload as indicated by the NASA-TLX questionnaire [28] (Section 4.3.2 and Figure 3). However, it has to be noted that participants' perceived avatar realism, their avatar preference when observing the interactions, and the qualitative feedback suggest that the more realistic avatars were preferred over the abstract avatar. Below, we discuss our findings' implications in more detail.

5.1 So What? Which Avatar to Use?

The answer is: *It depends.* It has always been argued that there is no "swiss-knife" evaluation and that the research questions should inform the choice of the research method [26, 66]. In a similar vein, we argue that it is important to consider the research questions and the aim of the research when making a decision about the avatar design. For example, prior research showed that the avatar realism plays an important role in social settings. There is a large body of work that showed that the avatar fidelity can have a significant impact on social interactions (e.g., [16]) and that more realistic avatars evoke stronger acceptance in terms of virtual body ownership [36]. It is important to note that the gained benefits of using abstract avatars might not outweigh the advantages of more realistic avatars in settings where social interactions and virtual body ownership play an important role (e.g., in social VR [63, 88]). However, as showed in our work, abstract avatars can already provide researchers with valid research findings when it comes to bystanders' interaction identification performance. This finding can be particularly beneficial for situations where, for example, researchers evaluate their systems against observations in virtual environments. VR comes with many advantages and ground-breaking opportunities (e.g., using VR for education purposes [65], for socialising [47], or for the health sector [20]), but there is also a need to secure users' interactions from a security and privacy perspective. For example, what we know from the real world when another human is looking over someone's shoulder, also defined as *"shoulder-surfing"* [12], could happen in a similar way in virtual environments. Researchers recently looked into adapting established real-world authentication schemes for VR [19, 51] or proposed novel VR authentication schemes [18, 43]. Being able to use abstract avatars instead of more realistic avatars can be particularly helpful for such *"shoulder-surfing"* investigations in VR: they likely reduce the required effort, expertise, and amount of additional hardware that is required.

Our **key message** here is that there is **great potential for abstract avatar designs in research where social factors play only a minor role (or can even be ignored)**. For example, when investigating users' identification performance when observing gestures performed by a virtual avatar, or when pointing or other gestures are fundamental components of an investigation (e.g., in the context of collaborative assembly tasks on shared surfaces [38]). In such research contexts, more abstract avatars can already be sufficient for researchers' investigation. Below, we discuss studying VR security systems as one potential application area for which abstract avatar designs can be particularly helpful.

5.2 Application: Studying VR Security Systems

There is a large body of work which studied VR security systems such as PIN/pattern authentication systems [19, 51] or 3D authentication systems [18, 43]. The aim of these systems is to protect VR users from observations by non-VR users with whom they often share the same physical space (e.g., see also [27]). In many of these systems, the resistance to observations, where a bystander attempts to identify the user's secret input through observations, is key to understanding the system's security. However, there is surprisingly little research that focused on observations that happen in the virtual environment rather than in the real world. With the works mentioned above that focus mainly on observations from a non-VR user, it remains unclear whether the proposed systems are secure in a shared virtual environment with co-located VR users. Protecting users' privacy and security against observations within virtual environments seems to be an important next step as the surge of VR applications and the availability of high-end untethered VR headsets (e.g., the Oculus Quest 2) have lead to a notable increase of users' interest in collaborative and social virtual environments (e.g., to participate in workshops and conferences [47, 88]). While the virtual environment could freeze (or even hide) the user's VR avatar for the duration of sensitive input in such social virtual environments (e.g., when entering a PIN), doing this will likely affect other VR users' experience and co-presence. Although shedding light on this requires additional investigations, we argue that it is important to protect (rather than temporarily freeze/remove) users in virtual environments. Abstract avatars as used in our work could already enable security and privacy researchers to quickly evaluate their VR system's resistance against in-VR observations without spending significant time and effort on the avatar design, which in return makes leveraging VR for human-centred studies more attractive and more accessible to the broader research community.

5.3 Main Takeaways

There are three main takeaways from this work that we would like to emphasise.

(1) It is inevitable to guide the avatar design by the research questions and the overall aims of the planned research.
(2) Depending on the research questions, one or the other avatar design (i.e. abstract or realistic) is to be preferred and there is no *"all-in-one"* avatar that fits the purpose of all investigations.
(3) In situations where the research focus is on, for example, identifying different gestures performed by an avatar (e.g., left/right/up/down/front mid-air gestures), an abstract avatar design can already support researchers in answering their research questions without requiring them to rely on investigations with more realistic avatars.

6 LIMITATIONS AND FUTURE WORK

There were some decisions we made that are worth discussing. The covered input methods (touch, mid-air, eye gaze) and our studied gesture set of overall 18 gestures clearly do not cover the entire spectrum of different input methods and gestures. Although we selected those that are frequently used in VR, we cannot claim that our findings will generalise to input methods or gestures beyond the ones studied in our work. Future work could extend over our findings by studying other gestures (e.g., user-defined touch gestures) or other input methods (e.g., head-based interaction [58]). We also used pre-defined observation angles throughout the study and did not introduce any potential distractions (e.g., additional bystanders, or in general more vivid environments), see Figure 2. While such pre-defined observation angles and simulating a so called "best-case" scenario for observers are commonly used when studying observations on user inputs (e.g., [4, 10]), different perspectives or even providing users with the opportunity to define their own observation angle might impact our findings. We encourage future work to investigate a) how a scenario where observers can change their position (i.e., distance and angle) impacts identifying interactions performed by avatars and b) to what extent more vivid contexts (e.g., multiple avatars) impact bystanders' observation performance. For example, at the point where some avatar body parts are hidden from direct eye gaze (e.g., the avatar's body covers their hands), corresponding arm movements could support observers in identifying gestures. Furthermore, although Thornton et al. [79] argued that the ability to process human action remains functional over a large range of distances, their reported human action identification accuracy indeed decreased when actors were positioned at a larger distance. As a result, it could be interesting to also incorporate such extreme distances as studied by Thornton et al. [79] in a follow-up work and shed further light on the impact of distance and avatar fidelity on an observers interaction identification performance. However, it is important to keep in mind that such extreme distances as studied by Thornton et al. [79] (up to 1000 m) are rather unlikely to be experienced in virtual environments. Finally, although we used a highly realistic avatar by Microsoft Research [24], we did not consider blinking eyes and/or facial expressions in this work. Future work could investigate to what extent such additional factors (e.g., facial expressions captured through VIVE's Facial Tracker [85]) affect bystanders' interaction identification performance and perceived workload when observing different avatar designs.

7 CONCLUSION

Through an online user study (N=28), we explored the impact of different avatars (one abstract avatar and two more realistic avatars) on bystanders' interaction identification performance when observing touch, mid-air, and eye gaze gestures. The difference of participants' performance and perceived workload when observing a user in the real world, an abstract avatar, and two more realistic avatars was found to be negligible. However, it is important to note that more realistic avatars are required in situations where, for example, realism and other social factors play an essential role. That being said, if the research focus is solely on the interaction, abstract avatars can already represent users in virtual environments; therefore, researchers do not necessarily need to rely on resource-heavy and expensive avatar implementations that often require additional hardware (e.g., motion tracking systems).

ACKNOWLEDGMENTS

We thank all participants for taking part in our research. We also thank all reviewers whose comments significantly improved the paper. This publication was supported by the University of Edinburgh and the University of Glasgow jointly funded PhD studentships. This work was also partially supported by the Royal Society of Edinburgh (award number #65040) and an EPSRC New Investigator Award (EP/V008870/1).

REFERENCES

[1] Yasmeen Abdrabou, Mohamed Khamis, Rana Mohamed Eisa, Sherif Ismail, and Amrl Elmougy. 2019. Just Gaze and Wave: Exploring the Use of Gaze and Gestures for Shoulder-Surfing Resilient Authentication. In *Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications* (Denver, Colorado) *(ETRA '19)*. Association for Computing Machinery, New York, NY, USA, Article 29, 10 pages. https://doi.org/10.1145/3314111.3319837

[2] Sara Albakry, Kami Vaniea, and Maria K. Wolters. 2020. What is This URL's Destination? Empirical Evaluation of Users' URL Reading. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3313831.3376168

[3] R. Atienza, R. Blonna, M. I. Saludares, J. Casimiro, and V. Fuentes. 2016. Interaction techniques using head gaze for virtual reality. In *2016 IEEE Region 10 Symposium (TENSYMP)*. 110–114. https://doi.org/10.1109/TENCONSpring.2016.7519387

[4] Adam J. Aviv, John T. Davin, Flynn Wolf, and Ravi Kuber. 2017. Towards Baselines for Shoulder Surfing on Mobile Authentication. In *Proceedings of the 33rd Annual Computer Security Applications Conference* (Orlando, FL, USA) *(ACSAC 2017)*. Association for Computing Machinery, New York, NY, USA, 486–498. https://doi.org/10.1145/3134600.3134609

[5] B. Biguer, M. Jeannerod, and C. Prablanc. 1982. The coordination of eye, head, and arm movements during reaching at a single visual target. *Experimental Brain Research* 46, 2 (01 May 1982), 301–304. https://doi.org/10.1007/BF00237188

[6] Emilio Bizzi, Ronald E. Kalil, and Vincenzo Tagliasco. 1971. Eye-Head Coordination in Monkeys: Evidence for Centrally Patterned Organization. *Science* 173, 3995 (1971), 452–454. https://doi.org/10.1126/science.173.3995.452

[7] Doug Bowman, Chadwick Wingrave, Joshua Campbell, and Vinh Ly. 2001. Using pinch gloves (tm) for both natural and abstract interaction techniques in virtual environments. (2001).

[8] Juan Casanueva and Edwin Blake. 2001. The effects of avatars on co-presence in a collaborative virtual environment. (2001).

[9] Gabrielle Darville, Charkarra Anderson – Lewis, Michael Stellefson, Yu-Hao Lee, Jann MacInnes, Jr. R. Morgan Pigg, Juan E. Gilbert, and Sanethia Thomas. 2018. Customization of Avatars in a HPV Digital Gaming Intervention for College-Age Males: An Experimental Study. *Simulation & Gaming* 49, 5 (2018), 515–537. https://doi.org/10.1177/1046878118799472

[10] Alexander De Luca, Emanuel von Zezschwitz, Ngo Dieu Huong Nguyen, Max-Emanuel Maurer, Elisa Rubegni, Marcello Paolo Scipioni, and Marc Langheinrich. 2013. *Back-of-Device Authentication on Smartphones*. Association for Computing Machinery, New York, NY, USA, 2389–2398. https://doi.org/10.1145/2470654.2481330

[11] José Mario De Martino, Ivani Rodrigues Silva, Carmen Zink Bolognini, Paula Dornhofer Paro Costa, Kate Mamhy Oliveira Kumada, Luis Cláudius Coradine, Patrick Henrique da Silva Brito, Wanessa Machado do Amaral, Ângelo Brandão Benetti, Enzo Telles Poeta, et al. 2017. Signing avatars: making education more inclusive. *Universal Access in the Information Society* 16, 3 (2017), 793–808.

[12] Malin Eiband, Mohamed Khamis, Emanuel von Zezschwitz, Heinrich Hussmann, and Florian Alt. 2017. *Understanding Shoulder Surfing in the Wild: Stories from Users and Observers*. Association for Computing Machinery, New York, NY, USA, 4254–4265. https://doi.org/10.1145/3025453.3025636

[13] Laura Fademrecht, Isabelle Bülthoff, and Stephan de la Rosa. 2016. Action recognition in the visual periphery. *Journal of Vision* 16, 3 (02 2016), 33–33. https://doi.org/10.1167/16.3.33 arXiv:https://arvojournals.org/arvo/content_public/journal/jov/934914/i1534-7362-16-3-33.pdf

[14] Sascha Fahl, Marian Harbach, Yasemin Acar, and Matthew Smith. 2013. On the Ecological Validity of a Password Study. In *Proceedings of the Ninth Symposium on Usable Privacy and Security* (Newcastle, United Kingdom) *(SOUPS '13)*. Association for Computing Machinery, New York, NY, USA, Article 13, 13 pages. https://doi.org/10.1145/2501604.2501617

[15] Guo Freeman, Samaneh Zamanifard, Divine Maloney, and Alexandra Adkins. 2020. My Body, My Avatar: How People Perceive Their Avatars in Social Virtual Reality. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI EA '20)*. Association for Computing Machinery, New York, NY, USA, 1–8. https://doi.org/10.1145/3334480.3382923

[16] Maia Garau, Mel Slater, Vinoba Vinayagamoorthy, Andrea Brogni, Anthony Steed, and M. Angela Sasse. 2003. The Impact of Avatar Realism and Eye Gaze Control on Perceived Quality of Communication in a Shared Immersive Virtual Environment. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Ft. Lauderdale, Florida, USA) *(CHI '03)*. Association for Computing Machinery, New York, NY, USA, 529–536. https://doi.org/10.1145/642611.642703

[17] Dariu M Gavrila. 1999. The visual analysis of human movement: A survey. *Computer vision and image understanding* 73, 1 (1999), 82–98.

[18] C. George, M. Khamis, D. Buschek, and H. Hussmann. 2019. Investigating the Third Dimension for Authentication in Immersive Virtual Reality and in the Real World. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 277–285. https://doi.org/10.1109/VR.2019.8797862

[19] Ceenu George, Mohamed Khamis, Emanuel von Zezschwitz, Marinus Burger, Henri Schmidt, Florian Alt, and Heinrich Hussmann. 2017. Seamless and Secure VR: Adapting and Evaluating Established Authentication Systems for Virtual Reality. In *Proceedings of the Network and Distributed System Security Symposium (NDSS 2017)* (San Diego, CA, USA) *(USEC '17)*. NDSS. https://doi.org/10.14722/usec.2017.23028

[20] Kalman Glantz, Nathaniel I Durlach, Rosalind C Barnett, and Walter A Aviles. 1997. Virtual reality (VR) and psychotherapy: Opportunities and challenges. *Presence: Teleoperators & Virtual Environments* 6, 1 (1997), 87–105.

[21] B. Gleeson, K. MacLean, A. Haddadi, E. Croft, and J. Alcazar. 2013. Gestures for industry Intuitive human-robot communication from human observation. In *2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. 349–356. https://doi.org/10.1109/HRI.2013.6483609

[22] Mar Gonzalez-Franco, Seifi Hasti, and Luca Massimiliano Di. 2021. Locomotion Vault. https://locomotionvault.github.io/ accessed 15 January 2021.

[23] Mar Gonzalez Franco, Eyal Ofek, Ye Pan, Angus Antley, Anthony Steed, Bernhard Spanlang, Antonella Maselli, Domna Banakou, Nuria Pelechano, Sergio Orts-Escolano, Veronica Orvalho, Laura Trutoiu, Markus Wojcik, Maria V. Sanchez-Vives, Jeremy Bailenson, Mel Slater, and Jaron Lanier. 2020. The Rocketbox library and the utility of freely available rigged avatars. *Frontiers in Virtual Reality* (November 2020). https://www.microsoft.com/en-us/research/publication/the-rocketbox-library-and-the-utility-of-freely-available-rigged-avatars-for-procedural-animation-of-virtual-humans-and-embodiment/ TECHNOLOGY AND CODE ARTICLE Front. Virtual Real. | frvir.2020.561558.

[24] Mar Gonzalez-Franco, Markus Wojcik, Eyal Ofek, Anthony Steed, and Dave Garagan. 2020. Microsoft Rocketbox. https://github.com/microsoft/Microsoft-Rocketbox accessed 15 September 2020.

[25] M. González-Franco, D. Pérez-Marcos, B. Spanlang, and M. Slater. 2010. The contribution of real-time mirror reflections of motor actions on virtual body ownership in an immersive virtual environment. In *2010 IEEE Virtual Reality Conference (VR)*. 111–114. https://doi.org/10.1109/VR.2010.5444805

[26] Saul Greenberg and Bill Buxton. 2008. Usability Evaluation Considered Harmful (Some of the Time). In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Florence, Italy) *(CHI '08)*. Association for Computing Machinery, New York, NY, USA, 111–120. https://doi.org/10.1145/1357054.1357074

[27] Jan Gugenheimer, Evgeny Stemasov, Julian Frommel, and Enrico Rukzio. 2017. *ShareVR: Enabling Co-Located Experiences for Virtual Reality between HMD and Non-HMD Users*. Association for Computing Machinery, New York, NY, USA, 4021–4033. https://doi.org/10.1145/3025453.3025683

[28] SG Hart and LE Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Human mental workload*. http://humanfactors.arc.nasa.gov/groups/TLX/downloads/NASA-TLXChapter.pdf

[29] Paul Heidicker, Eike Langbehn, and Frank Steinicke. 2017. Influence of avatar appearance on presence in social VR. In *2017 IEEE Symposium on 3D User Interfaces (3DUI)*. 233–234. https://doi.org/10.1109/3DUI.2017.7893357

[30] Dave Hunt and Yang-Hai Eakes. 2020. *Introducing the Animation Rigging preview package for Unity 2019.1*. https://blogs.unity3d.com/2019/05/14/introducing-the-animation-rigging-preview-package-for-unity-2019-1/ accessed 05 September 2020.

[31] Dongsik Jo, Ki-Hong Kim, and Gerard Jounghyun Kim. 2017. Effects of avatar and background types on users' co-presence and trust for mixed reality-based teleconference systems. In *Proceedings the 30th Conference on Computer Animation and Social Agents*. 27–36.

[32] Mohamed Khamis, Ludwig Trotter, Ville Mäkelä, Emanuel von Zezschwitz, Jens Le, Andreas Bulling, and Florian Alt. 2018. CueAuth: Comparing Touch, Mid-Air Gestures, and Gaze for Cue-Based Authentication on Situated Displays. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 4, Article 174 (Dec. 2018), 22 pages. https://doi.org/10.1145/3287052

[33] Rabia Khan and Antonella De Angeli. 2009. The attractiveness stereotype in the evaluation of embodied conversational agents. In *IFIP Conference on Human-Computer Interaction*. Springer, 85–97.

[34] Pascal Knierim, Valentin Schwind, Anna Maria Feit, Florian Nieuwenhuizen, and Niels Henze. 2018. Physical Keyboards in Virtual Reality: Analysis of Typing Performance and Effects of Avatar Hands. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–9. https://doi.org/10.1145/3173574.3173919

[35] Olivia Morgan Lapenta, Ana Paula Xavier, Sônia Cavalcanti Côrrea, and Paulo Sérgio Boggio. 2017. Human biological and nonbiological point-light movements: Creation and validation of the dataset. *Behavior research methods* 49, 6 (2017), 2083–2092.

[36] Marc Erich Latoschik, Daniel Roth, Dominik Gall, Jascha Achenbach, Thomas Waltemate, and Mario Botsch. 2017. The Effect of Avatar Realism in Immersive Social Virtual Realities. In *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology* (Gothenburg, Sweden) *(VRST '17)*. Association for Computing Machinery, New York, NY, USA, Article 39, 10 pages. https://doi.org/10.1145/3139131.3139156

[37] Sohye Lim and Byron Reeves. 2010. Computer agents versus avatars: Responses to interactive game characters controlled by a computer or other player. *International Journal of Human-Computer Studies* 68, 1 (2010), 57 – 68. https://doi.org/10.1016/j.ijhcs.2009.09.008

[38] Christian Mai, Sarah Aragon Bartsch, and Lea Rieger. 2018. Evaluating Shared Surfaces for Co-Located Mixed-Presence Collaboration. In *Proceedings of the 17th International Conference on Mobile and Ubiquitous Multimedia* (Cairo, Egypt) *(MUM 2018)*. Association for Computing Machinery, New York, NY, USA, 1–5. https://doi.org/10.1145/3282894.3282910

[39] Ville Mäkelä, Sheikh Radiah Rahim Rivu, Saleh Alsherif, Mohamed Khamis, Chong Xiao, Lisa Marianne Borchert, Albrecht Schmidt, and Florian Alt. 2020. Virtual Field Studies: Conducting Studies on Public Displays in Virtual Reality.. In *Proceedings of the 38th Annual ACM Conference on Human Factors in Computing Systems* (Honolulu, Hawaii, USA) *(CHI '20)*. ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3313831.3376796

[40] Diogo Marques, Tiago Guerreiro, Luis Carriço, Ivan Beschastnikh, and Konstantin Beznosov. 2019. Vulnerability & Blame: Making Sense of Unauthorized Access to Smartphones. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3290605.3300819

[41] Florian Mathis, Kami Vaniea, and Mohamed Khamis. 2021. RepliCueAuth: Validating the Use of a lab-based Virtual Reality Setup for Evaluating Authentication Systems. In *Proceedings of the 39th Annual ACM Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. ACM, New York, NY, USA, 18 pages. https://doi.org/10.1145/3411764.3445478

[42] Florian Mathis, John Williamson, Kami Vaniea, and Mohamed Khamis. 2020. RubikAuth: Fast and Secure Authentication in Virtual Reality. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI EA '20)*. Association for Computing Machinery, New York, NY, USA, 1–9. https://doi.org/10.1145/3334480.3382827

[43] Florian Mathis, John Williamson, Kami Vaniea, and Mohamed Khamis. 2021. Fast and Secure Authentication in Virtual Reality using Coordinated 3D Manipulation and Pointing. *ACM Transactions on Computer-Human Interaction (ToCHI)* (Jan. 2021). https://doi.org/10.1145/3428121

[44] Florian Mathis, Xuesong Zhang, Joseph O'Hagan, Daniel Medeiros, Pejman Saeghe, Mark McGill, Stephen Brewster, and Mohamed Khamis. 2021. Remote XR Studies: The Golden Future of HCI Research?. In *Proceedings of the CHI 2021 Workshop on XR Remote Research*. http://www.mat.qmul.ac.uk/xr-chi-2021/

[45] Nicholas Merola and Jorge Peña. 2009. The effects of avatar appearance in virtual worlds. *Journal for Virtual Worlds Research* 2, 5 (2009).

[46] Aske Mottelson and Kasper Hornbæk. 2017. Virtual Reality Studies Outside the Laboratory. In *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology* (Gothenburg, Sweden) *(VRST '17)*. Association for Computing Machinery, New York, NY, USA, Article 9, 10 pages. https://doi.org/10.1145/3139131.3139141

[47] Mozilla. 2020. *Mozilla Labs - Hubs by Mozilla.* https://hubs.mozilla.com/ accessed 15 September 2020.

[48] Kristine L. Nowak and Christian Rauh. 2008. Choose your "buddy icon" carefully: The influence of avatar androgyny, anthropomorphism and credibility in online interactions. *Computers in Human Behavior* 24, 4 (2008), 1473 – 1493. https://doi.org/10.1016/j.chb.2007.05.005 Including the Special Issue: Integration of Human Factors in Networked Computing.

[49] Kristine L. Nowak and Christian Rauh. 2017. The Influence of the Avatar on Online Perceptions of Anthropomorphism, Androgyny, Credibility, Homophily, and Attraction. *Journal of Computer-Mediated Communication* 11, 1 (07 2017), 153–178. https://doi.org/10.1111/j.1083-6101.2006.tb00308.x

[50] N. Ogawa, T. Narumi, and M. Hirose. 2019. Virtual Hand Realism Affects Object Size Perception in Body-Based Scaling. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 519–528. https://doi.org/10.1109/VR.2019.8798040

[51] Ilesanmi Olade, Hai-Ning Liang, Charles Fleming, and Christopher Champion. 2020. Exploring the Vulnerabilities and Advantages of SWIPE or Pattern Authentication in Virtual Reality (VR). In *Proceedings of the 2020 4th International Conference on Virtual and Augmented Reality Simulations* (Sydney, NSW, Australia) *(ICVARS 2020)*. Association for Computing Machinery, New York, NY, USA, 45–52. https://doi.org/10.1145/3385378.3385385

[52] World Health Organization. 2020. *Coronavirus disease (COVID-19) pandemic.* https://www.who.int/emergencies/diseases/novel-coronavirus-2019 accessed 15 September 2020.

[53] Ye Pan and Anthony Steed. 2019. Avatar Type Affects Performance of Cognitive Tasks in Virtual Reality. In *25th ACM Symposium on Virtual Reality Software and Technology* (Parramatta, NSW, Australia) *(VRST '19)*. Association for Computing Machinery, New York, NY, USA, Article 6, 4 pages. https://doi.org/10.1145/3359996.3364270

[54] Alexandra Petrakou. 2010. Interacting through avatars: Virtual worlds as a context for online education. *Computers & Education* 54, 4 (2010), 1020–1027.

[55] T. Piumsomboon, G. Lee, R. W. Lindeman, and M. Billinghurst. 2017. Exploring natural eye-gaze-based interaction for immersive virtual reality. In *2017 IEEE Symposium on 3D User Interfaces (3DUI)*. 36–39. https://doi.org/10.1109/3DUI.2017.7893315

[56] Thammathip Piumsomboon, Gun A. Lee, Jonathon D. Hart, Barrett Ens, Robert W. Lindeman, Bruce H. Thomas, and Mark Billinghurst. 2018. Mini-Me: An Adaptive Avatar for Mixed Reality Remote Collaboration. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3173574.3173620

[57] Prolific. 2020. *Quickly find research participants you can trust.* https://www.prolific.co/ accessed 15 September 2020.

[58] Yuan Yuan Qian and Robert J. Teather. 2017. The Eyes Don't Have It: An Empirical Comparison of Head-Based and Eye-Based Selection in Virtual Reality. In *Proceedings of the 5th Symposium on Spatial User Interaction* (Brighton, United Kingdom) *(SUI '17)*. Association for Computing Machinery, New York, NY, USA, 91–98. https://doi.org/10.1145/3131277.3132182

[59] Qualtrics. 2005. *Qualtrics Experience Managements.* https://www.qualtrics.com accessed 15 September 2020.

[60] Jack Ratcliffe, Francesco Soave, Nick Bryan-Kinns, Laurissa Tokarchuk, and Ildar Farkhatdinov. 2021. Extended Reality (XR) Remote Research: a Survey of Drawbacks and Opportunities. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (2021-05-08) *(CHI '21)*. Association for Computing Machinery, New York, New York, USA. https://doi.org/10.1145/3411764.3445170 arXiv:2101.08046 [cs.HC]

[61] Elissa M. Redmiles, Ziyun Zhu, Sean Kross, Dhruv Kuchhal, Tudor Dumitras, and Michelle L. Mazurek. 2018. Asking for a Friend: Evaluating Response Biases in Security User Studies. In *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security* (Toronto, Canada) *(CCS '18)*. Association for Computing Machinery, New York, NY, USA, 1238–1255. https://doi.org/10.1145/3243734.3243740

[62] D. Roth, J. Lugrin, D. Galakhov, A. Hofmann, G. Bente, M. E. Latoschik, and A. Fuhrmann. 2016. Avatar realism and social interaction quality in virtual reality. In *2016 IEEE Virtual Reality (VR)*. 277–278. https://doi.org/10.1109/VR.2016.7504761

[63] David Saffo, Sara Bartolomeo, Caglar Yildirim, and Cody Dunne. 2021. Remote and Collaborative Virtual Reality Experiments via Social VR Platforms. In *Proceedings of the 39th Annual ACM Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. ACM, New York, NY, USA, 15 pages. https://doi.org/10.1145/3411764.3445426

[64] Gian-Luca Savino, Niklas Emanuel, Steven Kowalzik, Felix Kroll, Marvin C. Lange, Matthis Laudan, Rieke Leder, Zhanhua Liang, Dayana Markhabayeva, Martin Schmeißer, Nicolai Schütz, Carolin Stellmacher, Zihe Xu, Kerstin Bub, Thorsten Kluss, Jaime Maldonado, Ernst Kruijff, and Johannes Schöning. 2019. Comparing Pedestrian Navigation Methods in Virtual Reality and Real Life. In *2019 International Conference on Multimodal Interaction* (Suzhou, China) *(ICMI '19)*. Association for Computing Machinery, New York, NY, USA, 16–25. https://doi.org/10.1145/3340555.3353741

[65] Sharad Sharma, Ruth Agada, and Jeff Ruffin. 2013. Virtual reality classroom as an constructivist approach. In *2013 Proceedings of IEEE Southeastcon*. IEEE, 1–5.

[66] Ben Shneiderman, Catherine Plaisant, Maxine Cohen, Steven Jacobs, Niklas Elmqvist, and Nicholas Diakopoulos. 2016. *Designing the User Interface: Strategies for Effective Human-Computer Interaction* (6th ed.). Pearson.

[67] Ludwig Sidenmark and Hans Gellersen. 2019. Eye, Head and Torso Coordination During Gaze Shifts in Virtual Reality. *ACM Trans. Comput.-Hum. Interact.* 27, 1, Article 4 (Dec. 2019), 40 pages. https://doi.org/10.1145/3361218

[68] Mel Slater, John McCarthy, and Francesco Maringelli. 1998. The influence of body movement on subjective presence in virtual environments. *Human factors* 40, 3 (1998), 469–477.

[69] M. Slater, D. . Pertaub, and A. Steed. 1999. Public speaking in virtual reality: facing an audience of avatars. *IEEE Computer Graphics and Applications* 19, 2 (1999), 6–9. https://doi.org/10.1109/38.749116

[70] Mel Slater and Martin Usoh. 1994. Body centred interaction in immersive virtual environments. *Artificial life and virtual reality* 1, 1994 (1994), 125–148.

[71] Open Broadcaster SoftwareB. 2020. *OBS (Open Broadcaster Software) | free and open source software.* https://obsproject.com/de accessed 15 September 2020.

[72] Peng Song, Wooi Boon Goh, William Hutama, Chi-Wing Fu, and Xiaopei Liu. 2012. A Handle Bar Metaphor for Virtual Object Manipulation with Mid-Air Interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA) *(CHI '12)*. Association for Computing Machinery, New York, NY, USA, 1297–1306. https://doi.org/10.1145/2207676.2208585

[73] Misha Sra and Chris Schmandt. 2015. MetaSpace: Full-Body Tracking for Immersive Multiperson Virtual Reality. In *Adjunct Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology* (Daegu, Kyungpook, Republic of Korea) *(UIST '15 Adjunct)*. Association for Computing Machinery, New York, NY, USA, 47–48. https://doi.org/10.1145/2815585.2817802

[74] Richard Stoakley, Matthew J. Conway, and Randy Pausch. 1995. Virtual Reality on a WIM: Interactive Worlds in Miniature. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '95)*. ACM Press/Addison-Wesley Publishing Co., USA, 265–272. https://doi.org/10.1145/223904.223938

[75] Sabine Ströfer, Elze G Ufkes, Merijn Bruijnes, Ellen Giebels, and Matthijs L Noordzij. 2016. Interviewing suspects with avatars: Avatars are more effective when perceived as human. *Frontiers in psychology* 7 (2016), 545.

[76] Hemant Bhaskar Surale, Aakar Gupta, Mark Hancock, and Daniel Vogel. 2019. 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* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3290605.3300243

[77] Hemant Bhaskar Surale, Fabrice Matulic, and Daniel Vogel. 2019. Experimental Analysis of Barehand Mid-Air Mode-Switching Techniques in Virtual Reality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3290605.3300426

[78] Tobii Technology. 2020. *Tobii HTC VIVE Devkit.* https://vr.tobii.com/sdk/develop/unity/getting-started/tobii-htc-dev-kit/ accessed 26 August 2020.

[79] Ian M. Thornton, Zac Wootton, and Pille Pedmanson. 2014. Matching biological motion at extreme distances. *Journal of Vision* 14, 3 (03 2014), 13–13. https://doi.org/10.1167/14.3.13

[80] Tobii. 2020. *Tobii Pro VR Integration.* https://www.tobiipro.com/product-listing/vr-integration/ accessed 26 August 2020.

[81] Jintana Tongpeth, Hui Yun Du, and Robyn A. Clark. 2018. Development and feasibility testing of an avatar-based education application for patients with acute coronary syndrome. *Journal of Clinical Nursing* 27, 19-20 (2018), 3561–3571. https://doi.org/10.1111/jocn.14528

[82] Mélodie Vidal, Andreas Bulling, and Hans Gellersen. 2013. Pursuits: Spontaneous Interaction with Displays Based on Smooth Pursuit Eye Movement and Moving Targets. In *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (Zurich, Switzerland) *(UbiComp '13)*. Association for Computing Machinery, New York, NY, USA, 439–448. https://doi.org/10.1145/2493432.2493477

[83] Vinoba Vinayagamoorthy, Andrea Brogni, Marco Gillies, Mel Slater, and Anthony Steed. 2004. An investigation of presence response across variations in visual realism. In *The 7th Annual International Presence Workshop*. Citeseer, 148–155.

[84] Vinoba Vinayagamoorthy, Anthony Steed, and Mel Slater. 2005. Building characters: Lessons drawn from virtual environments. In *Proceedings of Toward Social Mechanisms of Android Science: A CogSci 2005 Workshop*. COGSCI 2005, 119–126.

[85] HTC VIVE. 2021. VIVE Facial Tracker: Express more than words in VR. https://www.vive.com/us/accessory/facial-tracker/ accessed 13 April 2021.

[86] Alexandra Voit, Sven Mayer, Valentin Schwind, and Niels Henze. 2019. Online, VR, AR, Lab, and In-Situ: Comparison of Research Methods to Evaluate Smart Artifacts. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3290605.3300737

[87] M. Weiß, K. Angerbauer, A. Voit, M. Schwarzl, M. Sedlmair, and S. Mayer. 2020. Revisited: Comparison of Empirical Methods to Evaluate Visualizations Supporting Crafting and Assembly Purposes. *IEEE Transactions on Visualization and Computer Graphics* (2020), 1–1. https://doi.org/10.1109/TVCG.2020.3030400

[88] Julie Williamson, Jie Li, Vinoba Vinayagamoorthy, David A. Shamma, and Pablo Cesar. 2021. Proxemics and Social Interactions in an Instrumented Virtual Reality Workshop. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, 13 pages. https://doi.org/10.1145/3411764.3445729