IADIS International Journal on Computer Science and Information Systems
Vol. 19, No. 2, pp. 102-114
ISSN: 1646-3692

THE ERGONOMICS OF MID-AIR GESTURES IN VIRTUAL REALITY - HAND CHARACTERISTICS, GESTURE RECOGNITION FAILURES AND THE FEELING OF IMMERSION

Lisa Graichen[1] and Matthias Graichen[2]
[1]TU Berlin, 10587 Berlin, Germany
[2]Independent Researcher, Germany

ABSTRACT

Together with new hardware solutions, such as virtual reality (VR) headsets, the use of innovative interaction modes, such as mid-air gestures, is increasing in various areas of research, industry, and everyday life. As such setups are complex and not as well established as traditional haptic or touch-based interfaces, there is a higher risk of users experiencing errors, failures, or technical malfunctions. As for why gesture recognition rates may not be as high as desired, it is not yet clear why this is the case. We conducted a study in a VR context, using an HTC Vive headset and the Leap Motion gesture recognition device. Participants performed basic tasks with a "blocks" application using a pre-defined set of gestures. Afterwards they were asked to rate their levels on trust, acceptance and subjective feeling of immersion. We also measured basic hand parameters. We examined the correlation between hand size and observed detection errors of the gesture recognition device. Moreover, we analyzed the influence of perceived errors on the subjective feeling of immersion. We found no significant correlation between hand measurements and error rates. However, there is some evidence that hand length has some effect, which means smaller hands seem to slightly increase the errors rate in interactions using the gesture recognition device. Perceived errors had a negative impact on the feeling of immersion.

KEYWORDS

User Experience, Virtual Reality, Gestures, Malfunction

1. INTRODUCTION

In recent years, extended reality applications have emerged rapidly, finding their way into users' everyday lives as well as professional applications, science and research. The term "extended reality" (XR) comprises different levels of immersion, including virtual reality (VR), augmented reality (AR) and mixed reality (MR). Although XR systems have existed for decades, recent technological advances have catalyzed the development of various applications and research questions. Consequently, low-cost systems, such as Oculus Rift and HTC Vive, have entered the market, making the VR experience affordable and available to almost everyone.

Typically, XR applications are implemented using head-mounted displays (HMDs). HMD systems usually come with a set of controllers that users can hold in their hands in order to control or manipulate the virtual environment. However, using such additional devices to interact with the virtual environment may not be suitable for several reasons. For example, holding the controllers for an extended period may cause discomfort for the user, some applications require that the user's hands are free, and if the controllers are used by many users and/or frequently changing users, there are additional challenges from a hygiene perspective. As an alternative to these controllers, gesture and voice control be used. However, these interaction types are not yet often used in combination with VR (Rakkolainen *et al.*, 2021). In addition, the use of gesture-based interaction may have some disadvantages in terms of accuracy and reliability of detection, depending on various circumstances, such as lighting conditions or physical characteristics of the user. Therefore, we built a prototype application that can operated using gesture-based interaction (GBI; see Graichen, Graichen and Nudzor, 2023 for details) to investigate the effects of hand characteristics on mid-air gesture recognition.

2. THEORY

In the domain of human–machine interaction, there is a wide range of interaction activities that are referred to as "gestures," beginning with clicking, pointing and moving a traditional computer mouse. In this area, the term "gesture" is most frequently employed to refer to the use of touch screens, with their well-established set of gestures including swiping, zooming and tapping in different contexts and dimensions (Saffer, 2008). However, bodily movements that are performed in the air without making contact with a surface and hence without haptic feedback are also considered as gestures. Such mid-air gestures have applications in vehicles, for interaction with in-vehicle information systems, as well as in smart homes and VR. Even systems that use gaze interaction (where gazes are defined as a directional movement that triggers a predefined action) are referred to as "gesture-based systems" (e.g. Rakkolainen *et al.*, 2021). Finally, use of the classical controllers that usually accompany VR HMDs is referred to as a "gesture-based interaction". In this paper, we use the term "GBI" to denote mid-air gestures that do not require physical contact with a surface (Saffer, 2008).

Of special interest in the area of human machine interaction is the idea of user experience and its implications. It is first necessary to define what user experience (UX) actually refers to, what it consists of and how we can demarcate it from other widespread concepts like usability. According to Hassenzahl & Roto (2007), UX consists of being (be goals) and doing (do goals). Do goals are related to the pragmatic task a person wishes to fulfill with an object or product. Be goals refer to the way one wishes to be the psychological desires behind the pragmatic task. People want to relate to others, have influence, feel competent and so on. Both goals are related, and the term UX connects to both goals, unlike the term usability. From a development point of view, it is important that a user is able to fulfill do-goals with a product. Otherwise, they will not use it. However, it is also important to fulfill be goals so that the user can feel attached to the product. Usability itself, in this sense, has no value other than enabling be goals for the user (Hassenzahl, 2008).

Moreover, Hassenzahl (2008) states that the use of technical devices goes beyond winning time for other things but is intrinsically desirable, as it causes stimulation, experiences and feelings. Following on that, he defines UX as "a momentary, primarily evaluative feeling (good-bad) while interacting with a product or service" and "good UX is the consequence of fulfilling the human needs for autonomy, competency, stimulation (self-oriented), relatedness, and popularity (others-oriented) through interacting with the product or service" (Hassenzahl, 2008). This definition shows that UX is more about humans and their feelings than about the product and its dynamic. It is about the continuing stream of thoughts, feelings and judgments. Thus, to create hedonic quality and good UX, developers must do more than put a product designed for functionality in a beautiful box, but must really put effort into understanding the true needs of the person using it.

Nonetheless, Diefenbach and Hassenzahl (2008) argue that hedonic qualities always need to be justified. Hedonic qualities are important, but if users have to choose, they may rather stick with a product that fulfills more pragmatic functions, even if they would have better appreciated a more hedonic product. Indeed, there are indications that a product's pragmatic quality remains relatively constant over time, while hedonic quality, in the form of stimulation, declines rapidly due to familiarization and identity, and, similarly, the experienced beauty of a product decreases due to comparison with other users and products (Wilamowitz-Moellendorff, Hassenzahl and Platz, 2006).

Forlizzi and Battarbee (2004) developed a theory on the dimensions of the terms interaction and experience. According to the researchers, an interaction between a user and a product can be fluent (i.e., tasks that are highly automated and do not require a high level of attention, such as riding a bicycle), cognitive (i.e., tasks that require the attention of the user and result in changes in the user (gaining knowledge) or context), for example, when people visit foreign countries and interact with products they do not have experience with, or expressive (i.e., adjusting a product to an individual need or user desire, such as customizing a piece of furniture or a car). Experience can also be divided into three types: 1) Experience is the "self-talk" when interacting with products and the adjustment of goals. It has a clear beginning and a clear end and may invoke changes in emotions or behavior in the user. 2) Co-experience is a shared experience in a social context with other persons. Moreover, emotions determine how a person constitutes plans for an interaction situation, how the actual interaction proceeds, and how the person feels about the experience once it is over. 3) Emotions shape the information we spread and the mental model users have about an interaction with a product.

Sharp, Rogers and Preece (2019) state that UX cannot be designed, but designers can only implement features that hopefully create a good UX. They also point out the difference between usability and UX. Usability refers to six goals: Effectiveness of use, efficiency of use, safety of use, good utility, high learnability and good memorability. Usability can be assessed by the extent to which a product improves the performance of its user. By contrast, UX refers to emotions and desires that can be positive (e.g., exciting, enjoyable, cognitively stimulating etc.) or negative (e.g., frustrating, boring, unpleasant etc.). To summarize, usability is more related to objective qualities and UX is more related to subjective qualities.

The evolution of UX over time has been investigated by Marti and Iacono (2016), who states there are four types of UX: 1) anticipated (referring to the expectation users have towards a product before they actually use it), 2) momentary (referring to the user's perception at every moment during usage), 3) episodic (referring to an episode of usage and the subsequent assessment of UX) and 4) remembered (referring to the memory of UX after a longer period of product usage). Thielsch, Engel and Hirschfeld (2015) investigated how expected and perceived UX are connected and found no strong correlation between what users expected before actual use and how they rated a website after usage. When asked about positive hedonic product quality, users seemed to mainly describe imaginative future products. However, when asked about negative hedonic quality, users referred to already existing products (Yogasara *et al.*, 2011).

Obrist *et al.* (2012) analyzed existing theories on UX and found a total of 56 theories that were assigned to nine disciplines: psychology, sociology, philosophy, marketing, art, communication, education, anthropology, and design, and seven categories: human/user (mostly psychological models that are concerned with cognitive processes, emotions, motivations or characteristics of the persons using a product), product/artifact (focusing on the aesthetics or semiotics of a product), relation of user and artifact/environment (correlation between user and product, main focus is on the context of the interaction), UX and its social nature (an interdisciplinary category concerned with social aspects of experience and interaction), design (focusing on aspects of art and design), several aspects of the mentioned topics (frameworks that integrate several disciplines and holistic approaches) and theories addressing human existence (philosophical aspects of the interplay between society and technology).

Finally, there also is an ISO norm defining UX and distinguishing the term from usability. It defines UX as a “user’s perceptions and responses that result from the use and/or anticipated use of a system, product or service” and usability as the “extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use” (ISO, 2019).

Currently, there are few related works that combine GBI and VR. Therefore, research in this field, especially with a focus on user interaction, seems to be relevant. Some studies have been conducted in which users were asked to interact with virtual avatars. In Narayana *et al.*, 2019), the authors built a prototype using an avatar. Users were asked to perform a task with blocks, and it was found that they considered social, deictic and iconic gestures to be important to employ when they could not use speech as a mode of interaction. Furthermore, users considered iconic gestures less important for communication to solve the task when they could use speech interaction. A similar approach is VoxWorld by Pustejovksy *et al.*, 2020), which also used a virtual avatar. Deictic, action and affordance gestures were used together with facial expressions and could be detected and performed by the virtual agent. Rodriguez *et al.*, 2017) implemented a VR system using data gloves for gesture detection. This study investigated how quickly users could adapt to this system. It was found that experienced users adapted quickly, but inexperienced users needed more time to learn.

With regard to technical failures and errors gesture recognition device, it is of particular interest how users react and how such situations affect acceptance and trust in the system. Without acceptance and trust, users will not be willing to use a system. It is known from human-robot interaction that trust is not very stable, but changes over time, and trust can be reduced after users experience system failures (Esterwood and Robert, 2022). However, trust seems to somehow stabilize after longer periods of use (Tolmeijer *et al.*, 2021; Yu *et al.*, 2018) reaching a level that matches the capabilities of the system. Thus, high system performance leads to high trust, and if performance falls below a certain threshold, trust would also decrease. The level of this threshold seems to be individual (Yu *et al.*, 2018). Dorton and Harper; Dorton, Harper and Neville(2022) found that when interacting with AI, users may increase or decrease their level of trust after certain experiences and make adaptations in their workflows based on that by adding or removing certain tasks. Studies show that when confronted with errors in an automated vehicle, users seem to decrease their level of trust, but this effect was of a temporal nature. After perceiving accurate system performance, trust would recover, which is called trust repair (Dorton and Harper, 2022; Kraus *et al.*, 2020; Lee *et al.*, 2021). Mishler and Chen (2023) found similar results in a similar setup, but the results show that trust recovered after an experienced failure, but trust levels measured before the failure were not reached again.

Despite its simplicity and usefulness, it is still a challenge to implement gestures in a way that every potential user with differing hand sizes, muscle tones or movement habits would experience perfect gesture detection. This led us to the question if parameters moderating gesture detection rates can be identified with the goal to mitigate potential issues in the future. In this study, we examine how hand characteristics such as length from wrist to middle finger tip, length from thumb to little finger, and middle finger diameter affect error rates to get an idea of why system errors occur. In addition, we compare error rates between male and female participants to investigate other reasons, such as differences in myotonus. We used a virtual reality setup in which participants were asked to perform basic tasks using mid-air gestures. Although this is well established and works well, it can still lead to errors, either because the system does not detect the hand movement correctly, or the user performs an incorrect gesture or the correct gesture in an imprecise way. We defined an error as any malfunction that did not result in the desired outcome. It is especially interesting how perceived errors influence aspects of user experience and general satisfaction of the users. To approach this question, we analyzed the correlation between the number of perceived errors and the feeling of immersion of the users.

We documented errors for each participant in order to examine the following research questions:

1) How do hand measures affect error rates? Is there a correlation between hand size and the number of errors documented?

2) Is there a difference between male and female participants?

3) Does the perception of errors have a negative influence on aspects of user experience, especially on the feeling of immersion?

3. METHODS

3.1 Design and Independent Variables

We chose a one-way repeated measures design with interaction mode (controller vs. mid-air gestures) as the factor. The controller interaction mode was used because it is the standard device that comes with the HTC Vive to familiarize participants with the VR settings and practice performing tasks in the environment. Using the standard controllers would also be considered gesture-based interaction, but performing mid-air gestures feels more natural to users because no tool needs to be touched, as in human communication, therefore we refer to this interaction mode as "gesture-based interaction".

3.2 Participants

An opportunity sample of 35 participants (21 females and 14 males, mostly students) was selected. The participants' mean age was 25.97 years (SD = 7.41, min = 15, max = 48). Four of them had no prior experience with gesture-based interaction, 31 already had some experience. None reported low interest in gesture-based interaction, 18 reported medium interest, and 17 reported high interest. There were no restrictions regarding VR experience or visual aids, however, it was not possible to wear glasses during the study. This research complied with the tenets of the Declaration of Helsinki, such that informed consent was obtained from each participant.

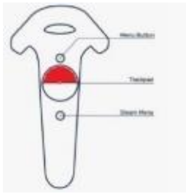
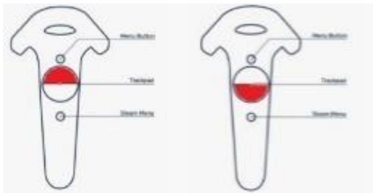
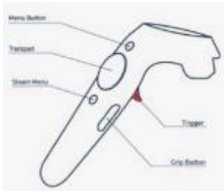
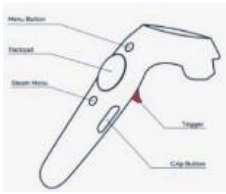
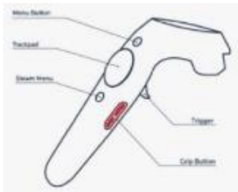
3.3 Facilities and Apparatus

We used an HTC Vive headset equipped with standard controllers to create VR experience. To implement GBI, we used a Leap Motion controller (https://www.ultraleap.com) that was connected to the HTC Vive and mounted on the front of the HTC Vive (see Figure 1). We used a basic application that allows users to perform basic tasks such as creating 3D blocks, resizing them, and grabbing and moving the created objects (see Figure 2). We used Unity to implement the block application. For details on the technical implementation. The experimenter was able to follow all the participants' actions on the connected laptop. We used a video camera to record the performed gestures and the laptop screen to be able to analyze the gesturing behavior and the reactions of the system.

3.4 Interaction Tasks

Six tasks were designed for both interaction types using the controller and the gesture recognition device. These tasks were derived from basic tasks that have the potential to cover a wide range of possible tasks (see Table 1 for details). To ensure that the gesture was appropriate for the selected interaction tasks, a pre-test was conducted with 17 participants using an online questionnaire. Participants were presented with images of different gestures and asked to rate the appropriateness of each gesture for each task. In the present study, we selected six gestures that were rated as highly appropriate for the selected interaction tasks. Table 1 shows the task descriptions and the corresponding gestures/controller actions.

Table 1. Tasks with Corresponding Mid-Air and Controller Gestures

Task	**Gesture**	**Controller**
Create Object		
Change Object Size		
Rotate Object		
Move Object		
Toss Object		
Remove Gravity		N/A

3.5 Procedure

Upon arrival, participants were introduced to the VR device and general usage of the system and devices (see Figure 1). Each participant was given as much time as necessary to learn and practice all interaction tasks. Participants were repeatedly tested on their interaction performance to reduce training effects during the experiment and to reduce the likelihood of making an incorrect gesture during an interaction task. The functionality of gesture recognition was demonstrated on-screen using an online visualization tool included with the gesture recognition device, which illustrated the device's tracking of the participant's fingers and palm. In the first phase, participants used a small application to familiarize themselves with the VR device. In the second phase, participants performed tasks using the controllers as a baseline or gestures using the gesture recognition device. In the third phase, participants used the other interaction mode to initiate tasks. See Table 1 for details on the tasks and gestures. Prior to each experimental trial, participants were instructed to perform the interaction task when they felt safe and comfortable doing so (see Figure 2).

Figure 1. Setup of HTC Vive with Leap Motion mounted to the front

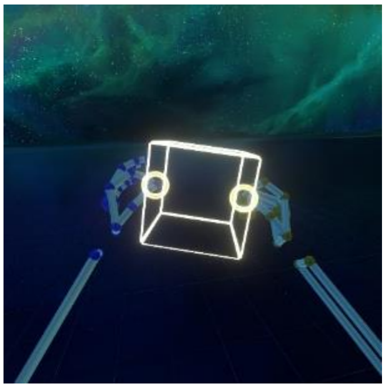


Figure 2. View of the participant in VR

After each phase, participants completed a questionnaire on their subjective impressions related to their most recent interaction. At the end of the experiment, participants completed additional demographic questionnaires, including hand measurements, sense of immersion, and simulator sickness. In total, the experiment took approximately 1.5 hours per participant.

3.6 Dependent Variables

Error rates were documented by the examiner during the study. A video recording was made in case an analysis was necessary after the study.

Hand measures have been documented by the participants with support of the examiner (see Figure 3). User experience, especially the subjective feeling of immersion was measured using the ITC-SOPI by Lessiter *et al.* (2001).

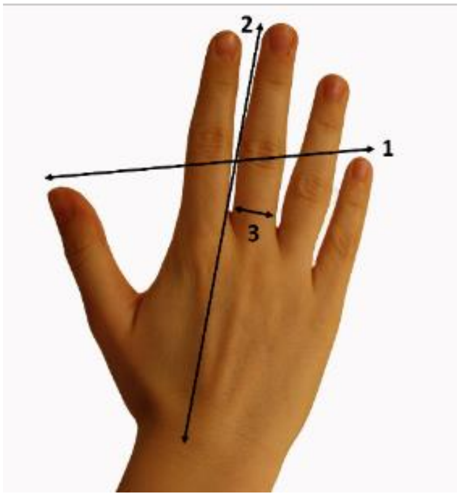


Figure 3. Hand measures used for the analysis: (1) Hand span, (2) hand length, (3) diameter of middle finger

4. RESULTS

One task from one participant had to be excluded from the analysis due to system crash. Since the data were not normally distributed, we calculated Spearman correlations. Correlations were interpreted according to (Cohen, 1988).

For hand measures, we found that females generally had smaller hands than males, as expected. For female hand length M = 17.7 and SD = 0.75, for male hand length M = 19.4 and SD = 1.2. For female hand span M = 18.9 and SD = 1.05 and for male hand span M = 21.1 and SD = 1.35. For female middle finger diameter M = 1.77 and SD = 0.12 and for male middle finger diameter M = 2.04 and SD = 0.23 (see Table 2 and Figures 4a-c).

Table 2. Hand measures for female and male participants

Hand measure	**Female**		**Male**	
Hand length in cm	M = 17.7	SD = 0.75	M = 19.4	SD = 1.2
Hand span in cm	M = 18.9	SD = 1.05	M = 21.1	SD = 1.35
Middle finger diameter in cm	M = 1.77	SD = 0.12	M = 2.04	SD = 0.23

For documented errors, we found M = 0.82 and SD = 1.18 for female participants and M = 0.78 and SD = 1.28 for male participants. We conducted a t-test and found no significant difference, $t(33) = 0.29$, $p = .77$. This shows that there is almost no systematic difference between male and female participants in observed errors. However, it also shows that each participant experienced almost one error per task performed (see Table 3 and Figure 5).

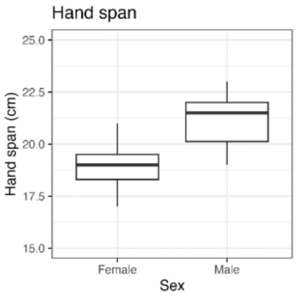


Figure 4a. Hand span for female and male participants

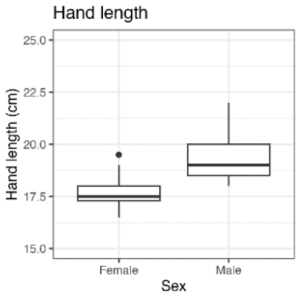


Figure 4b. Hand length for female and male participants

Figure 4c. Diameter of middle finger for female and male participants

Table 3. Error rates for female and male participants

Hand measure	**Female**		**Male**	
Error rate	*M* = 0.82	*SD* = 1.18	*M* = 0.78	*SD* = 1.28

Spearman correlations for hand span and error rates reveal a small, non-significant effect of ρ = -0.17, p = .331, for hand length and errors also a small, non-significant effect of ρ = -0.2, p = .241 and for middle finger diameter no correlation with ρ = -0.04, p = .813.

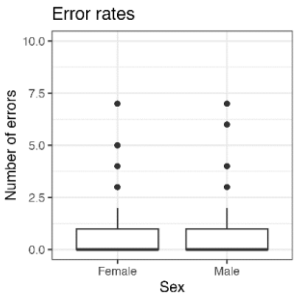


Figure 5. Error rates for female and male participants

Spearman correlations for errors and the subscale engagement of the ITC-SOPI reveal a moderate significant effect of ρ = -0.34, p =.04, of ecological validity/ naturalness and errors no correlation (p = .5), of negative effects and errors no correlation (p = .1), of spatial presence and errors a moderate significant effect of ρ = -0.43, p = .01 and between the overall ITC-SOPI score and errors a moderate significant effect of ρ = -0.34, p = 0.04.

ITC-SOPI subscale	**Values**	
Engagement	*M* = 52.7	*SD* = 6.89
Negative Effects	*M* = 10	*SD* = 3.98
Spatial Presence	*M* = 66.8	*SD* = 8.29
Ecological Validity/ Naturalness	*M* = 17.2	*SD* = 3.14
Overall Score	*M* = 147	*SD* = 16

5. DISCUSSION

In the presented study, we wanted to investigate whether we could identify systematic reasons for the errors in gesture recognition that we observed in previous studies and their influence on user experience, especially the subjective feeling of immersion measured by the ITC-SOPI and its subscales (Lessiter *et al.*, 2001). Therefore, we compared error rates between male and female participants to see if there were problems related to myotonus or similar anatomical characteristics. We also correlated various hand measures with observed error rates and subjective discomfort and error rates with ITC-SOPI results. We found that male participants had a significantly larger hand in all measures used. However, we did not find significant differences in error rates depending on the gender of the participants. We found a small to medium correlation between the size of the hand from the wrist to the middle finger and the error rates, but it was not significant. We found no correlation between hand size or middle finger diameter and error rates. Moreover, we found significant correlations between the overall score of the ITC-SOPI and error rates and between the subscales engagement and spatial presence and error rates.

The use of gestures for human-computer interaction is a complicated issue. In previous studies, we found that gestures are attractive, stimulating, well accepted, and trusted by users (anonymized). However, large anatomical differences in hand size and shape, myotonus, and movement performance, can make it difficult to ensure robust gesture recognition. With regard to technological advances and availability of the necessary hardware and software, GBI devices can be easily implemented in different contexts by developers and researchers. Nevertheless, it can be seen that gesture recognition rates are significantly lower than when using more traditional interaction modes such as interfaces with haptic elements or touchscreens. Interestingly, participants do not seem to be too bothered by these errors and still have a positive subjective feeling about gestures. Nevertheless, gestures require a longer learning time for users, they might have to consult a user manual carefully, and they have to remember the gestures used by the system. Since there is not yet an established set of gestures similar to applied standards in touch-based systems, these gestures may also differ in different systems or applications. Developers must choose between two possible paths when implementing GBI. Either they design an interaction system with a small number of gestures that users can remember more easily, but which are limited in terms of their adaptability to different tasks. Or they implement a larger number of gestures, which are more difficult for users to remember, but which are more adaptable to different tasks.

With this study, we aimed to contribute to this discussion by investigating possible reason for relatively low recognition rates in GBI and their influence and subjective assessments of the users. We could observe that hand size seems to have some effect, meaning that participants with larger hands from wrist to middle finger experienced slightly lower error rates. This makes sense if you consider that the used gesture recognition device uses an infrared camera system and image processing algorithms to build a 3D model of the user's hand. This model includes the wrist, palm, and joints. The larger the hand, the more accurate might be the model as the discriminatory power changes relative to the distance of these anatomic characteristics. Therefore, more effort should be put into algorithms that work robustly for smaller hands as well. It can also be shown that perceived errors or malfunctions have an influence on aspects of UX, especially the feeling of overall immersion, spatial presence and engagement of the uses.

The presented study is limited in scope due to its exploratory nature. Further research is needed to find the reasons why the gesture recognition rate is lower than desired and to find new technical ways to mitigate these problems. Since participants in such studies are often students, it would be interesting to see how attractive and acceptable gestures are to older potential users, and how gesture recognition works with these user groups. In addition, it would be interesting to gain insight into how hand muscle parameters affect gesture recognition and how gestures are generally performed by users, and how this compares to what is expected by gesture recognition developers, algorithms, and devices, and how improvements should be made based on these insights.

REFERENCES

Cohen, J., 1988. *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Hoboken, Taylor and Francis. Available at: http://gbv.eblib.com/patron/FullRecord.aspx?p=1192162.

Diefenbach, S. and Hassenzahl, M., 2008. Give Me a Reason. Hedonic Product Choice and Justification. *CHI '08 Extended Abstracts on Human Factors in Computing Systems, CHI '08: CHI Conference on Human Factors in Computing Systems*, Florence Italy. ACM, New York, NY, USA, pp. 3051-3056. doi: 10.1145/1358628.1358806

Dorton, S. L. and Harper, S. B., 2022. Self-Repairing and/or Buoyant Trust in Artificial Intelligence. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 66, No. 1, pp. 162-166. doi: 10.1177/1071181322661098

Dorton, S. L., Harper, S. B. and Neville, K. J., 2022. Adaptations to Trust Incidents with Artificial Intelligence. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 66, No. 1, pp. 95-99. doi: 10.1177/1071181322661146

Esterwood, C. and Robert, L. P., 2022. A Literature Review of Trust Repair in HRI. *2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, Napoli, Italy, pp. 1641-1646. IEEE. doi: 10.1109/RO-MAN53752.2022.9900667

Forlizzi, J. and Battarbee, K., 2004. Understanding experience in interactive systems. *Proceedings of the 2004 conference on Designing interactive systems processes, practices, methods, and techniques - DIS '04*, Cambridge, MA, USA. New York, NY, USA, p. 261. ACM Press. doi: 10.1145/1013115.1013152

Graichen, L., Graichen, M. and Nudzor, J., 2023. Integrating Mid-Air Gestures into a Virtual Reality Application: A Methodological Approach. *Proceedings of the 2023 7th International Conference on Virtual and Augmented Reality Simulations*, *International Conference on Virtual and Augmented Reality Simulations*, Sydney, Australia. New York, NY, United States, Association for Computing Machinery.

Hassenzahl, M., 2008. User Experience (UX): Towards an experiential perspective on product quality. *Proceedings of the 20th International Conference of the Association Francophone d'Interaction Homme-Machine on - IHM '08*, Metz, France, p. 11. New York, NY, USA, ACM Press. doi: 10.1145/1512714.1512717

Hassenzahl, M. and Roto, V., 2007. Being and doing: A perspective on User Experience and its measurement. *Interfaces*, Vol. 72, pp. 10-12.

ISO, 2019. *Ergonomics of human-system interaction*.

Kraus, J. *et al.*, 2020. The More You Know: Trust Dynamics and Calibration in Highly Automated Driving and the Effects of Take-Overs, System Malfunction, and System Transparency. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, Vol. 62, No. 5, pp. 718-736. doi: 10.1177/0018720819853686

Lee, J. *et al.*, 2021. Developing human-machine trust: Impacts of prior instruction and automation failure on driver trust in partially automated vehicles. *Transportation Research Part F: Traffic Psychology and Behaviour*, Vol. 81, pp. 384-395. doi: 10.1016/j.trf.2021.06.013

Lessiter, J. *et al.*, 2001. A Cross-Media Presence Questionnaire: The ITC-Sense of Presence Inventory. *Presence: Teleoperators and Virtual Environments,* Vol. 10, No. 3, pp. 282-297. doi: 10.1162/105474601300343612

Marti, P. and Iacono, I., 2016. Anticipated, Momentary, Episodic, Remembered: the many facets of User eXperience. *Proceedings of the 2016 Federated Conference on Computer Science and Information Systems*, pp. 1647-1655. IEEE. doi: 10.15439/2016F302

Mishler, S. and Chen, J., 2023. Effect of automation failure type on trust development in driving automation systems. *Applied Ergonomics*, Vol. 106, 103913. doi: 10.1016/j.apergo.2022.103913

Narayana, P. *et al.*, 2019. Cooperating with avatars through gesture, language and action. In K, Arai, S. Kapoor and R. Bhatia (eds.) *Intelligent Systems and Applications.* (Advances in Intelligent Systems and Computing). Cham: Springer International Publishing, pp. 272-293.

Obrist, M. *et al.*, 2012. In search of theoretical foundations for UX research and practice. *CHI '12 Conference Extended Abstracts on Human Factors in Computing Systems,* Austin, Texas, USA. New York, NY, USA, ACM, pp. 1979-1984. doi: 10.1145/2212776.2223739

Pustejovksy, J. *et al.*, 2020. Interpreting and generating gestures with embodied human computer interactions. *ACM IVA 2020 GENEA Workshop*.

Rakkolainen, I. *et al.*, 2021. Technologies for multimodal interaction in extended reality - A scoping review. *Multimodal Technologies and Interaction*, Vol. 5, No. 12, p. 81. doi: 10.3390/mti5120081

Rodriguez, G. *et al.*, 2017. Gestural interaction for virtual reality environments through data gloves. *Advances in Science, Technology and Engineering Systems Journal*, Vol. 2, No. 3, pp. 284-290. doi: 10.25046/aj020338

Saffer, D., 2008. *Designing gestural interfaces*. Sebastopol, CA: O'Reilly Media, Inc.

Sharp, H., Rogers, Y. and Preece, J., 2019. *Interaction Design: Beyond Human-Computer Interaction*: Wiley & Sons Canada, Limited, John.

Thielsch, M. T., Engel, R. and Hirschfeld, G., 2015. Expected usability is not a valid indicator of experienced usability. *PeerJ Computer Science*, Vol. 1, e19. doi: 10.7717/peerj-cs.19

Tolmeijer, S. *et al.*, 2021. Second Chance for a First Impression? Trust Development in Intelligent System Interaction. *Proceedings of the 29th ACM Conference on User Modeling, Adaptation and Personalization, UMAP '21,* Utrecht, Netherlands. New York, NY, USA: ACM, pp. 77-87. doi: 10.1145/3450613.3456817

Wilamowitz-Moellendorff, M. von, Hassenzahl, M. and Platz, A., 2006. Dynamics of user experience: How the perceived quality of mobile phones changes over time. *User Experience - Towards a unified view, Workshop at the 4th Nordic Conference on Human-Computer Interaction*, pp. 74-78.

Yogasara, T. *et al.*, 2011. General characteristics of anticipated user experience (AUX) with interactive products. *The 4th World Conference on Design Research: Diversity and Unity,* Delft, Netherlands, pp. 1-11.

Yu, K. *et al.*, 2018. Do I Trust a Machine? Differences in User Trust Based on System Performance. In J. Zhou and F. Chen (eds.) *Human and Machine Learning.* (Human–Computer Interaction Series). Cham: Springer International Publishing, pp. 245-264.