Journal of Environmental Psychology 92 (2023) 102137

Contents lists available at ScienceDirect

Journal of Environmental Psychology

journal homepage: www.elsevier.com/locate/jep

How good are virtual hands? Influences of input modality on motor tasks in virtual reality

Asim Hameed [a,*], Sebastian Möller [b], Andrew Perkis [a]

[a] Department of Electronic Systems, NTNU, NO-7491, Trondheim, Norway
[b] Quality & Usability Lab, TU Berlin, 10587, Berlin, DE, USA

ARTICLE INFO

Keywords:
Virtual reality
Input modalities
Gesture interaction
User experience
Performance metrics
Usability
Presence
Perceived realism

ABSTRACT

Hand-tracking enables controller-free interaction with virtual environments, which can make virtual reality (VR) experiences more natural and immersive. As naturalness hinges on both technological and human influence factors, fine-tuning the former while assessing the latter can be used to increase overall experience. This paper investigates a reach-grab-place task inside VR using two input modalities (hand-tracking vs. handheld-controller). Subjects (N = 33) compared the two input methods available on a consumer grade VR headset for their effects on objective user performance and subjective experience of the perceived sense of presence, cognitive workload, and ease-of-use. We found that virtual hands (with hand-tracking) did not influence the subjective feelings of perceived presence, naturalness, & engagement; neither did it inspire the overall ease-of-use while performing the task. In fact, subjects completed the task faster and felt a lower mental workload and higher overall usability with handheld-controllers. The result found that in this particular case, hand-tracking did not improve the psychological and emotional determinants of immersive VR experiences. The study helps expand on our understanding of the two input modalities in terms of their viability for naturalistic experiences in VR akin to real-world scenarios.

1. Introduction

Advances in consumer-grade VR devices are paving the way for natural interactions and direct manipulation of objects inside immersive virtual environments (VE). From a behavioral standpoint, VR is an advanced human-computer interface that allows users to ”immerse” into computer-generated environments and interact in a naturalistic manner within them (Slater, 2009). In general, the user-interaction paradigm for VR has relied on the use of head-mounted displays (HMD) and handheld controllers to freely move, look around, and respond to the various directional, visual, auditory, and haptic stimuli within the 360-degree omnidirectional VEs (Sheridan, 2016; Slater, 2018). More recently, improvements in input modalities on devices like HoloLens, Magic Leap, Oculus, etc. have now made non-mediated realistic interactions a possibility (LaViola Jr et al., 2017) – making VR further attractive for an array of training and learning applications (Liagkou et al., 2019; Thorsteinsson, 2013). VR applications are now useful in safety training in mining (Zhang, 2017), virtual assembly and manufacturing (Abidi et al., 2019; Palmas et al., 2019), medical training (Izard et al., 2018; Pottle, 2019), motor learning and rehabilitation (Crocetta et al., 2018; Mekbib et al., 2020), etc. Advances in the field have made it possible for users to apply their spatial awareness, literacies, and skills while performing in naturalistic real-world paradigms made possible by realistic, dynamic, and multi-sensory VEs (Pfeuffer et al., 2017). In addition to their real-world similitude, another benefit of VEs is that they are fully controllable. They allow perceptual modifications, task scaling, performance measurements, and behavioral observation of participants undertaking activities – well-suited for user experience studies and research.

Interactions within VR are predominantly mediated but natural and intuitive interaction has always been the goal (Regazzoni et al., 2018). Until recently, hand-controllers have been the primary means to interact within VEs but recent advances have made natural interactions a possibility (Kim et al., 2017; Masurovsky et al., 2020). The availability of hand-tracking (use of virtual hands) has been pivotal towards this end. In theory, at least, hand-tracking promises more realistic experiences by facilitating direct control of objects. Though still nascent, the possibility of closer-to-real interactions with hand-tracking will greatly influence the effectiveness and reach of VR. Hand-tracking technologies support gesture-based interactions with virtual objects/artifacts without the

---

* Corresponding author.
E-mail address: asim.hameed@ntnu.no (A. Hameed).

https://doi.org/10.1016/j.jenvp.2023.102137
Received 30 December 2022; Received in revised form 17 June 2023; Accepted 4 September 2023

need for handheld-controllers. With time, interactions inside VEs are expected to emulate the nuances of the real world with increased authenticity.

Given their potential, this work looks at hand-tracking technology. In particular, our study examines how hand-tracking modality compares to handheld-controllers in managing basic reach-and-grasp (or, grab-and-place) actions. For this purpose, we chose the commercially available Oculus Quest VR headset that supports inside-out hand-tracking. We intentionally used this standalone device since it offered affordable VR solutions for everyday entertainment and use. Our user study looks at how well the two input modalities available on the device, in comparison, support basic selection and manipulation operations.

Further, from a usability and user experience point of view, it is not always the case that higher naturalness may also lead to higher performance since the association between realism (i.e., the degree to which device interaction resembles features of real life) and the experience of naturalness is not a linear one (McMahan et al., 2016). Various other works have explored the influences of hand visualization on learning and task execution within VR (Pastel et al., 2022) (Punako Jr and Thropp) (Ricca et al., 2020; Shin et al., 2022). This makes it a point of interest to research the influences, and comparison, of input modalities on task performance and overall user experience inside VEs. Studies have previously used performance and quality metrics for quantitative assessments of VEs (Chiu et al., 2019; Hameed et al., 2021; Suznjevic et al., 2017; Wang et al., 2021). Subjective measures such as psychological aspects (Argelaguet et al., 2016; Lougiakis et al., 2020), cognitive loads (Luro & Sundstedt, 2019; Steed et al., 2016), and usability (Jankowski & Grabowski, 2015; Masurovsky et al., 2020; Voigt-Antons et al., 2020) are also commonly used evaluation methods. These studies indicate that user experience and performance inside VR are eventually influenced by a combination of technological determinants as well as human factors. That the two often overlap and one is not without the other. This paper presents, and discusses, the finding from our study that compared differences in user experience and user performance while performing a reach-grab-place task in VR using two different input modality types.

2. Related Works

2.1. Interactivity inside virtual environments

Interactivity is when users can respond, and make changes, to the contents of a mediated space. This is built upon three basic components: input devices that capture user actions (input modality); display devices that present the effect of these actions back to the user (HMDs); and, transfer functions that map device movements into movements of controlled display or interface elements (mapping method) (Bowman et al., 2001). VR systems translate our head and body positions, map them to the geometries/elements of the VE, and present it as immersive VR inside an HMD.

A similarity between our interactions within VEs to how we interact in the real-world can be marked on a continuum; on the higher end are realistic interactions that are loyal imitations of real-world interactions whereas non-realistic (symbolic/referential) interactions fall on the lower end of the spectrum (Bowman et al., 2012). Both have their uses depending on the nature of the task/challenge at hand. The objective degree of exactness with which real-world interactions can be reproduced inside VEs is defined as the *interaction fidelity* of the system (Ragan et al., 2015). In terms of experience, it is also understood as the degree to which actions involved in the performance of a virtual task correspond to the actions required for an equivalent real-world task (Jerald, 2015). The honesty, or congruence, of interactivity, has noted effects on the degree of perceived realism as well (Bowman et al., 2001). *Perceived realism* refers to how closely a virtual world resembles and feels like the real world (Weber et al., 2021). It is determined by evaluating both the subjective sense of reality (verisimilitude) of the environment and its overall perceived credibility (veridicality). Weber et al. (Weber et al., 2021) note that when using VR, a user will invariably judge the degree of realism of the virtual world in terms of its congruence: (1) its sights and sounds, and the virtual objects/artifacts contained therein, (2) the credibility and plausibility of the plot/scene/situation, and (3) the naturalness and ease of interaction within the VE.

It is important to note that these expectations will inevitably be different depending on the content presented to the user. For example, a fictitious story or experience in VR may warrant a very different set of expectations compared to a VRLE, such as a VR training simulator. The latter requires realistic interactions matched to the real-world due to the dexterity and precision involved. A user may judge if the size and proportion of the virtual body match their real body, and whether it corresponds well to the task at hand. It is intuitive to assume that higher interaction fidelity will naturally improve the training effectiveness of a VRLE (Hamblin, 2005). But interaction fidelity may also vary in requirement with the nature of the task at hand. For example, a higher interaction fidelity technique was found suitable for a Virtual Biopsy Trainer (Ricca et al., 2017), whereas a low-fidelity system was found sufficient for a Laparoscopic Surgical Trainer (Chellali et al., 2016; Kim et al., 2003).

2.2. Natural interaction paradigms for VR

Interactions inside VR comprise *selection* and *manipulation*. An entity can be selected via controller input, gestures, or gaze. It can then be manipulated via resizing, re-orienting, scaling, rotating, or translating the selected object.

VR systems employ direct and indirect methods for interaction. Indirect manipulation involves interacting with virtual objects through a proxy object like a controller and relying on symbolic referents to build an association with the virtual entities (e.g. pressing a button to move a box) (Holderied, 2017). Direct manipulation methods involve using our bodies to directly interact with entities. An entity can be grabbed naturally, using virtual hands, when it is close enough. If it is not in reach for direct grasping with the hand, ray-casting (laser-pointer) or gaze-based approaches can be used to create that association.

Current VR systems are increasingly employing interface paradigms suited for natural user interactions (NUI), such as, to support direct manipulation of objects (via gestures or body movements) without the need for communicating with intermedial devices. Hypothetically, direct interactions are considered to have a better sense of proximity or psychological closeness because of their use of action cues (Hutchins et al., 1985). This can bear a positive effect on a user's perceived realism of the VE. It is not to say however that direct manipulation methods do not have shortcomings. In fact, they can be exceedingly impractical due to either input device constraints like limited tracking range or any limitations of the human operators, for example, anatomical challenges. Generally speaking, either method appropriately matched to the interaction modality and corresponding to the demands of the task will yield better performance. Rieke et al. (Riecke et al., 2018) note that interaction effectiveness improves when the selection and manipulation options of a virtual interface correspond to the speed, range, and tracking capabilities of the system, and are complemented by the comprehension, skill, and learning capacity of the user (Myers et al., 2019). Unrealistically complicated interactions and/or expectation mismatches lead to adaptation problems and negative human factor implications (Nunnally & Bitan, 2006; Våpenstad et al., 2013). If, or not, this also increases user experience is another question and one which is not widely explored. In their experiments, Voigt-Antons et al. (Voigt-Antons et al., 2020) observed that even though participants felt lower control with hand-tracking, they still reported a positive user experience. In the case of Masurovsky et al. (Masurovsky et al., 2020), lower performance metrics were contrasted by high subjective scores reported by users for controller-free interaction. Given the novelty and interest in naturalistic interactions for VR, it is important to study the various influences input modalities may have on user performance and overall experience.

2.3. *Performing reach-grab-place tasks in VR*

Prehension is the ability to reach for and grasp onto an object. It is a fundamental motor skill that allows us to manipulate objects in our environment. Prehension involves the coordination of various muscles and joints in the hand, arm, and shoulder to position the hand and fingers around an object in a way that allows us to hold onto it firmly (van de Kamp & Zaal, 2007). It is an important skill for carrying out basic actions in daily life – for tasks such as holding utensils, picking up and carrying objects, or manipulating tools and other objects. Several classifications of human grasp types have emerged over the years (Feix et al., 2015; Kamakura et al., 1980; Schlesinger, 1919; Sollerman, 1980). Various taxonomies classify grasp types into broad categories of power, precision, and intermediate type. Further subcategories are based on thumb positions and finger/palm contacts. In short, the various prehension types depend on the specific movement and positioning of the hand and fingers.

In VR, prehension would relate to the ability of users to select and manipulate virtual objects using either:

- hand-held controllers, which mimic the movement and position of the user's hands in the VE. These controllers can include buttons, triggers, and other inputs that allow the user to perform a range of actions, such as grasping, releasing, and manipulating virtual objects; or
- hand gestures, also hand-tracking, which allows the VR system to detect and interpret the movement and position of the user's hands in the real world and map them onto corresponding hand gestures in the VE. This can be done using sensors on the user's hands or using cameras and other sensors to track the movement of the user's hands and fingers.

In theory, both these methods can enhance the sense of immersion in a VE and a user's perceived realism of the virtual world. But obvious caveats remain. For instance, handheld controllers are technically brilliant since they allow seamless real-time tracking of the hand location within virtual space, but there is still an intermediate device in between. Hand-tracking, on the other hand, promises a direct method but one can still not fully act upon virtual objects. One can pick them up by enveloping them, it is not possible to squeeze or lift them because virtual objects don't have weight, volume, or texture.

As such, there is good reason for understanding prehension for VR given its usefulness in a wide range of tasks, activities, and scenarios suited for simulation-based virtual training, assembly, prototyping, etc. The abundance of research around prehension in VR is concentrated in areas of motor therapy (Kaliki et al., 2012; Sveistrup, 2004; Viau et al., 2004) and rehabilitation (Grimm et al., 2016; Levin et al., 2015; Merians et al., 2006). Separately, in a study with 13 participants, Furmanek et al. (Furmanek et al., 2019) compared reach-to-grasp movement patterns inside VE to those performed in the physical environment. The comparison was based on established kinematic variables and carried out in three phases of initiation, shaping, and closure. They found that user performances remain similar in both environments with the exception of differences found in the closure phase, which was prolonged in VE. In another experiment, participants performed a reach-and-grasp task under monocular, motion parallax, and binocular viewing conditions using a telepresence system (Plooy & Wann, 2000). While a prehension-based activity was used, the study however focused on depth and distance judgements only.

2.4. *User experience and performance evaluations*

The relationship between user performance and user experience in VEs is complementary and multifaceted. Users who are able to effectively interact, perform actions, and achieve their goals are likely to have a more positive experience. Similarly, users who are more immersed, engaged, and satisfied are likely to perform better. The interplay of these aspects bears an influence on the overall quality of the immersive experience (Brunnström et al., 2013; Perkis et al., 2020). From a user perspective, a VR experience amounts to where they are, what they are doing, and how they are doing it. Also, otherwise theorized into concepts of immersion, presence, and immediacy. To have *presence* or a "sense of being there" in a mediated space is the success of a *place illusion* (Biocca, 2002; Slater, 2018). Such an *Immersion* into the medium is a user's response to either system characteristics or the content presented or both simultaneously (Nilsson et al., 2016). Finally, *immediacy*, speaks to the interaction fidelity of an immersive media experience (McMahan et al., 2016). Realistic interactions (closer to natural) have higher immediacy as opposed to unrealistic interactions (Liou et al., 2017).

Broadly speaking, an immersive and authentic VR experience (realism) may draw users in but their interest and motivation depend on the engagement/challenge offered by the content/task (involvement) and how readily and intuitively the system/interface allows them to perform it (usability). Both objective and subjective measures are common practices for assessing the aforementioned aspects. In most cases, ratings are conducted post-experience after participants have removed their head-mounted displays. More recently, researchers (Alexandrovsky et al., 2020; Feick et al., 2020; Graf & Schwind, 2020) have tried to optimize subjective tools for within-experience use as well.

For the purposes of this paper, we specifically look at the following:

- Performance data: provides game-based quantitative data to assess the quality of a VR product from the perspective of the user, for example, analyzing log data to identify patterns and trends in user behavior, and conducting user testing to measure task completion time, error rates, and other performance indicators. This data is useful for measuring a user's performance but it can also be crucial for understanding and optimizing learning and training environments (Loh et al., 2015).
- Psychological and emotional determinants: of the VE are measured with the widely used Igroup Presence Questionnaire (IPQ) (Schubert et al., 2001). It also broadly considers the verisimilitude and veridicality of the virtual world known in terms of its degree of perceived realism (discussed in an earlier sub-section). Buttussi et al. (Buttussi & Chittaro, 2019) used it in a user study to assess the effects of the three locomotion techniques (joystick, teleportation, and leaning) on participants. The measurement tool is useful for determining perceived interaction fidelity and visual render quality experienced by users (Berki, 2020; Blaga et al., 2020; Fromberger et al., 2015). The possibility of action is generally seen to influence VR experiences (Schubert et al., 2001; Slater, 2018).
- Perceived mental workloads: provide a good overview of a user's state of mind during the performance of tasks inside VEs (Feick et al., 2020; Lackey et al., 2016; Zheng et al., 2012). The multi-dimensional NASA Task Load Index (NASA-TLX) (Hart, 2006) has been extensively used for this purpose. High levels of cognitive workload can lead to frustration, reduced performance, and an overall negative experience. Whereas lower levels of workload allow users to effectively engage with the training material and achieve their learning objectives.
- Usability: is primarily the "ease of use" of the VR interface. Usability issues can include difficulty navigating the VR environment, difficulty understanding or interacting with the training material, or confusion with the interface. In other words, how easy or difficult is it to act in the VE. This is especially crucial for successfully using VRLEs. Toolkits like the AttrakDiff (Hassenzahl et al., 2008) have been used to good effect for measuring and identifying usability and desirability issues influencing the effectiveness of VR learning applications (Chen et al., 2016; Sassatelli et al., 2020).

3. Method

We conducted an empirical study to systematically compare two available input modalities on a consumer-grade VR system. It is widely hypothesized that the possibility of closer-to-real interactions using hand-tracking modality potentially improves our overall experience of VR (Haar et al., 2021; Kilteni et al., 2012; Perez-Marcos et al., 2017). So we investigate modality influences on user experience and performance in simulation-based reach-grab-place tasks in VR. The study additionally considers if a visual factor (e.g. changes in the environmental color) will also bear influence on users' perceptions while undertaking the task as previously observed (Billger et al., 2004; Felton, 2021). The collected results were assessed for the users' subjective experience (e.g. sense of presence, perceived cognitive workload, and ease of use) in correlation to their game-based performance metrics.

A repeated-measures 2 × 2 design was used to compare two input modalities across participants performing a reach-grab-place task inside a VE rendered at two levels of representational realism. The modality type (M) and representation level (R) form our two independent variables (IV). Each had two levels: M, hand-tracking (M1) and handheld controller (M2); and R, saturated (R1) and grayscaled (R2). Giving us four variations, see Fig. 1:

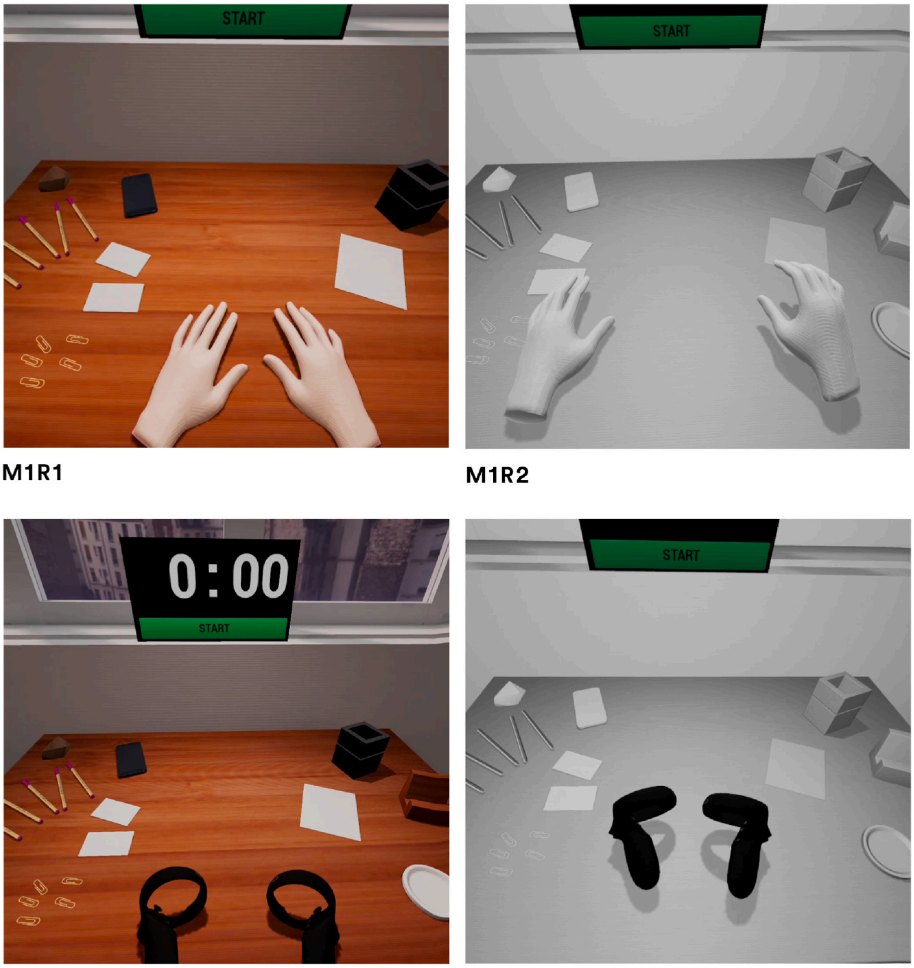
- M1R1 - hand-tracking x colors VE
- M1R2 - hand-tracking x grayscaled VE
- M2R1 - handheld-controller x saturated VE
- M2R2 - handheld-controller x grayscaled VE

Dependent variables (DV) included user performance metrics and scores for user-perceived overall presence, mental workload, and usability from self-reported questionnaires.

3.1. Participants

The sample consisted of a mixed demographic of N = 33 participants (15 male, 18 female, μ = 24.7 ± 2.3). Participants were recruited via an online volunteer portal of the university open to the general public. People from diverse backgrounds signed up for the study. They received gift cards for their participation. Participants reported their demographic data and knowledge/proclivity for immersive technology at the start of the test. Google Forms were used for the pre-study survey. In terms of user experience of VR: the majority of participants (n = 22) reported "no" experience; a few (n = 6) reported ”some” prior experience; others (n = 4) had ”intermediate/moderate” experience; and a single (n = 1) participant had ”good” prior experience of VR. Similarly, only 1 participant had used VR in a lab study before. All participants (n = 32) reported at least "some" prior experience with video game controllers and a single (n = 1) participant had previous experience with hand-tracking. All participants had normal visual acuity or normal corrected visual acuity.

The study was conducted at the VR/XR labs of the university, see Fig. 2. Informed consent was obtained from all subjects involved in the study, and institutional ethics were sought prior to the commencement of this study. All procedures performed in this study were in accordance with the ethical standards of the Institutional Review Board (or Ethics



M1R1 M1R2

M2R1 M2R2

Fig. 1. Four variations based on Modality Types (M1, M2) and Representation Levels (R1, R2). (Unreal Engine © Epic Games. Photo: Screenshot Image).

Fig. 2. A participant using hand-controller modality interacts with the environment displayed inside the HMD.

Committee), and with the 1964 Helsinki Declaration and its later amendments. Experiment protocols and data collection followed the guidelines described in ITU-T Recommendations P.809 (ITU-T, 2018), P.911 (ITU-T, 1998), and P.919 (ITU-T, 2020). Data presented in this study is not publicly available due to privacy protection and is available on request only. This research was funded by the NTNU IE Faculty, Project No. 63350581770958. The study was conducted during the second Covid-19 wave. Strict hygiene protocols were followed during test sessions based on the specified SARS-CoV-2 guidelines for educational institutions.

3.2. *Stimuli*

The VE comprised a virtual room with a table overlooking a window. Participants sat at the virtual table to carry out a simple reach-grab-place task. The VE was modeled on a 1:1 exactness to the physical laboratory where the experiment was conducted; size, 5.4 m by 4.4 m. The model was prepared in Sketchup Pro (Version 20.2.172). The door and windows inside the VE matched the physical lab. Similarly, the furniture location and orientation were also matched to the physical room. The virtual table overlapped the physical table, and the heights were matched so that participants would feel a surface under their arms while at the virtual table. Textures, lighting, interactivity features, and gamification elements were applied inside Unreal Engine (Version 4.26).

3.3. *Task*

The VR task required the participants to re-organize a number of assorted items (N = 15) on a virtual table. At the start of each variation, all 15 items would appear on the left-hand side of the virtual table. Participants were then required to reach-grab-place each of the 15 items one by one to their designated positions on the right-hand side. It was important that participants accurately place the objects in their desired spots. Participants could orient the items to their liking and ease. Figs. 3 and 4 show a few select screenshots of the two modalities in action (see captions for details). The task required grabbing and moving the

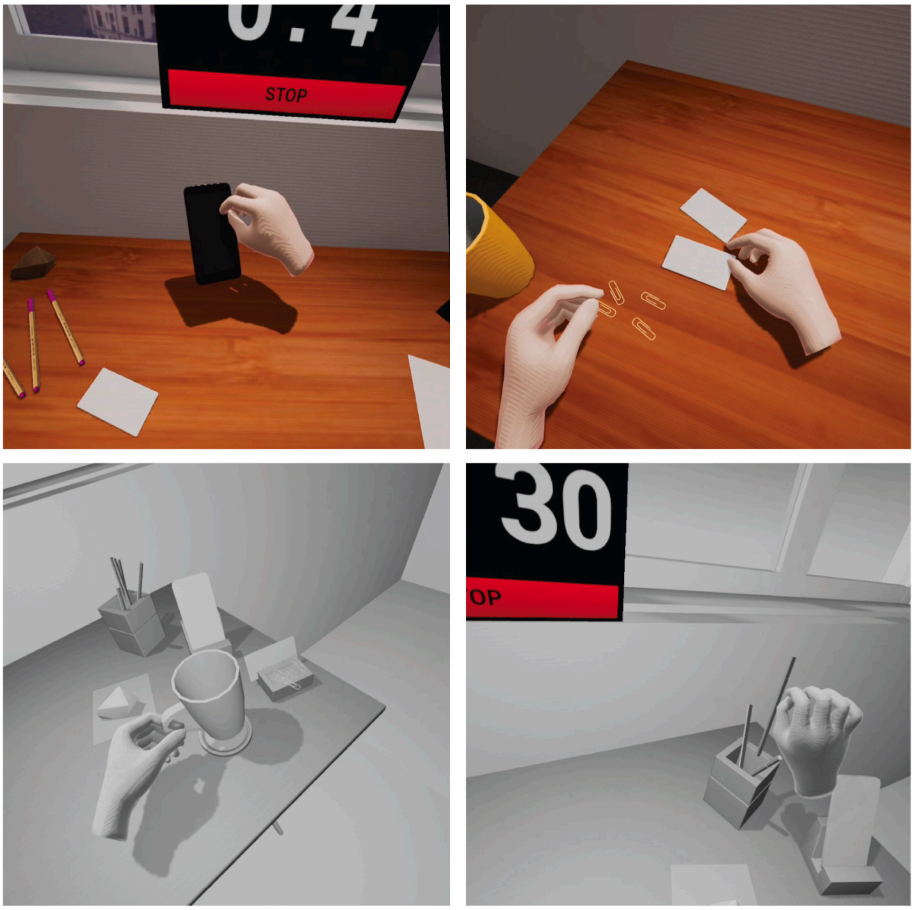


Fig. 3. Top Left: mobile phone grabbed with virtual hands in M1R1. Top Right: user picks a business card with virtual hands in M1R1. Bottom Left: user places the coffee mug with virtual hands in M1R2. Bottom Right: A marker is placed using virtual hands in M1R2. (Unreal Engine © Epic Games. Photo: Screenshot Image).

Fig. 4. Top Left: coffee mug placed on the coaster using controllers in M2R1. Top Right: marker placed in the holder using controllers in M2R1. Bottom Left: user picks a paper clip using controllers in M2R2. Bottom Right: paper weight placed on the paper in M2R2. (Unreal Engine © Epic Games. Photo: Screenshot Image).

following items:

- 5 pens into a pen holder
- 5 paperclips into the clip saucer
- 2 business cards into the cardholder
- 1 mobile phone to the mobile phone holder
- 1 paperweight on top of the paper
- 1 coffee mug onto a coaster

All six objects selected for the task required distinct modes of prehension (see Fig. 5). Each required a separate static grip, described below (Feix et al., 2015; Kapandji, 1987):

1. *Terminal opposition* (Fig. 5A) is a precision grip, which allows one to hold a thin object or to pick up a very fine object like a needle or a paperclip. The thumb and the tip of the index (or the middle finger) come into contact during opposition when fine objects are being grasped.
2. *Tetradigital grip* (Fig. 3B) is for holding larger objects firmly. With a tetradigital grip involving pulp contact, one can hold a pencil, a brush, or a pen. The pulp of the thumb presses the object firmly against the pulps of the index, middle, and ring fingers.

3. *Subterminal opposition* (Fig. 3C) is a grip that involves the sides of the fingers and thumb coming into contact with the object, rather than the fingertips or palms. It is often used for holding onto thin, elongated objects like paper.
4. *Tridigital grip* (Fig. 3D) involves the thumb, index, and middle finger. It is a popular grip used for bringing food to the mouth. Subterminal tridigital prehension in the case of small round or irregular objects.
5. *Panoramic pentadigital grip* (Fig. 3E) involves holding large, flat objects with the fingers widely separated and the thumb positioned in maximal counter opposition. This grip allows one to securely grasp such objects.
6. *Three-finger pinch-dorsal contact grip* (Fig. 3F) is used for holding objects like cups. It involves placing the pad of the index finger at the level of the middle phalanx and the radial aspect of the middle finger on the cup to provide balance and support.

A virtual timer kept time for each task. The timer could be stopped if the user got bored, frustrated, or annoyed with the input modality or the task itself. In such cases, the result would be reported as ”incomplete”.

3.4. Apparatus

The VR game was optimized for use in the Oculus Quest VR headset (Oculus VR, Inc.”, 2020). The system comprises of a standalone device capable of running games and software wirelessly under an Android-based operating system. The headset provided a stereoscopic viewing using OLED display for each eye, with an individual resolution of 1440 × 1600 and a refresh rate of 72 Hz. This headset comes with internal cameras for inside-out, positional tracking of movements that afford six-degrees-of-freedom, 6DOF. Oculus Quest uses both the Oculus Touch (Hand-controller) and also supports controller-free gestures. Fig. 6 shows the use of the two modalities.

Oculus hand-tracking analyzes discrete hand poses and tracks the

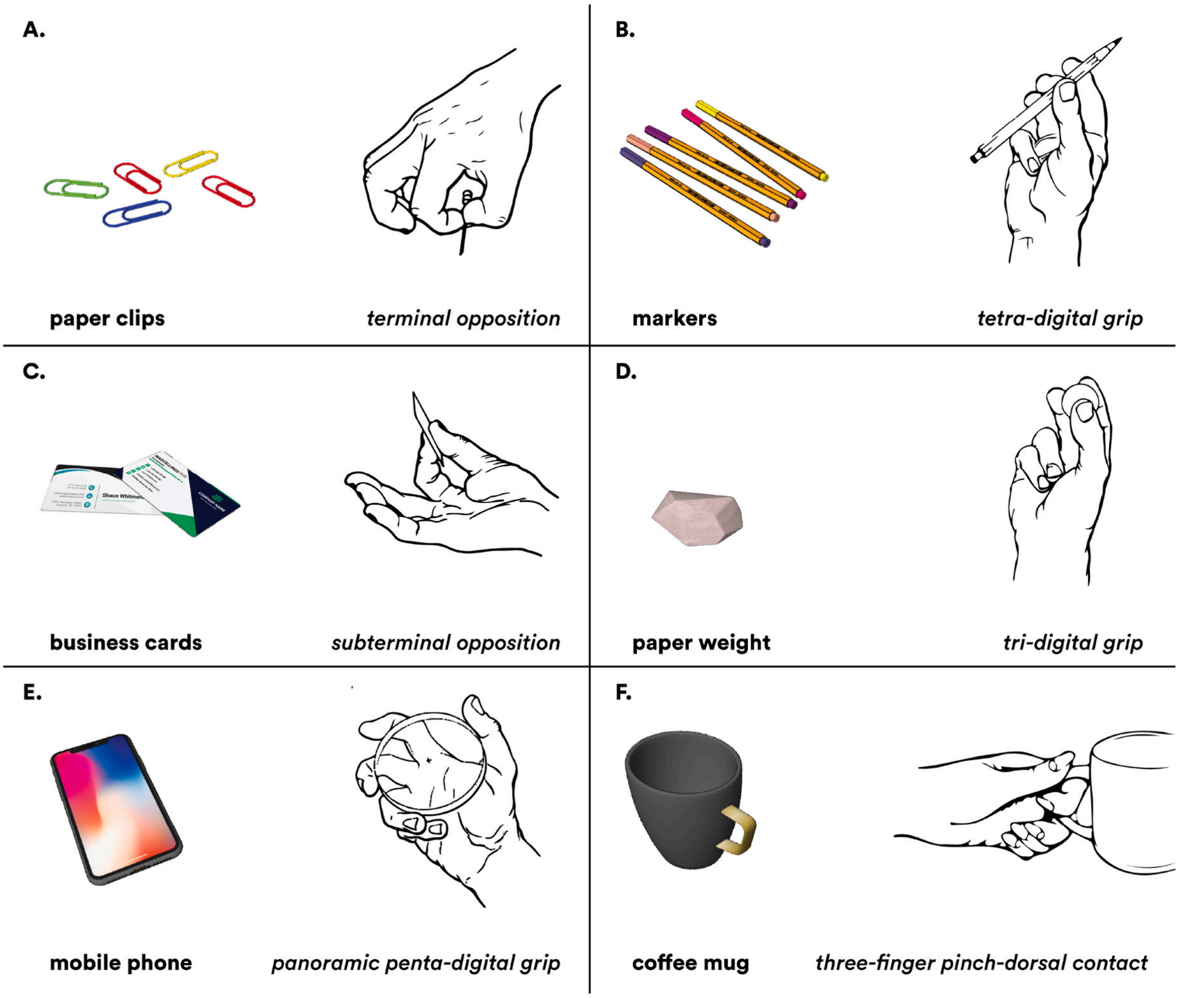


Fig. 5. Six objects used in the reach-grab-place task. Each object required a specific static grip to be handled. (prehension illustrations by I.A. Kapandji (Kapandji, 1987)).

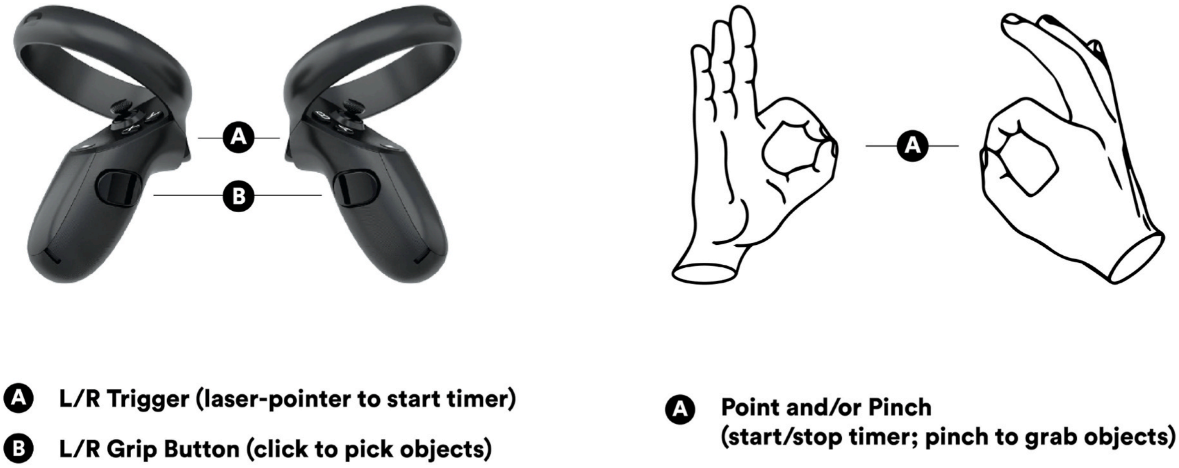


Fig. 6. Two input modalities available on Oculus Quest (Oculus Quest © Oculus VR, Inc.).

position of certain key points on hands in real-time, such as knuckles or fingertips. No official accuracy rate was specified for the Oculus Quest. Separately, AbdlKarim et al. (Abdlkarim et al., 2023) measured the overall performance of the hand-tracking system. At an average, their results showed a fingertip positional error of 1.1 cm, a finger joint angle error of 9.60, and a temporal delay of 38.0 ms.

On the handheld-controller, the left or right grip button could be pressed to lift any of the six objects. One object could be picked at a time by bringing the controller next to it. For hand-tracking, object interactions on the device are limited to basic gestures that include (Oculus, 2020):

- pointing, users extend their index finger and point it forward. The system recognizes it as a gesture for selecting items.
- pinching, when the thumb and index finger are together, the system recognizes it as a gesture used for grabbing the items.

3.5. Procedure

The study had four experimental conditions – M1R1, M1R2, M2R1, and M2R2. Each participant was to perform under all four conditions. The order of the experimental conditions was randomized across the participants to avoid sequence effects. The task requirement remained the same across the four conditions. However, the input modality and degree of representational realism varied in each case (refer to Fig. 2).

Upon arrival at the lab, participants were briefed on the purpose of the study. After signing the consent form, participants filled out a demographic survey. Afterward, participants were seated at a table facing a window. The virtual table within the game was matched at 1:1 to the physical table. All participants first attempted a physical version of the task to familiarize themselves with it so that while performing the task within the virtual game they could focus on *how to do* it rather than *what to do*. Each condition was preceded by a brief instructional phase explaining how to use the input modality. Participants were familiarized with the manipulation techniques through written instructions, which were complemented by a live presentation by the experimenter. The procedure followed throughout the test.

A typical testing session lasted 60 min – divided into four separate condition runs. All four experimental conditions were conducted back-to-back. A 5–7 min resting period between each test condition allowed the participants to recover and avoid eye strain. Each condition proceeded as follows:

- A participant would put on the HMD.
- A task would begin as soon as a participant ”started” the timer.
- The Participant would perform the reach-grab-place task.
- Once finished, the user would ”stop” the timer.
- The Participant would take off the HMD.
- They would rest for a few mins.
- The participant then fills out the post-game questionnaires.

3.6. Measures

Performance Metrics: refer to the patterns and trends identified in user behavior while performing tasks. User performance was judged on the basis of a game log generated at the end of each completed task, which marked:

1. Active-Time (AT), the time taken by a user to complete the task.
2. Object-Pick (OP), number of objects grabbed in the task
3. Click-Frequency (CF), number of attempted clicks to grab the objects.

Igroup Presence Questionnaire (IPQ): The IPQ is a self-report questionnaire to measure the sense of presence in virtual reality environments (Schubert et al., 2001). It contains 14 items rated on a seven-point Likert scale ranging from 0 to 6. The IPQ contains three sub-scales that measure:

1. Spatial Presence (SP), related to the sense of physically being in the VE.
2. Involvement (INV), is meant to evaluate the attention devoted to the VE.
3. Realness (REAL), evaluates the sense of reality attributed to the VE.

These concepts have been covered in the ’Related Works’ section of this paper. Additionally, the IPQ contains one general item which assesses the general “sense of being there” (GP), and has high loadings on all three factors, with an especially strong loading on Spatial Presence. The IPQ has a high reliability (Cronbach’s $\alpha = 0.87$).

Cognitive Load (NASA-TLX): The NASA task load index (NASA TLX) is a tool for measuring and conducting a subjective mental workload (MWL) or cognitive load assessment (Hart, 2006). It allows to determine the perceived cognitive load of users while they are performing a task. The index rates performance across six dimensions to determine an overall workload rating. These are the requirements of:

1. Mental demand (MD), thinking, decisions, or calculations.
2. Physical demand (PD), the amount and intensity of physical activity.
3. Temporal demand (TD), time pressure involved for completion.
4. Effort (ED), how hard is it to maintain performance?
5. Performance (RD), the level of success in completing the task.
6. Frustration level (FD), do you feel secure/insecure or discouraged/ content.

Each question has a rating from 1 to 10, where 1 represents the lowest task demand, and 10 represents the highest, with the exception of the performance question, where 1 indicates the highest, and 10 indicates the lowest.

UX Evaluation (AttrakDiff): The Attrakdiff assesses UX-related quality perceptions of the game application (Hassenzahl et al., 2008). The tool applies a hedonic/pragmatic model of UX. This questionnaire evaluates the perceived pragmatic (4 items) and hedonic (4 items) attributes of the interface, as well as its overall attractiveness (2 items), using a 7-point semantic differential scale. The tool is effective for surveys due to its easy and intuitive handling. Furthermore, AttrakDiff offers fast and well-visualized illustrations of the results of comparisons of different products in terms of the user experience. The use of semantic differential makes it possible to narrow down how strongly a user might connect a survey term with a certain property of the application. With the help of opposite adjectives, users can indicate their perception of the application. The word-pairs make a collation of the evaluation dimensions possible; these being:

1. Pragmatic Quality (PQ), describes the usability of a given product and how successfully users achieve their goals using it.
2. Hedonic Quality (HQ), measures emotional reactions of how stimulating/inspiring a product is and whether users identify with it.
3. Attractiveness (ATT), describes the global value of the product based on quality perception, either positive or negative.

3.7. Analysis

The IBM SPSS Statistics (Version 28.0.1.0) software was used for analysis. Two-way MANOVA (Multivariate Analysis of Variance) was used to examine the effects of two independent variables (IV), modality type and representation level, on user performance metrics and the user-perceived sense of presence, mental workload, and usability. Independent observations were collected from a sufficient sample size ($N = 33$) as per ITU-T recommendations (ITU-T, 1998; ITU-T, 2020). Participants were equally distributed across all conditions (X4). Scores for $N = 1$ participants were excluded on account of being ”incomplete”. Multivariate normality, outliers, linearity, multicollinearity, and homogeneity of variance were tested before analysis to ensure the assumptions of MANOVA (Pallant, 2020). In the event of significant results, further univariate analyses were conducted. All analyses considered two independent variables (IV x 2) with two categorical groups: M-type (M1, M2), and R-Level (R1, R2). Subjects evaluated all four variations of the application in a randomized order. 128 data entries (4 per subject X 32) were received and analyzed. Below we look at the various dependent variables (DV) that included quantitative game metrics (x 3 DVs) and the different subscales for IPQ (x 4 DVs), NASA-TLX (x 6 DVs), and AttrakDiff (x 3 DVs). The AttrakDiff scores were also additionally analyzed using the official online eSurvey tool for AttrakDiff with outputs of portfolio-presentation, diagram of average values, and description of word-pairs. (User Interface Design - UID GmbH”).

4. Results

In this section, we look at the results of our reach-grab-place task in VR. A total of 33 adults were randomly assigned to the four conditions of the 2×2 design. Scores for only 32 subjects were considered for results since $N = 1$ subjects failed to complete the task. We evaluated the main effects and interaction of M (M1, M2) and R (R1, R2) based on the performance of the users and their self-reported responses to the subjective measures. Below, we present our results.

4.1. Performance data

A two-way MANOVA examined the effects of M and R on the measures of user performance: AT, OP, and CF. The dataset for AT and OP had a positively skewed distribution so logarithmic transformation was applied uniformly. Results of Box’s M Test (Equality of Covariance Matrices) showed no violation of the assumption ($p = 0.26$). Significance values for Levene’s Test (Equality of Error Variances) ($P_{AT} = 0.73$, $OP_{OP} = 0.85$, and $P_{CF} = 0.37$) exceeded 0.05.

The MANOVA revealed a significant main effect of modality type, F (3, 122) $= 195.8$, $p = < 0.001$, Wilk’s $\lambda = 0.172$, $\eta^2 p = 0.83$. There was no significant main effect of representation level, F (3, 122) $= 0.85$, $p = 0.47$, Wilk’s $\lambda = 0.98$, $\eta^2 p = 0.02$. Nor was there a significant interaction effect between modality and representation, F (3, 122) $= 0.20$, $p = 0.90$, Wilk’s $\lambda = 0.99$, 0.005. Follow-up univariate analyses indicated that modality type produced significant differences across all three performance scores,

AT: $F(1, 124) = 62.74$, $p = < 0.001$, $\eta^2 p = 0.34$

OP: $F(1, 124) = 20.1$, $p = < 0.001$, $\eta^2 p = 0.14$

CF: $F(1, 124) = 374.8$, $p = < 0.001$, $\eta^2 p = 0.75$

Descriptive statistics for the dependent variables, presented in Table 1, shows the number of observations(N), the mean(μ), and the standard deviation (SD.) in each group. The table shows both the log-transformed and the original data. As shown, users using hand-tracking took longer to complete the task (μ = 314s, SD. = 169.42; μ = 272s, SD. = 116.58) compared to those using handheld-controllers (μ = 152s, SD. = 67.52; μ = 154s, SD. = 104.12). There was no significant separation for the mean scores on the basis of levels of representation. The mean score for a number of objects grabbed OP, was more in hand-tracking (μ = 32, SD. = 17.04; μ = 29, SD. = 11.82) compared to handheld-controller (μ = 21, SD. = 9.20; μ = 18, SD. = 8.3). Similarly, the number of attempted grabs, CF, was exponentially higher for hand-tracking (μ = 399, SD. = 243.65; μ = 371, SD. = 92.1) as well. It indicated that subjects found it harder to grab objects using hand-tracking. This difficulty, therefore, can be one reasonable explanation for the longer completion times.

Table 1
The mean and standard deviation for the four different dependent variables, which have been split by the independent variables for all $N = 33$ participants.

Item	Condition	Mean(μ)		Std.Dev (SD.)		N
		org	*log*	*org*	*log*	
AT	M1R1	314s	2.45	169.43	0.21	32
	M2R1	152s	2.14	67.52	0.19	32
	M1R2	272s	2.40	116.58	0.19	32
	M2R2	154s	2.12	104.12	0.24	32
OP	M1R1	32	1.44	17.04	0.25	32
	M2R1	21	1.26	9.20	0.31	32
	M1R2	29	1.43	11.82	0.17	32
	M2R2	18	1.20	8.30	0.28	32
CF	M1R1	399	2.54	243.65	0.23	32
	M2R1	58	1.70	37.15	0.23	32
	M1R2	371	2.52	192.1	0.20	32
	M2R2	66	1.70	65.10	0.29	32

4.2. Sense of presence

A two-way MANOVA with covariate was conducted to control for AT. This was done in consideration of the possible effects the duration of time spent within the VE may have on the four IPQ items. The maximum Mahalanobis distance value was checked for assumption testing. It was 15.55, which is less than the critical value of 18.47 ($df = 4$) required for multivariate normality (refer to Table 2). Results of Box’s M Test of Equality of Covariance Matrices showed no violation of the assumption ($p = 0.59$). All significance values ($P_{GP} = 0.64$, $P_{SP} = 0.59$, $P_{INV} = 0.47$, and $P_{REAL} = 0.55$) in Levene’s Test of Equality of Error Variances were more than 0.05. The scores for the IPQ-items were compiled on a likert scale. No statistically significant effects were obtained from the MANOVA results. There was no difference in the means of the four dependent variables of IPQ (GP, SP, INV, and REAL) (see Table 3).

Firstly, there is a non-significant effect of modality type on IPQ scores, M: F (4, 120) $= 0.734$, $p = 0.57$, Wilk’s $\lambda = 0.98$, $\eta^2 p = 0.024$. Secondly, there is a non-significant effect of representation level on IPQ scores, R: F (4, 120) $= 0.670$, $p = 0.61$, Wilk’s $\lambda = 0.98$, $\eta^2 p = 0.022$. Finally, there is a non-significant interaction effect between modality and representation on IPQ scores, M x R: F (4, 120) $= 0.158$, $p = 0.96$, Wilk’s $\lambda = 0.995$, $\eta^2 p = 0.005$.

We, therefore, fail to reject the null hypothesis, and that neither input modality nor visual representation bore significant influence when considered jointly for the user-reported IPQ scores. Table 3 shows Means (μ) and SDs of the IPQ sub-scales as a function of M-type and R-Level; charts are shown in Fig. 5.

4.3. Mental WorkLoad – MWL

The perceived cognitive workload, or MWL, for the virtual reach-grab-place task was evaluated using the NASA-TLX. A two-way MANOVA assessed the effects of modality and representation on users’ NASA-TLX scores. As a preliminary step, we checked for MANOVA assumptions. Logarithmic transformation was applied so that all data presented normal distribution. The result of Box’s M Test of Equality of Covariance Matrices was $p = 0.50$. Values for the Levene’s Test of Equality of Error Variances were: $P_{MD} = 0.31$, $P_{PD} = 0.48$, $P_{TD} = 0.53$, $P_{RD} = 0.16$), $P_{ED} = 0.42$), and $P_{FD} = 0.05$.

All four conditions were analyzed across the six sub-scales. No significant effects for representation levels were found on the NASA-TLX scores, R: textitF (6, 119) $= 0.136$, $p = 0.99$, Wilk’s $\lambda = 0.99$, $\eta^2 p = 0.007$. There were also no significant interaction effects between modality and representation on the indeces, M x R: textitF (6, 119) $= 0.684$, $p = 0.66$, Wilk’s $\lambda = 0.97$, $\eta^2 p = 0.033$.

The MANOVA revealed a significant main effect of modality type on the NASA-TLX indeces, M: textitF (6, 119) $= 8.374$, $p = < 0.001$, Wilk’s $\lambda = 0.703$, $\eta^2 p = 0.30$. Univariate ANOVAs were conducted to examine the specific effects of modality type on each subscale item. Below, we look at the test for between-subject effects:

Table 2
The critical chi-square values for evaluating Mahalanobis Distance at a critical alpha of 0.001 are shown below (Pallant, 2020). Values are shown from 2 to 10 degrees of freedom.

df	critical value	df	critical value	df	critical value
2	13.82	5	20.52	8	26.13
3	16.27	6	22.46	9	27.88
4	18.47	7	24.32	10	29.59

Table 3
The mean(μ) and standard deviation (SD.) for the four different dependent variables of IPQ, which have been split by the independent variables for all $N = 32$ participants.

IPQ	M1R1		M1R2		M2R1		M2R2	
	μ	SD.	μ	SD.	μ	SD.	μ	SD.
GP	4.63	1.62	4.31	1.67	4.93	1.54	4.77	1.62
SP	4.21	0.71	4.18	0.77	4.27	0.78	4.20	0.89
INV	4.50	1.37	4.37	1.21	4.81	1.20	4.52	1.23
REAL	3.61	0.65	3.42	0.70	3.49	0.79	3.35	0.63

Table 4
The mean and standard deviation for the six different dependent variables of NASA-TLX, which have been split by the independent variables for all $N = 32$ participants.

Item	Condition	Mean(μ)		Std. Dev (SD.)		N
		org	*log*	*org*	*log*	
MD	M1R1	4.06	1.82	1.86	0.87	32
	M2R1	2.53	1.05	1.65	0.92	32
	M1R2	4.06	1.89	1.52	0.70	32
	M2R2	2.38	1.06	1.24	0.75	32
PD	M1R1	4.06	1.88	1.66	0.69	32
	M2R1	2.59	1.17	1.39	0.79	32
	M1R2	4.19	1.95	1.49	0.60	32
	M2R2	2.53	1.20	1.08	0.66	32
TD	M1R1	4.25	1.92	1.90	0.73	32
	M2R1	2.81	1.257	1.73	0.85	32
	M1R2	4.13	1.91	1.56	0.70	32
	M2R2	2.78	1.27	1.48	0.80	32
RD	M1R1	3.03	1.39	1.51	0.84	32
	M2R1	3.16	1.45	1.51	0.84	32
	M1R2	3.22	1.55	1.31	0.70	32
	M2R2	2.84	1.26	1.65	0.89	32
ED	M1R1	3.60	1.62	1.90	0.85	32
	M2R1	2.75	1.21	1.72	0.85	32
	M1R2	3.97	1.86	1.62	0.62	32
	M2R2	2.47	1.04	1.63	0.88	32
FD	M1R1	4.84	2.15	1.72	0.68	32
	M2R1	3.34	1.53	1.62	0.85	32
	M1R2	4.81	2.18	1.47	0.55	32
	M2R2	3.06	1.47	1.30	0.69	32

- MD: $F\,(1, 124) = 30.5, p = < 0.001, \eta^2 p = 0.20$
- PD: $F\,(1, 124) = 36.41, p = < 0.001, \eta^2 p = 0.23$
- TD: $F\,(1, 124) = 23.2, p = < 0.001, \eta^2 p = 0.16$
- RD: $F\,(1, 124) = 0.60, p = 0.44, \eta^2 p = 0.005$
- ED: $F\,(1, 124) = 18.8, p = < 0.001, \eta^2 p = 0.132$
- FD: $F\,(1, 124) = 28.9, p = < 0.001, \eta^2 p = 0.19$

Significant difference was observed across all indices but PD ($p = 0.44$). The results indicate that the overall perceived workload was higher for hand-tracking (M1) compared to hand-controllers (M2) but barely diverged between saturated (R1) and grayscaled (R2) representations. Table 4 shows both the log-transformed and original means(μ) and standard deviations (SD.) in each group. Less cognitive workload was required when subjects used the handheld-controller to complete the task. Handheld-controllers were least imposing as evidenced by the means for M2R2 across all indices: MD (μ = 2.38, SD = 1.24), PD (μ = 2.53, SD = 1.08), TD (μ = 2.78, SD = 1.48), RD (μ = 2.84, SD = 1.65), ED (μ = 2.47, SD = 1.63), and FD (μ = 3.06, SD = 1.30).

4.4. User experience – UX

The results from the AttrakDiff were compiled into three dimensions of pragmatic (PQ) and hedonic quality (HQ), and attractiveness (ATT).

Prior to a two-way MANOVA, the Mahalanobis distance was checked for assumption testing. It maximum value was 14.95, which is less than the critical value of 16.27 (df = 3) required for multivariate normality (refer to Table 2). Results of Box's M Test of Equality of Covariance Matrices showed no violation of the assumption ($p = 0.57$). All significance values $P_{PQ} = 0.16$, $P_{HQ} = 0.41$, and $P_{ATT} = 0.15$) in Levene's Test of Equality of Error Variances were more than 0.05.

No significant main effects were noticed for R-Level: textitF (3, 122) = 1.953, $p = 0.125$, Wilk's $\lambda = 0.95$, $\eta^2 p = 0.05$. Nor were any interaction effect revealed for, (M x R): textitF (3, 122) = 0.335, $p = 0.80$, Wilk's $\lambda = 0.99$, $\eta^2 p = 0.08$. The two-way MANOVA did however reveal statistically significant the means of PQ, HQ and ATT when based on M-type: $F\,(3, 122) = 7.953, p = < 0.001$, Wilk's $\lambda = 0.84$, $\eta^2 p = 0.16$.

Follow-up univariate ANOVAs examined the specific differences for M and R on the three dependent variable individually to identify specific differences. Modality type had a significant influence on the usability (PQ: $p = < 0.001$) and desirability (ATT: $p = 0.033$) of the VR application. Whereas representation levels significantly affected intrigue (HQ: $p = 0.025$) and desirability (ATT: $p = 0.038$) only.

We further examined the results with the online AttrakDiff tool that outputs the following "Result Diagrams": portfolio-presentation, diagram of average values, and description of word-pairs. Fig. 7 depicts the overlapped portfolio-presentation results for the four conditions. M1R1 is located in "neutral" edging towards "self-oriented"; M1R2 is within "neutral"; M2R2 falls within "task-oriented"; and M2R1 can be seen in the "desired" position with a slight tendency towards "task-oriented". All confidence intervals are pretty similar in size, indicating that there was a general agreement amongst participants in terms of the hedonic and pragmatic qualities of the conditions. However, M2R1 visibly has the smallest confidence rectangle in the group implying higher reliability and a less coincidental result. The overlapped portfolio-presentation clearly demonstrates that confidence intervals are overlapping based on M-type, for example, blue over red, and green over yellow. This corroborates with the results from our analysis above. Subjects found the handheld-controller modality (M2R1 and M2R2) better oriented for performing tasks than the hand-tracking modality

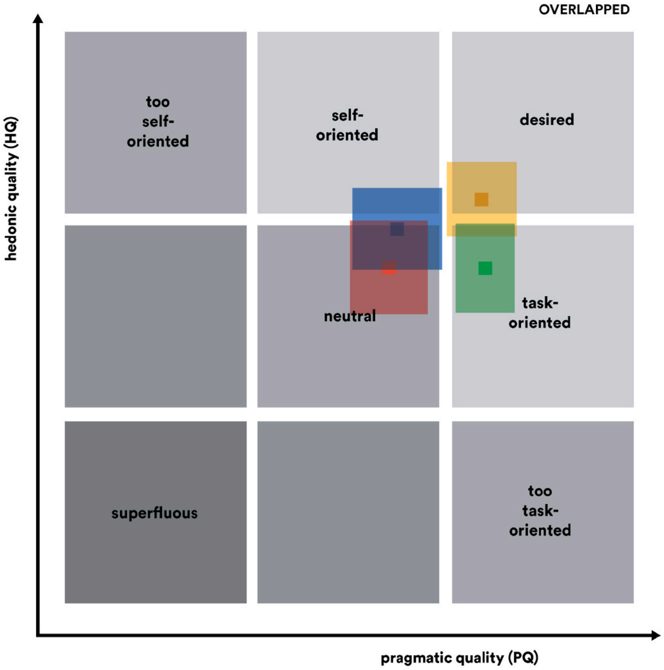


Fig. 7. An overlapped Portfolio-Presentation for the four experimental conditions. Color legend: M1R1 (blue), M1R2 (red), M2R2 (green), M2R1 (yellow). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

(M1R1 and M1R2). Overall, M2R1 exhibited higher perceived quality for participants.

Findings from the portfolio-presentation can be compared against the diagram of average values (Fig. 8) for further insight into perceived user experience. The results confirm that M2R1 (PQ = 1.40, HQ = 1.24, ATT = 1.85) generally does better (evident by the yellow line). However, M2R2 scores higher at PQ. M1R2 has the lowest performance (PQ = 0.43, HQ = 0.53, ATT = 0.90). The scores for M2R2 oscillate from a high for perceived pragmatic quality to a low score for hedonic quality, (PQ = 1.44, HQ = 0.86, ATT = 1.28). Overall, perceived usability remains higher than emotional response across all conditions.

Finally, (Fig. 9) shows the results of the description of word-pairs with all scores of the 10 different word-pairs. From the figure, the higher scores for M2R1 in terms of user experience are obvious. This condition shows superior performance in almost all word pairs (except confusing—clearly structured). An encouraging takeaway is that all conditions are within the positive user experience range. We can deduce that hand-controllers engender higher perceived usability for participants while at the same time visual realism had a positive emotional impact. This trend is most evident in M2R1, which also had the highest global value for quality perception in terms of appeal and pleasantness.

5. Discussion

For a reach-grab-place task in immersive VR, we systematically compared two input modalities (hand-tracking vs.handheld-controller) in two visual representation levels (saturated vs. grayscaled). We measured objective user performance metrics and subjective user experiences of perceived sense of presence, mental workload, and ease-of-use.

Generally speaking, the two input modalities can be used in VR to different effects. Hand-tracking allows users to interact with the virtual environment using natural hand gestures whereas a handheld-controller uses buttons and triggers to perform actions. Since virtual hands enable a more naturalistic interaction compared to handheld-controllers, we hypothesized higher ratings on naturalness and intuitiveness.

5.1. *On performance and effectiveness*

We found that the handheld-controller input modality was more effective for completing the reach-grab-place task, as it resulted in faster task completion times and better object manipulation compared to the hand-tracking. This was reflected in the mental workload scores where subjects recorded lower loads for handheld-controllers, indicating that it may be less cognitively demanding to use. Similarly, subjects also ranked the handheld-controller modality higher on task-orientedness and the overall appeal of the modality for the said task. However, subjects did not find any significant differences vis-a-vis immersion, involvement, and realness between the two input modalities inside VR. This suggests that while the handheld-controller may be more effective and preferred in terms of practical performance, it does not necessarily have a greater impact on the subjective experience of VR for the user. Surprisingly, the level of visual representation (saturated to grayscaled) also did not seem to have influenced the subjective experience of VR for the subjects.

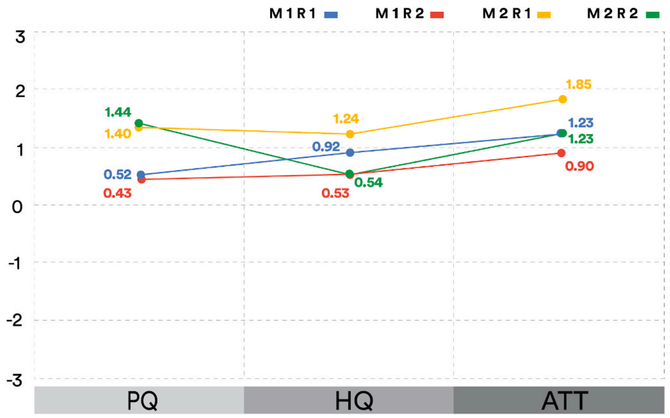


Fig. 8. Diagram of Average Values for the four experimental conditions. Color legend: M1R1 (blue), M1R2 (red), M2R2 (green), M2R1 (yellow). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

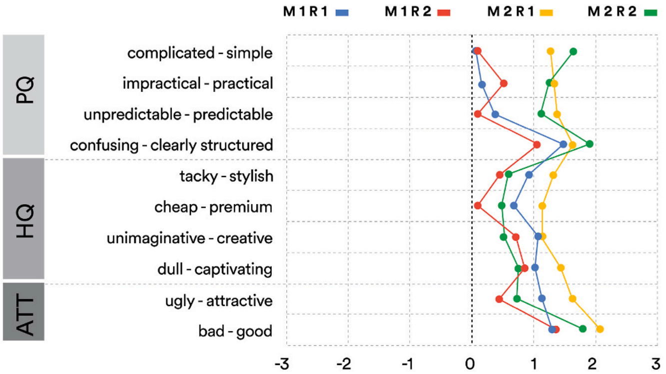


Fig. 9. Description of Word-Pairs for the four experimental conditions.

It is generally considered that virtual experiences that involve physical movements can enhance both spatial presence and mental immersion. In particular, using interfaces that recall hand gestures from daily life is poised to carry a higher sensation (De Paolis & De Luca, 2020). But our findings confirm the hesitation expressed in earlier works that found hand-tracking modules to be perceptively more difficult for performance (Caggianese et al., 2018; Ricca et al., 2020; Voigt-Antons et al., 2020). This could be because of unfulfilled user expectations about the capabilities of hand-tracking (Myers et al., 2019) spurlock-2019. There are several factors to consider when comparing hand-tracking and handheld-controller input modalities. One important factor is the accuracy and reliability of the tracking. Hand-tracking systems may sometimes have difficulty accurately tracking the user's hands, particularly if the user is making fast or complex gestures. There is no official hand-tracking accuracy rate specified for the Oculus Quest device but previous works have found it suitable for a wide range of applications Holzwarth et al. (2021) (Carnevale et al., 2022). Another common limitation with VR headsets (like the Oculus Quest) that provide inside-out tracking is having trouble detecting physical hands due to self-occlusion (Pacchierotti et al., 2016; Rehg & Kanade, 1995). Several subjects faced this issue while performing the task which hampered their experience. Handheld-controllers, on the other hand, offered a more precise and reliable means of input, as they are not reliant on the headset's tracking capabilities.

5.2. *On actions and expectations*

It is also important to consider the type of tasks that the user will be performing. In our case, the reach-grab-place task used in this experiment required the subjects to use different prehension types to grab the specific objects. We had considered that hand-tracking may be more suitable for the reach-grab-place task – due to the fine motor skills or precise hand movements involved – as it will allow users to use their own hands rather than relying on a physical controller. Here it is worthwhile to revisit the grips illustrated in (Fig. 3) of this document. The illustrations show six distinct ways in which the six respective objects are to be held. This is contrasted by the simplistic single "pinch" gesture currently available on the VR system. For example, lifting the paper clips using the *terminal opposition* grip had the best coincidence with the "pinch" gesture supported by the VR headset. A close second was the *subterminal opposition* grip used for grabbing the two business cards. However, subjects had to use the same pinch gesture to grab and move the virtual mobile phone. This is a sharp departure from the real-world operation, or natural prehension, where the panoramic penta-digital grip is often used to lift objects of that shape. Isomorphic mapping is responsible for establishing a relationship between elements in one system (such as physical movements or actions) and corresponding elements in another system (such as virtual movements or actions) (Hutchins et al., 1985). This can be useful for creating a sense of immersion or presence in a VE but as we found out a mismatch could result in the opposite. Technological limitations with isomorphic mapping of real hands may contribute to lower scores. This is because in real-world conditions an activity like holding a mug would prompt an instantaneous, automatic response to the stimuli by the subjects. The grip mismatch in VR suddenly demanded more mental processing for the same task. Something that should have been *ready-at-hand* (transparent) was all at once *present-at-hand* (opaque) (Coyne, 1994). It is imaginable why such interactions might be perceived as counter-intuitive and why such operational mismatches might negatively affect the user-perceived realism or naturalness of the environment (Hameed & Perkis, 2021; Weber et al., 2021).

5.3. On usability and demand

One more factor to consider is the user's level of comfort and ease of use. In our analysis, we found a significant difference between the handheld-controller and hand-tracking in terms of usability – the former was perceived as being more useable. We found that the controller-free modality continuously performed in the "neutral" zone, which is not discouraging but did not receive the high user experience valuation we had hypothesized (Masurovsky et al., 2020). Overall, the controller-in-a-realistic-scenario (M2R1 variation) outperformed across the AttrakDiff dimensions showing that user experience draws a sensitive balance between achieving goals, maintaining desirability, and remaining instinctive at the same time.

The preference for the controller modality was most pronounced across the mental workload indices of NASA-TLX where users indicated a higher effort of the mental processing of information and individual reactions when using virtual hands. Mental-, physical-, and temporal-demands almost halved when users switched from virtual hands to controller-based interactions. This suggests that controller-based interactions had better intuitiveness, contrary to our initial hypothesis, and hand-tracking was perceived as significantly frustrating and demanding. This can also be attributed to learned digital literacies (Riecke et al., 2018), since all ($N = 32$) subjects had reported prior knowledge of controllers. It would appear that though closer-to-natural, hand-tracking still required some getting-used-to before requisite proficiency for the VR task could be met. So a learnability component may have influenced perception (Drew et al., 2018). Conversely, handheld-controllers offered a more familiar, stable, and ergonomic means of grabbing the objects; and, because subjects using controllers did not expect to grab virtual objects with natural prehension, this is why effects of expectation mismatch were also absent. The results indicate that unreliable behavior of the input modality can adversely impact user performance and overall experience.

5.4. Limitations

In our study we only assessed the wireless Oculus Quest 1 due to its affordability and wide access. This can be seen as a limitation. Also because this was not the most advanced, or current, inside-out system at the time. The use of a high performing and stable system may have provided a more responsive experience. Another noticeable limitation was the requisite digital literacy of the participants. Hand-tracking technology is still in its nascency. Familiarity with the technology, understanding of the interface, and comfort of use amongst users is still far limited compared to the ubiquitous handheld-controller devices. This limitation should be met as the technology becomes mainstream. Finally, and in retrospect, only 2 out of 6 grip types coincide with the gesture supported by the VR system. The number and/or variety of prehension types used for the task could be reduced for efficiency purposes. Especially because despite the simplicity of the reach-grab-place task, an obvious limitation was the availability of only a single selection and manipulation gesture, *pinch*, on the Oculus Quest system.

6. Conclusion

In summary, our study looked at the potential for common implementation of hand-tracking VR interface by comparing it to handheld-controller in a virtual reach-grab-place task representative of real-world motor performance. We also investigated if enhancing the visual realism (level of representation) of the environment alongside natural gestures improved subjective evaluations of presence, mental workload, and ease-of-use. The results of the statistical analysis show that visual realism had no effect on user performance and surprisingly nor did it have an effect on their subjective experience of the VE. Regarding interaction, we found that input modality did bear significant influence on user performance and overall experience. Subjects took lesser attempts at grabbing virtual objects using handheld-controllers and reported efficient completion times. With hand-tracking they took longer to complete the same task and reported higher perceived mental workload scores. There was a significant difference in the ease-of-use of the two modalities. Subjects found handheld-controllers to be more task-oriented and appealing compared to hand-tracking. Lastly, the subjective feeling of immersion, perceived realism, and engagement within the VE did not differ much across the four experimental variations. Our results do not support the hypothesis of higher naturalness and user experience for hand-tracking in its current state. However, as familiarity with hand-tracking increases and technical issues are progressively overcome, this may change. Just like touch interfaces were inferior to mouse point-and-click for many years before becoming commonplace. From a research perspective, it would be interesting to see if iterative improvements in hand-tracking technologies – enhanced scope and range of available gestures – may come to surpass handheld-controllers in the future. The results of this study make a good case for taking a closer look at performative and experiential aspects of gesture-based modalities. The authors are currently investigating the particularities of manipulation in hand-tracking with respect to its action possibilities, or affordances. For future research, the perception of object affordances within VR has been highlighted as an area for investigation, because understanding object manipulation from an affordance point-of-view can help achieve interfaces and mechanisms that are effective and efficient for VR interactions.

Author agreement statement

We declare that this manuscript is original, has not been published before and is not currently being considered for publication elsewhere. We confirm that the manuscript has been read and approved by all named authors and that there are no other people who satisfied the criteria for authorship but are not listed. We further confirm that the order of authors listed in the manuscript has been approved by all of us. We understand that the Corresponding Author is the sole contact for the Editorial process. He/she is responsible for communicating with the other authors about progress, submissions of revisions and final approval of proofs.

References

Abdlkarim, D., Di Luca, M., Aves, P., Yeo, S. H., Miall, R. C., Holland, P., & Galea, J. M. (2023). A methodological framework to assess the accuracy of virtual reality hand-tracking systems: A case study with the Meta Quest 2. *Behavior Research Methods*, 1–12.

Abidi, M. H., Al-Ahmari, A., Ahmad, A., Ameen, W., & Alkhalefah, H. (2019). Assessment of virtual reality-based manufacturing assembly training system. *The International Journal of Advanced Manufacturing Technology, 105*, 3743–3759.

Alexandrovsky, D., Putze, S., Bonfert, M., Höffner, S., Michelmann, P., Wenig, D., Malaka, R., & Smeddinck, J. D. (2020). Examining design choices of questionnaires in vr user studies. In *Proceedings of the 2020 CHI conference on human factors in computing systems* (pp. 1–21).

Argelaguet, F., Hoyet, L., Trico, M., & Lécuyer, A. (2016). The role of interaction in virtual embodiment: Effects of the virtual hand representation. In *2016 IEEE virtual reality (VR)* (pp. 3–10). IEEE.

Berki, B. (2020). Experiencing the sense of presence within an educational desktop virtual reality. *Acta Polytechnica Hungarica, 17*, 255–265.

Billger, M., Heldal, I., Stahre, B., & Renstrom, K. (2004). Perception of color and space in virtual reality: A comparison between a real room and virtual reality models. In *Human vision and electronic imaging IX* (pp. 90–98). SPIE.

Biocca, F. (2002). The evolution of interactive media. Narrative impact. *Social and cognitive foundations*, 97–130.

Blaga, A. D., Frutos-Pascual, M., Creed, C., & Williams, I. (2020). Too hot to handle: An evaluation of the effect of thermal visual representation on user grasping interaction in virtual reality. In *Proceedings of the 2020 CHI conference on human factors in computing systems* (pp. 1–16).

Bowman, D. A., Kruijff, E., LaViola, J. J., & Poupyrev, I. (2001). An introduction to 3-d user interface design. *Presence, 10*, 96–108.

Bowman, D. A., McMahan, R. P., & Ragan, E. D. (2012). Questioning naturalism in 3d user interfaces. *Communications of the ACM, 55*, 78–88.

Brunnström, K., Beker, S. A., De Moor, K., Dooms, A., Egger, S., Garcia, M. N., Hossfeld, T., Jumisko-Pyykkö, S., Keimel, C., Larabi, M. C., et al. (2013). *Qualinet white paper on definitions of quality of experience*.

Buttussi, F., & Chittaro, L. (2019). Locomotion in place in virtual reality: A comparative evaluation of joystick, teleport, and leaning. *IEEE Transactions on Visualization and Computer Graphics, 27*, 125–136.

Caggianese, G., Gallo, L., & Neroni, P. (2018). The vive controllers vs. leap motion for interactions in virtual environments: A comparative evaluation. In *International conference on intelligent interactive multimedia systems and services* (pp. 24–33). Springer.

Carnevale, A., Mannocchi, I., Sassi, M. S. H., Carli, M., De Luca, G., Longo, U. G., Denaro, V., & Schena, E. (2022). Virtual reality for shoulder rehabilitation: Accuracy evaluation of oculus quest 2. *Sensors, 22*, 5511.

Chellali, A., Mentis, H., Miller, A., Ahn, W., Arikatla, V. S., Sankaranarayanan, G., De, S., Schwaitzberg, S. D., & Cao, C. G. (2016). Achieving interface and environment fidelity in the virtual basic laparoscopic surgical trainer. *International Journal of Human-Computer Studies, 96*, 22–37.

Chen, F., Lin, Y. C., Chien, J. W., & Tsai, C. E. (2016). Virtual reality for digital user experience and interactive learning based on user satisfaction: A pilot study. In *2016 international conference on computational science and computational intelligence (CSCI)* (pp. 374–377). IEEE.

Chiu, P., Takashima, K., Fujita, K., & Kitamura, Y. (2019). Pursuit sensing: Extending hand tracking space in mobile vr applications. In *Symposium on spatial user interaction* (pp. 1–5).

Coyne, R. (1994). Heidegger and virtual reality: The implications of heidegger's thinking for computer representations. *Leonardo, 27*, 65–73.

Crocetta, T. B., de Araújo, L. V., Guarnieri, R., Massetti, T., Ferreira, F. H. I. B., De Abreu, L. C., & de Mello Monteiro, C. B. (2018). Virtual reality software package for implementing motor learning and rehabilitation experiments. *Virtual Reality, 22*, 199–209.

De Paolis, L. T., & De Luca, V. (2020). The impact of the input interface in a virtual environment: The vive controller and the myo armband. *Virtual Reality, 24*, 483–502.

Drew, M. R., Falcone, B., & Baccus, W. L. (2018). What does the system usability scale (sus) measure?. In *International conference of design, user experience, and usability* (pp. 356–366). Springer.

Feick, M., Kleer, N., Tang, A., & Krüger, A. (2020). The virtual reality questionnaire toolkit. In *Adjunct publication of the 33rd annual ACM symposium on user interface software and technology* (pp. 68–69).

Feix, T., Romero, J., Schmiedmayer, H. B., Dollar, A. M., & Kragic, D. (2015). The grasp taxonomy of human grasp types. *IEEE Transactions on human-machine systems, 46*, 66–77.

Felton, W. M. (2021). *Presence in virtual environments: Visual factors and measure convergence*. Ph.D. thesis. University of Idaho.

Fromberger, P., Meyer, S., Kempf, C., Jordan, K., & Müller, J. L. (2015). Virtual viewing time: The relationship between presence and sexual interest in androphilic and gynephilic men. *PLoS One, 10*, Article e0127156. https://doi.org/10.1371/journal.pone.0127156

Furmanek, M. P., Schettino, L. F., Yarossi, M., Kirkman, S., Adamovich, S. V., & Tunik, E. (2019). Coordination of reach-to-grasp in physical and haptic-free virtual environments. *Journal of NeuroEngineering and Rehabilitation, 16*, 1–14.

Graf, S., & Schwind, V. (2020). Inconsistencies of presence questionnaires in virtual reality. In *26th ACM symposium on virtual reality software and technology* (pp. 1–3).

Grimm, F., Naros, G., & Gharabaghi, A. (2016). Closed-loop task difficulty adaptation during virtual reality reach-to-grasp training assisted with an exoskeleton for stroke rehabilitation. *Frontiers in Neuroscience, 10*, 518.

Haar, S., Sundar, G., & Faisal, A. A. (2021). Embodied virtual reality for the study of real-world motor learning. *PLoS One, 16*, Article e0245717.

Hamblin, C. J. (2005). *Transfer of* training from virtual reality environments*. Wichita State University.

Hameed, A., & Perkis, A. (2021). A subjective and behavioral assessment of affordances in virtual architectural walkthroughs. *Applied Sciences, 11*, 7846.

Hameed, A., Perkis, A., & Möller, S. (2021). Evaluating hand-tracking interaction for performing motor-tasks in vr learning environments. In *2021 13th international conference on quality of multimedia experience (QoMEX)* (pp. 219–224). IEEE.

Hart, S. G. (2006). Nasa-task load index (nasa-tlx); 20 years later. In *Proceedings of the human factors and ergonomics society annual meeting* (pp. 904–908). Los Angeles, CA: Sage publications Sage CA.

Hassenzahl, M., Koller, F., Burmester, M., 2008. Der user experience (ux) auf der spur: Zum einsatz von www. attrakdiff. de. Tagungsband UP08 ..

Holderied, H. (2017). Evaluation of interaction concepts in virtual reality applications. *Informatik, 2017*.

Holzwarth, V., Gisler, J., Hirt, C., & Kunz, A. (2021). Comparing the accuracy and precision of steamvr tracking 2.0 and oculus quest 2 in a room scale setup. In *2021 the 5th international conference on virtual and augmented reality simulations* (pp. 42–46).

Hutchins, E. L., Hollan, J. D., & Norman, D. A. (1985). Direct manipulation interfaces. *Human-Computer Interaction, 1*, 311–338.

ITU-T. (1998). *Recommendation P.911. Subjective audiovisual quality assessment methods for multimedia applications*. https://handle.itu.int/11.1002/1000/4538.

ITU-T. (2018). *Recommendation P.809. Subjective evaluation methods for gaming quality*. https://handle.itu.int/11.1002/1000/13626.

ITU-T. (2020). *Recommendation P.919. Subjective test methodologies for 360° video on head-mounted displays*. https://handle.itu.int/11.1002/1000/14429.

Izard, S. G., Juanes, J. A., Peñalvo, F. J. G., Estella, J. M. G., Ledesma, M. J. S., & Ruisoto, P. (2018). Virtual reality as an educational and training tool for medicine. *Journal of Medical Systems, 42*, 1–5.

Jankowski, J., & Grabowski, A. (2015). Usability evaluation of vr interface for mobile robot teleoperation. *International Journal of Human-Computer Interaction, 31*, 882–889.

Jerald, J. (2015). *The VR book: Human-centered design for virtual reality*. Morgan & Claypool.

Kaliki, R. R., Davoodi, R., & Loeb, G. E. (2012). Evaluation of a noninvasive command scheme for upper-limb prostheses in a virtual reality reach and grasp task. *IEEE Transactions on Biomedical Engineering, 60*, 792–802.

Kamakura, N., Matsuo, M., Ishii, H., Mitsuboshi, F., & Miura, Y. (1980). Patterns of static prehension in normal hands. *American Journal of Occupational Therapy, 34*, 437–445.

van de Kamp, C., & Zaal, F. T. (2007). Prehension is really reaching and grasping. *Experimental Brain Research, 182*, 27–34.

Kapandji, I. A. (1987). *The physiology of the joint*. lower limb. 2.

Kilteni, K., Groten, R., & Slater, M. (2012). The sense of embodiment in virtual reality. *Presence: Teleoperators and Virtual Environments, 21*, 373–387.

Kim, M., Jeon, C., & Kim, J. (2017). A study on immersion and presence of a portable hand haptic system for immersive virtual reality. *Sensors, 17*, 1141.

Kim, H. K., Rattner, D. W., & Srinivasan, M. A. (2003). The role of simulation fidelity in laparoscopic surgical training. In *International conference on medical image computing and computer-assisted intervention* (pp. 1–8). Springer.

Lackey, S., Salcedo, J., Szalma, J., & Hancock, P. (2016). The stress and workload of virtual reality training: The effects of presence, immersion and flow. *Ergonomics, 59*, 1060–1072.

LaViola, J. J., Jr., Kruijff, E., McMahan, R. P., Bowman, D., & Poupyrev, I. P. (2017). *3D user interfaces: Theory and practice*. Addison-Wesley Professional.

Levin, M. F., Magdalon, E. C., Michaelsen, S. M., & Quevedo, A. A. (2015). Quality of grasping and the role of haptics in a 3-d immersive virtual reality environment in individuals with stroke. *IEEE Transactions on Neural Systems and Rehabilitation Engineering, 23*, 1047–1055.

Liagkou, V., Salmas, D., & Stylios, C. (2019). Realizing virtual reality learning environment for industry 4.0. *Procedia CIRP, 79*, 712–717.

Liou, H. H., Yang, S. J., Chen, S. Y., & Tarng, W. (2017). The influences of the 2d image-based augmented reality and virtual reality on student learning. *Journal of Educational Technology & Society, 20*, 110–121.

Loh, C. S., Sheng, Y., & Ifenthaler, D. (2015). Serious games analytics: Theoretical framework. *Serious games analytics: Methodologies for performance measurement, assessment, and improvement, 3–29*.

Lougiakis, C., Katifori, A., Roussou, M., & Ioannidis, I. P. (2020). Effects of virtual hand representation on interaction and embodiment in hmd-based virtual environments using controllers. In *2020 IEEE conference on virtual reality and 3D user interfaces* (pp. 510–518). IEEE: VR).

Luro, F. L., & Sundstedt, V. (2019). A comparative study of eye tracking and hand controller for aiming tasks in virtual reality. In *Proceedings of the 11th ACM symposium on eye tracking research & applications* (pp. 1–9).

Masurovsky, A., Chojecki, P., Runde, D., Lafci, M., Przewozny, D., & Gaebler, M. (2020). Controller-free hand tracking for grab-and-place tasks in immersive virtual reality: Design elements and their empirical study. *Multimodal Technologies and Interaction, 4*, 91.

McMahan, R. P., Lai, C., & Pal, S. K. (2016). Interaction fidelity: The uncanny valley of virtual reality interactions. In *International conference on virtual, augmented and mixed reality* (pp. 59–70). Springer.

Mekbib, D. B., Han, J., Zhang, L., Fang, S., Jiang, H., Zhu, J., Roe, A. W., & Xu, D. (2020). Virtual reality therapy for upper limb rehabilitation in patients with stroke: A meta-analysis of randomized clinical trials. *Brain Injury, 34*, 456–465.

Merians, A. S., Poizner, H., Boian, R., Burdea, G., & Adamovich, S. (2006). Sensorimotor training in a virtual reality environment: Does it improve functional recovery poststroke? *Neurorehabilitation and Neural Repair, 20*, 252–267.

Myers, C. M., Furqan, A., & Zhu, J. (2019). The impact of user characteristics and preferences on performance with an unfamiliar voice user interface. In *Proceedings of the 2019 CHI conference on human factors in computing systems* (pp. 1–9).

Nilsson, N. C., Nordahl, R., & Serafin, S. (2016). Immersion revisited: A review of existing definitions of immersion and their relation to different theories of presence. *Human Technology, 12*.

Nunnally, M. E., & Bitan, Y. (2006). Time to get off this pig's back?: The human factors aspects of the mismatch between device and real-world knowledge in the health care environment. *Journal of Patient Safety, 2*, 124–131.

Oculus, V. (2020). Hand tracking in unity | oculus developers. URL: https://developer.oculus.com/documentation/unity/unity-handtracking/.

Oculus, V. R. (2020). *Oculus device specifications | Oculus developers*. Inc.". URL: https://developer.oculus.com/resources/oculus-device-specs/.

Pacchierotti, C., Salvietti, G., Hussain, I., Meli, L., & Prattichizzo, D. (2016). The hring: A wearable haptic device to avoid occlusions in hand tracking. In *2016 IEEE haptics symposium (HAPTICS)* (pp. 134–139). IEEE.

Pallant, J. (2020). *SPSS survival manual: A step by step guide to data analysis using IBM SPSS*. Routledge.

Palmas, F., Labode, D., Plecher, D. A., & Klinker, G. (2019). Comparison of a gamified and non-gamified virtual reality training assembly task. In *2019 11th international conference on virtual worlds and games for serious applications (VS-games)* (pp. 1–8). IEEE.

Pastel, S., Petri, K., Bürger, D., Marschal, H., Chen, C. H., & Witte, K. (2022). Influence of body visualization in vr during the execution of motoric tasks in different age groups. *PLoS One, 17*, Article e0263112.

Perez-Marcos, D., Chevalley, O., Schmidlin, T., Garipelli, G., Serino, A., Vuadens, P., Tadi, T., Blanke, O., & Millán, J.d. R. (2017). Increasing upper limb training intensity in chronic stroke using embodied virtual reality: A pilot study. *Journal of NeuroEngineering and Rehabilitation, 14*, 1–14.

Perkis, A., Timmerer, C., Baraković, S., Husić, J. B., Bech, S., Bosse, S., Botev, J., Brunnström, K., Cruz, L., De Moor, K., et al. (2020). *Qualinet white paper on definitions of immersive media experience (imex)*. arXiv preprint arXiv:2007.07032.

Pfeuffer, K., Mayer, B., Mardanbegi, D., & Gellersen, H. (2017). Gaze+ pinch interaction in virtual reality. In *Proceedings of the 5th symposium on spatial user interaction* (pp. 99–108).

Plooy, A. M., & Wann, J. P. (2000). Supporting natural prehension in virtual environments. In *Stereoscopic displays and virtual reality systems VII* (pp. 362–368). SPIE.

Pottle, J. (2019). Virtual reality and the transformation of medical education. *Future healthcare journal, 6*, 181.

Punako Jr, R., Thropp, S., . Grasping the future: Virtual hands control for fine motor tasks ..

Ragan, E. D., Bowman, D. A., Kopper, R., Stinson, C., Scerbo, S., & McMahan, R. P. (2015). Effects of field of view and visual complexity on virtual reality training effectiveness for a visual scanning task. *IEEE Transactions on Visualization and Computer Graphics, 21*, 794–807.

Regazzoni, D., Rizzi, C., & Vitali, A. (2018). Virtual reality applications: Guidelines to design natural user interface. In *International design engineering technical conferences and computers and information in engineering conference*. American Society of Mechanical Engineers, Article V01BT02A029.

Rehg, J. M., & Kanade, T. (1995). Model-based tracking of self-occluding articulated objects. In *Proceedings of IEEE international conference on computer vision* (pp. 612–617). IEEE.

Ricca, A., Chellali, A., & Otmane, S. (2017). Study of interaction fidelity for two viewpoint changing techniques in a virtual biopsy trainer. In *2017 IEEE virtual reality (VR)* (pp. 227–228). IEEE.

Ricca, A., Chellali, A., & Otmane, S. (2020). Influence of hand visualization on tool-based motor skills training in an immersive vr simulator. In *2020 IEEE international symposium on mixed and augmented reality (ISMAR)* (pp. 260–268). IEEE.

Riecke, B. E., LaViola, J. J., Jr., & Kruijff, E. (2018). 3d user interfaces for virtual reality and games: 3d selection, manipulation, and spatial navigation. In *ACM SIGGRAPH 2018 courses* (pp. 1–94).

Sassatelli, L., Winckler, M., Fisichella, T., Dezarnaud, A., Lemaire, J., Aparicio-Pardo, R., & Trevisan, D. (2020). New interactive strategies for virtual reality streaming in degraded context of use. *Computers & Graphics, 86*, 27–41.

Schlesinger, G. (1919). The mechanical structure of the artificial limbs. In *Replacement links and work aids* (pp. 321–661). knight.

Schubert, T., Friedmann, F., & Regenbrecht, H. (2001). The experience of presence: Factor analytic insights. *Presence: Teleoperators and Virtual Environments, 10*, 266–281.

Sheridan, T. B. (2016). Recollections on presence beginnings, and some challenges for augmented and virtual reality. *Presence: Teleoperators and Virtual Environments, 25*, 75–77.

Shin, Y., Lim, J., Kim, Y., Seo, D. G., & Ihm, J. (2022). Effects of virtual body-representation on motor skill learning. *Scientific Reports, 12*, Article 15283.

Slater, M. (2009). Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*, 3549–3557.

Slater, M. (2018). Immersion and the illusion of presence in virtual reality. *British Journal of Psychology, 109*, 431–433.

Sollerman, C. (1980). *The use of eight main hand grips in activities of daily living*. Thesis: Handens greppfunktion.

Steed, A., Pan, Y., Zisch, F., & Steptoe, W. (2016). The impact of a self-avatar on cognitive load in immersive virtual reality. In *2016 IEEE virtual reality (VR)* (pp. 67–76). IEEE.

Suznjevic, M., Mandurov, M., & Matijasevic, M. (2017). Performance and qoe assessment of htc vive and oculus rift for pick-and-place tasks in vr. In *2017 Ninth international conference on quality of multimedia experience (QoMEX)* (pp. 1–3). IEEE.

Sveistrup, H. (2004). Motor rehabilitation using virtual reality. *Journal of NeuroEngineering and Rehabilitation, 1*, 1–8.

Thorsteinsson, G. (2013). Developing an understanding of the pedagogy of using a virtual reality learning environment (vrle) to support innovation education. In *The Routledge International handbook of innovation education* (pp. 486–500). Routledge.

User interface design - UID GmbH". AttrakDiff. URL: http://www.attrakdiff.de/index-en.html.

Våpenstad, C., Hofstad, E. F., Langø, T., Mårvik, R., & Chmarra, M. K. (2013). Perceiving haptic feedback in virtual reality simulators. *Surgical Endoscopy, 27*, 2391–2397.

Viau, A., Feldman, A. G., McFadyen, B. J., & Levin, M. F. (2004). Reaching in reality and virtual reality: A comparison of movement kinematics in healthy subjects and in adults with hemiparesis. *Journal of NeuroEngineering and Rehabilitation, 1*, 1–7.

Voigt-Antons, J. N., Kojic, T., Ali, D., & Möller, S. (2020). Influence of hand tracking as a way of interaction in virtual reality on user experience. In *2020 twelfth international conference on quality of multimedia experience (QoMEX)* (pp. 1–4). IEEE.

Wang, Y., Wu, Y., Jung, S., Hoermann, S., Yao, S., & Lindeman, R. (2021). Enlarging the useable hand tracking area by using multiple leap motion controllers in vr. *IEEE Sensors Journal*.

Weber, S., Weibel, D., & Mast, F. W. (2021). How to get there when you are there already? Defining presence in virtual reality and the importance of perceived realism. *Frontiers in Psychology, 1538*.

Zhang, H. (2017). Head-mounted display-based intuitive virtual reality training system for the mining industry. *International Journal of Mining Science and Technology, 27*, 717–722.

Zheng, B., Jiang, X., Tien, G., Meneghetti, A., Panton, O. N. M., & Atkins, M. S. (2012). Workload assessment of surgeons: Correlation between nasa tlx and blinks. *Surgical Endoscopy, 26*, 2746–2750.