




# A comparative analysis of two immersive virtual reality systems in the integration and visualization of natural hand interaction

Guillermo Molina<sup>1</sup> · Jesús Gimeno<sup>1</sup> · Cristina Portalés<sup>1</sup>  · Sergio Casas<sup>1</sup>

Received: 19 May 2021 / Revised: 16 September 2021 / Accepted: 22 November 2021 /

Published online: 28 January 2022

© The Author(s) 2022

## Abstract

It is generally accepted that the use of natural interaction provides a positive impact in Virtual Reality (VR) applications. Therefore, it is important to understand what is the best way to integrate and visualize this feature in VR. For this reason, this paper presents a comparative study of the integration of natural hand interaction in two immersive VR systems: a Cave Audio Visual Experience (CAVE) system –where users’ real hands are visible– and a non-see-through Head-Mounted Display (HMD) system –where only a virtual representation of the hands is possible–. In order to test the suitability of using this type of interaction in a CAVE and compare it to an HMD, we raise six research questions related to task performance, usability and perception differences regarding natural hand interaction with these two systems. To answer these questions, we designed an experiment where users have to complete a pick-and-place task with virtual balls and a text-typing task with virtual keyboards. In both systems, the same tracking technology, based on a Leap Motion device, was used. To the best of our knowledge this is the first academic work addressing a comparison of this type. Objective and subjective data were collected during the experiments. The results show that the HMD has a performance, preference and usability advantage over the CAVE with respect to the integration of natural hand interaction. Nevertheless, the results also show that the CAVE system can be, as well, successfully used in combination with an optical hand tracking device.

**Keywords** Virtual reality · CAVE · HMD · HTC vive · Leap motion · Hand tracking · Natural interface · Comparative analysis

---

✉ Cristina Portalés  
Cristina.Portales@uv.es

<sup>1</sup> Department of Computer Science, University of Valencia (UV), Valencia, Spain

# 1 Introduction

It is generally accepted that the use of natural interfaces in Virtual Reality (VR) systems provides a positive impact in the use of these technologies. With these interfaces users can perform direct manipulations of virtual objects without the need of additional tools, bridging the human-computer barrier [33]. Recent advances in hand gesture recognition make it possible to consider user's hands to manipulate virtual objects in different VR systems. The integration of this technology is becoming common in those VR systems based on non-see-through Head-Mounted Displays (HMD). The integration in VR setups based on Cave Audio Visual Experience (CAVE) systems [11] is also possible, although it is more challenging due to the differences between the two systems. Whereas users in a CAVE directly see their real hands, in non-see-through HMDs only a virtual representation is possible. Therefore, the question arises as to which one of the two systems is better for manipulating virtual objects, since the visualization of the real hands is expected to enhance the sense of embodiment. However, the real hands will partially occlude the virtual objects, making the interaction potentially challenging.

There is a lack of research works comparing such systems. For this reason, this paper presents a comparative study of two modes of visualizing user's hands (virtual vs. real) by means of HMD and CAVE systems integrating the same tracking technology, a Leap Motion device, in both immersive setups. To the best of our knowledge this is the first academic work reporting the use of this hand tracking device in a CAVE system and the first comparison of this type. This reflects the novelty of our work and is one of its contributions.

The objective of this work is two-fold: (i) test the suitability of using optical-based natural hand interaction in the CAVE; (ii) compare natural hand interaction between a CAVE (where real hands are visible) and an HMD (where virtual hands are used instead). In this regard, we address six research questions related to task performance, usability and perception differences between the two systems. We also study how the sense of embodiment changes when users see their own real hands (in the CAVE) instead of using virtual hands (in the HMD). To answer these questions, we designed an experiment where users have to complete two different tasks with virtual objects: a pick-and-place task, and a text-typing task. These tasks correspond with two of the most common interactions in VR systems: pick and touch. Different sizes are also considered for the virtual objects.

The results are analyzed in both objective (measuring errors and completion times) and subjective ways (with questionnaires), proving that both systems can be successfully used in combination with a hand tracking device. However, the HMD has a performance and usability advantage over the CAVE with respect to the integration of natural hand interaction, and users show a clear preference for the HMD-based system. This is the main contribution of this work.

The main challenges of this work have been the design and development of all the necessary elements to integrate the hand tracking device (Leap Motion) in a CAVE system, and the design and execution of an experiment to compare both systems. All this is described in this paper, which is organized as follows. Section 2 reviews related works about comparative studies in VR, with special focus on hand interaction, CAVEs and HMDs. Section 3 describes the materials and methods utilized to develop the tools employed in the experiments. Section 4 details the experimental study. In section 5, the results of these experiments are presented and discussed. Finally, section 6 outlines the future work and shows the conclusions of the paper.

## 2 Related work

Both CAVE and HMD systems have been widely used to provide visual cues and interaction in virtual environments (VE), but sometimes they are used indiscriminately regardless of the research objective and it is not well known if they are equally effective [29]. Their different performance in specific setups and experiments has been reported before, but few works make use of natural hand interaction to manipulate the virtual objects and do not focus on studying the impact of visualizing virtual vs. real hands. There are other works reporting this, i.e., virtual vs. real hands, although not with the consideration of both CAVE and HMD systems as in our work. In this section a review of such related works is provided, which are further summarized in Table 1 for the sake of clarity.

An early study comparing CAVE vs. HMD systems is presented in [7], which showed an experiment addressing human behavior and performance when doing turns in a virtual environment. Authors found that users showed a significant preference for real turns in the HMD and for virtual turns in the CAVE, concluding that the HMD produces higher levels of spatial orientation and may make navigation more efficient. In [18] a comparative study of the levels of presence and anxiety in an acrophobic environment is presented, demonstrating that the CAVE induces a higher level of presence and more anxiety in users than the HMD. The work described in [19] aims to investigate the effects of these systems on emotional arousal and task performance in a modified Stroop task presented under low- and high-stress conditions. The experiments showed that the fully immersive system (a six-wall CAVE) induced the highest sense of presence, while the HMD elicited the highest amount of simulator sickness. In [13] a comparison of the influence of the two systems in the specific task of selecting a preferred seat in an opera theater is presented. Results show that the subjects taking part in the experiment could select similar seats, but their decisions differed between both environments, so no advantage between the two systems could be defined. In [32] it is investigated the nature of the user experience of viewing panoramic videos in these systems. The authors studied which viewing condition users preferred and concluded that their opinion was equally split between the two systems. In [24, 29] different studies are presented to study the behavior of street crossing, both finding the HMD setup more advantageous over the CAVE. Other works that bring comparative studies between CAVE- and HMD-based systems can be found in [1, 15, 16, 26, 27, 34, 37].

The question of which system should be recommended (a CAVE or an HMD) is too broad to get a conclusive answer, as it depends on many factors. For instance, as can be seen from the above reported papers, the pros of one or other system might depend on the purpose of the experiment. This fact is reported for instance in [19], where authors argue that CAVEs can be useful to evoke happy and positive emotions, whereas HMDs can be advantageous to evoke negative emotions. Another example can be found in [15], which aims at measuring the accuracy of perceived distances in a virtual environment, where the results show that the HMD provides the best results for distances above 8 m whereas the CAVE provides the best results for close distances.

In addition, the maturity of one or the other system over time is important. In 2015, [3] reported that current HMDs were able to compete with many CAVEs and had started to take over them, something that was not true only a few years before. Also, in 2017, [10] reported that HMD displays will soon be commodity devices that may represent a low-cost alternative to the CAVE system. This can be considered true, now. Additionally, the technological features of the systems taking part in the experiments might be of relevance (e.g., high- or low-cost devices). For instance [27] used in the experiments a six-sided CAVE,

**Table 1** Summary of the related work papers. For those that compare HMD and CAVE systems, an asterisk (\*) means that such system is found to be advantageous over the other

Paper	Research purpose	HMD	CAVE	Hand tracking and/or interaction with virtual objects
[29]	Compare street-crossing behavior	Yes*	Yes	HMD: Vive controllers CAVE: None / not relevant
[7]	Behavior and performance when doing turns	Yes*	Yes	None / not relevant
[18]	Levels of presence and anxiety in an acrophobic VE	Yes	Yes*	None / not relevant
[19]	Emotional arousal and task performance in a Stroop task	Yes	Yes	HMD: wand CAVE: integrated hand tracking
[13]	Influence in selecting a seat in an opera theater	Yes	Yes	HMD: gaze-based raycasting. CAVE: virtual wand, headtracking
[32]	User experience in viewing panoramic videos	Yes	Yes	None / not relevant
[24]	Compare crossing traffic-filled roadways behavior	Yes*	Yes	None / not relevant
[27]	Compare presence in a Virtual Game-Day Experience	Yes	Yes	HMD: game pad controller CAVE: not specified
[37]	User experience in an edutainment VE	Yes	Yes*	HMD: crosshair CAVE: flystick
[15]	Accuracy of perceived distances in a VE	Yes	Yes	HMD: keyboard CAVE: flystick
[34]	Compare performance for evacuation research	Yes*	Yes	HMD & CAVE: controller
[16]	Training in fighting enclosure fires	Yes*	Yes	HMD & CAVE: Vive controllers
[11]	Depth and distance perceptions in architectural design	Yes	Yes*	None / not relevant
[26]	Pros and cons of the technology (e.g., field of view, latency)	Yes	Yes	None / not relevant
[10]	Collaborative analysis of network connectivity	Yes	Yes	HMD: hand gesture (Leap Motion on user's head) CAVE: wand
[14]	Exergaming for adults with and without disabilities	Yes*	Yes	HMD: Vive controllers CAVE: Optitrack wrist markers
[30]	Neurological effects of visualizing actions made by a virtual vs. a real hand	Yes	No	HMD: markers (to produce an animation)
[31]	Body continuity and arm alignment on the virtual hand illusion	Yes	No	HMD: wand
[6]	Thresholds for fixed offsets for virtual hands in VE	Yes	No	HMD: Oculus touch controllers
[22]	Impact related to the size of users virtual and real hands	Yes	No	HMD: tracked gloves and touch controllers
[2]	Study sense of embodiment. Pick-and-place task	Yes	No	HMD: Hand gesture (Leap Motion on a fixed position)
[4]	Manipulate multiple virtual objects with passive haptics	Yes	No	HMD: Hand gesture (Kinect on a fixed position)
[20]	Develop a framework for interacting in CAVE	No	Yes	CAVE: Hand gesture (Kinect on a fixed position)
[35]	Accuracy of commodity finger tracking systems	Yes	No	HMD: Vive controllers and hand gesture (Leap Motion on user's head)
[23]	Comparing controllers for HMDs	Yes	No	HMD: different controllers
[9]	Distance perception	Yes	Yes	HMD: Vive controllers CAVE: flystick
Our work	Performance in pick-and-place and text-typing tasks. Size of virtual objects. Usability	Yes	Yes	HMD & CAVE: Hand gesture (Leap Motion on user's head)

whereas [24, 32, 37] used four-sided CAVEs, being all these works reported in 2017. Also for the same year, the experiments in [13, 32] were done with a Samsung Gear VR, [24] used an HTC Vive and in [27] a VR smartphone-driven HMD was also used.

In [26] the pros and cons of both systems are addressed regarding different aspects, such as field of view, temporal delay and latency, or spatial and temporal resolution. Authors argue that it is difficult to say which one is best and that both technologies have limitations. Other researchers suggested that current technological features (e.g. in displays and tracking) might lead to the need to re-study old experiments in VR [12]. In this sense, it is important to highlight that our experiments focus on general aspects that differ in both kinds of systems (e.g., usability or visual perception of real vs. virtual hands) rather than on specific issues related to technological solutions, in order to make the results as general as possible, with the unavoidable limitations that this might have.

None of the works reviewed above focus on the perception of user's virtual vs. real hands with these two systems, which is one of the contributions of our paper. This can be explained because in many works the interaction with the virtual objects is performed by means of an artifact instead of by means of the hands. This is evidenced in Table 1—where the technology used for hand interaction with the virtual objects is reported in the last column— and represents a clear gap in the research. For instance, in [19] participants interacted with the virtual objects through the use of a wand in the HMD setup, and this is also the case for many CAVEs. In [14], the participants using the CAVE setup wore head and wrist Optitrack markers, whereas the participants in the HMD setup used the Vive controllers. In other works it is not clear how users interact with the virtual objects, or they only use hand tracking for one of the setups, as in [10], where an Oculus Rift —with a Leap Motion device attached to it— is compared to a CAVE, but the users in the CAVE-based setup use 6-DOF controllers instead of hand tracking. Even in this case, users are not asked about the different experiences related to their own hands' perception. Therefore, we believe that our contribution is relevant to the field.

There are other works that deal with perception issues regarding the visualization of virtual and/or real hands, although not with the consideration of the VR systems studied in this work (CAVE vs. HMD). For instance, in [30] the neurological effects of visualizing actions made by a virtual vs. a real hand are investigated. Results showed that only the perception of actions made in reality maps onto existing action representations, while VR conditions do not access the full motor knowledge. However, in this work the experiment does not consider user's own hands in neither of the two visual representations. In [31] an experiment is presented on a virtual hand illusion. Results showed that the subjective illusion of ownership over the virtual arm and the time to evoke this illusion are highly dependent on synchronous visuo-tactile stimulation and on the connectivity of the virtual arm with the rest of the body. They also found that the alignment between the real and virtual arms and the distance between these were less important. Differently from our work, in this experiment only the visualization of the virtual arms is considered.

Another interesting experiment is conducted in [6], where the authors studied the effect of applying fixed positional offsets to the location of a virtual hand. The study provides estimations for detection thresholds for positional hand offsets. Although the experiment used only one device (an Oculus Rift HMD) and a semi-natural interaction by means of the Oculus Touch controllers, the results are interesting because significant differences in offset detection were identified based on offset direction. A somehow similar experiment is presented in [22] but focusing on the size of the virtual hands. As before, an Oculus Rift HMD is used, although the authors use a pair of motion capture gloves for hand tracking.

Another kind of experiment is conducted in [2], where the authors studied the sense of embodiment when interacting with three different virtual hand representations, one of them being highly realistic. Results showed that the sense of agency was stronger for less realistic virtual hands, whereas the sense of ownership was increased for the most realistic representation of the hand. It is worth mentioning that in that paper, the experiment consisted in a pick-and-place task and they use a Leap Motion device, as in our case. However, they only use an HMD and therefore they do not compare with the visualization of users' real hands, as in a CAVE-based system. A similar work can be found in [4]. Only one work has been found reporting the use of the CAVE with a Kinect sensor in order to include hand gesture recognition for controlling the CAVE [20]. However, in [20] the sensor is not placed on the user's head but in a fixed place on one of the walls, and thus, this limits the ability to use it for natural hand interaction due to occlusion, low resolution and other problems.

Other related works can be found comparing or analyzing the accuracy of hand tracking in immersive VR systems [35], comparing controllers for HMDs [23] or comparing distance perception between a CAVE and a HMD [1, 9], but to the best of our knowledge, no previous works exist performing the kind of experimental research that we propose here.

### 3 Materials and methods

As previously mentioned, two different hardware setups are used in this research: an HMD-based system and a CAVE-based system. Both are enhanced with natural hand interaction by means of a Leap Motion device. These two setups represent the most common immersive systems used in VR applications.

#### 3.1 HMD-based system

The first hardware setup utilizes an HTC Vive. This is an HMD featuring two  $1080 \times 1200$  pixels resolution AMOLED displays, with a 90 Hz refresh rate and  $110^\circ$  field of view with stereoscopic vision. This device includes also two wireless controllers to handle the interaction with the virtual world (which are not used in this research) and two lighthouse base stations that allow tracking both the user head and the wireless controllers by sending pulsed infrared light. To perform a fair comparison, the HTC Vive was configured so that the dimensions of the virtual space match the space available in the CAVE used in the experiments ( $6.25 \text{ m}^2$ ). This way, both setups provide the same amount of navigation space.

In order to provide hand-tracking capabilities, we attached a Leap Motion device (a robust state-of-the-art hand tracking device [40]) to the HMD (see Fig. 1) in a known position, close to the user's eyes. The final setup can be observed in Fig. 2. Differently from the CAVE system, the users of this setup are not able to see their real hands while using the HMD because their visual senses are completely extracted from the real world and users cannot see their own body.

#### 3.2 CAVE-based system

The second hardware setup utilizes the CAVE system installed at the Institute of Robotics and Information and Communications Technologies (IRTIC [39]) of the University of Valencia. This is a four-sided CAVE with four  $2.5 \times 2.5 \text{ m}$  screens, four  $1050 \times 1050$  pixels resolution



**Fig. 1** An HTC Vive enhanced with a Leap Motion device

projectors, four infrared cameras and a pair of stereoscopic glasses with infrared markers used for tracking the position of the user (see Fig. 3).

In order to provide also hand-tracking capabilities for this system, we designed and 3D-printed a specific mount to attach the Leap Motion to the pair of glasses used in the CAVE for providing stereoscopic rendering (see Fig. 4). This allows also to place the Leap Motion in a



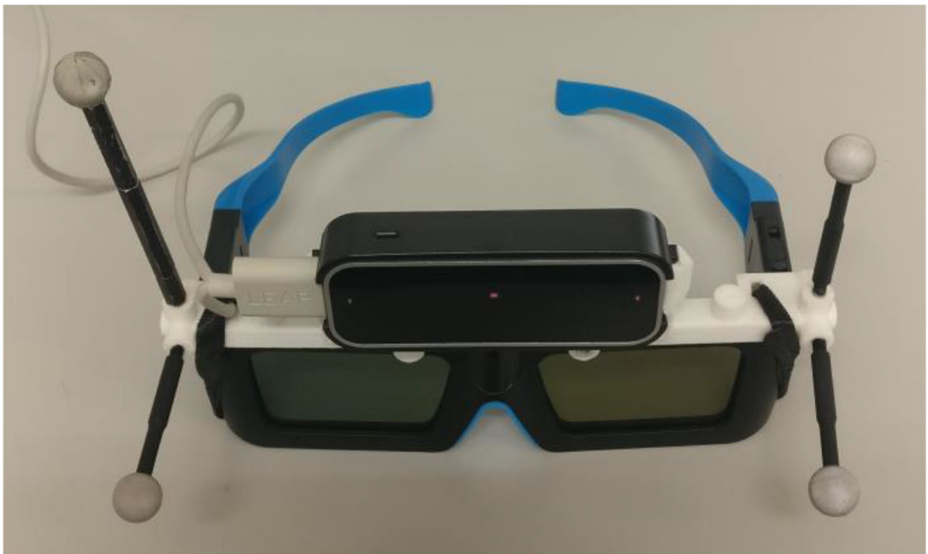
**Fig. 2** A person using the HMD-based setup during the experiments



**Fig. 3** A person using the CAVE system of the University of Valencia during the experiments

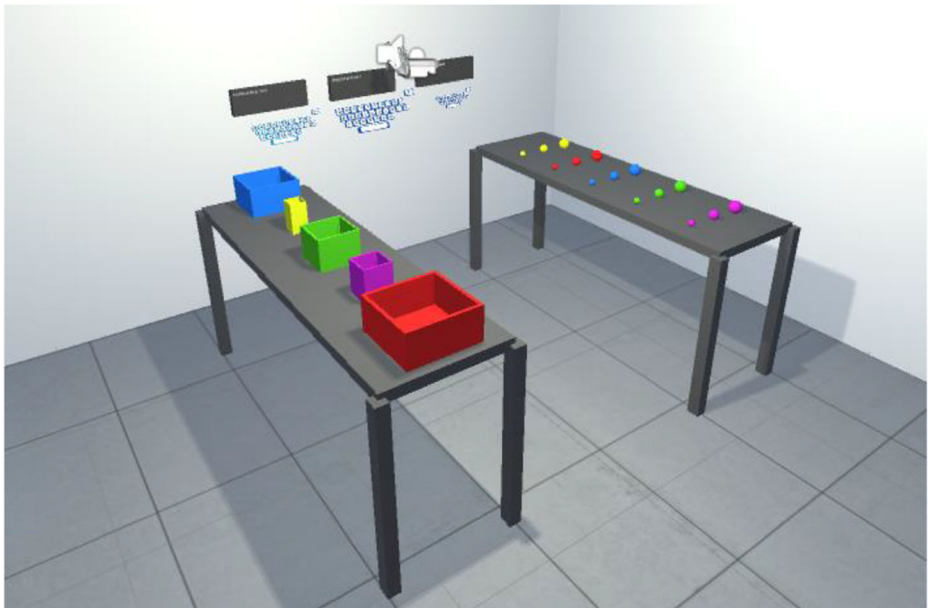
known position with respect to the user's point of view. This particular solution was required because no previous works (or hardware) were found reporting the use of a Leap Motion device with a CAVE system.

One of the features of this setup is that it is possible that users see both their virtual hands (rendered with the data provided by the Leap Motion) and their real hands at the same time. This happens, because, unlike in an HMD, the CAVE allows seeing both a virtual scene and



**Fig. 4** CAVE glasses with a Leap Motion device attached to them





**Fig. 5** The Unity3D scene designed for the experiments

one's real body. Nevertheless, we decided not to show the virtual hands in this setup, since the goal of the experiment is to compare the two visualization modes (virtual hands vs. real hands).

### 3.3 Software

The software application has been implemented using Unity3D [38] for both setups, since the same Unity scene can be shared and deployed with different hardware. Figure 5 shows the virtual scenario created for both setups. In the HMD-based system, Unity offers an automatic Mixed Reality mode where stereo view and tracking are implemented. However, in the CAVE-based setup, it is necessary to implement a custom viewer with eight virtual cameras. In a CAVE, each wall is like a window to the virtual world, so each virtual camera renders what each of the two eyes would see through the corresponding window (hence  $4 \times 2 = 8$  cameras are needed). Stereo active glasses are tracked by a Natural Point infrared system by means of a custom marker structure attached to the glasses (see Fig. 4), so the position of each eye and the position of each screen are used to calculate the correct view using an asymmetric frustum.

In the case of the Leap Motion, the manufacturer provides all the packages needed to use this device in Unity. In this regard, it is very important to create the spatial relationship between the sensor and the user's eyes, because the sensor is moving fixed to the user's head. A key aspect in this research is how the virtual hands interact with the virtual elements, to press the keys or to grab the balls. In the first case, each key detects the collision with the mesh of the virtual hand, so the virtual hand literally pushes the virtual key. In the second case, the virtual hands do not grab the balls, because the physics simulation gets very unstable when a rigid sphere is compressed by two or more fingers. To grab objects, the manufacturer recommends using a pinch gesture, so when the virtual hand makes a pinch gesture close to the ball, this object is attached –with a virtual joint– to the hand until the pinch gesture ends. Using this technique, the physics simulation is more stable.

In order to offer a better user experience, visual and acoustic feedback was implemented. Each action –grab or release a ball, press a key, place a ball correctly or not– is associated to a convenient sound. For instance, a sound is reproduced when any key is pressed. A different sound is reproduced when the desired text is correctly typed by the user. Visual feedback has been implemented as well: each key turns gray when any part of the virtual hands is less than 1 cm away from the key, and turns green when it is in contact with the fingers. This is very useful because users can see the key in gray color before making the push gesture on the key. In this case, the sound feedback reveals when the key has been pushed with enough force to write a character.

## 4 Experiments

As previously explained, the goal of the experiment is to study the suitability of enhancing two existing VR systems –a CAVE and an HMD– with hand tracking capabilities and test if there are statistically significant differences in the use of the two different visualization modes (virtual vs. real hands) that these devices represent. Several research questions are formulated and tested in this research, which are enunciated next.

Research Question 1 (RQ1): Which system will allow a faster completion of the tasks?

Research Question 2 (RQ2): With what system will users complete the tasks with fewer mistakes?

Research Question 3 (RQ3): Will the size of the virtual elements affect the performance of the users?

Research Question 4 (RQ4): Which system will be more usable?

Research Question 5 (RQ5): Which system will users prefer?

Research Question 6 (RQ6): Which system will users report to have a higher sense of embodiment with respect to their hands?

In order to answer these questions, a series of experiments were designed and carried out at the IRTIC laboratory. The procedure follows a two-treatment counterbalanced within-subjects design and it is explained in the following subsections.

### 4.1 Participants and procedure

17 adult people (i.e., over 18 years-old) –recruited through social media– participated in the experiment. 9 were women (52.94%) and 8 were men (47.06%), with ages ranging from 19 to 58 (mean  $36.88 \pm 15.45$ ). The age distribution is shown in Table 2.

Before the start of the experiment, users were asked to rate in a 1–5 Likert scale [21] their previous experience with these technologies, with 1 meaning no previous experience, 2 little experience, 3 some experience, 4 enough experience and 5 meaning expert. Table 3 summarizes the results and shows that there was no experience bias in favor of either system and the majority of the participants had none or little previous experience in the use of these two systems.

In order to avoid carryover effects and biases for either of the systems, users were counterbalanced and randomly assigned to two groups of roughly equal size: group A (9 participants) and group B (8 participants). Users in group/treatment A tested first the CAVE-based system and then the HMD, whereas users in group/treatment B tested first the HMD-based system and then the CAVE.

**Table 2** Age distribution of the participants of the experiment

Age group	<20	20–29	30–39	40–49	50–59
%	5.88	41.18	11.76	5.88	35.29
Number	1	7	2	1	6

The experiment was divided into two similar sections, one section for each hardware setup. Therefore, users tested first one system, filled a questionnaire about that system, then tested the other system, filled a similar questionnaire about the use of the second system and finally filled a comparative questionnaire.

The experiment consisted in eight separate phases that are explained next.

**Phase 1** Presentation and description. Before proceeding with the experiment, the participants were provided with a short description about the tasks they had to complete and the maximum time available to complete them (10 min in total). Then, users were required to sign an informed consent where they declared to agree with the terms of the experiment. They also filled an anonymous questionnaire in order to provide some basic demographics (gender, age and previous experience with VR), as shown in Tables 2–3.

**Phase 2** Instruction and practice. Before the start of each section, users received a short briefing on how to use the CAVE system or the HMD-based system (depending on which one they were required to use). Then, they tested the system for 5 min, where they were free to practice the different actions allowed by the system (touch, pick, release, walk, look, type, etc.).

**Phase 3** Experiment. The experiment consisted in two different tasks, which are explained in the next subsection, that the participants needed to complete in the corresponding hardware setup.

**Phase 4** Setup evaluation. After users finished the experiment in the first device, they were asked to fill a System Usability Scale (SUS) questionnaire (see Table 4). SUS is a de facto standard questionnaire which consists of a series of five-response questions about the usability of a system [8]. In addition, during the experiment, the software application collected objective data about the two tasks, which was saved to a CSV file when the user finished the experiment.

**Phases 5–7** Once the participant finished the experiment with the first hardware setup, all the previous phases of the process were repeated (except from the presentation phase) using the other hardware setup.

**Phase 8** Final comparative evaluation. Finally, when the participants finished the experiment with the second hardware setup, they were asked to fill the comparative questionnaire shown

**Table 3** Previous experience of the participants

Experience	No	Little	Some	Enough	Expert
CAVE (%)	47.06	17.65	17.65	5.88	11.76
CAVE (number)	8	3	3	1	2
HMD (%)	52.94	17.65	11.76	11.76	5.88
HMD (number)	9	3	2	2	1

**Table 4** Questions of the System Usability Scale (SUS)

#	Question
1	I think that I would like to use this system frequently.
2	I found the system unnecessarily complex.
3	I thought the system was easy to use.
4	I think that I would need the support of a technical person to be able to use this system.
5	I found the various functions in this system were well integrated.
6	I thought there was too much inconsistency in this system.
7	I would imagine that the most people would learn to use this system very quickly.
8	I found the system very cumbersome to use.
9	I felt very confident using the system.
10	I needed to learn a lot of things before I could get going with this system.

in Table 5. The questionnaire included also an open-ended question “Additional comments and explanations” for the users to leave their comments, impressions and explanations about the experiments, the systems or about the different answers provided for the three two-choice questions of this questionnaire.

## 4.2 Tasks

In Task 1, users were instructed to grab a series of colored balls from a table and place them in different colored square boxes (see Fig. 5) placed on another table a few feet away. Fifteen balls with three different ball sizes (with diameters of 7 cm, 5.25 cm and 3.5 cm, respectively) in five colors (red, blue, green, purple and yellow) and different box sizes (32 cm, 25 cm, 18 cm, 12 cm and 8 cm, respectively) are used. The goal of this task is to move all colored balls from the table to the corresponding colored box. Users can choose any ball, so the order in which the boxes are filled is not important. The colored boxes and the ball sets were randomly presented to the users and no participants were color-blinded.

In order to provide feedback to the user, different sounds are played when a sphere is touched, when it is successfully picked, when it is released in the correct box and when it is released in the wrong place. Table 6 summarizes the objective data that the software application collects during the execution of this task. The time dataset measures the number of seconds elapsed between the moment when the user first picked a sphere and the moment when it was successfully placed in the right box. As there are three ball sizes, this dataset is divided in datasets 1.S (small size), 1.M (medium size) and 1.L (large size). Each of these contains five values per user, since there are five balls for each size (one per color).

The errors dataset counts the total number of times the user released a sphere outside the right target box. This dataset is also split in datasets 2.S, 2.M and 2.L (one per size). Finally,

**Table 5** Comparative questionnaire

#	Question
1	Which of the two systems have you found more comfortable to use?
2	Which of the two systems did you like more?
3	With which of the two systems did you have a greater sense of embodiment with respect to your hands?

**Table 6** Objective data collected during Task 1

<i>Ids.</i>	<i>Type</i>	<i>Description</i>
1.S 1.M 1.L	float [18]	Time needed for the user to correctly place each of the small, medium-sized and large spheres
2.S 2.M 2.L	int	Total number of errors for the small, medium-sized and large spheres
3.S 3.M 3.L	int	Total number of pick-up gestures for the small, medium-sized and large spheres

S = Small. M = Medium-sized. L = Large

the pick-up gestures dataset counts the total number of pinch gestures produced by the user (regardless of the success of the action). This dataset is also split in three: 3.S, 3.M and 3.L.

In Task 2, users had to type a 20-character text string using a virtual keyboard. Three virtual keyboards with 29 keys (the 27 letters of the Spanish alphabet plus the space key and the backspace key) and different sizes (letter keys of size 38 mm, 28 mm and 18 mm, respectively) were created for this task. As in Task 1, there is not a predefined order in which the three keyboards should be typed and users were presented the three keyboards randomly sorted. Table 7 summarizes the objective data that the software application collects during the execution of this task.

All these datasets were chosen because they should reflect well the performance of the users in an objective way.

## 5 Results and discussion

This section presents the results of the analysis performed using the data obtained in the different experiments. The first three research questions can be answered with the objective data gathered during the experiments. The rest of the research questions are analyzed by means of the SUS questionnaire and the two-choice comparative evaluation.

### 5.1 Statistical analysis of objective data

In order to compare the two systems and test if there are objective significant differences between the two VR systems, statistical analysis with the objective data collected in the

**Table 7** Objective data collected during Task 2

<i>Ids.</i>	<i>Type</i>	<i>Description</i>
4.S 4.M 4.L	float	Time needed for the user to correctly type the text in the small, medium-sized and large virtual keyboard
5.S 5.M 5.L	int	Number of errors made by the user before successfully typing the text in the small, medium-sized and large virtual keyboard

S = Small. M = Medium-sized. L = Large

experiments were performed, using IBM SPSS 26 software [17]. All the statistical tests were two-tailed and were conducted at the 0.05 significance level.

First, it was checked if the different datasets follow a normal distribution. For datasets with less than 50 samples, the Shapiro-Wilk test [36] was used, whereas the Kolmogorov-Smirnov test [25] was used for datasets with 50 or more than 50 samples (as suggested in [28]). On the one hand, with those datasets that follow a normal distribution, parametric paired two-sample t-test were applied. On the other hand, non-parametric tests were used for non-normally distributed datasets: the Wilcoxon signed-rank (for paired datasets with two variables) and the Friedman test (for paired datasets with more than two variables).

**Task 1** None of the paired datasets obtained for the pick-and-place task passed the normality tests. Therefore, Wilcoxon signed-rank tests were used to compare the performance of the participants. Table 8 and Fig. 6 summarizes these tests, where mean time (averaged for all fifteen spheres of each user), total number of errors (of each user) and total number of pick-up gestures (of each user) are compared between the two VR systems. Results show that the null hypothesis of these tests (“the median of the differences between the paired groups is zero”) should be rejected. Thus, it can be assumed that there are statistically significant differences in both time, total number of errors and total number of pick-up gestures in favor of the HMD-based system. These results provide answers to RQ1, since users clearly need less time when they use the HMD-based system, and to RQ2, since there are statistically significant differences in the number of errors and in the number of pick-up gestures needed to complete the task in favor of the HMD-based setup.

In order to study the differences between the two systems for each sphere size and analyze the effect of size in the performance of the participants, Friedman tests (since the datasets grouped by sphere size were not normally distributed) were performed for the same variables: time needed for each user and each sphere size ( $16 \times 5 = 80$  values for each size), total number of errors (16 values per size) and total number of pick-up gestures (16 values per size).

Results are depicted in Table 9 (where time is measured), Table 10 (where errors are shown) and Table 11 (showing pick-up gestures). In five of the six tests, there are statistically significant differences between, at least, two sphere sizes. Figures 7, 8, and 9 show the box plots corresponding with these experiments. As can be seen, the smaller the size of the spheres, the more time, errors and pick-up gestures are generally needed to complete this pick-and-place task, although in the case of the HMD, the significance level is slightly above the 0.05 limit for the errors datasets.

These results suggest a possible answer to RQ3, since the size of the spheres has a significant effect in the performance of the users for this task.

**Table 8** Results of the Wilcoxon signed-rank test for Task 1

<i>Parameter</i>	<i>System</i>	<i>Median ± IQR</i>	<i>Z (sig.)</i>
<i>1 – Time (s)</i>	<i>CAVE</i>	28.56 ± 67.66	−3.206
	<i>HMD</i>	12.41 ± 9.77	(.001)
<i>2 – Errors</i>	<i>CAVE</i>	22.00 ± 26.00	−2.096
	<i>HMD</i>	14.50 ± 16.75	(.036)
<i>3 – Pick-up gestures</i>	<i>CAVE</i>	87.00 ± 85.50	−3.258
	<i>HMD</i>	54.50 ± 41.25	(.001)

Number of samples = 16 (one user did not finish this task)

Degrees of freedom = 15

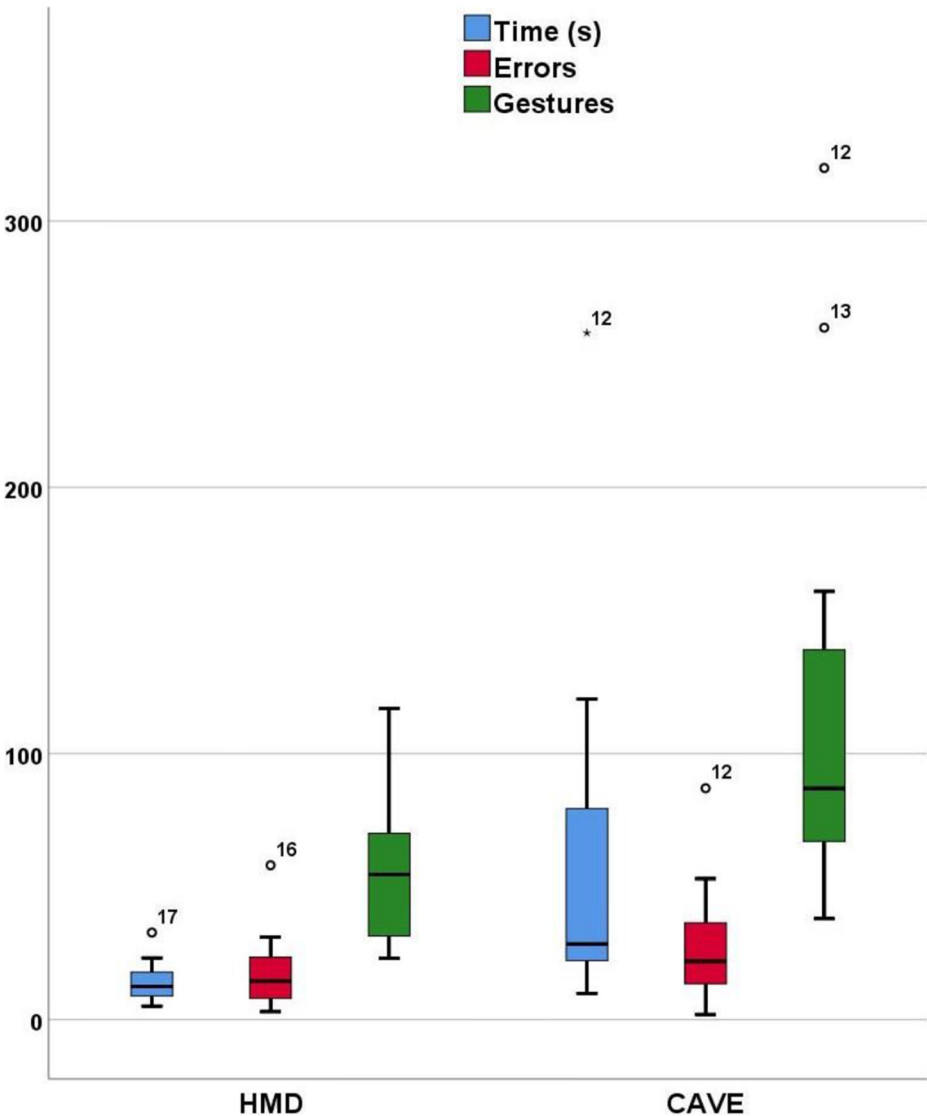


Fig. 6 Box plots for the datasets of Task 1

**Task 2** A similar analysis is performed for Task 2. In this case, however, the Shapiro-Wilk test reveals that the dataset with the total number of errors for each participant (regardless of the size of the virtual keyboards –all the errors made in all three virtual keyboards are accounted and summed for each participant–) can be assumed to follow a normal distribution in both systems (CAVE:  $W = .893, p = .052$ ; HMD:  $W = .951, p = .474$ ). The same happens with the dataset measuring time for the HMD system (mean time needed for each participant to successfully type the same text, averaged for the three keyboard sizes) ( $W = .913, p = .111$ ), although the same dataset in the CAVE system does not pass this normality test by a very small margin ( $W = .890, p = .047$ ). For this reason, paired t-tests were applied for both variables and the Wilcoxon signed-rank test was also applied for the time dataset.

**Table 9** Results of the Friedman test for time in Task 1 (1.S, 1.M, 1.L)

<i>System</i>	<i>Size</i>	<i>Median±IQR</i>	<i>Chi-square</i>	<i>Significance</i>
<i>CAVE</i>	S	33.45 ± 12.60	28.075	< .001
	M	15.65 ± 50.00		
	L	8.10 ± 25.33		
<i>HMD</i>	S	9.45 ± 17.43	10.959	.004
	M	6.25 ± 9.08		
	L	5.95 ± 8.13		

Number of samples = 80. Degrees of freedom = 2

The results of these tests are depicted in Table 12 and in Fig. 10. As can be seen, there are statistically significant differences, in favor of the HMD, for the time dataset for both tests (parametric and non-parametric). These results further reaffirm the answer to RQ1, since users clearly need less time when they use the HMD-based system. The number of errors with the HTC Vive is also substantially smaller than with the CAVE-based system, reaffirming also the answer to RQ2.

As in Task 1, Friedman tests were performed in order to analyze the effect of size in both systems, since the corresponding datasets did not follow a normal distribution. Table 13 and Table 14 show the results, which in this case do not show statistically significant differences between the groups. Figures 11–12 show the box plots corresponding with these experiments. These plots show that, despite the absence of statistical significance with these samples, the time and (especially) the number of errors in Task 2 are likely to be reduced with a larger keyboard size.

These results are different to those of Task 1, probably because the complexity of Task 2 is lower, and maybe also because of a certain amount of learning effect, since the text to type is shared between the different sizes. In any case, the box plots of both time and errors datasets suggest a certain size effect. Although with the performed test the null hypothesis (“there are no differences between the three groups”) cannot be rejected, this does not mean that the null hypothesis should be accepted. It is possible that the test has not sufficient power to prove the size effect. Therefore, the previous answer to RQ3 can be still considered partially valid.

## 5.2 Usability test

The results of the usability test are shown in Table 15. As can be seen, the HMD-based system obtains a better SUS score than the system using the CAVE. The difference in the SUS score is higher than 10 points. Considering that a value of 68 is usually established as the average

**Table 10** Results of the Friedman test for number of errors in Task 1 (2.S, 2.M, 2.L)

<i>System</i>	<i>Size</i>	<i>Median±IQR</i>	<i>Chi-square</i>	<i>Significance</i>
<i>CAVE</i>	S	15.00 ± 18.00	21.233	< .001
	M	2.50 ± 10.00		
	L	0.50 ± 3.00		
<i>HMD</i>	S	8.00 ± 5.00	5.934	.051
	M	3.00 ± 5.00		
	L	2.50 ± 10.00		

Number of samples = 16. Degrees of freedom = 2



**Table 11** Results of the Friedman test for number of pick-up gestures in Task 1 (3.S, 3.M, 3.L)

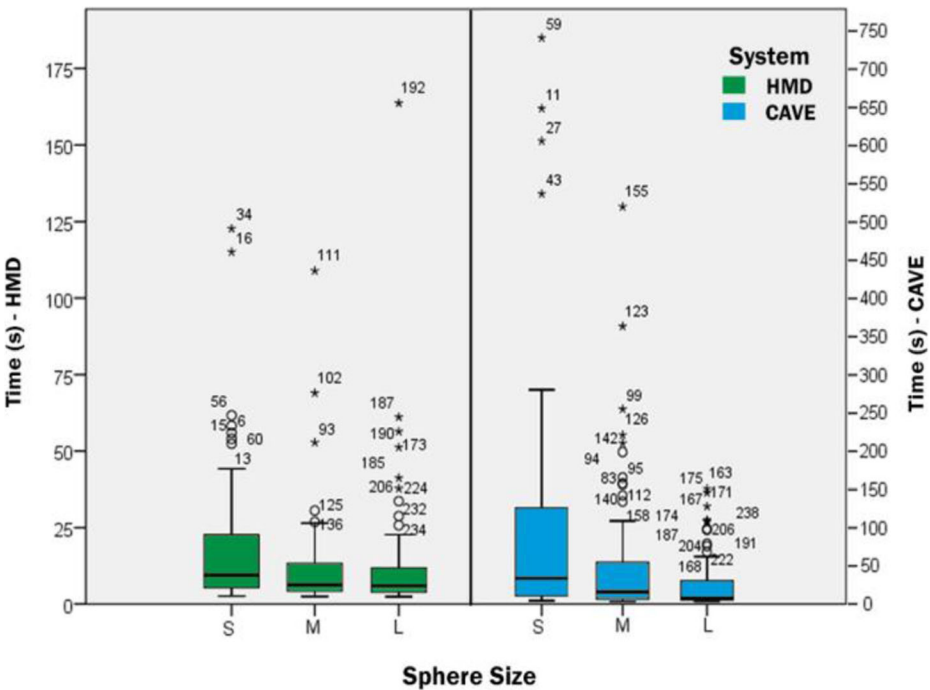
System	Size	Median ± IQR	Chi-square	Significance
CAVE	S	49.00 ± 50.00	18.032	< .001
	M	24.00 ± 17.00		
	L	17.50 ± 20.00		
HMD	S	24.00 ± 24.00	11.586	.002
	M	15.00 ± 11.00		
	L	12.50 ± 16.00		

Number of samples = 16. Degrees of freedom = 2

usability score [5, 8], the CAVE-based system falls just slightly below the 50th percentile, whereas the combination of HTC Vive and Leap Motion is well considered by the users with respect to usability and would be in the top-most quartile of usability [5] (between “good” and “excellent”). These results provide an answer to RQ4, since users clearly feel that natural hand interaction with Leap Motion in the HMD-based system is more usable than in the CAVE-based system.

### 5.3 Comparative evaluation

Finally, the comparative questionnaire is analyzed. The results are shown in Table 16. Question 1 (which of the two systems have you found more comfortable to use?) and Question 2 (which of the two systems did you like more?) have a clear winner: the HMD system.



**Fig. 7** Box plot of the distribution of the time dataset in Task 1 for the different sphere sizes (S, M, L) and both systems (HMD in green, CAVE in blue)

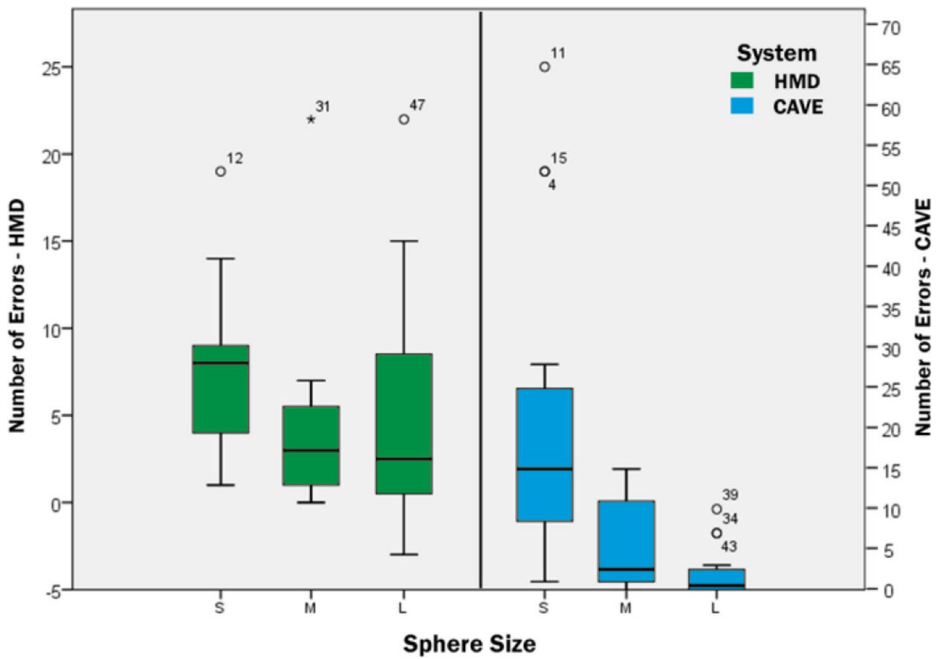


Fig. 8 Box plot of the distribution of the errors dataset in Task 1 for the different sphere sizes and both systems

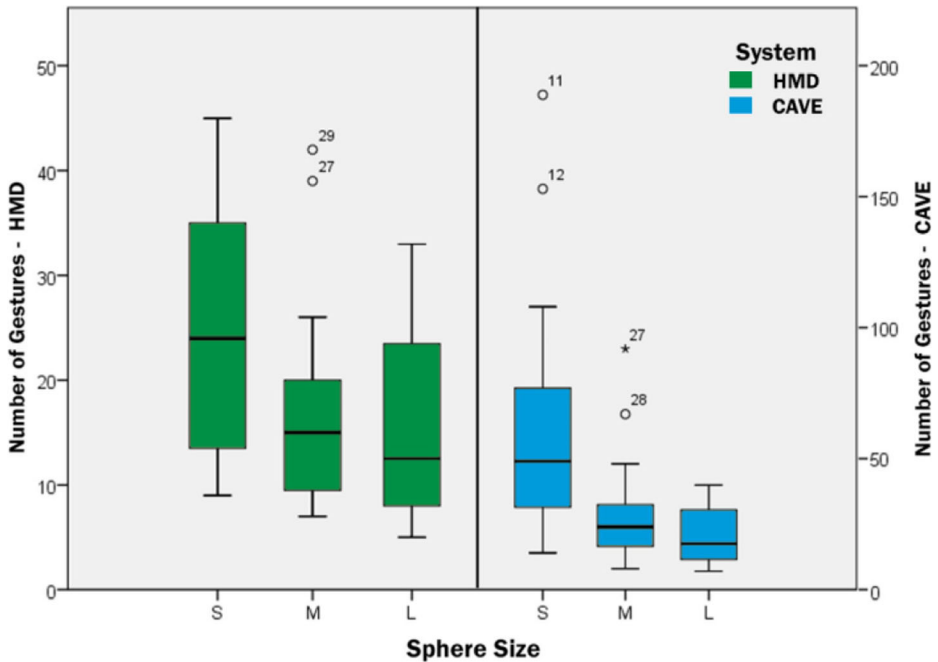
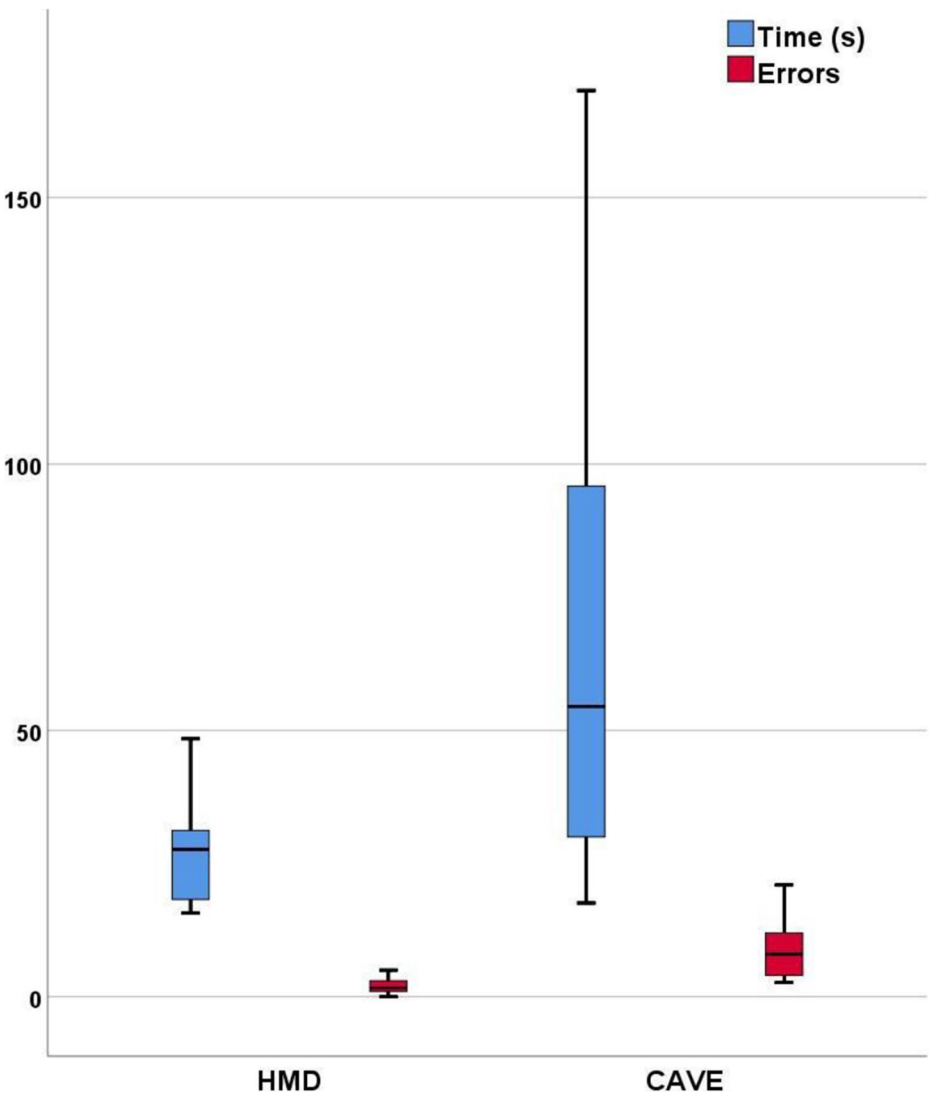


Fig. 9 Box plot of the distribution of the pick-up gestures dataset in Task 1 for the different sphere sizes and both systems

**Table 12** Results of the Wilcoxon signed-rank test and T-tests for Task 2

Test	Parameter	System	Median±IQR	Z (sig.)
Wilcoxon signed-rank	4 – Time (s)	CAVE	54.50±73.22	-3.621 (< .001)
		HMD	27.70±14.25	
Paired t-test	4 – Time (s)	CAVE	64.48±43.66	4.254 (.001)
		HMD	26.72±9.79	
Paired t-test	5 – Errors	CAVE	9.02±5.80	5.727
		HMD	1.98±1.45	(< .001)

Number of samples = 17. Degrees of freedom = 16



**Fig. 10** Box plots for the datasets of Task 1

**Table 13** Results of the Friedman test for the time necessary to complete Task 2 (4.S, 4.M, 4.L)

<i>System</i>	<i>Size</i>	<i>Median±IQR</i>	<i>Chi-square</i>	<i>Significance</i>
<i>CAVE</i>	S	60.10 ± 86.55	5.059	.085
	M	42.80 ± 69.10		
	L	53.50 ± 55.15		
<i>HMD</i>	S	29.20 ± 20.40	1.882	.415
	M	20.00 ± 15.55		
	L	23.20 ± 7.70		

Number of samples = 17. Degrees of freedom = 2

However, the results of Question 3 (with which of the two systems did you have a greater sense of embodiment with respect to your hands?) give a divided answer. For this reason, it can be said that this questionnaire provides an answer for RQ5, since users clearly show a preference for the HMD-based system and overwhelmingly declare that the HMD-based system is more comfortable. However, a clear answer to the research question about hand representation (RQ6) cannot be given, as responses to Question 3 are fairly divided, and the result is statistically inconclusive.

The explanation for these answers and for the results of the statistical analysis of objective data may be that visual disparities, such as hand tracking errors, are more clearly noticeable in the CAVE-based system, as the comments of the participants for the open-ended question reveal. Indeed, some users commented that “it was sometimes difficult in the CAVE system to understand where the position of my hands were with respect to the balls” and “although it is much more interesting and appealing to see my own hands and feel that I can use them to pick objects, the interaction in this system is sometimes complex and frustrating”. Apparently, the fact that users’ real hands are visible is positive for the sense of embodiment. However, as the mismatch (tracking is not perfect) between real and virtual hands is tangible in the CAVE-based system (precisely because real hands are visible), the HMD-based system feels more comfortable and usable. With the HMD, the virtual hand is not exactly in the expected place either (tracking errors exist as well) but at least the visual feedback is consistent and users are able to adapt to this perceptual error. This is also consistent with the findings of [31].

In addition, it is not possible in a CAVE to fully blend the real hand and the virtual objects, since there exists a mismatch between the accommodation distance of the real hand and the accommodation distance of the virtual content. For this reason, when a ball is picked, users do not really see their real hand in direct contact with the virtual ball, as they are observed in different physical planes. In the HMD-based setup, however, all the content has the same accommodation distance. Thus, users can adjust and adapt their hands’ movements in order to

**Table 14** Results of the Friedman test for number of errors during Task 2 (5.S, 5.M, 5.L)

<i>System</i>	<i>Size</i>	<i>Median±IQR</i>	<i>Chi-square</i>	<i>Significance</i>
<i>CAVE</i>	S	12.00 ± 13.00	3.600	.169
	M	5.00 ± 9.00		
	L	5.00 ± 8.00		
<i>HMD</i>	S	3.00 ± 6.00	4.607	.101
	M	0.00 ± 4.00		
	L	1.00 ± 2.00		

Number of samples = 17. Degrees of freedom = 2

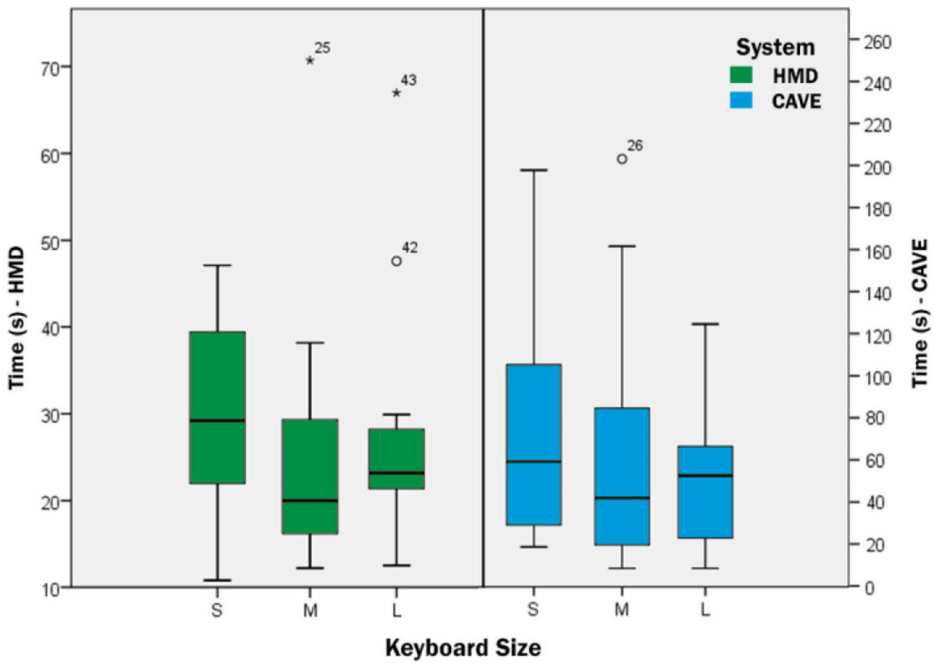


Fig. 11 Box plot of the distribution of the time dataset in Task 2 for the different keyboard sizes and both systems

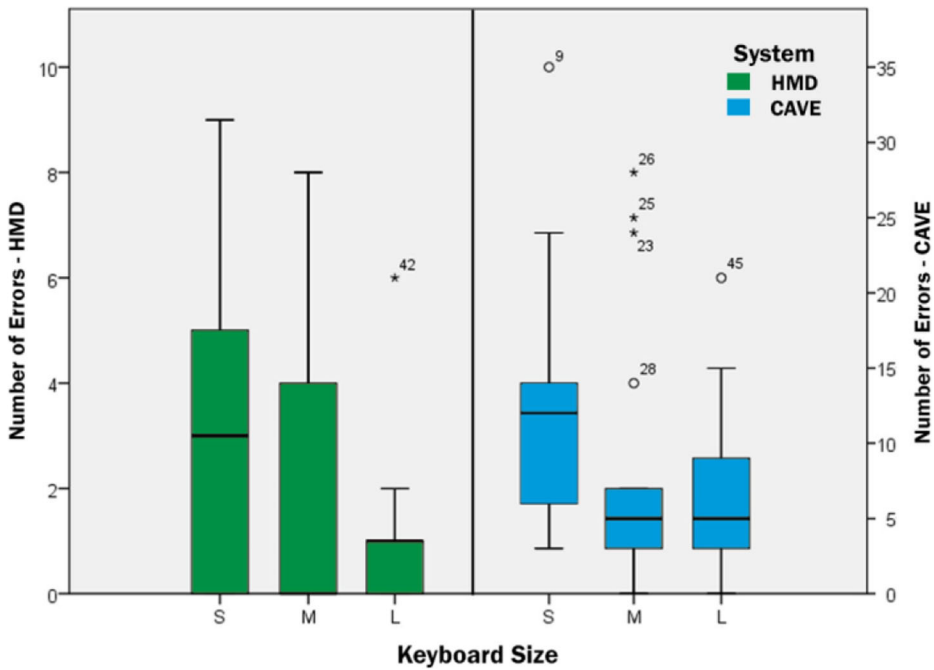


Fig. 12 Box plot of the distribution of the errors dataset in Task 2 for the different keyboard sizes and both systems

**Table 15** Results of the System Usability Scale (SUS) for both systems

Question	CAVE			HMD		
	Mean	SD	Range	Mean	SD	Range
1	3.00	1.17	1–5	3.29	1.05	1–5
2	2.18	0.73	1–4	1.71	0.92	1–4
3	3.71	0.85	2–5	3.94	1.09	1–5
4	2.29	1.21	1–5	2.00	1.06	1–5
5	3.71	1.16	1–5	4.53	1.01	1–5
6	2.12	0.86	1–4	1.88	0.99	1–4
7	3.71	0.99	1–5	4.24	0.56	3–5
8	2.47	1.33	1–5	1.71	0.77	1–3
9	3.82	1.01	1–5	4.06	0.83	3–5
10	1.94	1.14	1–5	1.65	0.79	1–3
<b>SUS Score</b>		<b>67.35</b>			<b>77.79</b>	

interact with virtual objects, so that they accomplish the tasks faster and with fewer mistakes, even if hand tracking errors exist. They can also see the virtual hands touching the virtual objects. These reasons probably explain why users prefer the HMD-based system, despite they are unable to see their real hands.

#### 5.4 Summary of the results

In summary, the results show that users need more time when they use the CAVE and they make fewer mistakes with the HMD, for both the pick-and-place task and the text-typing task. There is also a partial indication that the size of the virtual objects has an important effect on the performance. This effect is clear and significant for the first task although is smaller and not statistically significant in the second task. The usability question has a clear answer, since the SUS scores reveal that the system with the HMD is more usable than the CAVE-based system. A similar answer is given to the preference question since users overwhelmingly prefer the HMD-based system and perceive it as the more comfortable of the two. Finally, the question about the sense of embodiment has not a clear answer, since although the HMD-based system is clearly the preferred choice of the users, only about half of the users reported a higher sense of embodiment when using the HMD.

Thus, questions RQ1, RQ2, RQ4 and RQ5 share the same answer: the HMD-based system. The answer to RQ3 is (partially) yes. Finally, it is not possible to answer RQ6 with these results.

**Table 16** Results of the comparative questionnaire

Question	Choice		Binomial test p value
	CAVE	HMD	
1	Answers (%)	Answers (%)	
1	1 (5.88%)	16 (94.12%)	0.0003
2	2 (11.76%)	15 (88.24%)	0.0023
3	8 (47.06%)	9 (52.94%)	1.0

## 6 Conclusion and future work

This paper presents a comparative study of the integration of natural hand interaction in two immersive VR systems: a CAVE system (where users' real hands are visible) and a non-see-through HMD system (where only a virtual representation of the hands is possible), using the same tracking technology, a Leap Motion device, in both cases. The motivation is to understand and test the suitability of using optical-based natural hand interaction in a CAVE and compare it to an HMD system. For that purpose, we raise six research questions related to task performance, usability and perception differences with these two systems. To answer these questions, we designed an experiment with 17 participants where users have to complete a pick-and-place task with virtual balls and a text-typing task with virtual keyboards.

The results of the experiment show that users prefer the HMD-based setup and also perform better with this setup, which also reaches a higher usability score than the CAVE-based system. Therefore, the most important conclusion of this research is that the HMD-based system provides a performance and usability advantage over the CAVE system regarding natural hand interaction with a Leap Motion device. This means that the visualization of users' real hands does not imply a performance or usability advantage in virtual environments or, at least, it is not a strong argument if other issues are present (e.g. visual disparities).

Delving into this last question, it seems reasonable to assume that seeing one's real hands would enhance the sense of embodiment (although this is neither denied nor validated by the results of the experiments). However, the fact that this makes tracking errors more noticeable and the difficulty of seeing the real hands and the virtual objects simultaneously, may induce a preference and/or a performance advantage for the visualization mode in which real hands are not visible, as it is the case of the setups tested in this experiment.

Another important conclusion is that, despite both objective and subjective data point out that the HMD-based system is a better option for this particular purpose, the CAVE-based system still obtains 67.35 points in the SUS scale, and is usable in combination with a Leap Motion, a result that, to the best of our knowledge, is reported here for the first time. However, if hand tracking is to be used with the CAVE system, larger virtual objects are needed compared to the HMD-based system. On the other hand, the usability of the HMD-based system is rather high, with almost 78 points (between "good" and "excellent"), which means that this setup is more than suitable for providing natural hand interaction in VR applications.

The experiment and the conclusions drawn open research avenues with respect to the use of natural hand interaction in VR. Future work includes the extension of the experiments to other HMDs with built-in hand tracking (such as an Oculus Quest 2) or VR gloves (such as Sensory VRfree gloves). It would be very interesting to analyze if different hand tracking technologies and methods, which are expected to provide different tracking errors, imply different results in the use and visualization of hand interaction. This will allow us to clearly identify if the lower ratings obtained with the CAVE system are caused mainly by the awareness of tracking errors or because of the accommodation problem. It might be also interesting to test different setups, such as an Augmented Virtuality setup in which real hands are introduced in the virtual scene by means of an HMD with a video camera. Finally, more tasks could be added to the experiment, such as throwing objects or more precision-demanding tasks such as hand drawing.

**Funding** Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature. Cristina Portalés is supported by the Spanish government postdoctoral grant Ramón y Cajal, under grant No. RYC2018–025009-I.

**Data availability** Not applicable.

**Code availability** Not applicable.

## Declarations

**Ethics approval** All the participants in this research volunteered for the experiment and signed an informed consent regarding their participation in the study, which was unpaid and completely anonymous.

**Conflict of interest** The authors declare no conflict of interest.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Ahmadzadeh Bazzaz S (2020) Depth and distance perceptions within virtual reality environments - a comparison between HMDs and CAVEs in architectural design. In: Werner, L and Koering, D (eds.), Anthropologic: architecture and fabrication in the cognitive age - proceedings of the 38th eCAADE conference - volume 1, TU Berlin, Berlin, Germany, 16–18 September 2020, pp. 375–382. CUMINCAD
2. Argelaguet F, Hoyet L, Trico M, Lecuyer A (2016) The role of interaction in virtual embodiment: effects of the virtual hand representation. In: 2016 IEEE virtual reality (VR). Pp 3–10
3. Aykent B, Merienne F, Kemeny A (2015) Effect of VR device – HMD and screen display – on the sickness for driving simulation
4. Azmandian M, Hancock M, Benko H, et al (2016) Haptic Retargeting: Dynamic Repurposing of Passive Haptics for Enhanced Virtual Reality Experiences. In: Proceedings of the 2016 CHI conference on human factors in computing systems. ACM, New York, NY, USA, pp. 1968–1979
5. Bangor A, Kortum P, Miller J (2009) Determining what individual SUS scores mean: adding an adjective rating scale. *J Usability Stud* 4:114–123
6. Benda B, Esmacili S, Ragan ED (2020) Determining detection thresholds for fixed positional offsets for virtual hand remapping in virtual reality. In: 2020 IEEE international symposium on mixed and augmented reality (ISMAR). IEEE, pp 269–278
7. Bowman DA, Datey A, Ryu YS, Farooq U, Vasnaik O (2002) Empirical comparison of human behavior and performance with different display devices for virtual environments. Proceedings of the Human Factors and Ergonomics Society Annual Meeting 46:2134–2138. <https://doi.org/10.1177/154193120204602607>
8. Brooke J (1996) SUS-A quick and dirty usability scale. Usability evaluation in industry 189:4–7
9. Combe T, Chardonnet J-R, Merienne F, Ovtcharova J (2021) CAVE vs. HMD in distance perception. In: 2021 IEEE conference on virtual reality and 3D user interfaces abstracts and workshops (VRW). IEEE, pp 448–449
10. Cordeil M, Dwyer T, Klein K, Laha B, Marriott K, Thomas BH (2017) Immersive collaborative analysis of network connectivity: CAVE-style or head-mounted display? *IEEE Trans Vis Comput Graph* 23:441–450. <https://doi.org/10.1109/TVCG.2016.2599107>
11. Cruz-Neira C, Sandin DJ, DeFanti TA, et al (1992) The CAVE: audio visual experience automatic virtual environment. *Communications of the ACM* 35:64–



12. Cruz-Neira C, Fernández M, Portalés C (2018) Virtual reality and games. *Multimodal Technologies and Interaction* 2:8. <https://doi.org/10.3390/mti2010008>
13. Dorado JL, Figueroa P, Chardonnet J, et al (2017) Comparing VR environments for seat selection in an opera theater. In: 2017 IEEE symposium on 3D user interfaces (3DUI). Pp 221–222
14. Elor A, Powell M, Mahmoodi E, Hawthorne N, Teodorescu M, Kurniawan S (2020) On shooting stars: comparing cave and hmd immersive virtual reality exergaming for adults with mixed ability. *ACM Transactions on Computing for Healthcare* 1:1–22
15. Ghinea M, Frunzã D, Chardonnet J-R, et al (2018) Perception of absolute distances within different visualization systems: HMD and CAVE. In: De Paolis LT, Bourdot P (eds) *augmented reality, virtual reality, and computer graphics*. Springer international publishing, pp 148–161
16. Grabowski A (2021) Practical skills training in enclosure fires: an experimental study with cadets and firefighters using CAVE and HMD-based virtual training simulators *Fire Safety Journal* 103440. <https://doi.org/10.1016/j.firesaf.2021.103440>
17. International Business Machines Corporation (IBM) IBM SPSS Software. <https://www.ibm.com/analytics/spss-statistics-software>. Accessed 3 Sep 2021
18. Juan MC, Pérez D (2009) Comparison of the levels of presence and anxiety in an acrophobic environment viewed via HMD or CAVE. *Presence Teleop Virt* 18:232–248. <https://doi.org/10.1162/pres.18.3.232>
19. Kim K, Rosenthal MZ, Zielinski D, Bradi R (2012) Comparison of desktop, head mounted display, and six wall fully immersive systems using a stressful task. In: *virtual reality conference, IEEE(VR)*. Pp 143–144
20. Leite DQ, Duarte JC, Neves LP, de Oliveira JC, Giraldo GA (2017) Hand gesture recognition from depth and infrared Kinect data for CAVE applications interaction. *Multimed Tools Appl* 76:20423–20455
21. Likert R (1932) A technique for the measurement of attitudes. *Archives of psychology*
22. Lin L, Normoyle A, Adkins A, et al (2019) The effect of hand size and interaction modality on the virtual hand illusion. In: 2019 IEEE conference on virtual reality and 3D user interfaces (VR). IEEE, pp 510–518
23. Lindsey S (2017) Evaluation of low cost controllers for Mobile based virtual reality headsets
24. Mallaro S, Rahimian P, O’Neal EE, et al (2017) A comparison of head-mounted displays vs. large-screen displays for an interactive pedestrian simulator. In: *proceedings of the 23rd ACM symposium on virtual reality software and technology*. ACM, New York, NY, USA, p 6:1-6:4
25. Massey FJ Jr (1951) The Kolmogorov-Smirnov test for goodness of fit. *J Am Stat Assoc* 46:68–78
26. Mestre DR (2017) CAVE versus head-mounted displays: ongoing thoughts. *Electronic Imaging* 2017:31–35. <https://doi.org/10.2352/ISSN.2470-1173.2017.3.ERVR-094>
27. Miller J, Baiotto H, MacAllister A, et al Comparison of a Virtual Game-Day Experience on Varying Devices. <https://www.ingentaconnect.com/contentone/ist/ei/2017/00002017/00000016/art00006>. Accessed 3 Sep 2021
28. Mishra P, Pandey CM, Singh U, Gupta A, Sahu C, Keshri A (2019) Descriptive statistics and normality tests for statistical data. *Ann Card Anaesth* 22:67
29. Pala P, Cavallo V, Dang NT, Granié MA, Schneider S, Maruhn P, Bengler K (2021) Analysis of street-crossing behavior: comparing a CAVE simulator and a head-mounted display among younger and older adults. *Accid Anal Prev* 152:106004. <https://doi.org/10.1016/j.aap.2021.106004>
30. Perani D, Fazio F, Borghese NA, Tettamanti M, Ferrari S, Decety J, Gilardi MC (2001) Different brain correlates for watching real and virtual hand actions. *NeuroImage* 14:749–758. <https://doi.org/10.1006/nimg.2001.0872>
31. Perez-Marcos D, Sanchez-Vives MV, Slater M (2012) Is my hand connected to my body? The impact of body continuity and arm alignment on the virtual hand illusion. *Cogn Neurodyn* 6:295–305. <https://doi.org/10.1007/s11571-011-9178-5>
32. Philpot A, Glancy M, Passmore PJ, et al (2017) User experience of panoramic video in CAVE-like and head mounted display viewing conditions. In: *Proceedings of the 2017 ACM international conference on interactive experiences for TV and online video*. ACM, New York, NY, USA, pp. 65–75
33. Rautaray SS, Agrawal A (2015) Vision based hand gesture recognition for human computer interaction: a survey. *Artif Intell Rev* 43:1–54
34. Ronchi E, Mayorga D, Lovreglio R, Wahlqvist J, Nilsson D (2019) Mobile-powered head-mounted displays versus cave automatic virtual environment experiments for evacuation research. *Computer Animation and Virtual Worlds* 30:e1873. <https://doi.org/10.1002/cav.1873>
35. Schneider D, Otte A, Kublin AS, et al (2020) Accuracy of commodity finger tracking systems for virtual reality head-mounted displays. In: 2020 IEEE conference on virtual reality and 3D user interfaces abstracts and workshops (VRW). IEEE, pp 805–806
36. Shapiro SS, Wilk MB (1965) An analysis of variance test for normality (complete samples). *Biometrika* 52: 591–611

37. Tcha-Tokey K, Loup-Escande E, Christmann O, Richir S (2017) Effects on user experience in an edutainment virtual environment: comparison between CAVE and HMD. In: Proceedings of the European conference on cognitive ergonomics 2017. ACM, New York, NY, USA, pp 1–8
38. Unity Technologies Unity Real-time Development Platform. In: Unity 3D. Accessed 3 Sep 2021
39. University of Valencia IRTIC. In: Institute on Robotics and Information and Communication Technologies. <https://www.uv.es/uvweb/university-research-institute-robotics-information-communication-technologies/en/institute-robotics-information-technology-communications-1285895240418.html>. Accessed 3 Sep 2021
40. Weichert F, Bachmann D, Rudak B, Fisseler D (2013) Analysis of the accuracy and robustness of the leap motion controller. *Sensors* 13:6380–6393

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.