ORIGINAL ARTICLE

The effects of touchless interaction on usability and sense of presence in a virtual environment

Lucio Tommaso De Paolis[1] · **Valerio De Luca**[1]

Abstract

For software applications with a significant level of user involvement, the traditional concept of usability has evolved into the more complex idea of user experience, which also covers emotional, cognitive or physical responses. In virtual reality, user experience also depends on the user perception related to some peculiarities of immersive environments, where also the devices employed for user interaction play a determinant role. This has led to the design of the Presence Questionnaire (PQ) for the evaluation of the effectiveness of virtual environments. This work analyzes the effects of two different interaction modalities on usability and sense of presence: in particular, the Myo armband, a gesture-based device for touchless interaction, is compared with the Vive handheld controller bundled with the HTC Vive headset. A total of 84 subjects were recruited to test the virtual environment and asked them to fill in a questionnaire obtained by combining the Usability Metric for User eXperience (UMUX) questionnaire, the System Usability Scale (SUS) and the presence questionnaire (PQ), which was specifically designed for virtual environments. A comparison between the scores obtained for the two interaction modalities revealed which questionnaire items are significantly influenced by the input interface and deduce some insights about the consequences on human factors.

Keywords Handheld controller · Touchless interaction · Usability · User experience · Presence · Virtual environment

1 Introduction

User experience (Tullis and Albert 2013) goes beyond the traditional concept of usability by including emotional and aesthetical factors (Bachmann et al. 2018). As defined by ISO 9241-210 (ISO 2010), it encompasses "perceptions and responses that result from the use or anticipated use of a product, system or service." User task performance concerns specific tasks such as navigation, object placement and object selection. An important factor influencing the user performance is the mental workload (Cain 2004), also known as cognitive workload, which represents the mental effort required to perform the tasks. It can be assessed by means of quantitative performance tests, physiological measures or subjective feedback collected through questionnaires (Moustafa et al. 2017).

A key aspect for the effectiveness of virtual environments is the sense of presence. It denotes the subjective experience of being in one place or environment, independently of where a subject is actually located (Witmer and Singer 1998). Another important aspect that characterizes the experience in virtual environments is user interaction. Touchless interfaces enable new forms of natural interaction, where well-defined positions or movements of some parts of the human body, known as gestures (Garber 2013), are associated with meaningful commands. For instance, gestures allow doctors to consult medical data during interventions (De Paolis 2016). Moreover, in image-guided navigation systems (De Mauro et al. 2012) and augmented reality applications (De Paolis and De Luca 2019; De Paolis and Ricciardi 2018) for the intraoperative support of surgical procedures, a touchless interaction with virtual organs (De Paolis 2018; Indraccolo and De Paolis 2017) is fully compliant with sanitary regulations. In military operations, it can increase the situational awareness by providing a faster and more natural way to issue commands (Zocco et al. 2015). In

✉ Valerio De Luca
valerio.deluca@unisalento.it

Lucio Tommaso De Paolis
lucio.depaolis@unisalento.it

1   Department of Engineering for Innovation, University of Salento, Lecce, Italy

addition to offering greater intuitiveness compared to traditional WIMP (Windows Icon Mouse Pointer) interfaces, gesture-based interaction can increase motivation in the educational field (De Paolis et al. 2019) by fostering kinesthetic learning (Kosmas et al. 2018), where the cognitive process is enhanced through the physical involvement of the human body.

In a virtual environment, gesture-based interaction can improve immersion as long as no conscious attention is required to the specific gestures being performed (Deller et al. 2006). The concept of *guessability* (Wobbrock et al. 2005), which is a stronger concept than immediate usability, represents the ability to guess symbolic input without any prior knowledge or learning phase. Unfortunately, the design of gesture detection systems is often based more on ease of technology implementation than on naturalness of gestures: a typical example is the Myo armband, a wearable device that can detect five basic gestures based on the arm electromyographic impulses it can recognize (Rawat et al. 2016).

This paper investigates the differences between handheld controller and gesture-based interaction in a virtual reality experience. In particular, it aims to evaluate the impact that abstract gestures based on electromyography (EMG) technology can have on usability and sense of presence. It is expected that the main obstacles may derive either from the fact that gestures are abstract/symbolic (as they do not deal with grasping or directly manipulating objects), or from any problems of accuracy in detection that are intrinsic to the device. This will allow us to understand whether future research should focus more on improving the accuracy of detection, expanding and diversifying the vocabulary of abstract gestures, or introducing the possibility of detecting 'fine' movements, also involving the fingers, to aim at the recognition of less abstract gestures and closer to manipulative ones.

The considered scenario is a virtual navigator for the HTC Vive headset (HTC 2021) that allows the user to explore the organs of the human body and navigate inside them. The study started with earlier work (De Paolis and De Luca 2020), which compared the performance of the Myo armband with that of the Vive controller, a handheld device bundled with the HTC Vive headset. In such context, data on user impressions were collected through the Usability Metric for User eXperience (UMUX) (Finstad 2010), the System Usability Scale (SUS) (Brooke 1996) and the presence questionnaire (PQ) (Witmer and Singer 1998). While the previous work compared various factors of the user experience between the two devices, trying to identify any interdependence between several components (*involvement*, *interface quality*, *adaptation/immersion*, *visual fidelity*), this work goes deeper in the search for the possible causes of the differences in terms of user experience taking into account the peculiarities of the two interfaces.

The rest of the paper is structured in this way: Sect. 2 presents the related work about presence and interaction in virtual environments; Sect. 3 introduces the main factors of user experience in a virtual environment; Sect. 4 briefly summarizes the target experimental scenario, which was introduced and described more in detail in our previous work (De Paolis and De Luca 2020); Sect. 5 presents the questionnaires chosen for our study, after a brief introduction on the most frequently used questionnaires for the evaluation of usability and user experience; Sects. 6, 7, 8 and 9 present the results; Sect. 10 discusses and summarizes the main findings; Sect. 11 concludes the paper.

2 Related work

Several studies in the literature addressed presence in virtual environments. Mutual relations between presence and emotions were discussed in Diemer et al. (2015) with a special focus on the perception of fear. The technological, cognitive and emotional factors that increase the sense of presence were analyzed in Gorini et al. (2011), which considers a virtual hospital scenario. The effects of immersive VR on empathy were evaluated in terms of immersion-presence, illusion of body ownership, illusion of agency, engagement and mind-wandering (Barbot and Kaufman 2020): the results of the study highlighted the role of VR as a "perspective taking-machine". Usability, presence and perceived workload were assessed in a virtual operating room (Li et al. 2020) recreated through an Oculus Rift VR headset and a LapMentor III laparoscopic simulator: the high mental demand and low frustration revealed by the experimental tests suggested that users tend to enjoy intellectual challenges provided by the virtual environment. Presence, workload, usability and flow were evaluated in a study comparing a video display terminal and a head-mounted display in walking and driving scenarios (Rhiu et al. 2020): the former is better in terms of workload, while the latter provides a higher sense of presence and flow in direct interaction; on the other hand, in mediated interaction, such as in driving a vehicle, a head-mounted display has also a lower usability. An older study (Livatino et al. 2015) assessed the usability of a stereoscopic bronchoscope in terms of depth impression, presence and comfort. It compared seven systems (from laptops to wall screen and head-mounted display), four approaches to stereo viewing (colored Anaglyph, polarized filters, shutter glasses, and separated displays), and five types of display technologies (digital light processing, cathode ray tube, liquid crystal display, LED, and organic LED). Another comparative study between a head-mounted display and a desktop LCD screen in driving simulation focused on sickness, subjective eye symptoms, game engagement and

game performance (Cao et al. 2020): it revealed a moderately higher simulator sickness for the head-mounted display. Other usability studies comparing various visualization systems for virtual environments can be found in Yu et al. (2019), Somrak et al. (2019), Webster and Dues (2017), Tcha-Tokey et al. (2017).

The influence of navigation control and screen size was assessed in Clemente et al. (2014) by recording brain activation with an EEG. Three different VR systems for firefighter training were compared in terms of usability, ergonomics and learning effectiveness (Corelli et al. 2020): in such context, the locomotion management is a key aspect that influences accuracy and timing of the execution due to the need to cover long distances and exert a fine movement control. Teleportation and natural walking were studied as interaction modalities for abstract data exploration in a virtual environment enhanced with haptic feedback (Zenner et al. 2020): a comparison with a traditional 2D tablet interface revealed a tradeoff between efficiency and user interest, but no difference in model understandability.

In most virtual environments, an eye-centered interaction principle has been adopted, since the user's visual searching is typically considered more efficient and comfortable than hand interaction. The influence of arm movements was studied in a free hand target selection experiment in a virtual environment (Lou et al. 2020), with a special focus on the hand choice and on the hand position. In Pai et al. (2019), gaze tracking and forearm contractions, detected through electromyography, were associated with cursor movements and selections: this combined modality for user interaction was proved to perform better than HTC Vive controller, Xbox gamepad, dwelling time and eye-gaze dwelling time.

Other experimental studies evaluated the impact of touchless and gesture-based interaction on the user experience.

The Leap Motion controller (Leap Motion 2021) and the Myo armband were compared with a standard game controller in a ball-balancing maze-like game scenario (Chen et al. 2018). Despite overall good performance, Leap Motion and Myo provide a lower level of control, but the gap with the traditional controller tends to narrow as users become more familiar with the devices. In particular, the understanding of Myo's physical dynamics gained gradually through practice gives better control than Leap Motion, which makes it more difficult to move objects accurately due to the lack of feedback. Users said they found interaction through devices such as Leap Motion and Myo more interesting than through traditional controllers, which they considered "aged."

Another work (Wirth et al. 2018) also revealed that the interaction through EMG devices is considered to be more interesting and exciting than interaction through handheld controllers. On the other hand, users involved in this experiment declared the EMG experience was more stressful, as it requires more thinking. Similarly, experimental tests on

virtual manipulation based on Leap Motion highlighted some interaction fidelity issues in grabbing virtual objects, caused by a difficulty in perceiving the position of their fingers (Gieser et al. 2016).

Two interaction modes for teleportation based on one hand and two other based on two hands are compared in Schäfer et al. (2021): results revealed that single hand gestures generally provide a more comfortable and effective way to move in a virtual environment, whereas two hands require a higher workload.

The voice and motion gesture controller presented in Niño et al. (2019), which also provides tactile feedback, proved that interaction has a significant influence on the sense of presence, embodiment and immersion.

The effects of gesture-based interaction in a computer-based science lesson were studied in Bailey et al. (2019): experimental tests proved that gestures perceived as natural contribute in increasing not only the sense of control, but also the feeling of presence; on the contrary, gestures perceived as unnatural and not very usable can generate interferences that may distract the user from a task.

3 The factors of user experience in a virtual environment

The following subsections define the main factors characterizing the user experience in a virtual environment that outline the framework of this study: presence, immersion and interaction.

3.1 Presence

Presence refers to the sensation of being in a place other than the actual physical location, achieved by deceiving the user's cognitive and perceptual systems (Slater et al. 1994).

Riva et al. (2004) identified three layers of presence:

– proto presence, based on kinesthetic information about the relative position of the user's body;
– core presence, representing the process of selective attention applied to perceptions;
– extended presence, which enhances the significance of external events and reinforces the presence of the subject in significant experiences.

According to Slater (2009), two perceptual illusions simultaneously contribute to the sense of presence: the *place illusion* of being really in a place and the *plausibility illusion* that events are really happening.

Place illusion is not closely related to perceptual realism (Sanchez-Vives and Slater 2005), which indicates the extent to which virtual representations appear similar to the

corresponding real objects. Plausibility illusion, on the other hand, shares some aspects with social realism (Nilsson et al. 2016), which is concerned with the level of real-life adherence of a multimedia representation.

Even though place illusion is a subjective experience, it partially depends on system immersion, which can be described in terms of objective properties (Slater 2003) (such as frame rate, fidelity of tracking and field of view) and measured in quantitative terms. An immersive system can be described through its sensorimotor contingencies (Slater 2009), which are the actions performed to explore the environment and inspect the objects in it. On the contrary, the sense of presence is not an objectively measurable quantity: it can vary between different subjects according to the different actions they perform, as one individual, for example, may explore the environment more widely than another.

Plausibility illusion requires that (Rovira et al. 2009):

– each action performed by the user corresponds to a specific reaction in the virtual environment;
– the environment always responds directly to the user;
– the events within the environment correspond to the user's expectations deriving from everyday life.

Waterworth and Waterworth (2001) described presence as "a conscious emphasis on direct perception of currently present stimuli rather than on conceptual processing." In particular, he identified three dimensions in the allocation of attention in a virtual environment:

– focus of attention is directly related to the sense of presence, which increases when a low degree of conceptual and abstract reasoning is required; a technologically immersive system plays an important role in this dimension insofar as it provides concrete information that can be processed directly by perceptual-motor systems;
– locus of attention deals with the attention paid to the virtual environment rather than to the external environment;
– sensus of attention represents the degree of conscious arousal, i.e., the basic psychological response to external stimuli.

3.2 Immersion

According to Agrawal et al. (2020), immersion is a state of deep mental involvement characterized by a shift of attention from the awareness of the physical world and can be described through three main factors. The first one is the subjective sense of being surrounded, known as *perceptual immersion*. Another factor is the absorption in the narrative, which causes a shift of attention in the player: it consists of *spatial immersion*, which concerns the sense of space and the act of exploration, *temporal immersion*, which derives

from the curiosity to know how a story evolves, and *emotional immersion*, which refers to the emotional attachment with a story character (Ryan et al. 2003). The third factor is made up of *strategic immersion* and *tactical immersion*: the former derives from the absorption when a player is making choices, while the latter derives from the attention on tasks requiring quick reactions. Opposed to the concept of immersion based on strategic or tactical challenges, *systematic immersion* (Arsenault 2005), which is more suitable to non-participatory activities, simply focuses on the acceptance of the game's system instead of real-world laws of physics.

In the various papers presented so far, immersion has been described sometimes as the ability of technology to faithfully reproduce real-world perception and actions, known as system immersion, and other times as a subjective experience induced by specific devices (Nilsson et al. 2016).

According to Witmer and Singer (1998), who designed the Presence Questionnaire employed in this study, the psychological state deriving from directing attention to a stimulus is called involvement, while immersion is considered as the subjective experience of being surrounded in an interactive environment. According to Witmer's view, the main factors influencing immersion are:

– the level of isolation from the physical environment;
– the sense of self-inclusion in the virtual environment;
– egocentric motion perception and the natural interaction with the environment.

A concept close to immersion is that of flow, defined as "the state in which people are so involved in an activity that nothing else seems to matter" (Csikszentmihalyi 1991). It was described through the following eight components: balance between ability and challenge, concentration/attention, clear goals, immediate feedback, escape from everyday life, sense of personal control, loss of self-consciousness and altered sense of time. Passive activities, in which some of these factors (balance between ability and challenge, clear goals, immediate feedback) are absent, cannot be considered as flow experiences, although they may induce immersion.

3.3 Interaction

An important concept underlying the interaction in a virtual environment is *discoverability* (Norman 2013), which refers to the possibility of exploring how the environment works and what operations are possible. Norman identified three phases in interaction (Norman 2013): forming the goal, executing the action and evaluating the results. *Execution* consists of planning the action, specifying the action sequence and performing it. *Evaluation* consists of perceiving the state of the world, interpreting it and comparing it with the original goal. The user is only vaguely aware during

the execution and evaluation phases. Only the appearance of something new or an obstacle awakens conscious attention. In particular, the processes of specifying actions and interpreting perception occur in a semiconscious manner; on the contrary, the processes of performing actions and perceiving are usually automatic and subconscious unless some action requires special attention. *Gulf of execution* refers to the situation where it is unclear how to achieve an objective that is known, while *gulf of evaluation* refers to not understanding the results of an action (Jerald 2015).

The degree to which physical actions required for a task in a virtual environment correspond to the physical actions required by the equivalent real-world task is known as *interaction fidelity* (Bowman et al. 2012). *Realistic interactions* avoid adaptation problems, which may have a negative impact on the effectiveness of a training procedure. However, even non-realistic *magic interactions*, based on metaphors that enable gestural interaction with distant objects, can improve the user experience by overcoming the limitations of the real world. In both realistic and non-realistic interactions, the intuitiveness of the interface is crucial. Interaction metaphors (Jerald 2015) are interaction modalities that refer to specific knowledge deriving from other domains. An ideal virtual environment should be consistent with the conceptual model in the mind of each user, without any explanation.

In general, two key interaction fidelity factors that are crucial to achieve good performance and avoid user frustration are *input veracity* and *control symmetry* (McMahan et al. 2015): the former represents the ability of an input device to capture user actions in terms of accuracy, precision, and latency, while the latter is the level of control in the interaction compared to the equivalent real-world task.

3.3.1 Touchless interaction with virtual objects

Natural User Interfaces (NUIs) were introduced to allow users to interact with a computer system through actions similar to those performed in the real world. These devices include touchless interfaces, which allow commands to be given in the form of hand *gestures*. In virtual environments, they also offer the possibility to move or manipulate virtual objects by means of gestures or by simulating the act of grasping them.

Various criteria for gesture classification were discussed in Li et al. (2019):

– from the point of view of spatiotemporal status, gestures can be static or dynamic: the former are static poses at a given instant of time that do not include information about time series, while the latter represent changes in poses over a certain period of time;

– from a semantic point of view, manipulative gestures allow moving or rotating objects through arm or hand movements, whereas communicative gestures have a specific information function;

– from the point of view of the scope of interaction, stroke gestures consist in moving the hand on a support surface such as a touchscreen, whereas mid-air gestures consist in movements in a free space without any support surface. The former exhibit a better ability to capture the details of actions, while the latter allow a larger volume of interaction.

Another systematic review (Vuletic et al. 2019) classified hand gestures into the following categories:

– deictic gestures, i.e., pointing gestures used to indicate a direction or a point selection or to move an object along a path created by the hand;

– pantomimic gestures, which emulate the real-life actions performed to pick up, pull and modify parts of an objects;

– free-form gestures, used to move virtual objects, windows or pointers;

– manipulative gestures, associated with translation, rotation, scaling/zooming or object size manipulation;

– semaphoric gestures, which consist in abstract and predefined hand motions representing concepts.

Other studies analyzed which gestures are preferred by users (Chen et al. 2015) and perceived as natural (Grandhi et al. 2011).

Visual affordances play a critical role on several gestures operating on virtual objects (Kang et al. 2020): they concern the direct perception of the potential actions that can be performed with them without the need for high-level processes such as reasoning about object properties (Thill et al. 2013). Electroencephalographic (EEG) signals, gathered through some proper headsets (Invitto et al. 2015, 2016), were analyzed to study the activation of brain areas for affordances during gesture-based tasks.

On the contrary, the abstract gestures detected by the Myo armband, which are the subject of this study (described in Sect. 4), fall into the category of semaphoric gestures and are not related to the concept of grasping an object. Myo is based on electromyography (EMG), which enables the possibility to recognize gestures by detecting the user's motion from changes in the muscle physiological signals: EMG-based devices are able to detect changes of the electric current by means of some electrodes placed on the skin surface (Kim et al. 2008).

Several studies have been carried out to find the best ways to arrange sensors on the user's arm for gesture recognition (Saponas et al. 2008, 2009; Benko et al. 2009; Saponas et al. 2010), and several methods (Jo and Oh 2020; Lu et al.

2020; Benalcazar et al. 2018; Wan et al. 2017; Kratz et al. 2012; Huang et al. 2015; Dai et al. 2021) have been proposed to improve detection and classification accuracy. The Myo-based gesture classification was compared in terms of accuracy with that based on a more expensive EMG system that provides a higher sampling rate (Gieser et al. 2017): the experimental tests showed that similar results can be achieved for the two devices through KNN classification; moreover, feature extraction gives better classification results than the use of raw data. Other studies (Dolopikos et al. 2021) proved that a model calibration process is able to produce an improvement of 24% in gesture classification accuracy.

4 The considered scenario

The considered application, developed for the HTC Vive headset (HTC 2021) in Unity (Unity Technologies 2021) with the SteamVR API (Steam 2021), is a virtual navigator of the human body: it provides a combined visualization where 3D models of organs are intersected by the plane of the CT images used to reconstruct them.

A total of 84 subjects were recruited to test the virtual reality application: half of them used the Vive controller, while the other ones used the Myo armband.

As explained in our previous work (De Paolis and De Luca 2020), commands for zooming/rotating the 3D models and for translating the CT slices were associated with buttons on the Vive controller and gestures detected by the Myo armband (Fig. 1). A preliminary calibration process was necessary for each user wearing the Myo armband to adjust the device to the forearm's morphology and improve gesture detection.

After the virtual reality experience, each users was requested to fill some questionnaires as described in the following section.

5 Evaluation of usability and user experience for Vive and Myo interaction modalities

Several questionnaires were proposed to measure the perceived usability of a system (Assila et al. 2016; Lewis 2018; Hajesmaeel-Gohari and Bahaadinbeigy 2021).

The Software Usability Measurement Inventory (SUMI) (Kirakowski and Corbett 1993) is a long and complex questionnaire designed to assess "software quality from the end user's point of view."

The System Usability Scale (SUS) (Brooke 1996) is a simpler questionnaire, made up of 10 items, forming up two subscales (Lewis and Sauro 2009): *learnability*, which refers to the ability of quickly and independently learning how to use a system, and *usability* in a more strict sense.

The more recent Usability Metric for User Experience (UMUX) questionnaire (Finstad 2010), made up of 4 items, was designed to make usability evaluation compliant with ISO 9241-11 (International Organization For Standardization 1998; ISO 2018) by including effectiveness, efficiency and satisfaction. Users are requested to express their opinion on each of the 4 items according to a 7-point Likert scale from "Strongly Disagree" to "Strongly Agree."

Fig. 1 Myo gestures for activation, zoom, rotation and CT slice control

Alongside the classic concept of usability, user experience was introduced to cover additional factors such as usefulness, emotional factors and design elegance (Vosinakis and Koutsabasis 2018). Although the questionnaires such as SUMI and UMUX took a first step toward user experience by including a user satisfaction component, other specific questionnaires were designed for user experience evaluation. A first example was the User Experience Questionnaire (UEQ) (Laugwitz et al. 2008), made up of 26 items grouped into 6 scales (attractiveness, perspicuity, efficiency, dependability, stimulation and novelty). The more recent User eXperience Context Scale (UXCS) (Lallemand and Koenig 2020), made up of 30 items, focuses on the evaluation of objective and perceived contextual dimensions in user experience.

In this study, the UMUX and SUS questionnaires were employed to evaluate the perceived usability, while the Presence Questionnaire (PQ) (Witmer and Singer 1998) was used to assess the user experience in the virtual environment. For each questionnaire, a comparative analysis was carried out among three datasets:

- the *Vive* dataset, made up of the item scores collected for the 42 subjects using the Vive controller;
- the *Myo* dataset, made up of the item scores collected for the 42 subjects using the Myo armband;
- the *All* dataset, made up of the item scores collected for all the 84 subjects involved in the test.

The difference between this work and the previous one, described in the introductory section, results in a different type of approach adopted in the analysis of the data collected through the three questionnaires. Whereas in the previous work, Principal Component Analysis was employed to identify interdependencies between the factors into which the items of the questionnaire are grouped, this new analysis addresses the variability of the answers given to each individual item. The first part of the analysis considers the overall scores of the questionnaires in Sect. 6. Then, the differences between the factors/components and the effects on the individual items are evaluated in the subsequent sections.

Besides the mean values, we computed the coefficient of variation, which is the ratio of the scores' standard deviation to the mean value, to evaluate the level of dispersion, which is an indication of how much users' opinions differ. A significantly higher coefficient of variation for an interaction modality suggests users' impressions about a questionnaire item are more in disagreement. Moreover, a high coefficient of variation in the *All* dataset and a low coefficient of variation in the *Vive* and in the *Myo* datasets would suggest a clear influence of the interaction modality on users' opinions, which in this case appear quite different between the two datasets while remaining consistent within each of them.

The datasets collected for the Vive controller and for the Myo armband were compared also by means of a Kruskal–Wallis nonparametric test (Kruskal and Wallis 1952): p values lower than the 0.05 threshold (highlighted in bold in the tables of the following sections) suggest significant differences between the two interaction modalities.

6 Overall scores of UMUX, SUS and PQ questionnaires

Table 1 reports the mean values and the coefficients of variation computed on UMUX, SUS and PQ scores computed on the three datasets. For a correct comparison between scores computed on different scales (as SUS scores varied between 0 and 4 while UMUX and PQ scores varied between 0 and 6), the mean values shown in the table were normalized by dividing SUS scores by 4 and UMUX and PQ scores by 6.

Kruskal–Wallis p values suggest a significant difference between the two interaction modalities on all the three questionnaire scores. The UMUX score is characterized by the highest difference between Vive's and Myo's mean values.

The charts in Fig. 2 compare the probability density functions between *Vive* and *Myo* datasets for UMUX, SUS and PQ scores.

7 UMUX items

The scores of odd items, which measure the user's satisfaction and the perceived ease of use, are computed by subtracting 1 from the scale value chosen by the user. The scores of even items, which represent the perceived frustration and the system reliability, are computed by subtracting the scale value chosen by the user from 7. In this way, all the scores increase with users' positive impressions: high scores for items 1 and 3 denote high levels of satisfaction and ease of use, whereas high scores of items 2 and 4 denote low levels of frustration and the absence of system malfunctions.

For each UMUX item, Table 2 reports the mean values and the coefficients of variation computed over the scores

Table 1 Normalized mean values, coefficients of variation and Kruskal–Wallis p values for the scores of the questionnaires

Questionnaire	All	Vive	Myo	Kruskal–Wallis p value
UMUX	0.8090278	0.8740079	0.7440476	**0.0001141**
	0.1970018	0.1445568	0.220063	
SUS	0.8032738	0.8345238	0.7720238	**0.01724**
	0.1569781	0.1409709	0.1656139	
PQ	0.7974102	0.8241437	0.7706767	**0.03064**
	0.1574491	0.1469569	0.1629878	
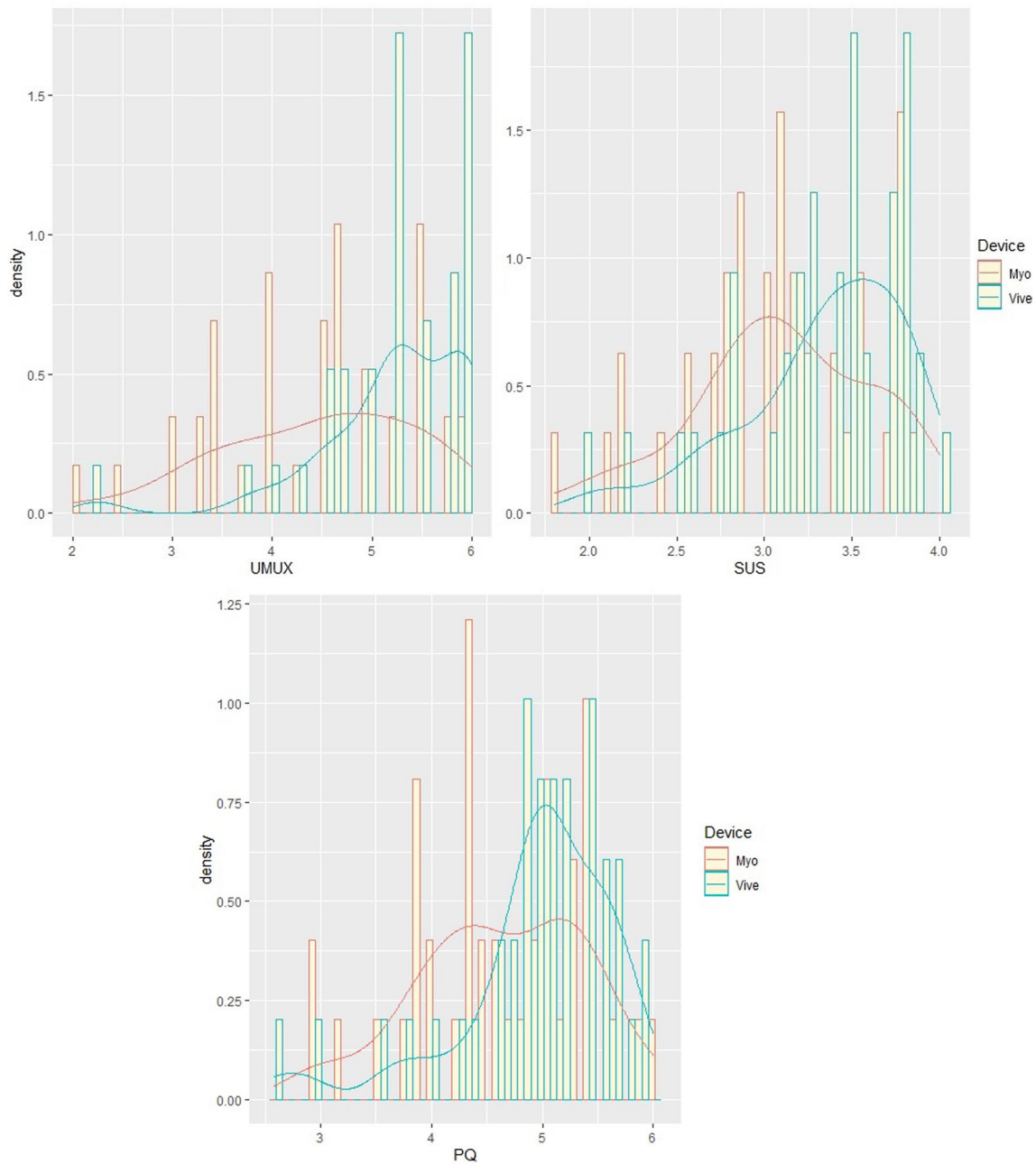
Fig. 2 Density plots for the UMUX, SUS and PQ scores

of the three datasets presented in the previous section. The numbers in brackets specify the ranking in decreasing order based on the mean scores and on the coefficients of variation.

To focus only on the significant differences among such values, the k-means method was employed to cluster the UMUX items according to mean scores and coefficients of variation. The relevant number of clusters was estimated by means of the *NbClust* package (Charrad et al. 2014), available for *R* software. Three clusters for mean scores and three other clusters for coefficients of variation were obtained:

each cluster corresponds to a score level, represented in Table 2 by a number in brackets ranging from 1 to 3. The last column in the table contains the p values obtained through a Kruskal–Wallis test: the p values lower than the 0.05 threshold, which reveal a significant difference between the two datasets, are highlighted in bold.

The mean scores computed on all the 84 subjects are in cluster 2, which represents an intermediate level: this dataset represents a mean between the scores collected for the Vive controller and the scores collected for the Myo armband.

Table 2 Mean values, coefficients of variation and Kruskal–Wallis *p* values for the UMUX questionnaire items

UMUX item	All	Vive	Myo	Kruskal–Wallis *p* value
1. This system's capabilities meet my requirements (UMUX1)	4.976190 (2)	5.095238 (1)	4.857143 (2)	0.575
	0.1922360 (3)	0.1357634 (3)	0.2387656 (2)	
2. Using this system is a frustrating experience (UMUX2)	4.880952 (2)	5.119048 (1)	4.642857 (2)	0.08101
	0.3356783 (1)	0.3026404 (1)	0.3679308 (1)	
3. This system is easy to use (UMUX3)	4.726190 (2)	5.380952 (1)	4.071429 (3)	**5.111e–07**
	0.2774376 (2)	0.1923113 (3)	0.3039427 (1)	
4. I have to spend too much time correcting things with this system (UMUX4)	4.833333 (2)	5.380952 (1)	4.285714 (3)	**0.0003315**
	0.3272635 (1)	0.2320147 (2)	0.3968029 (1)	

The mean scores obtained for the Vive controller experience are in cluster 1, which represents the highest level, whereas those obtained for subjects using the Myo armband are in lower rank clusters. In particular, the most important difference can be noticed on items 3 and 4, which fall into the lowest level for the Myo armband, represented by cluster 3: this suggests both the perceived ease of use and reliability are clearly worse for the Myo armband. This significant difference is confirmed by the Kruskal–Wallis nonparametric test, whose *p* values are lower than the 0.05 threshold. On the other hand, the test reveals no significant difference on items 1 and 2, which express the user's satisfaction and level of frustration: in this case, when comparing the Vive controller and the Myo armband, the average scores simply move from the highest to the intermediate level.

The coefficients of variation, reported in the second row within each cell of Table 2, represent the level of disagreement of users' opinions. In the dataset of the scores of all the 84 subjects, the coefficients of variation of items 2 and 4 are in the highest level cluster: this reveals a significant disagreement on level of frustration and reliability. Users' opinions are highly discordant on these two factors especially in the Myo armband scenario, where there is also a high variability in the perceived ease of use. A high variability about the level of frustration can be noticed also in the Vive scenario, but in this context users seem to strongly agree on a good ease of use. Probably the discordant opinions on ease of use and reliability in the Myo armband scenario can be explained by the different performance of the device in terms of responsiveness, which tends to vary according to the forearm morphology, unlike a handheld device like the Vive controller that exhibited almost the same performance for all the users. On the other hand, the high variability on the level of frustration in both the Vive controller and the Myo armband scenarios suggests there are also other factors, not related to the perceived ease of use and not depending on the interaction modality, that contribute in generating different opinions in users.

The significant differences on items UMUX3 and UMUX4 seem to be confirmed by the charts in Fig. 3, which combine density and histogram plots to compare the distributions of item scores in the *Vive* and *Myo* datasets: high scores (5 and 6) are predominant in both *Vive* and *Myo* datasets for items UMUX1 and UMUX2; on the other hand, the maximum score (6) is highly predominant in *Vive* datasets for items UMUX3 and UMUX4, whereas in the *Myo* dataset a more moderate score (4) is predominant for item UMUX3 and no particular score is clearly predominant for item UMUX4.

8 SUS factors: usability and learnability

Table 3 reports the mean values and the coefficients of variation computed for SUS *usability* and *learnability* components on the scores collected for all the 84 users, for subjects using the Vive controller and for subjects using the Myo armband. For both interaction modalities, *usability* has a higher mean score than *learnability*. Moreover, the higher coefficient of variation for *learnability* suggests that users have more discordant opinions about the ability to learn how to use the system quickly and independently. The last column of the table reports the *p* values of the Kruskal–Wallis test carried out to compare the *Vive* and the *Myo* datasets: only for *usability* the *p* value reveals a significant difference between the two interaction modalities.

The k-means method applied to SUS factors detected two clusters for mean scores and two other clusters for coefficients of variation, represented in Table 3 by the numbers in brackets. In both interaction modalities, the *usability* mean scores fall in the higher level (represented by cluster 1), whereas the *learnability* mean scores are in the lower level (represented by cluster 2).

Coefficients of variation highlight a significant variability in *learnability* scores, especially in the Myo scenario,
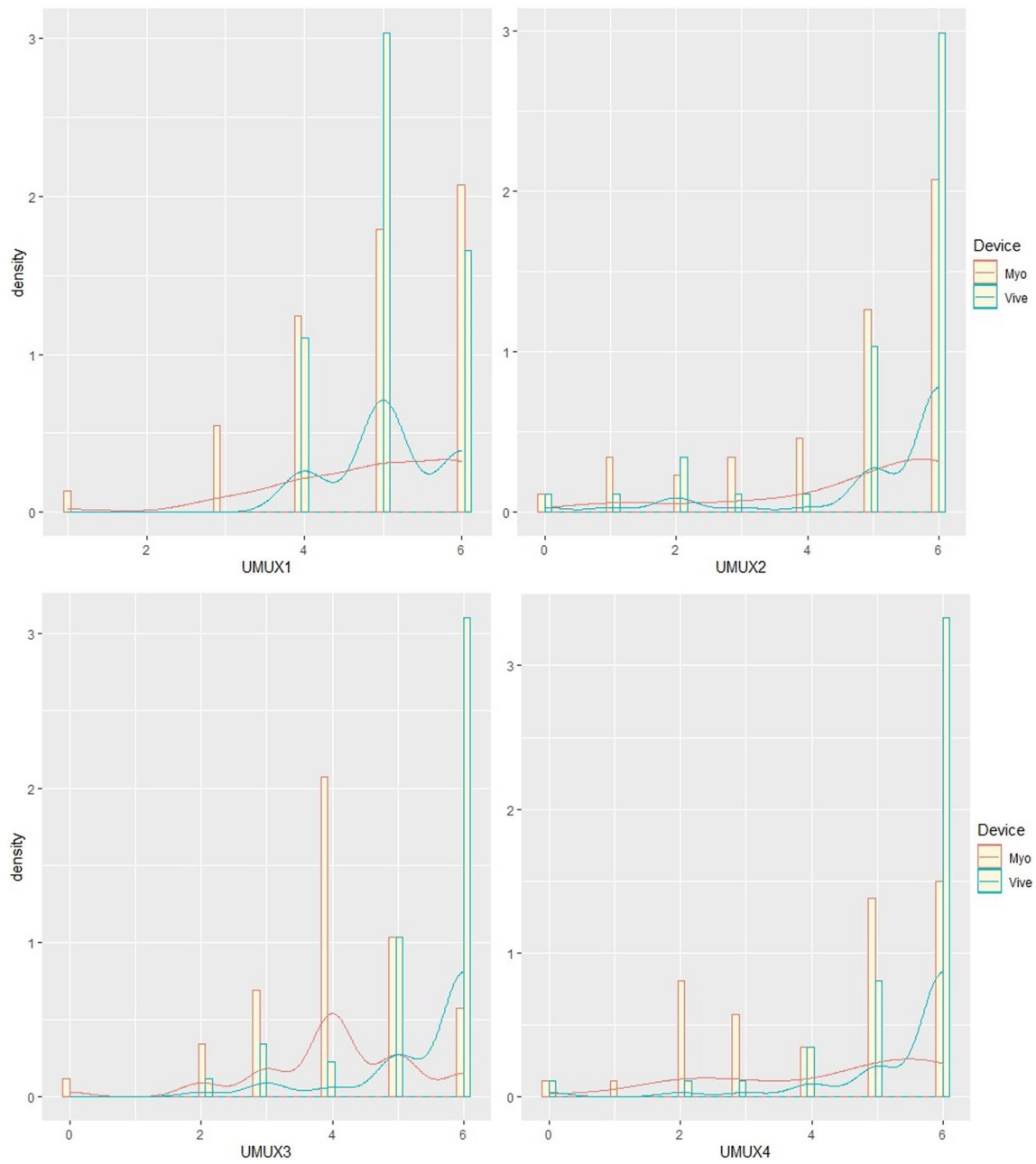
Fig. 3 Density plots for the four UMUX items

Table 3 Mean values, coefficients of variation and Kruskal–Wallis *p* values for the SUS components	SUS component	All	Vive	Myo	Kruskal–Wallis *p* value
	Usability	3.330357 (1)	3.443452 (1)	3.217262 (1)	**0.04096**
		0.1557542 (2)	0.1436906 (2)	0.1626035 (2)	
	Learnability	2.744048 (2)	2.916667 (2)	2.571429 (2)	0.06737
		0.3039724 (1)	0.2509117 (1)	0.3503987 (1)	
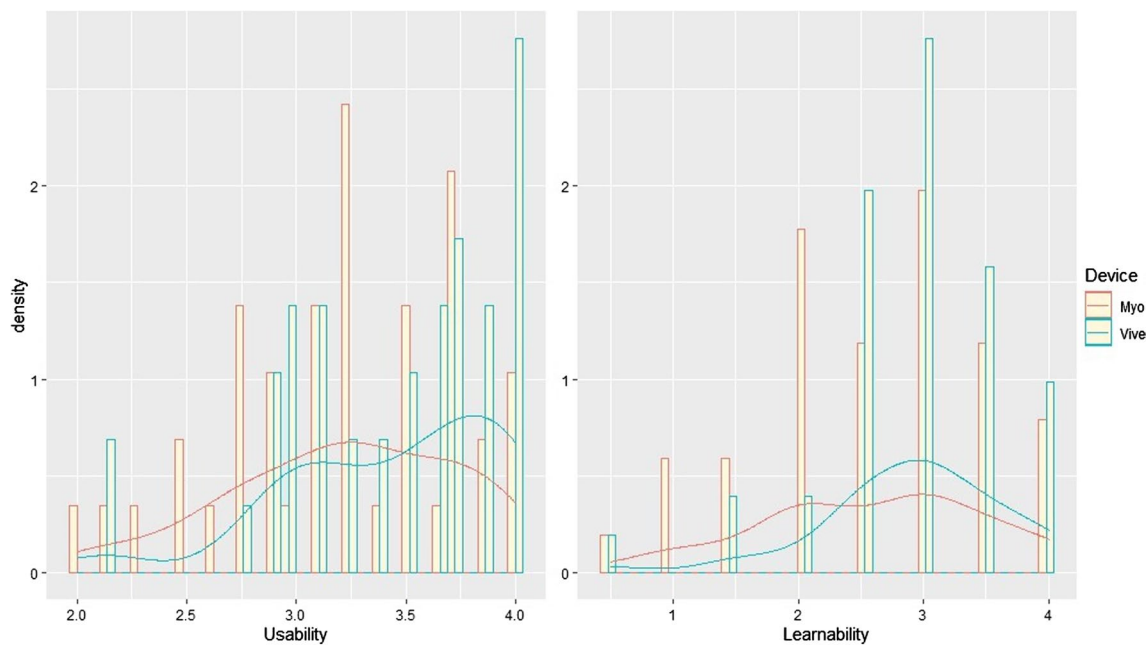
Fig. 4 Density plots for usability and learnability SUS components

where some users probably supposed that responsiveness and accuracy issues in gesture detection were caused by a poor familiarity with the device.

The charts in Fig. 4 compare the distributions of *Vive* and *Myo* datasets for each SUS component: the maximum score (4) is predominant in Vive usability, while there is no clear score dominance in Myo usability; on the other hand, in Vive learnability a moderate score (3) is predominant, followed by an intermediate score (2.5), while Myo learnability scores are mainly divided between moderate (3) and intermediate (2) values.

8.1 Usability and learnability items

Table 4 reports the mean values and the coefficients of variation computed for each single item of the SUS questionnaire.

For items 4 and 10, which make up the *learnability* component, a significant difference can be noticed only on the former, which has also the lowest mean values among all the SUS items in both the interaction modalities: users think there are not many things to learn to use the system, but they think they need the support of a technical person, especially to interact through the Myo armband; however, the very high coefficient of variation ($\simeq$ 0.689) suggests a great disagreement on the latter point.

The mean values in Table 4 were grouped in 4 clusters, as suggested by the *NbClust* package, through the k-means method. Table 5 reports the 4 clusters in decreasing order, corresponding to 4 score levels. The items for which the Kruskal–Wallis test revealed a significant difference between Vive and Myo are highlighted in bold.

Item 3, which represents the ease of use in the strictest sense just as item 3 of the UMUX questionnaire analyzed in the previous section, is greatly influenced by the interaction modality: it achieves the highest score level (1) for the Vive controller, but it drops two levels for the Myo armband, when it reaches level 3. The corresponding *p* value, which is lesser than the 0.05 threshold, confirms there is a significant difference in terms of ease of use between the two interaction modalities. This difference is probably related to a better familiarity with handheld controllers and some reliability issues of the Myo armband, which showed a different accuracy in gesture detection based on the forearm morphology of users wearing it. However, the second aspect seems to take a back seat for users, since no significant difference between the two interaction modalities can be noticed in terms of system consistency, represented by item 6.

Also, the mean values of items 8, 9 and 10, which deal with encumbrance, the confidence acquired with the system and the practice needed to become confident with it, are influenced by the interaction modality, although to a lesser extent, since each of them drops just one score level: however, only on item 9 the Kruskal–Wallis *p* value suggests a significant difference between the two interaction modalities.

For the coefficients of variation of SUS items, the *NbClust* package suggested 3 clusters, corresponding to 3 score levels reported in decreasing order in Table 6. The k-means method produced the 3 clusters represented in Table 6. The items in bold are those with a different score level for the coefficient of variation between the *Vive* and *Myo* datasets. They suggest that opinions on the

Table 4 Mean values, coefficients of variation and Kruskal–Wallis test for SUS scores and their relative errors

SUS item	All	Vive	Myo	Kruskal–Wallis p value
1. I think that I would like to use this system frequently	3.261905	3.333333	3.190476	0.2289
	0.2697031	0.2862712	0.2518722	
2. I found the system unnecessarily complex	3.500000	3.571429	3.428571	0.9825
	0.2449389	0.1869916	0.2962045	
3. I thought the system was easy to use	3.297619	3.642857	2.952381	**1.941e–05**
	0.2653694	0.1994450	0.2987882	
4. I think that I would need the support of a technical person to be able to use this system	2.130952	2.404762	1.857143	**0.0381**
	0.5632216	0.4410381	0.6890689	
5. I found the various functions in this system were well integrated	3.214286	3.238095	3.190476	0.7716
	0.2268855	0.2242522	0.2320673	
6. I thought there was too much inconsistency in this system	3.547619	3.452381	3.642857	0.5656
	0.2603573	0.3147266	0.1994450	
7. I would imagine that most people would learn to use this system very quickly	3.571429	3.690476	3.452381	0.138
	0.1908762	0.1524401	0.2234738	
8. I found the system very cumbersome to use	3.214286	3.333333	3.095238	0.1588
	0.3132457	0.3084185	0.3175894	
9. I felt very confident using the system	3.035714	3.285714	2.785714	**0.003779**
	0.3129106	0.2874413	0.3224022	
10. I needed to learn a lot of things before I could get going with this system	3.357143	3.428571	3.285714	0.3317
	0.2721529	0.2667175	0.2794711	

Table 5 Score levels for the SUS questionnaire items

Score level	All	Vive	Myo
1 (highest)	2, 6, 7	2, **3**, 6, 7, 10	2, 6, 7
2	1, 3, 5, 8, 10	1, 5, 8, **9**	1, 5, 10
3	9	–	**3**, 8, **9**
4 (lowest)	4	**4**	**4**

Table 6 Levels of disagreement among users about the SUS questionnaire items

Score level	All	Vive	Myo
1 (highest)	4	–	**4**
2	1, 3, 6, 8, 9, 10	**1**, **4**, **6**, 8, 9, 10	**2**, **3**, 8, 9, 10
3 (lowest)	2, 5, 7	**2**, **3**, 5, 7	**1**, 5, **6**, 7

consistency of the system (item 6) and on its frequent use (item 1) in the Myo scenario are more in agreement than in the Vive scenario. On the other hand, opinions on system complexity (item 2), ease of use (item 3), need for support (item 4) in the Vive scenario are more in agreement than in the Myo scenario. In particular, the huge disagreement on the need for support is the main cause of the high variability in *learnability* scores highlighted in the previous subsection.

9 PQ: presence

The Presence Questionnaire (PQ) (Witmer and Singer 1998) focuses on the evaluation of the sense of *presence* in virtual environments. It consists of 24 items with seven response options, but only 19 items were considered in this study,

while items about audio and haptic interaction, which are not included in the current version of our simulator, were discarded.

The scores of items 14, 17 and 18 are computed by subtracting the scale value chosen by the user from 7 (since for such items higher values denote worse opinions expressed by the users). The scores of all the remaining items are computed by subtracting 1 from the scale value chosen by the user.

9.1 PQ factors

PQ items can be grouped into four factors (Witmer et al. 2005) representing *involvement* (items 1, 2, 3, 4, 5, 6, 7, 10 and 13), *interface quality* (items 17 and 18), *adaptation/immersion* (items 8, 9, 14, 15, 16 and 19) and *visual fidelity* (items 11 and 12).

Table 7 Mean values, coefficients of variation and Kruskal–Wallis *p* values for PQ components

PQ component	All	Vive	Myo	Kruskal–Wallis *p* value
Involvement	4.895503 (2)	5.079365 (1)	4.711640 (2)	**0.01463**
	0.1608888 (3)	0.1477554 (3)	0.1675680 (3)	
Interface quality	4.446429 (3)	4.309524 (3)	4.583333 (2)	0.4235
	0.3107968 (2)	0.3447037 (1)	0.2777802 (2)	
Adaptation/immersion	4.488095 (3)	4.722222 (2)	4.253968 (3)	**0.01472**
	0.2028900 (3)	0.1842476 (3)	0.2114712 (3)	
Visual fidelity	5.273810 (1)	5.190476 (1)	5.357143 (1)	0.3103
	0.1743863 (3)	0.1817212 (3)	0.1678301 (3)	

As suggested by the *NbClust* package, three clusters for mean scores and three other clusters for coefficients of variation of PQ components were created through the k-means method. In this way, three different score levels were obtained: they are represented by the numbers in brackets in Table 7.

A Kruskal–Wallis nonparametric test was performed on these four PQ components: the *p* values lower than 0.05 for *involvement* and *adaptation/immersion* revealed a significant difference for these two factors between the two interaction modalities.

The charts in Fig. 5 compare the distributions of *Vive* and *Myo* datasets for each PQ factor. In the *Vive* scenario, most *involvement* scores achieve high values in the range between 5 and 5.8, while in the *Myo* scenario most of the scores are distributed within a wider range between 4.2 and 5.5. In the *Vive* scenario, *adaptation/immersion* scores are mainly concentrated around the values of 5.5 and 4.8, while in the *Myo* scenario there are no particular dominant values or ranges. These differences in the charts reflect the significant difference between Vive and Myo for *involvement* and *adaptation/immersion* components. Apart from *visual fidelity*, which is obviously not affected by the interaction mode, *interface quality* also has very similar distributions between the two interaction modes, except for a clear predominance in the *Myo* scenario of values around 5.

9.2 PQ items

Tables 8 and 9 report the mean scores and the coefficients of variation computed for the 19 items of the PQ questionnaire. The *p* values of the Kruskal–Wallis test, reported in the last column of the tables, revealed a significant difference between Vive and Myo on items 1, 2, 3, 5, 14, 15, 16 and 19. In particular, items 1, 2, 3 and 5 belong to the *involvement* factor, while items 14, 15, 16 and 19 belong to the *adaptation/immersion* factor.

To focus only on significant differences among mean values, PQ items with similar mean scores and with similar coefficients of variation were grouped together through k-means clustering. The *NbClust* package was employed to estimate the relevant number of clusters. Two different clusterings, reported in Tables 10 and 11, were obtained: the former groups mean values into four score levels, whereas the latter groups coefficients of variation into three score levels.

The items for which the Kruskal–Wallis test suggested a significant difference are highlighted in bold in Table 10.

PQ items 1, 2, 3, 5, 14 and 16 are the most influenced by the interaction modality, since their scores downgrade by two levels when the Myo armband is employed in place of the Vive controller: in particular, items 1, 2, 3, 5 and 14 downgrade from level 2 to level 4, while item 16 downgrades from level 1 to level 3. On the other hand, the scores of items 15 and 19 downgrade only by one level: the former from level 1 to level 2, the latter from level 3 to level 4: therefore, the interaction modality has only a moderate influence on the adaptation in a strict sense (i.e., the ability to adjust to the virtual experience, represented by item 15) and on the ability to concentrate on tasks and activities (item 19). However, the *adaptation/immersion* factor in a wider sense also includes items 14 and 16, which are heavily influenced: the former represents the delay experienced between actions and outcomes, whereas the latter refers to the proficiency in movements and interactions. Items 8 and 9, which also belong to the *adaptation/immersion* factor, do not present significant differences according to the Kruskal–Wallis test: the former expresses the ability of users to anticipate the response to their actions, whereas the latter represents the ability to survey and search the environment.

PQ items 4, 6, 10, 11, 12 and 13 remain in the highest score level for subjects using the Myo armband, despite the more difficult interaction experienced with such device. In particular, it is important to underline that the sense of object moving through space (item 6) and the ability to examine objects closely (item 11) and from multiple viewpoints (item 12) are not influenced by Myo's usability issues, even though such activities can be carried out also by giving commands through the Vive controller or by performing Myo's gestures
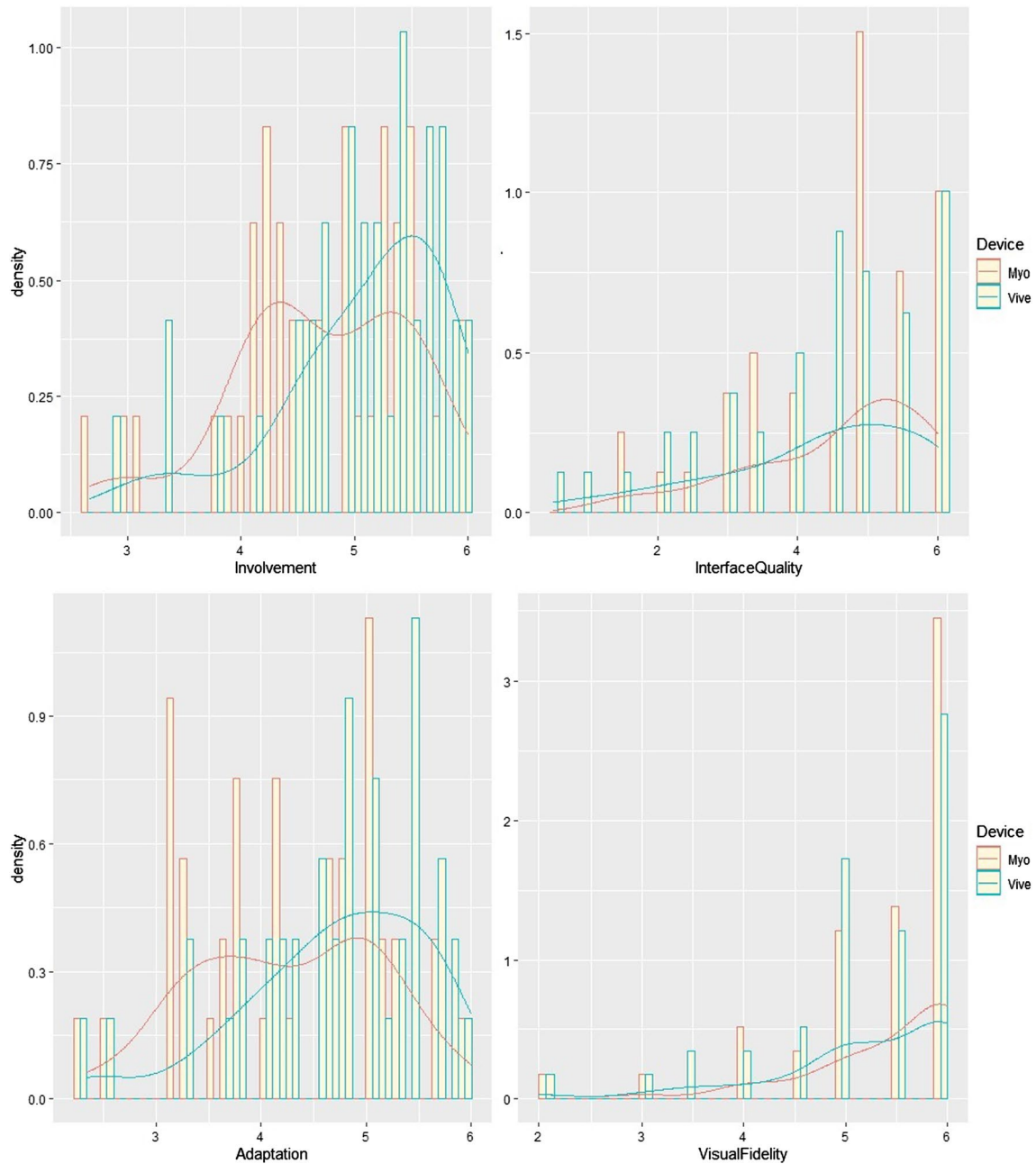
Fig. 5 Density plots for the four PQ components

to zoom or rotate the objects. A possible explanation is that users were more impressed by the possibility of walking toward the virtual objects to get closer and take a look inside them: this modality of examining the virtual objects probably prevailed over the interaction through the Vive controller or the Myo armband.

PQ item 13, representing involvement in a strict sense together with item 4, is not influenced by the interaction modality, but it is worth highlighting that in a wider sense the *involvement* factor also includes items 1, 2, 3 and 5.

Therefore, the *involvement* factor is affected only on its two subcomponents dealing with *control*, made up of items 1 and 2, and *natural interaction*, made up of items 3, 5 and 7. A heavy influence can be noticed on the *control* subcomponent, since the scores of both its items downgrade by two levels. The influence is strong also on the *natural interaction* subcomponent, due to the scores of items 3 and 5 downgrading by two scales, even though item 7 is not significantly influenced by the interaction modality: naturalness in interaction and in movement control is

Table 8 Mean values, coefficients of variation and Kruskal–Wallis *p* values for the PQ questionnaire (items 1–10)

PQ item	All	Vive	Myo	Kruskal–Wallis *p* value
1. How much were you able to control events?	4.523810	4.833333	4.214286	**0.009418**
	0.2921254	0.2661233	0.3079877	
2. How responsive was the environment to actions that you initiated (or performed)?	4.595238	4.976190	4.214286	**0.01622**
	0.3042491	0.2368960	0.3575134	
3. How natural did your interactions with the environment seem?	4.511905	4.833333	4.190476	**0.006175**
	0.2675915	0.2457256	0.2750731	
4. How much did the visual aspects of the environment involve you?	5.226190	5.333333	5.119048	0.5548
	0.2022186	0.1639127	0.2371125	
5. How natural was the mechanism which controlled movement through the environment?	4.452381	4.833333	4.071429	**0.001136**
	0.2709688	0.2326301	0.2890552	
6. How compelling was your sense of objects moving through space?	5.285714	5.357143	5.214286	0.444
	0.1982525	0.1969524	0.2010777	
7. How much did your experiences in the virtual environment seem consistent with your real world experiences?	4.702381	4.809524	4.595238	0.4824
	0.2487453	0.2260842	0.2721408	
8. Were you able to anticipate what would happen next in response to the actions that you performed?	4.214286	4.500000	3.928571	0.07532
	0.3495791	0.3074896	0.3870302	
9. How completely were you able to actively survey or search the environment using vision?	5.047619	5.023810	5.071429	0.9308
	0.2400021	0.2618974	0.2194565	
10. How compelling was your sense of moving around inside the virtual environment?	5.488095	5.452381	5.523810	0.8637
	0.1399860	0.1526573	0.1279364	

Table 9 Mean values, coefficients of variation and Kruskal–Wallis *p* values for the PQ questionnaire (items 11-19)

PQ item	All	Vive	Myo	Kruskal–Wallis *p* value
11. How closely were you able to examine objects?	5.345238	5.238095	5.452381	0.2033
	0.1754365	0.1923449	0.1579402	
12. How well could you examine objects from multiple viewpoints?	5.202381	5.142857	5.261905	0.5804
	0.2083980	0.2128795	0.2058845	
13. How involved were you in the virtual environment experience?	5.273810	5.285714	5.261905	0.6349
	0.1798886	0.1927807	0.1682079	
14. How much delay did you experience between your actions and expected outcomes?	4.523810	4.928571	4.119048	**0.009855**
	0.3335284	0.2665998	0.3874105	
15. How quickly did you adjust to the virtual environment experience?	4.988095	5.190476	4.785714	**0.03413**
	0.2233159	0.2179626	0.2238927	
16. How proficient in moving and interacting with the virtual environment did you feel at the end of the experience?	4.750000	5.142857	4.357143	**0.0005198**
	0.2507540	0.2171681	0.2625199	
17. How much did the visual display quality interfere or distract you from performing assigned tasks or required activities?	4.619048	4.380952	4.857143	0.2814
	0.3325259	0.3903532	0.2712020	
18. How much did the control devices interfere with the performance of assigned tasks or with other activities?	4.273810	4.238095	4.309524	0.9707
	0.4001131	0.4162892	0.3885821	
19. How well could you concentrate on the assigned tasks or required activities rather than on the mechanisms used to perform those tasks or activities?	3.880952	4.452381	3.309524	**0.001816**
	0.4443279	0.3549571	0.5103772	

Table 10 Score levels for PQ items

Score level	All	Vive	Myo
1 (highest)	4, 6, 10, 11, 12, 13	4, 6, 10, 11, 12, 13, **15**, **16**	4, 6, 10, 11, 12, 13
2	7, 9, **15**, **16**	**1**, **2**, **3**, **5**, 7, 9, **14**	9, **15**, 17
3	**1**, **2**, **3**, **5**, **14**, 17, 18	8, 17, **19**	7, **16**, 18
4 (lowest)	8, **19**	18	**1**, **2**, **3**, **5**, 8, **14**, **19**

Table 11 Levels of disagreement among users about PQ items (based on the coefficients of variation)

Level of disagreement	All	Vive	Myo
1 (highest)	8, 14, 17, 18, 19	**17**, 18, 19	**2**, **8**, **14**, 18, 19
2	1, 2, 3, 5, 7, 9, 16	1, **2**, 3, 5, **8**, **9**, **14**	1, 3, **4**, 5, **7**, **16**, 17
3 (lowest)	4, 6, 10, 11, 12, 13, 15	**4**, 6, **7**, 10, 11, 12, 13, 15, **16**	6, **9**, 10, 11, 12, 13, 15

considerably lower in the Myo armband scenario, but the consistency of the virtual experience seems not greatly affected by Myo's usability issues.

The *interface quality* factor includes item 18 about the interferences of the control devices: however, the Myo armband did not generate a higher interference than the Vive controller, even though the high coefficients of variation suggest that users have discordant opinions on both the Vive's and the Myo's interferences, as well as on the ability to concentrate on tasks, represented by item 19.

In general, the differences in terms of coefficients of variation between the two interaction modalities (Table 11) are not so pronounced, since all the items remain at the same level or downgrade at most by one level. The variability of the scores for PQ items 2 and 14, which concern the responsiveness of the environment and the delay between the user's actions and the expected outcomes, is slightly higher in the Myo scenario. This was probably caused by the unpredictable behavior of the Myo device, whose performance varies depending on the morphology of the user's forearm: in particular, during the tests the Myo device sometimes failed in detecting some gestures (especially *spread fingers*, the one on the bottom right in Fig. 1, used to move the plane of the CT slices) when it was worn by thin forearms, with a low muscle and tendon tone. Users experiencing such issues often tried to repeat the same gesture or thought there was a delay in gesture detection. For this reason, a higher coefficient of variation can be noticed on item 16, which suggests a higher variability in the interaction proficiency perceived by Myo's users.

Even though the Kruskal–Wallis *p* values do not highlight any significant difference between the two interaction modalities on the level of consistency of the virtual experience (item 7) and on the ability of users to anticipate the responses to their actions (item 8), a slightly higher variability can be noticed in the Myo scenario also for these two items.

10 Discussion

The interaction modality influences several aspects of the user experience in a virtual environment. The *involvement* factor is the one with the highest number of items influenced by the interaction mode, even though the bar charts in Fig. 5 show a greater difference in the distribution of scores for the *adaptation/immersion* factor.

In the Myo armband scenario, users perceived a lower ability to control events (PQ item 1), a lower responsiveness of the system (PQ items 2 and 14), less natural and less proficient forms of interaction (PQ items 3, 5 and 16), a greater difficulty in adapting to the virtual experience (PQ item 15) and a greater distraction generated by the mechanisms used to perform activities (PQ item 19). Nevertheless, they did not perceive any sensible degradation in task performance.

The SUS questionnaire can be decomposed into two factors (Lewis and Sauro 2009), namely *usability* and *learnability*, which are weakly correlated (Borsci et al. 2009). In particular, the *learnability* factor is made up of items 4 ("I think that I would need the support of a technical person to be able to use this system") and 10 ("I needed to learn a lot of things before I could get going with this system"), whereas item 7 ("I would imagine that most people would learn to use this system very quickly"), contrary to what it may seem, belongs to the *usability* factor together with the remaining items, probably because it refers to the skills of other users (Lewis and Sauro 2009). A previous work (De Paolis and De Luca 2020) highlighted Vive's controls and Myo's gestures have a similar level of intuitiveness, with no significant difference between the two devices in terms of *learnability*, despite Myo's inferior usability. Now, the more detailed analysis presented in this paper on the two learnability items and usability item 7 reveals interesting insights into the structure of the learnability component and how the user sees his/her performance in relation to that of other users. A user generally thinks he/she is the

one who has the main problems with Myo, because he/she believes that there are not many things to learn, although he/she feels the need to be helped by an expert, and thinks that it is easy for other users to learn how to use the system. This also suggests that the learnability component of SUS can be further decomposed into two sub-components: one component, closer to item 10, represents the amount of things to be learned while the other, closer to item 4, concerns the practice needed under the guidance of an expert to master and apply them correctly. There are not many gestures that need to be learned to interact with Myo, but sometimes they require practice before they can be performed accurately. During the experiments, it was actually the device that sometimes exhibited problems in terms of *input veracity*: in particular, the "spread your fingers" gesture, which allows translating the plane of the CT images, was sometimes interpreted by Myo as "wave left" or "wave right" because some users did not stretch their fingers enough when opening their hand. An earlier study (Dolopikos et al. 2021) had explained this ambiguity by pointing out that finger abduction, required to spread fingers, involves the flexor carpi radialis muscles, on which the wrist flexion necessary for the wave-in gesture is also based. In light of this, to avoid usability problems, touchless interaction should be based on a vocabulary of gestures that are well differentiated and clearly distinguishable: the design of the device should be careful to avoid ambiguities that could derive from particularities in the conformation of the user's limbs (e.g., too thin arms in the case of devices based on electromyographic pulses).

However, the significantly lower score on PQ item 19 for the Myo device, which indicates a reduced level of concentration, could also refer to a decrease in focus of attention, caused by certain gestures requiring a higher degree of abstract reasoning. The gestures detected by the Myo armband, designed according to the principle of operation of the device (electromyographic pulse detection), might not be perceived as a natural form of interaction. Ideally, however, gesture design should be oriented toward the principle of *guessability* (Wobbrock et al. 2005), which would allow users to master their use without the need for prior knowledge or training. This could include the definition of gesture vocabularies based on user preferences and attitudes (Dong et al. 2016).

Also during the tests, some of the users experiencing problems using the Myo armband usually tended to repeat gestures, probably because they could not see the results of their actions (the so-called *gulf of evaluation* mentioned in Sect. 3.3) and thought these issues were related to a delay in gesture detection (as reflected in the lower average score for PQ item 14).

Anyway, the lower ease of use (UMUX item 3 and SUS item 3) and reliability (UMUX item 4) of Myo do not generate frustration among users (UMUX item 2), who do not perceive any serious inconsistencies (SUS item 6) and remain generally satisfied with the system (UMUX item 1), probably because the fascination with virtual reality outweighed the difficulties related to the use of the device. In the experimental scenario considered in this work, the user was left completely free to explore the environment without time constraints, trying all the available interaction modes. It would be interesting to assess how the opinions expressed by the users would change in the presence of tasks to be performed within certain timeframes (as in the case of action video games) and how haste might amplify the perception of problems.

It is possible to speculate that leaving the user free to explore the environment autonomously at will, without any particular task to be completed, allowed curiosity and attraction toward the VR environment to take over any difficulties. Such activities can fall within the context of informal learning (Lin et al. 2012): unlike institutional education, it relies on intrinsic motivations, such as personal curiosity, which make it an enjoyable activity.

Moreover, it is likely that the possibility of approaching the organs by walking and exploring them from the inside partially compensates the possible shortcomings of the input device. In general, the possibility to walk in a virtual environment with the same movements as in the real world improves presence, ease of navigation, spatial orientation and movement understanding (Usoh et al. 1999; Chance et al. 1998). Therefore, in the considered scenario, the device used for interaction affects the ability to concentrate but not performance, as users still manage to complete the tasks they undertake. On the contrary, the unreliable behavior of the Myo armband could interfere also with the performance of tasks or the achievement of set goals, generating a greater sense of frustration which would also have a negative impact on user involvement.

11 Conclusions and future work

This paper focused on the usability and the perceived presence in a virtual environment: starting from the findings of a previous work (De Paolis and De Luca 2020), the differences between the HTC Vive handheld controller and the Myo armband touchless device were investigated. The analysis considered both traditional usability, covered by the items of UMUX and SUS questionnaires, and user experience in a virtual environment, assessed by means of the PQ questionnaire. The tests revealed a significant difference in terms of user experience between the two devices and a relevant impact on the sense of presence: we analyzed in detail which components are influenced by

the mode of interaction and how much the users' opinions agree on each of them.

In the considered scenario, the user was left free to interact as he/she liked, combining the interaction through Vive controller or Myo armband with walking in the virtual environment. In this context, it has been seen that the use of the Myo armband partially distracts users from the tasks of exploring the virtual environment, but without generating frustration, probably because the fascination of the experience in the virtual environment tends to prevail over possible problems related to the functioning of the interaction device. Compared to the interaction based on a traditional controller, the Myo-based interaction suffers from problems in both the detection accuracy and the naturalness of abstract gestures.

Another interesting finding is that users feel that they do not have much to learn in order to use the Myo armband, but need the support of an expert, so they probably have more difficulty practising with the device than understanding the gestures: this may suggest that there is a slight prevalence of problems related to accuracy of detection over problems related to intuitiveness of abstract gestures.

For future experiments, we will evaluate the possibility of assigning precise tasks to the user, defining a workflow of actions that he/she will have to perform through certain gestures. In this way, it will be possible to force the user to try at least once with all the interaction possibilities, avoiding that he/she adopts, for example, the walk to approach a virtual object when he/she is not able to perform the zoom gesture.

Then, it will also be possible to establish timeframes within which the user will have to perform the various actions or to assign scores after the achievement of certain objectives. In this way, it will be possible to evaluate how the impact of the different interaction modalities varies on the user experience both in case of guided actions and in presence of gaming policies.

The study could be extended to a wider range of more articulated gestures through the use of camera-based touchless devices, such as the Leap Motion controller (Leap Motion 2021), which can track hands and fingers, or Microsoft Azure Kinect (Microsoft 2021a), which can detect movements of arms and legs. Furthermore, the considered experimental scenario could be ported in a mixed reality environment by adopting Microsoft Hololens (Microsoft 2021b), a headset able to detect hand gestures. These devices will enable the possibility of implementing gestures to grasp virtual objects, which will be compared with the abstract semaphoric gestures addressed in this study in terms of usability and impact on the sense of presence. Moreover, by recording signals through an EEG headset, we could study the perception of affordances and make a comparison with the impressions gathered through the questionnaires.

Declarations

References

Agrawal S, Simon A, Bech S, BÆrentsen K, Forchhammer S (2020) Defining immersion: Literature review and implications for research on audiovisual experiences. https://doi.org/10.17743/jaes.2020.0039

Arsenault D (2005) Dark waters: Spotlight on immersion. pp 50–52

Assila A, Marçal De Oliveira K, Ezzedine H (2016) Standardized usability questionnaires: features and quality focus. J Comput Sci Inf Tech (eJCSIT) 6(1)

Bachmann D, Weichert F, Rinkenauer G (2018) Review of three-dimensional human-computer interaction with focus on the leap motion controller

Bailey S, Johnson C, Sims V (2019) Using natural gesture interactions leads to higher usability and presence in a computer lesson. Adv Intell Syst Comput 826:663–671. https://doi.org/10.1007/978-3-319-96065-4_70

Barbot B, Kaufman JC (2020) What makes immersive virtual reality the ultimate empathy machine? Discerning the underlying mechanisms of change. Comput Hum Behav 111:106431. https://doi.org/10.1016/j.chb.2020.106431

Benalcazar M, Motoche C, Zea J, Jaramillo A, Anchundia C, Zambrano P, Segura M, Benalcazar Palacios F, Perez M (2018) Real-time hand gesture recognition using the Myo armband and muscle activity detection. Institute of Electrical and Electronics Engineers Inc., vol 2017-January, pp 1–6, https://doi.org/10.1109/ETCM.2017.8247458

Benko H, Saponas T, Morris D, Tan D (2009) Enhancing input on and above the interactive surface with muscle sensing. pp 93–100, https://doi.org/10.1145/1731903.1731924

Borsci S, Federici S, Lauriola M (2009) On the dimensionality of the System Usability Scale: a test of alternative measurement models. Cogn Process 10(3):193–197. https://doi.org/10.1007/s10339-009-0268-9

Bowman DA, McMahan RP, Ragan ED (2012) Questioning naturalism in 3D user interfaces. Commun ACM 55(9):78–88. https://doi.org/10.1145/2330667.2330687

Brooke J (1996) SUS—a quick and dirty usability scale

Cain B (2004) A Review of the Mental Workload Literature. NATO RTO-TR-HFM-121-Part-II

Cao S, Nandakumar K, Babu R, Thompson B (2020) Game play in virtual reality driving simulation involving head-mounted display and comparison to desktop display. Virtual Real 24(3):503–513. https://doi.org/10.1007/s10055-019-00412-x

Chance SS, Gaunet F, Beall AC, Loomis JM (1998) Locomotion mode affects the updating of objects encountered during travel: the contribution of vestibular and proprioceptive inputs to path integration. Presence: Teleoperators Virtual Environ 7(2):168–178. https://doi.org/10.1162/105474698565659

Charrad M, Ghazzali N, Boiteau V, Niknafs A (2014) Nbclust: an R package for determining the relevant number of clusters in a data set. J Stat Softw

Chen MY, Tung YC, Wu PJ, Hsu CY, Chyou S, Valstar A, Wang HY, Lin JW (2015) User-Defined Game Input for Smart Glasses in Public Space

Chen K, Liang HN, Yue Y, Craig P (2018) Infrared motion detection and electromyographic gesture recognition for navigating 3D environments. Computer Animation and Virtual Worlds 29(3-4), https://doi.org/10.1002/cav.1829

Clemente M, Rodríguez A, Rey B, Alcañiz M (2014) Assessment of the influence of navigation control and screen size on the sense of presence in virtual reality using EEG. Expert Syst Appl 41(4, Part 2):1584–1592. https://doi.org/10.1016/j.eswa.2013.08.055

Corelli F, Battegazzorre E, Strada F, Bottino A, Cimellaro GP (2020) Assessing the usability of different virtual reality systems for fire-fighter training. In: VISIGRAPP 2020 - Proceedings of the 15th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, INSTICC, SciTe-Press, pp 146–153, https://doi.org/10.5220/0008962401460153

Csikszentmihalyi M (1991) Flow: the psychology of optimal experience. Harper Perennial, New York

Dai Q, Li X, Geng W, Jin W, Liang X (2021) CApG-Myo: A muscle-computer interface supporting user-defined gesture recognition. Association for Computing Machinery, pp 52–58, https://doi.org/10.1145/3479162.3479170

De Mauro A, Mazars J, Manco L, Hernández Fernández A, De Paolis LT (2012) Intraoperative navigation system for spine surgery. Int J Comput Assist Radiol Surg 7:19–22. https://doi.org/10.1007/s11548-012-0700-y

De Paolis LT (2016) A Touchless Gestural Platform for the Interaction with the Patients Data. In: XIV Mediterranean Conference on Medical and Biological Engineering and Computing 2016 (MEDICON 2016), March 31st-April 2nd 2016, Paphos, Cyprus, Springer International Publishing, IFMBE Proceedings, vol 57, pp 880–884

De Paolis LT (2018) Augmented Visualization and Touchless Interaction with Virtual Organs. In: International Conference on Bioinformatics and Biomedical Engineering (IWBBIO 2018) Granada, Spain, April 25-27, 2018, Springer International Publishing, Lecture Notes in Bioinformatics, LNBI 10814, pp 118–127

De Paolis LT, De Luca V (2019) Augmented visualization with depth perception cues to improve the surgeon's performance in minimally invasive surgery. Med Biol Eng Comput 57(5):995–1013

De Paolis LT, De Luca V (2020) The impact of the input interface in a virtual environment: the Vive controller and the Myo armband. Virtual Real 24(3):483–502. https://doi.org/10.1007/s10055-019-00409-6

De Paolis LT, Ricciardi F (2018) Augmented visualisation in the treatment of the liver tumours with radiofrequency ablation. Comput Methods Biomech Biomed Eng: Imaging Vis 6(4):396–404. https://doi.org/10.1080/21681163.2017.1287598

De Paolis LT, De Luca V, Paladini GI (2019) Touchless Navigation in a Multimedia Application: The Effects Perceived in an Educational Context. In: Sixth International Conference Augmented and Virtual Reality, and Computer Graphics (AVR 2019), Santa Maria al Bagno, Italy, June 24-27, 2019, Lecture Notes in Computer Science, LNCS 11614, Springer International Publishing, pp 348–367

Deller M, Ebert A, Bender M, Hagen H (2006) Flexible gesture recognition for immersive virtual environments. In: Tenth International Conference on Information Visualisation (IV'06), pp 563–568, https://doi.org/10.1109/IV.2006.55

Diemer J, Alpers GW, Peperkorn HM, Shiban Y, Mühlberger A (2015). The impact of perception and presence on emotional reactions: a review of research in virtual reality. https://doi.org/10.3389/fpsyg.2015.00026

Dolopikos C, Pritchard M, Bird J, Faria D (2021) Electromyography signal-based gesture recognition for human-machine interaction in real-time through model calibration. Adv Intell Syst Comput 1364 AISC:898–914. https://doi.org/10.1007/978-3-030-73103-8_65

Dong H, Figueroa N, El Saddik A (2016) An Elicitation Study on Gesture Attitudes and Preferences Towards an Interactive Hand-Gesture Vocabulary

Finstad K (2010) The usability metric for user experience. Interact Comput 22(5):323–327

Garber L (2013) Gestural technology: moving interfaces in a new direction [technology news]. Computer. https://doi.org/10.1109/mc.2013.352

Gieser S, Gentry C, LePage J, Makedon F (2016) Comparing objective and subjective metrics between physical and virtual tasks. Lecture Notes Comput Sci 9740:3–13. https://doi.org/10.1007/978-3-319-39907-2_1

Gieser S, Kanal V, Makedon F (2017) Evaluation of a low cost EMG sensor as a modality for use in virtual reality applications. Lecture Notes Comput Sci (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 10280:97–110. https://doi.org/10.1007/978-3-319-57987-0_8

Gorini A, Capideville CS, De Leo G, Mantovani F, Riva G (2011) The role of immersion and narrative in mediated presence: the virtual hospital experience. Cyberpsychol Behav Social Networking 14(3):99–105. https://doi.org/10.1089/cyber.2010.0100

Grandhi SA, Joue G, Mittelberg I (2011) Understanding Naturalness and Intuitiveness in Gesture Production - Insights for Touchless Gestural Interfaces. In: Proceedings of the International Conference on Human Factors in Computing Systems (CHI'11), New York, NY, USA, pp 821–824

Hajesmaeel-Gohari S, Bahaadinbeigy K (2021) The most used questionnaires for evaluating telemedicine services. BMC Med Inf Decis Making. https://doi.org/10.1186/s12911-021-01407-y

HTC (November, 2021) HTC Vive. Retrieved from https://www.vive.com/us/product/vive-virtual-reality-system/

Huang D, Zhang X, Saponas T, Fogarty J, Gollakota S (2015) Leveraging dual-observable input for fine-grained thumb interaction using forearm EMG. Association for Computing Machinery, Inc, pp 523–528, https://doi.org/10.1145/2807442.2807506

Indraccolo C, De Paolis LT (2017) Augmented reality and MYO for a touchless interaction with virtual organs. In: Fourth International Conference Augmented and Virtual Reality, and Computer Graphics (AVR 2017), Ugento, Italy, June 12-15, 2017, Lecture Notes in Computer Science, LNCS 10325, pp 63–73

International Organization For Standardization (1998) Ergonomic requirements for office work with visual display terminals (VDTs) - Part 11: Guidance on usability (ISO 9241-11)

Invitto S, Faggiano C, Sammarco S, De Luca V, De Paolis LT (2015) Interactive Entertainment, Virtual Motion Training and Brain

Ergonomy. In: 7th International Conference on Intelligent Technologies for Interactive Entertainment (INTETAIN 2015), Torino, Italy, June 10-12, 2015, pp 88–94

Invitto S, Faggiano C, Sammarco S, De Luca V, De Paolis LT (2016) Haptic, virtual interaction and motor imagery: entertainment tools and psychophysiological testing. Sensors 16(3):394

ISO (2010) Human-centred design for interactive systems. Ergonomics of human system interaction Part 210 (ISO 9241-210). Iso 9241210 https://doi.org/10.1039/c0dt90114h

(ISO) IOfS (2018) Ergonomics of human-system interaction - Part 11: Usability: Definitions and Concepts. Iso 9241-11:2018(E)

Jerald J (2015) The VR book: human-centered design for virtual reality. Association for Computing Machinery and Morgan & Claypool

Jo YU, Oh DC (2020) Real-Time Hand Gesture Classification Using Crnn with Scale Average Wavelet Transform. Journal of Mechanics in Medicine and Biology 20(10), https://doi.org/10.1142/S021951942040028X

Kang HJ, Shin J, Ponto K (2020) A Comparative Analysis of 3D User Interaction: How to Move Virtual Objects in Mixed Reality. In: 2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), pp 275–284, https://doi.org/10.1109/VR46266.2020.00047

Kim J, Mastnik S, André E (2008) EMG-based hand gesture recognition for realtime biosignal interfacing. pp 30–39, https://doi.org/10.1145/1378773.1378777

Kirakowski J, Corbett M (1993) SUMI: the software usability measurement inventory. Br J Educ Technol

Kosmas P, Ioannou A, Retalis S (2018) Moving bodies to moving minds: a study of the use of motion-based games in special education. TechTrends 62(6):594–601. https://doi.org/10.1007/s11528-018-0294-5

Kratz L, Saponas T, Morris D (2012) Making gestural input from arm-worn inertial sensors more practical. pp 1747–1750, https://doi.org/10.1145/2207676.2208304

Kruskal WH, Wallis WA (1952) Use of ranks in one-criterion variance analysis. J Am Stat Assoc 47(260):583–621

Lallemand C, Koenig V (2020) Measuring the Contextual Dimension of User Experience: Development of the User Experience Context Scale (UXCS). In: ACM International Conference Proceeding Series, New York, NY, USA, https://doi.org/10.1145/3419249.3420156

Laugwitz B, Held T, Schrepp M (2008) Construction and evaluation of a user experience questionnaire. In: Lecture Notes in Computer Science, vol 5298 LNCS, pp 63–76, https://doi.org/10.1007/978-3-540-89350-9_6

Leap Motion (November, 2021) Leap Motion. Retrieved from https://www.leapmotion.com/

Lewis JR (2018) Measuring perceived usability: the CSUQ, SUS, and UMUX. Int J Hum-Comput Interact 34(12):1148–1156

Lewis JR, Sauro J (2009) The factor structure of the system usability scale. In: Lecture Notes in Computer Science, pp 94–103

Li Y, Huang J, Tian F, Wang HA, Dai GZ (2019) Gesture interaction in virtual reality. Virtual Real Intell Hardw 1(1):84–112. https://doi.org/10.3724/SP.J.2096-5796.2018.0006

Li M, Ganni S, Ponten J, Albayrak A, Rutkowski AF, Jakimowicz J (2020) Analysing usability and presence of a virtual reality operating room (VOR) simulator during laparoscopic surgery training. In: 2020 IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2020, pp 566–572, https://doi.org/10.1109/VR46266.2020.1581301697128

Lin AC, Fernandez WD, Gregor S (2012) Understanding web enjoyment experiences and informal learning: a study in a museum context. Decis Support Syst 53(4):846–858

Livatino S, De Paolis LT, D'Agostino M, Zocco A, Agrimi A, De Santis A, Bruno LV, Lapresa M (2015) Stereoscopic visualization and 3-D technologies in medical endoscopic teleoperation. IEEE Trans Ind Electron 62(1):525–535. https://doi.org/10.1109/TIE.2014.2334675

Lou X, Li X, Hansen P, Du P (2020) Hand-adaptive user interface: improved gestural interaction in virtual reality. Virtual Real. https://doi.org/10.1007/s10055-020-00461-7

Lu L, Mao J, Wang W, Ding G, Zhang Z (2020) A study of personal recognition method based on EMG signal. IEEE Trans Biomed Circuits Syst 14(4):681–691. https://doi.org/10.1109/TBCAS.2020.3005148

McMahan R, Ragan E, Bowman D, Tang F, Lai C (2015) Framework for Interaction Fidelity Analysis. TechReport UTDCS-06-15. The University of Texas at Dallas

Microsoft (November, 2021a) Azure Kinect DK. Retrieved from https://azure.microsoft.com/en-us/services/kinect-dk/

Microsoft (November, 2021b) Microsoft Hololens 2. Retrieved from https://www.microsoft.com/en-IE/hololens

Moustafa K, Luz S, Longo L (2017) Assessment of mental workload: a comparison of machine learning methods and subjective assessment techniques. In: Communications in computer and information science

Nilsson N, Nordahl R, Serafin S (2016) Immersion revisited: a review of existing definitions of immersion and their relation to different theories of presence. Hum Technol 12(2):108–134. https://doi.org/10.17011/ht/urn.201611174652

Niño J, Kiss J, Edwards G, Morales E, Ochoa S, Bernier B (2019) Enhancing mobile VR immersion: a multimodal system of neural networks approach to an IMU gesture controller. Soc Imaging Sci Technol. https://doi.org/10.2352/ISSN.2470-1173.2019.2.ERVR-184

Norman D (2013) The design of everyday things: revised and expanded edition. Basic Books

Pai Y, Dingler T, Kunze K (2019) Assessing hands-free interactions for VR using eye gaze and electromyography. Virtual Real 23(2):119–131. https://doi.org/10.1007/s10055-018-0371-2

Rawat S, Vats S, Kumar P (2016) Evaluating and exploring the MYO ARMBAND. In: 2016 International Conference System Modeling Advancement in Research Trends (SMART), pp 115–120, https://doi.org/10.1109/SYSMART.2016.7894501

Rhiu I, Kim YM, Kim W, Yun MH (2020) The evaluation of user experience of a human walking and a driving simulation in the virtual reality. Int J Ind Ergon 79:103002. https://doi.org/10.1016/j.ergon.2020.103002

Riva G, Waterworth JA, Waterworth EL (2004) The layers of presence: a bio-cultural approach to understanding presence in natural and mediated environments. CyberPsychol Behav 7(4):402–416. https://doi.org/10.1089/cpb.2004.7.402

Rovira A, Swapp D, Spanlang B, Slater M (2009) The use of virtual reality in the study of people's responses to violent incidents. Front Behav Neurosci 3:59. https://doi.org/10.3389/neuro.08.059.2009

Ryan M, Press JHU, Firm P (2003) Narrative as virtual reality: immersion and interactivity in literature and electronic media. Parallax. Johns Hopkins University Press, Baltimore

Sanchez-Vives M, Slater M (2005) From presence to consciousness through virtual reality. Nat Rev Neurosci 6(4):332–339. https://doi.org/10.1038/nrn1651

Saponas T, Tan D, Morris D, Balakrishnan R (2008) Demonstrating the feasibility of using forearm electromyography for muscle-computer interfaces. pp 515–524, https://doi.org/10.1145/1357054.1357138

Saponas T, Tan D, Morris D, Balakrishnan R, Turner J, Landay J (2009) Enabling always-available input with muscle-computer interfaces. pp 167–176, https://doi.org/10.1145/1622176.1622208

Saponas T, Tan D, Morris D, Turner J, Landay J (2010) Making muscle-computer interfaces more practical. 2:851–854. https://doi.org/10.1145/1753326.1753451

Schäfer A, Reis G, Stricker D (2021) Controlling teleportation-based locomotion in virtual reality with hand gestures: a comparative evaluation of two-handed and one-handed techniques. Electronics (Switzerland) 10(6):1–21. https://doi.org/10.3390/electronics10060715

Slater M (2003) A note on presence terminology. Presence Connect 3(3)

Slater M (2009) Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. Philos Trans R Soc B: Biol Sci 364(1535):3549–3557. https://doi.org/10.1098/rstb.2009.0138

Slater M, Usoh M, Steed A (1994) Depth of presence in virtual environments. Presence: Teleoperators Virtual Environ 3(2):130–144. https://doi.org/10.1162/pres.1994.3.2.130

Somrak A, Humar I, Hossain MS, Alhamid MF, Hossain MA, Guna J (2019) Estimating VR sickness and user experience using different HMD technologies: an evaluation study. Future Gener Comput Syst 94:302–316

Steam (November, 2021) SteamVR. Retrieved from https://steamcommunity.com/steamvr

Tcha-Tokey K, Loup-Escande E, Christmann O, Richir S (2017) Effects on user experience in an edutainment virtual environment

Thill S, Caligiore D, Borghi AM, Ziemke T, Baldassarre G (2013). Theories and computational models of affordance and mirror systems: An integrative review. https://doi.org/10.1016/j.neubiorev.2013.01.012

Tullis T, Albert B (2013) Measuring the user experience: collecting, analyzing, and presenting usability metrics, 2nd Edn

Unity Technologies (November, 2021) Unity. Retrieved from https://unity.com/

Usoh M, Arthur K, Whitton MC, Bastos R, Steed A, Slater M, Brooks FP (1999) Walking > Walking-in-Place > Flying, in Virtual Environments. In: Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques, ACM Press/Addison-Wesley Publishing Co., USA, SIGGRAPH '99, pp 359–364, https://doi.org/10.1145/311535.311589

Vosinakis S, Koutsabasis P (2018) Evaluation of visual feedback techniques for virtual grasping with bare hands using Leap Motion and Oculus Rift. Virtual Real 22(1):47–62

Vuletic T, Duffy A, Hay L, McTeague C, Campbell G, Grealy M (2019) Systematic literature review of hand gestures used in human computer interaction interfaces. Int J Hum Comput Stud 129:74–94. https://doi.org/10.1016/j.ijhcs.2019.03.011

Wan B, Wu R, Zhang K, Liu L (2017) A new subtle hand gestures recognition algorithm based on EMG and FSR. Institute of Electrical and Electronics Engineers Inc., pp 127–132, https://doi.org/10.1109/CSCWD.2017.8066682

Waterworth EL, Waterworth JA (2001) Focus, locus, and sensus: the three dimensions of virtual experience. CyberPsychol Behav 4(2):203–213. https://doi.org/10.1089/109493101300117893

Webster R, Dues J (2017) System Usability Scale (SUS): Oculus Rift® DK2 and Samsung Gear VR®. In: 2017 ASEE Annual Conference & Exposition, ASEE Conferences, Columbus, Ohio

Wirth M, Grad S, Poimann D, Richer R, Ottmann J, Eskofier B (2018) Exploring the Feasibility of EMG Based Interaction for Assessing Cognitive Capacity in Virtual Reality. Institute of Electrical and Electronics Engineers Inc., vol 2018-July, pp 4953–4956, https://doi.org/10.1109/EMBC.2018.8513213

Witmer BG, Singer MJ (1998) Measuring presence in virtual environments: a presence questionnaire. Presence: Teleoperators Virtual Environ 7(3):225–240

Witmer BG, Jerome CJ, Singer MJ (2005) The factor structure of the Presence Questionnaire

Wobbrock JO, Aung HH, Rothrock B, Myers BA (2005) Maximizing the guessability of symbolic input. New York, pp 1869–1872

Yu M, Zhou R, Wang H, Zhao W (2019) An evaluation for VR glasses system user experience: the influence factors of interactive operation and motion sickness. Appl Ergon 74:206–213

Zenner A, Makhsadov A, Klingner S, Liebemann D, Krüger A (2020) Immersive process model exploration in virtual reality. IEEE Trans Vis Comput Graphics 26(5):2104–2114. https://doi.org/10.1109/TVCG.2020.2973476

Zocco A, Zocco MD, Greco A, Livatino S, de Paolis LT (2015) Touchless interaction for command and control in military operations. In: Lecture Notes in Computer Science, Springer International Publishing, pp 432–445, https://doi.org/10.1007/978-3-319-22888-4_32