Check for updates

Factors Influencing Engagement in Hybrid Virtual and Augmented Reality

YUE LI, Department of Computing, School of Advanced Technology, Xi'an Jiaotong-Liverpool University, China
EUGENE CH'NG, NVIDIA Joint-Lab on Mixed Reality, University of Nottingham, China
SUE COBB, Human Factors Research Group, Faculty of Engineering, University of Nottingham, United Kingdom

Hybridity in immersive technologies has not been studied for factors that are likely to influence engagement. An noticeable factor is the spatial enclosure that defines where users meet. This involves a mutual object of interest, contents that the users may generate around the object, and the proximity between users. This study examines these factors, namely how object interactivity, user-generated contents (UGC) and avatar proximity influence engagement. We designed a Hybrid Virtual and Augmented Reality (HVAR) environment that supports paired users to experience cultural heritage in both Virtual Reality (VR) and Augmented Reality (AR). A user study was conducted with 60 participants, providing assessments of engagement and presence via questionnaires, together with mobile electroencephalogram (mEEG) and user activity data that measures VR user engagement in real-time. Our findings provide insights into how engagement between users can occur in HVAR environments for the future hybrid reality with multi-device connectivity.

CCS Concepts: • **Human-centered computing → HCI design and evaluation methods**; **Mixed / augmented reality**; **Virtual reality**; **User studies**.

Additional Key Words and Phrases: virtual reality, augmented reality, mixed reality, engagement, presence, electroencephalogram, mEEG

1 INTRODUCTION

Hybridity in immersive technologies will be a future trend for multi-device connectivity as users adopt preferred devices between levels of Extended Reality (XR) experiences. The two most prominent technologies in the spectrum of reality are Virtual Reality (VR) and Augmented Reality (AR). The former immerses users completely in a simulated environment, whereas the latter augments information and virtual objects onto the real world. We propose that hybrid environments that support both VR and AR is possible and will be a valid space from which connections between users and between users and objects are possible and necessary. Separate studies have found that both VR and AR can provide engaging user experiences for education, exhibition enhancement, and exploration, etc. [2]. Studies have also found that engagement is a significant quality of user experience [51] and an important component of presence [35, 76]. Facilitating engagement is a desired goal of VR and AR. However, the majority of research focusing on engagement uses post-experiment questionnaires as subjective measures. A

Authors' addresses: Yue Li, Department of Computing, School of Advanced Technology, Xi'an Jiaotong-Liverpool University, 111 Renai Road, Suzhou, China, 215123, yue.li@xjtlu.edu.cn; Eugene Ch'ng, NVIDIA Joint-Lab on Mixed Reality, University of Nottingham, 199 Taikang East Road, Ningbo, China, 315100, eugene.chng@nottingham.edu.cn; Sue Cobb, Human Factors Research Group, Faculty of Engineering, University of Nottingham, University Park, Nottingham, United Kingdom, NG7 2RD, sue.cobb@nottingham.ac.uk.

1073-0516/2023/3-ART $15.00
https://doi.org/10.1145/3589952

lack is that the subjective assessment of an overall evaluation is not able to explain user engagement in relation to interactions within a session, such as minute or second intervals or at different distances where users have engaged with objects.

In this study, we investigate the factors that have necessary influences on user engagement in Hybrid Virtual and Augmented Reality (HVAR) environments. The three aspects are object interactivity, user-generated contents (UGC), and avatar proximity. We examined the effects of these factors within HVAR populated with virtual objects of high, medium, and low interactivity by connecting paired users in VR and AR respectively, each embodied in a full-body avatar. Similar to previous works, we collected subjective assessment of engagement and presence using established questionnaires. In the meantime, we validated an objective measure of engagement with the combined use of 1) psychophysiological measure using a mobile electroencephalogram (mEEG), and 2) user activity monitoring at system runtime. We recorded the brain signal data of each user via mEEG that continuously measures brain signals that yield real-time indices of user engagement, arousal, and valence at $t=1$ second intervals; we also implemented a user activity monitoring function that records users' gaze and controller interactions at the intervals. The combined objective approach has allowed us to obtain the real-time engagement index throughout the VR session, contributing to an objective measure of user engagement in relation to interaction in HVAR.

Our data analysis of 60 participants in 30 pairs confirmed positive correlations between perceived presence and engagement, and that both levels of presence and engagement are greater in VR than with AR. We demonstrated that objects of high interactivity are able to sustain users' attention for longer periods of time, whereas users' real-time engagement index was higher for objects of low interactivity. Our analysis further showed that users' gaze on UGC and information labels attached to objects resulted in significantly higher engagement and arousal indices than viewing and interacting with objects alone. We observed that users spent more time reading labels around objects of low interactivity (74.36%), whereas more time was spent viewing and interacting with objects of medium and high interactivity. This can account for the higher engagement index for low interactivity objects. mEEG data indicated that engagement and arousal indices were significantly higher when UGC was presented rather than not, and significantly higher for UGC labels as compared to object information labels. These findings demonstrated the positive effects of UGC on user engagement in hybrid environments. For avatar proximity, users maintained an average distance of around 3.03 meters between each other. Engagement and arousal indices were the greatest when users were in the wider range of social distance (between 2 to 4 meters). It decreased in a closer proximity (between 1 to 2 meters), due to user expectations on the visual appearances of avatars and social interactions.

Our research offers three contributions to the community. First, our results provide a fundamental understanding of the factors influencing engagement in hybrid virtual and augmented reality environments. Second, our findings provide insights into how user engagements in hybrid environments can be supported through the use of virtual object, UGC, and avatar proximity. Third, we highlight the need for objective measures on top of subjective questionnaires, that using a combined psychophysiological and user activity monitoring is a necessary complement to ground-truthing subjective measures for user engagement and interaction in hybrid realities.

2 RELATED WORK

2.1 Conceptualisations of Engagement and Presence

In the community of VR and AR, presence refers to the sense of 'being there' [25, 26, 41, 63, 68], and engagement is considered as a determinant of presence [35, 39, 60, 67, 76]. It is worth noting that *engagement* and *involvement* are often used interchangeably to refer to a state of focused attention or interest [64]; *presence* is sometimes used interchangeably with *immersion* [43], but can refer to different meanings. We would like to make a distinction here based on our observations. When presence is used synonymously with immersion, it often refers to spatial presence, or telepresence [47]. The scope of presence, in this case, often excludes perception mediation to denote the sense of the physical space as determined by media characteristics and external factors [35, 67] and emphasises place illusion as constrained by the sensorimotor contingencies afforded by the virtual reality systems [65]. However, a broader scope of presence also includes plausibility illusion, an experience that makes users believe that the scenario is actually happening, such as the feeling of an avatar 'looking' at you [65]. In this case, immersion indicates the technical capabilities of a system and is part of presence [6, 66, 68].

The broader scope of presence takes into account more than immersion, but also user characteristics and internal factors [29, 35, 67], with consideration of users' mental models and perception mediation. In this regard, engagement (or involvement) is considered as a component of presence that accounts for users' cognition, perception and emotion. This conceptualisation of presence as determined by both media characteristics and user characteristics is more comprehensive and well acknowledged in recent literature. However, some researchers have considered engagement and presence as two logically orthogonal components. For example, Cummings and Bailenson [14] stated that 'a user can feel spatially present in a virtual environment designed to be boring without feeling engaged in it or cognitively involved'. Nevertheless, the orthogonal relationship was hinting more on the relationship between engagement and presence in the sense of physical space. In fact, many factor analyses of presence [35, 60, 75] included both immersion and engagement as factors of presence, where the immersion factor encapsulates the sense of the physical space and the engagement factor indicates the psychological state experienced as a consequence of focused attention and enjoyment. Therefore, we hold the stance that engagement is a component of presence, and presence includes both media and user characteristics. It indicates more than the illusion of being there, but also users' perceived realness of a mediated or virtual experience [64].

2.2 Evaluation of Engagement and Presence in VR and AR

Engagement and presence have been used to evaluate games and they were often found to be positively correlated [9, 44]. However, comparisons between virtual and real environments [7, 10] found that spatial presence and ecological validity were higher in the real environment, whereas engagement and negative effects such as disorientation, tiredness, eyestrain and nausea were greater in the virtual environment. This made us question whether there are positive correlations between engagement and presence in hybrid VR and AR environments. Tang et al. [71] compared the use of VR and AR in two sessions for paired users discussing personal preferences of two cellular phone models. They found no significant differences in engagement, but less spatial presence for users in immersive VR as compared to AR because the condition of AR was only partially mediated. Dow's [17] study on an interactive drama found that increased presence did not result in increased sense of engagement in AR experience. He explained that an increase in presence can make users feel too close to the action, whereas lower presence can create a sense of distance which some users found more comfortable to engage with. We observed that the comparisons in aforementioned studies were conducted for two separate single-user sessions. It is not clear what the relationship between engagement and presence will be in situations where multiple users connect, such as the hybrid use of VR and AR. In addition, the metrics differ in previous studies in relation to presence for AR and the real environments due to the lack of clarity. For example, in Tang et al.'s study, the spatial presence felt in AR was greater than VR due to the fact that the AR users evaluated presence in terms of the real world they were situated at, where interactions were largely unmediated. Busch et al. and Brade et al.'s comparisons of spatial presence in the virtual and real environments have demonstrated similar results. However, Dow's evaluation of presence for AR was based on the simulated narrative of the interactive drama. Researchers have also studied augmented virtual objects in AR and evaluated the level of presence felt for these objects as compared to real physical objects [58, 70]. Therefore, it is important to clarify what the measure is targeting: the virtual objects or the entire environment, the simulated environment or the real world. The questionnaires need to specify the target subject, otherwise results will be ambiguous.

2.3 Factors Affecting Engagement around Virtual Objects

2.3.1 Object Interactivity. The ability to interact with the mediated environment is believed to be the most important factor in perceived presence [39]. Interactivity is 'the extent to which users can participate in modifying the form and content of a mediated environment in real-time' [69]. We therefore define object interactivity as the extent of interaction possibilities to which users can perform with a virtual object in real-time. Object interactivity is thus related to the affordances of an object. Previous research has demonstrated that the richness of control over environments and objects can trigger positive factors of user engagement such as enjoyment, fun, and physiological arousal [50, 59]. Overall, interactivity can contribute to user engagement through real-time actions provided to users in both the environment and the objects as contents of the environment. Objects in virtual environments are by nature digital. Virtual objects can therefore provide a richer set of interaction possibilities beyond that of our natural physical constraints. Users have reported greater user experience and less simulator sickness when free controls for movements and manipulating information are provided in a virtual environment [38]. As user interactions around a virtual object are determined by the amount and variety of controls they have, object interactivity is arguably the most fundamental aspect of affordances virtual objects can have. However, the effects of object interactivity on user engagement around virtual objects have not been examined for hybrid environments. Research is needed to understand how object interactivity and of what type of interactions can contribute to improve user engagement.

2.3.2 User-Generated Contents for Virtual Objects. We propose that hybird virtual and augmented reality environment is essentially social, for which user-generated contents (UGC) such as comments can add to the affordances of virtual objects. Users develop engagement with the platform and with other users [30], and the engagement of one user can drive the engagement of the other [53]. A hybrid multiuser environment that supports both VR and AR is becoming a necessary trend for future social interactions. Li et al. [36] studied the technology acceptance of hybrid VR and AR environments and demonstrated that virtual objects can be the shared focus of interest in mediating communication between users in VR and AR. It is reasonable to speculate that endowing virtual objects with UGC can facilitate better user engagement, especially for hybrid VR and AR environments. For objects in museums and exhibitions specifically, UGC presents the user subjective interpretations of objects and provides them with a different perspective other than what was intended by the curators [27]. Shaby et al. [62] argued that UGC is able to facilitate social interactions and increase user engagement with museum objects. Based on the previous work, we are interested to know if UGC can facilitate user engagement around virtual objects in hybrid environments.

2.3.3 Avatar Proximity. Hall [23] introduced the proxemics theory to study how people perceive and manage their spatial relationship with others in order to achieve communication goals. Hall correlated the physical distance to social distance between people, and regarded 0.5-1 meter as personal distances, and 1-4 meters as social distances. The proxemics theory has been used to inform the design of interactive systems with five key dimensions: distance, orientation, movement, identity and location [42]. Researchers have applied the theory in the system design of a shared display between multiple users, for which system can proactively react to user interactions around the display based on user distances and movements [73]. It has also been used to indicate interaction possibilities, such as by sensing the presence of a user around an object to enable interactions within close range [36]. We believe that effects of virtual proximity are similar to the physical world. Therefore, our present research measures and records distances between virtual avatars in order to understand how proximity influences engagement within mediated environments.

2.4 Experimental Methods for Engagement and Presence

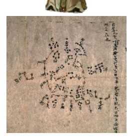
Engagement and presence can be measured subjectively and objectively [16, 61]. These two types of measures are complementary and will provide more comprehensive evaluations. Subjective measures include questionnaires, interviews, observations and other self-reported measures that provide such data. Self-reported questionnaire is the most frequently used measure for engagement [8, 50, 52] and presence [35, 40, 67, 76]. Questionnaire measures yield subjective perspectives on user experience, but the retrospective nature of asking questions post-experiment only allows an overall evaluation on what was felt in the past and not during the process. Previous research has also used psychophysiological metrics for objectively measuring engagement [5] and presence [46, 66]. These brain signal data from EEG devices was shown to be effective in accessing the engagement level of a person with indices construed from alpha (8-13 Hz), beta (13-30 Hz), and theta (4-8 Hz) waves [19, 45, 48, 56]. In addition, users' left and right frontal activity indicated in the alpha and beta waves can demonstrate a person's arousal and valence [22]. Data measured with psychophysiological devices shows continuous real-time data at $t=1$ second intervals, which can help to gain an objective understanding of real-time user engagement. Standard EEG devices with electrodes that cover the scalp can collect more information than mobile EEG devices, which helps avoid the loss of data and perform better in detecting important clinical signals. However, these devices are cumbersome and require extensive experimental settings, time, and expertise [20]. Recent work has validated the use of the Muse headband as a viable tool for research and medical diagnosis, such as for the study of visuospatial attention [33], well-being [11], and the prediction of stroke [74]. Some showed data collection results that are comparable to those recorded by research grade products [32], demonstrating its valid measure of neural responses associated with the engagement of cognitive control and perceptual processing. The monitoring of participant experience in VR with mEEG shows two strengths [55]: it does not interfere with the immersive experience, and it provides continuous measures versus a one-off, overall measure, making it possible to investigate user experience of specific activities with a high granularity. Despite these great potentials, researchers have also identified that outcomes from such consumer-grade EEG devices should be interpreted with caution, especially for symptom diagnosis [21] and the monitoring of mental health [1].

Another objective measure of engagement is user activity monitoring, also known as behaviour tracking [34]. Online user activity data have been used to analyse user engagement in online communities, such as blog visiting [28]. This approach has two significant advantages: scale and validity [34]. The continuous monitoring of user activity can generate data at large scales, which records actual user interactions as opposed to getting the users to recall what was felt using questionnaires and interviews. User activity monitoring at system runtime is applicable in VR and AR environments as they are by nature digital. In fact, objective data from the logging of user's real-time positions, gaze directions, and controller interactions can complement questionnaires and reveal information that is otherwise impossible with subjective post-experiment questionnaires. For example, Li et al. [38] recorded user gaze time on each painting in a virtual exhibition and compared it with the length of audio information, indicating how much they have engaged with each painting. Ch'ng and Cooke [12] analysed gaze patterns with multimodal behaviour for natural and task-based activities for users interacting with virtual objects and found that it can reveal interaction intention in terms of position and duration of visual attention. Similarly, Tennent et al. [72] analysed user activities within a virtual exhibition by keeping logs of the headset positions and directions. Doherty and Doherty [16] argued that the combination of subjective and objective measures can provide a better understanding of engagement. Therefore, in addition to using self-reported questionnaires to provide retrospective and subjective evaluations, we propose an objective measure of engagement by combining psychophysiological measure using mEEG and user activity monitoring. The combined approach supports continuous real-time data collection and provides objective measures of engagement at $t=1$ second intervals. As a consequence, the merging of mEEG data and user activity data allows us to define and classify user engagement for different object interactivity, avatar proximity, and user interactions in VR. To the best of our knowledge, this combined approach has not been explored in previous work.

3 HYBRID VR AND AR ENVIRONMENTS

We begin our experiment by developing a Hybrid Virtual and Augmented Reality (HVAR) environment populated with virtual objects of high, medium, and low interactivity and by connecting paired users with full-body avatars. Details of the virtual objects are shown in Table 1.

Table 1. Overview of seven virtual objects.

#	Name	Picture	Museum	Object interactivity
1	*Xie Zhi* (Pottery Unicorn)		Shaanxi History Museum	High: Grabbable; can be used to 'attack' a menacing face with the horn
2	Bronze Music Instrument		Tianjin Museum	High: Can be drummed to trigger sound effects
3	Bronze Mask with Protruding Pupils		Sanxingdui Museum	Medium: Grabbable
4	Tri-coloured Camel		Nanjing Museum	Medium: Grabbable
5	Pottery Figure of a Standing Lady		National Palace Museum, Taipei	Medium: Grabbable
6	Figure of an Assistant to the Judge of Hell		British Museum	Low
7	Chinese Star Chart		British Library	Low

3.1 Materials

Our HVAR system supports use of VR head-mounted display (HMD) and allows users to explore the virtual environment paired with AR users that have access to augmented virtual objects using smartphones. The experiment used the HTC Vive system for VR and the Android OS with the Samsung Galaxy S7 smartphone for AR. The system was developed in Unity with the SteamVR SDK and Google ARCore. Virtual avatars and objects were synchronised in the hybrid space via a Wireless Local Area Network. We used anthropomorphic full-body virtual avatars (a male and a female, see Fig. 1). Information about the objects on display were collected from the museums and their websites, and were made available in both English and Chinese in view of the demographic of our users. The UGC labels of objects #1-6 contain user comments collected from the interviews, questionnaires, and conversations during our previous user studies [36, 37] and our three public exhibitions in Beijing, Shanghai and Hong Kong. Comments for object #7 were randomly collected from social media. We designed a tutorial room to allow users to practice navigation and interaction controls prior to the actual study.

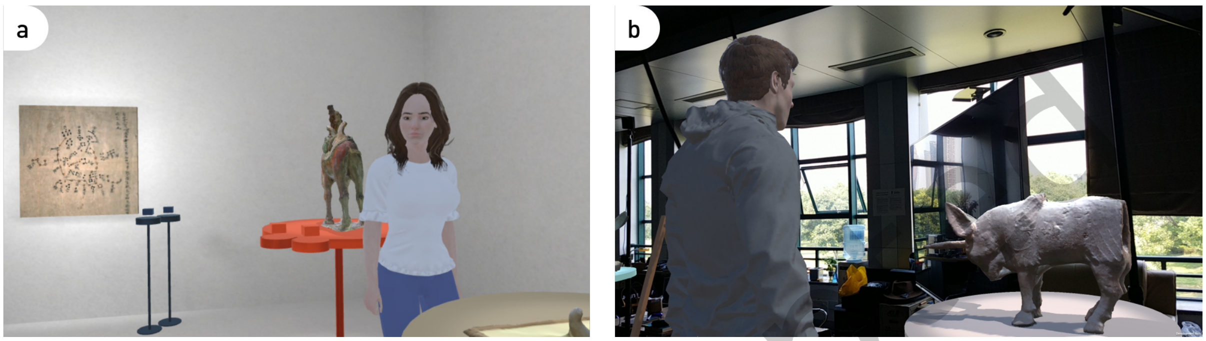


Fig. 1. Hybrid VR and AR environments: a) VR environment, b) AR environment.

3.2 VR Controls

The virtual environment is a simulated exhibition room with seven virtual objects (see Table 1 and Fig. 2a). Users in the virtual environment are given control of 1) movements and 2) virtual objects. Users can walk around physically in the tracking boundary (3.5m x 3.5m) as well as using handheld controllers to teleport. Teleportation points were preset around virtual objects, allowing users to quickly and precisely position themselves on the four sides (front, back, left, right) of the display stand (see Fig. 2b). Information and UGC labels of each object can be accessed by pressing down the two virtual buttons using either handheld controllers (see Fig. 3a). The labels are dismissed once users have released the virtual buttons. VR users can see their virtual hands and controllers but not their own full-body avatars. However, they can see the full-body avatar of the AR users around the observed virtual object (see Fig. 1a).

3.3 AR Controls

The AR application starts with ARCore's ground detection algorithm using the smartphone camera. It recognises a flat surface where virtual objects can be placed onto (see the white grids in Fig. 2c). By tapping on the detected ground on the screen, the virtual objects are placed on location that mirrors the layout as in VR, with the only difference being that the room enclosing the objects were not presented in the AR environment. Objects in AR were also placed on the virtual tables (see Fig. 2c). The application tracks the camera position relative to the objects so that users' movements are mapped within the augmented environment. Users in AR can interact with the objects the same way as VR users, except for the grabbing of objects. Virtual objects displayed on the screen can be rotated by the 'tap and drag' gesture on the touch screen to orientate and view the virtual objects. Sound effects can also be triggered by tapping on the music instrument. Information and UGC labels can be displayed on the user interface by tapping on the information and comments icons (see Fig. 3b and 3c). Labels can be dismissed by tapping on the label or the cancel icon. The full-body avatar of the VR user around the observed virtual object can be seen by the AR user (see Fig. 1b).

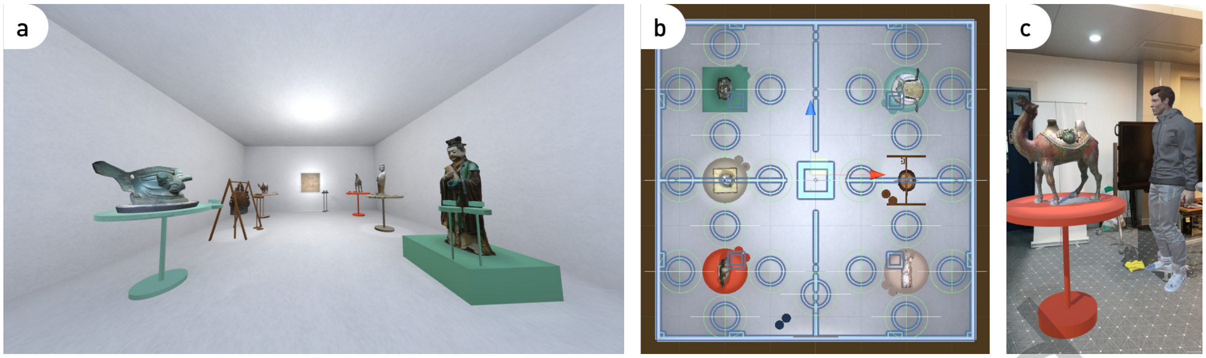


Fig. 2. Virtual exhibition room layout in VR and AR: a) Layout of the seven virtual objects, which is identical in VR and AR environments; b) Teleport points and area (top-down view) in VR; c) A virtual object on a virtual table in AR.

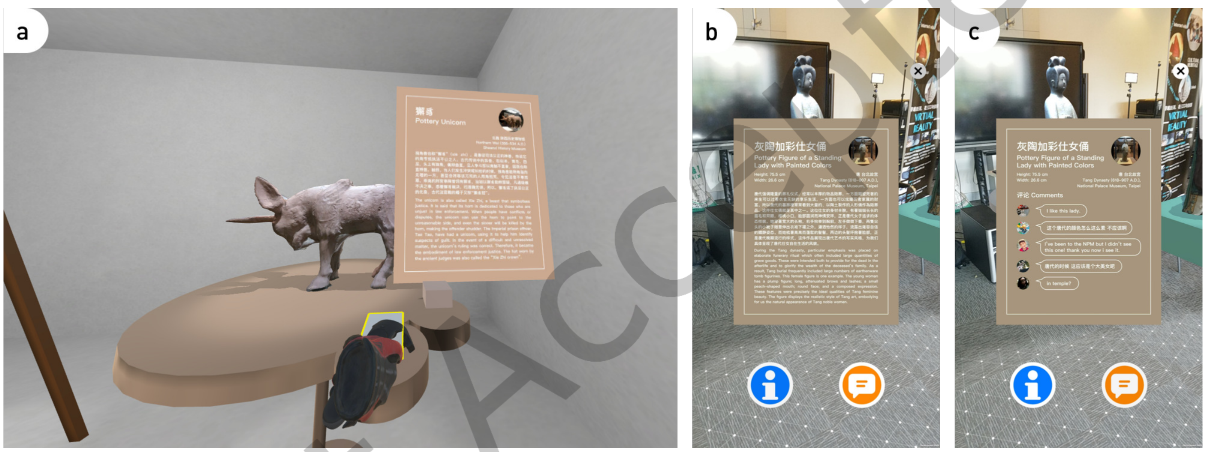


Fig. 3. Interactions with labels in VR and AR: a) VR information label; b) AR information label; c) AR UGC label.

4 EXPERIMENTAL METHODS

4.1 Research Questions and Hypotheses

We carried out an experimental study to investigate this research question: *what are the factors influencing user engagement in hybrid VR and AR environments?* Our prior discussion of related work demonstrated the need to examine the effects of object interactivity, user-generated contents (UGC) and avatar proximity on user engagement. Presence was also investigated for two reasons. First, users' perceived presence is a strong indicator of the overall user experience in VR and AR. Second, the perceived presence and its relationship with engagement in hybrid VR and AR environments are not well understood. Specifically, we investigated if perceived presence and engagement are positively correlated in in HVAR (**H1**). By comparing two sessions: 1) objects with information labels (HVAR-Info), and 2) objects with both information labels and UGC labels (HVAR-UGC), we investigated the effects of UGC on the subjective assessment of engagement (**H3**). Furthermore, the combined psychophysiological measure and user activity monitoring allow us to study the effects of object interactivity

(**H2**), UGC (**H3**), and avatar proximity (**H4**) on real-time user engagement. We propose and test the following hypotheses:

H1. Engagement is greater when higher presence is perceived in HVAR.
H2. Engagement is greater for objects of higher interactivity in HVAR.
H3. Engagement is greater when UGC on virtual objects is presented in HVAR.
H4. Engagement is greater when user avatars are in close proximity in HVAR.

4.2 Data Collection: Questionnaires, mEEG and User Activity Monitoring

We adopted two established questionnaires to measure presence and engagement in HVAR. ITC-Sense of Presence Inventory (ITC-SOPI) provides data on perceived presence with cross-media comparison [35]. This questionnaire includes four scales: spatial presence, engagement, ecological validity, and negative effects. For user engagement, we used the short form of the User Engagement Scale (UES), including constructs of focused attention, perceived usefulness, aesthetic appeal, and reward (novelty, felt involvement, and endurability) [50]. However, the two questionnaires are retrospective in that questions are asked after each session. Therefore, we used the Muse headband to record brain signal data to obtain a psychophysiological measure of engagement as comparisons. Pike and Ch'ng [55] supported the use of Muse headband together with VR head-mounted displays, for that users reported no discomfort and its use does not intrude nor influence participant activities. The brain signal data is time-stamped at $t=1$ second intervals and can be used to calculate the indices of real-time engagement, arousal, and valence [45]. We also implemented a function with C# in Unity to record user activity data at system runtime, including real-time position, direction of gaze, controller interactions, and relative distances of virtual avatars and objects. These are cross-referenced with the real-time engagement index measured by the mEEG. The function starts to record interaction data at session initiation, and produces sequences of interaction data per second. Data generated during each session was tracked and saved to a CSV file once the program is shut down.

We tested the measures in a pilot study. The Muse headband design makes it prone to movements when using AR, as there is the need for our participants to physically move around the room in the AR environment. On the contrary, the VR experience with the teleportation navigation required less physical movements. Using the headband inside the VR headset also ensured a better fit. Thus it has yielded better data quality in VR. In addition, we found that since users were already wearing the headset in VR, the addition of the mEEG was minimally invasive. However, wearing the mEEG was noticeable and an invasive measure for the AR condition. Regarding the recording of user activity data, our main purpose was to map it to the psychophysiological data for a comprehensive view of objective measures. As the mEEG measure was not ideal for our AR condition, the objective measures were used only for VR.

4.3 Independent Variables: Object Interactivity, User-Generated Contents and Avatar Proximity

4.3.1 Object Interactivity. VR and AR environments are often populated with virtual objects of different levels of interactivity. For some scenarios, high object interactivity is a basic functional requirement, such as in shooting games. However, in most scenarios, different levels of object interactivity can be designed in their digital forms to achieve those that may not be easily afforded in the real world, such as museum exhibits, room decorations, and construction sites. In our study, we present seven virtual objects with three levels of interactivity – high, medium, and low. These are based on the amount of controls implemented for each object. Generally, static objects are of low interactivity; objects that can be grabbed are of medium interactivity; and objects with additional context-specific interactions are of high interactivity. For example, the *Chinese Star Chart* is an archive document, and is virtually represented as a static poster on the wall. Another example is the *Figure of Assistant to the Judge of Hell*, the large size and the fact that it was on the ground floor of the British Museum, is reflected in it being immovable in the virtual environment. These two are categorised as low interactivity objects due to the limited controls they afford. The *Pottery Figure of a Standing Lady*, the *Tri-coloured Camel*, and the *Bronze Mask with Protruding Pupils* are of handheld size. Such affordances indicate that they can be grabbed, picked up, and placed on pedestals. We implemented additional interaction possibilities based on perceived affordances addressed by users in our previous studies [36, 37]. For example, users expected sound effects from the *Bronze Music Instrument*. We considered the user feedback in our design of the environment to allow some flexibility with interactions: the *Bronze Music Instrument* can be drummed with a stick to trigger sound effects (see Fig. 4a). As the history of the *Pottery Unicorn* suggests that it is a symbol of justice and law enforcement, we display a menacing face on top of the pedestal when the object is grabbed. Users can then use the horn of the object to 'attack' the face (see Fig. 4b). The intersection of the horn and the face will dismiss the display of the menacing face and trigger a 'victory' sound effect. These two objects are thus categorised as high interactivity.

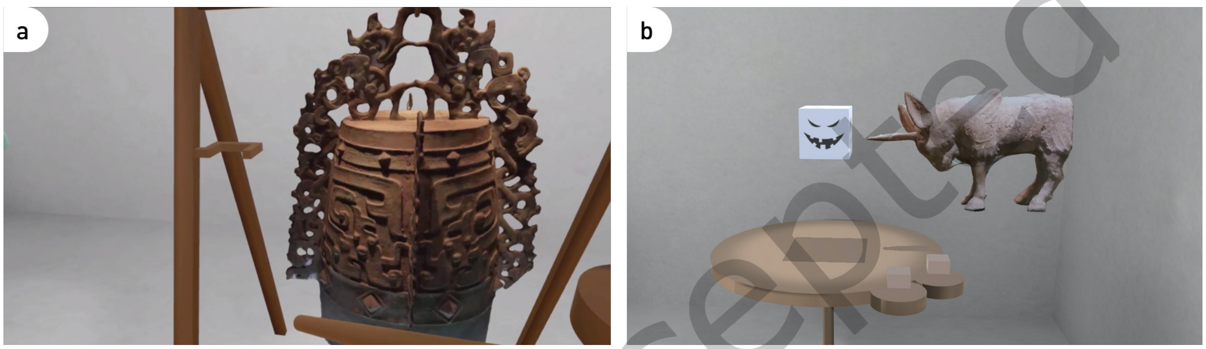


Fig. 4. Virtual objects of high interactivity: a) Triggering a sound effect for the *Bronze Music Instrument*; b) Using the horn of *Pottery Unicorn* to 'attack' a menacing face.

4.3.2 User-Generated Contents. User-generated contents support synchronous and asynchronous communications. These often take place within scenarios such as virtual classrooms, collaborative design spaces, and any social occasions when users need to leave a message. In our study, user-generated contents are presented to users through the UGC labels containing previous comments made on each virtual object. The comparison of HVAR-UGC and HVAR-Info allows us to investigate the effects of UGC on the subjective assessment of engagement. In addition, we capture real-time engagement index for gaze on object information labels and UGC labels. This objective data can provide a further understanding of the effects of UGC (subjective information) on engagement as compared to information only labels (objective information).

4.3.3 Avatar Proximity. Virtual avatars will play a fundamental role in future social activities in any virtual world, such as the envisioned Metaverse where simulations mimic aspects of the physical world and involve users in social activities and digital economies. Avatar proximity is an important factor of user experience in social scenarios. In our study, the real-time position of avatars was recorded in the user activity data at *t=1* second intervals. We calculated the Euclidean distances between virtual avatars using the x and z values of avatar position and analysed the real-time distances and engagement indices to study the effects of avatar proximity on user engagement.

4.4 Dependent Variables: Engagement and Presence

4.4.1 Engagement. We are able to obtain three types of engagement metrics via our dataset: 1) subjective assessment of engagement as indicated in the engagement scale of the ITC-SOPI and the UES score, 2) the real-time engagement index calculated from the brain signal data measured by the mEEG at *t=1* second intervals, and 3) the total engagement index of each virtual object calculated by summing all real-time engagement index on an object. In addition to the engagement index, brain signal data provides us the indices of arousal and valence, calculated using the formula below [see 22, 24, 45, 57].

engagement index = Beta / (Alpha + Theta)
arousal index = (Beta_AF7 + Beta_AF8) / (Alpha_AF7 + Alpha_AF8)
valence index = (Alpha_AF8 / Beta_AF8) - (Alpha_AF7 / Beta_AF7)

Specifically, the engagement index is calculated based on power bands that reflect vigilance and attention [57]. The advantage of using the established frequency band analyses of engagement is that it can be linked to, and compared with results reported in previous EEG research, such as [3].

Arousal indicates the perceived intensity of how calming or exciting the experience is. The *Beta* frequency indicates the state of alert, normal alert consciousness, and active thinking, whereas the *Alpha* frequency indicates physically and mentally relaxed states. These two waves were shown to be effective to demonstrate the index of arousal [22, 45].

Valence indicates whether the emotion is positive or negative. The mEEG data can be used to distinguish brain-state discrimination that depends on asymmetry. Positive emotions are associated with the right frontal inactivation (AF8) whereas the negative emotions are associated with the left frontal inactivation (AF7). Therefore, the emotional valence index is calculated using signals from these electrodes [22, 45].

4.4.2 Presence. We applied the full ITC-SOPI [35], an established cross-media measure of presence using four scales (see Table 2). The questionnaire also includes additional participant background information on their prior use of media. All questions were answered on a five-point Likert scale.

Table 2. ITC-SOPI scales and definitions.

Scale	Definition
Spatial presence	User's sense of being located within a spatially contiguous physical environment.
Engagement	User's involvement and interest in the content of the displayed environment, and their general enjoyment of the media experience.
Ecological validity	The believability and the realism of the content as well as the naturalness of the environment.
Negative effects	Adverse physiological effects of exposure to media, such as disorientation, tiredness, eyestrain and nausea.

4.5 Setup and Experimental Procedure

The user study took place at the NVIDIA Joint-Lab on Mixed Reality, an NVIDIA Technology Centre at the University of Nottingham's China campus. Each study involved 2 participants in a pair and lasted for an hour. The study consisted of three parts: 1) briefing, 2) main study, and 3) interview and debriefing. Detailed experimental procedure is shown in Fig. 5. Demographic background of participants was collected at the briefing session. The main study included two sessions (HVAR-Info and HVAR-UGC) with counterbalanced sequence. The first session included one participant in VR and the other in AR; in the second session the participants swapped the use of VR and AR. We engaged participants in conversations by encouraging them to 1) figure out the historical chronological order of the exhibits, and 2) initiate dialogues on objects that they find interesting. Participants were able to freely explore the exhibition room either independently or in parallel. During each session, the user in VR wore the Muse headband to record their brain signal data. After each session, both participants were asked to fill in the ITC-SOPI and the UES questionnaires. Once both sessions ended, the paired participants were invited to a semi-structured interview to discuss 1) their subjective feelings of the two sessions, 2) how they liked VR and AR in the hybrid environments, and 3) if they had intended to leave any comments on the exhibits. The study was approved by the University of Nottingham Ningbo China's Ethics Committee prior to data collection. Participants voluntarily signed up for the experiments.

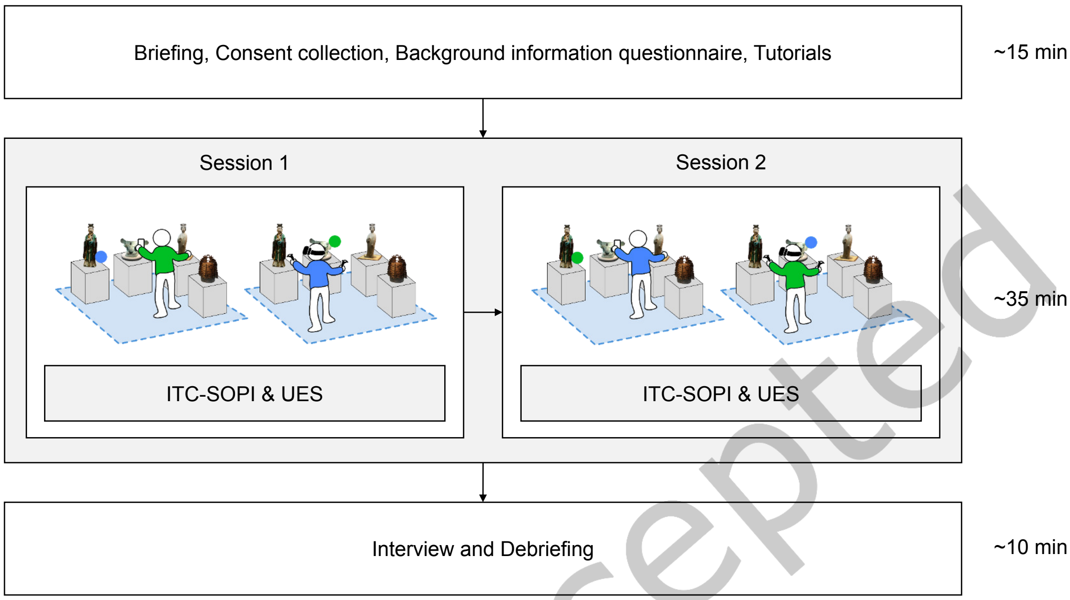


Fig. 5. Experimental procedure.

5 RESULTS

In total, we obtained 60 responses of the background information questionnaire, 120 responses on the ITC-SOPI questionnaire, 120 responses of the UES questionnaire and 60 CSV files of user activity data generated during each session. Two participants declined to be recorded wearing the Muse headband. We also encountered several issues such as interrupts of data recording and saving. These issues yielded a total of 47 mEEG data files from the 60 participants. We processed the mEEG data to obtain engagement, arousal, and valence indices using signals of the alpha, beta, and theta waves at *t=1* second intervals. We removed signals marked as bad data quality by Muse and applied a *z*-score measure to filter the outliers. In total, we collated 12,882 lines of real-time data (mEEG x user activity) from 47 participants: 25 VR-Info, and 22 VR-UGC. Details of the dataset are summarised in Table 3.

Table 3. Summary and sample size of the collected data.

	VR-Info	AR-Info	VR-UGC	AR-UGC	Total	Raw file
Background information					60	
ITC-SOPI	30	30	30	30	120	
UES	30	30	30	30	120	
mEEG data	25		22		47	17,755
User activity data	30		30		60	24,126

We explored the distribution of the ITC-SOPI and the UES questionnaire data and confirmed parametric test assumptions. Paired-samples t tests were conducted to analyze questionnaire data. As the sample size of mEEG data is sufficiently large (>10,000), the Central Limit Theorem ensures that the distribution of disturbance term will approximate normality. We conducted t tests and one-way ANOVA to analyze mEEG data. The interview data were analyzed using theme-based content analysis [49], structured by the questions asked.

5.1 Sample and Background Information

The majority of our participants are university students and staff aged between 18 and 51 inclusive ($M = 22.45, SD = 7.11$). Gender is not evenly distributed (13 males, 47 females), but indicates a representative ratio of the gender distribution of our university international community (3:7). Our further analysis showed that gender has no effect on measured variables. Most pairs signed up and appeared together knowing each other, with only two pairs of exceptions. Background information on participant use of media in the past is shown in Table 4. Most participants (93.33%) have viewed stereoscopic 3D images. 61.67% of the participants have used VR and 40% of the participants have used AR. Participants' self-evaluated knowledge of 3D image ($M = 1.83, SD = 0.67$), VR ($M = 1.72, SD = 0.75$), and AR ($M = 1.52, SD = 0.58$) was lower than their level of general computer experience ($M = 2.5, SD = 0.69$). Our analysis shows that these factors have no effect on engagement or presence.

Table 4. Participants' prior experience with 3D image, VR and AR.

	Item	Frequency	Percentage
3D image	Yes	56	93.33%
	No	4	6.67%
Virtual reality	Yes	37	61.67%
	No	23	38.33%
Augmented reality	Yes	24	40.00%
	No	36	60.00%

5.2 Psychometric Properties of the Scales

Cronbach's alpha (CA) was calculated to measure the internal consistency of the psychometric scales. The scales of both ITC-SOPI and UES are reliable as most of the CA values are greater than 0.70 (see Table 5), except for negative effects (0.65) and perceived usefulness (0.52). Reasons for the low CA can be the limited items presented on the scale. For example, there were only 3 items used for measuring perceived usefulness. These two scales will be interpreted with caution in the following data analysis.

5.3 Comparisons of Engagement and Presence in VR and AR

Fig. 6 illustrates user responses to the ITC-SOPI questionnaire. Results of paired-samples t tests indicated significantly higher levels of spatial presence, engagement, and ecological validity in VR as compared to AR (see Table 6). However, the negative effects were also higher in VR.

Fig. 7 illustrates user responses to the UES questionnaire. Results of paired-samples t tests indicated that VR users ($M = 4.32, SD = 0.43$) reported significantly greater engagement than AR users ($M = 3.66, SD = 0.78$), $t(59) = 6.10, p < .001$. The differences in all scales of the UES were significant (see Table 7).

Spearman correlation analysis found significant positive correlations between spatial presence, engagement, ecological validity, and the UES score (see Table 8). There were no significant correlations between negative effects

Table 5. Cronbach's alpha measures for the ITC-SOPI and the UES psychometric scales.

	Scale	CA
ITC-SOPI	Spatial presence (SP)	0.95
	Engagement (EG)	0.96
	Ecological validity (EV)	0.86
	Negative effects (NE)	0.65
UES	Focused attention (FA)	0.94
	Perceived usefulness (PU)	0.52
	Aesthetic appeal (AE)	0.88
	Reward (RW)	0.90

Table 6. Analysis results showing means (standard deviations) of the ITC-SOPI scales (rated from 1 to 5) for the VR and AR environments.

	VR	AR	Significance
Spatial presence	3.92 (0.47)	3.19 (0.64)	$t(59) = 8.90, p < .001$
Engagement	4.24 (0.49)	3.33 (0.79)	$t(59) = 9.20, p < .001$
Ecological validity	3.65 (0.77)	3.24 (0.79)	$t(59) = 4.21, p < .001$
Negative effects	2.33 (0.83)	1.98 (0.80)	$t(59) = 2.83, p < .01$

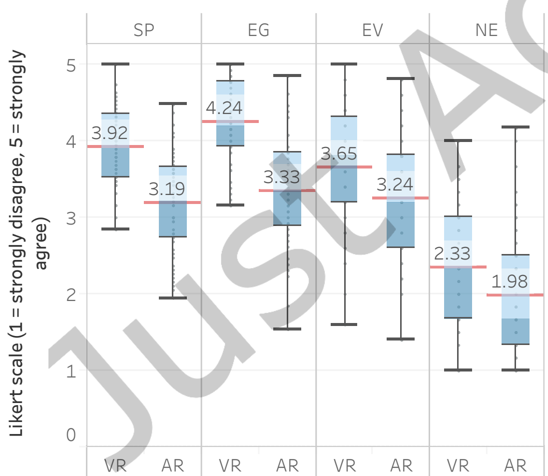


Fig. 6. Box plots and means of the ITC-SOPI.

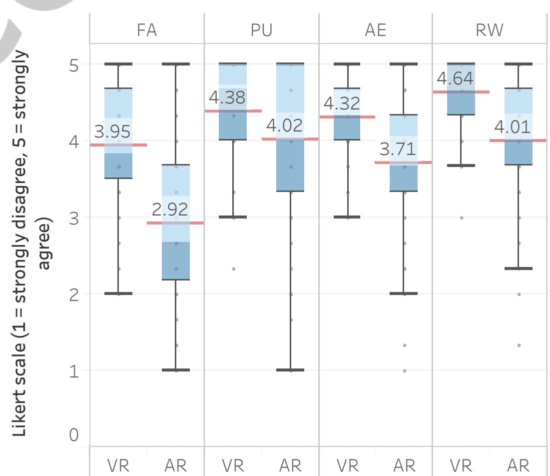


Fig. 7. Box plots and means of the UES.

and the other scales. Subjective assessments of engagement are greater when spatial presence and ecological validity are higher in HVAR. Hence, **H1** is supported.

Table 7. Analysis results showing means (standard deviations) of the UES scales (rated from 1 to 5) for the VR and AR environments.

	VR	AR	Significance
Focused attention	3.95 (0.72)	2.92 (1.01)	$t(59) = 6.97, p < .001$
Perceived usefulness	4.38 (0.67)	4.02 (0.96)	$t(59) = 2.49, p < .05$
Aesthetic appeal	4.32 (0.46)	3.71 (0.93)	$t(59) = 5.43, p < .001$
Reward	4.64 (0.46)	4.01 (0.83)	$t(59) = 5.47, p < .001$

Table 8. Spearman correlations of the ITC-SOPI scales and the UES score.

	SP	EG	EV	NE	UES
Spatial presence	1				
Engagement	.80**	1			
Ecological validity	.76**	.66**	1		
Negative effects	.12	.05	.09	1	
UES	.73**	.77**	.58**	-.12	1

**. Correlation is significant at the 0.01 level (2-tailed).

5.4 Effects of Object Interactivity on Engagement

Here, we determine the levels of object interactivity based on the amount of control implemented for each object. Details of the categories of interactivity have been listed in Table 1 and explained in Section 4.3.1.

5.4.1 Total Engagement Index. For engagement index of each second, it is counted to an object if at that second the object was the nearest to the user among all objects. The total engagement index of an object is calculated by summing the real-time engagement index for the object (see Table 9). Results of the Spearman correlation tests found significant positive correlations between object interactivity and time spent on the nearest object, $rs = .95, p < .01$, and between object interactivity and total engagement index, $rs = .76, p < .05$. Therefore, **H2** is supported in terms of the total engagement index of objects: total engagement index is greater for objects of higher interactivity in HVAR.

Table 9. Time spent and total engagement index for each object in VR.

	Nearest object	Time (s)	Total engagement index
High	Xie Zhi (Pottery Unicorn)	2554	1488.97
	Bronze Music Instrument	2189	1253.04
Medium	Bronze Mask with Protruding Pupils	1852	1166.78
	Pottery Figure of a Standing Lady	1752	1083.39
	Tri-coloured Camel	1677	1053.08
Low	Figure of an Assistant to the Judge of Hell	1605	1109.23
	Chinese Star Chart	1253	839.91

5.4.2 Real-time Engagement Index. We mapped mEEG data with activity data at system runtime and calculated the real-time indices of engagement, arousal, and valence for objects of high, medium, and low interactivity (see Table 10). Results of the Spearman correlation tests found negative correlations between object interactivity and real-time engagement index, $rs = -0.13, p < .001$. This was not what we expected. Therefore, **H2** is not supported in terms of real-time engagement index: real-time engagement index did not show positive correlations with object interactivity in HVAR.

Table 10. Analysis results showing means (standard deviations) of engagement, arousal, and valence index for high, medium and low object interactivity.

	High	Medium	Low	Significance
Engagement index	0.58 (0.27)	0.63 (0.27)	0.68 (0.32)	$F(2, 12879) = 122.71, p < .001$
Arousal index	1.05 (0.75)	1.16 (0.77)	1.28 (0.85)	$F(2, 12879) = 72.34, p < .001$
Valence index	0.09 (1.08)	0.13 (1.05)	0.11 (1.02)	$F(2, 12879) = 1.88, p = .153$

One-way ANOVA revealed significant differences in real-time engagement and arousal indices among objects of low, medium, and high interactivity (see Table 10). Post hoc tests indicated that the differences in real-time engagement and arousal indices among the three interactivity groups were statistically significant ($p < .001$). Engagement and arousal indices were the highest around objects of low interactivity, followed by objects of medium and high interactivity.

5.5 Effects of User-Generated Contents on Engagement

We investigated the effects of user-generated contents on engagement by comparing subjective assessments of engagement and real-time engagement index between HVAR-Info and HVAR-UGC, and by comparing the real-time engagement index between reading information and UGC labels.

5.5.1 Comparison of Engagement between HVAR-Info and HVAR-UGC. There were significant differences in real-time engagement and arousal indices between VR-Info and VR-UGC (see Table 11). Users found VR-UGC to be more engaging and arousing than VR-Info. Thus, **H3** is supported by mEEG data: engagement is greater when UGC of virtual objects is presented in HVAR. However, the analysis of ITC-SOPI and UES showed no significant difference in engagement between HVAR-Info and HVAR-UGC. The continuous signals provided some markers of engagement but not at a level that was perceptible and recognizable by the participant while filling out the questionnaire post hoc.

Table 11. Analysis results showing means (standard deviations) of engagement, arousal and valence index for VR-Info and VR-UGC.

	VR-Info	VR-UGC	Significance
Engagement index	0.53 (0.18)	0.67 (0.21)	$t(45) = 2.43, p < .05$
Arousal index	0.92 (0.44)	1.22 (0.65)	$t(45) = 1.85, p < .05$
Valence index	0.19 (0.42)	0.05 (0.14)	$t(45) = 1.50, p = .07$

5.5.2 Comparison of Engagement between Reading Labels of Information and UGC. We also analyzed the 3,559 lines of data of label-reading activities. Independent-samples t tests found significant differences in real-time engagement and arousal indices between reading information and UGC labels (see Table 12). Users found reading UGC labels more engaging and arousing than reading object information labels. This further supports **H3**.

Table 12. Analysis results showing means (standard deviations) of engagement, arousal and valence index for information and user-generated content labels.

	Information labels	UGC labels	Significance
Engagement index	0.68 (0.30)	0.77 (0.33)	$t(3557) = 6.47, p < .001$
Arousal index	1.32 (0.77)	1.49 (0.85)	$t(3557) = 4.47, p < .001$
Valence index	0.12 (0.84)	0.04 (0.68)	$t(3557) = 2.11, p = .035$

5.6 Effects of Avatar Proximity on Engagement

User activity data records the real-time position of users within the virtual environment at *t=1* second intervals. Fig. 8 presents a heatmap of user positions within the virtual environment. Positions were found to be mostly at the front of each object and around the four teleport points pre-defined in the system.

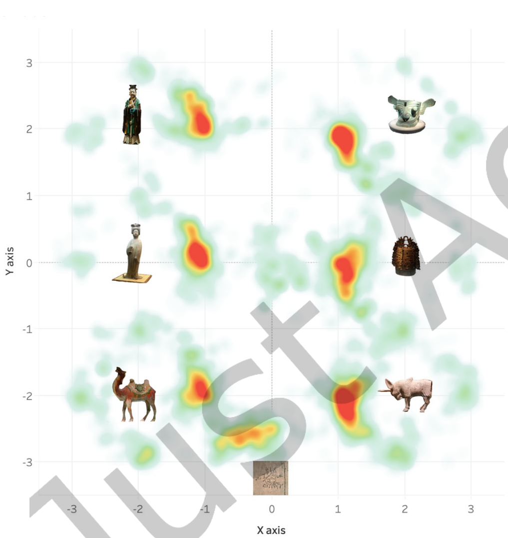


Fig. 8. Heatmap of user positions within the virtual environment.

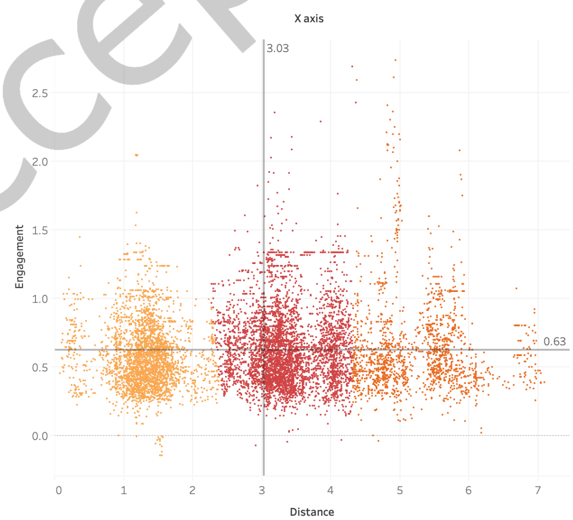


Fig. 9. Cluster analysis on relationship between distance and engagement.

5.6.1 Comparison of Engagement between Avatar Shown and Not Shown. We first investigated the effect of the avatar visibility. Independent-samples t tests found a significant difference in real-time engagement index between avatar shown and not shown, $t(12880) = 5.50, p < .001$. Engagement index with avatar shown ($M = 0.63, SD = 0.28$) was significantly higher than avatar not shown ($M = 0.60, SD = 0.29$). There was no significant difference shown for the arousal or valence index.

5.6.2 Cluster Analysis of Engagement and Avatar Distance. When avatars of both VR and AR users were shown, users kept an average social distance of 3.03 meters. Fig. 9 shows the results of cluster analysis on the relationship between engagement index and avatar distance. A one-way ANOVA revealed a significant difference in engagement index among the three clusters, $F(2, 9590) = 55.27, p < .001$. Post hoc tests showed a significant difference ($p < .001$) in engagement index between cluster 1 and 2 and between cluster 1 and 3. The difference between cluster 2 and 3 was insignificant. Engagement index was lower when the avatar distance was closer (see Table 13). This is not expected and therefore **H4** is not supported: engagement in close proximity in HVAR was not greater.

Table 13. Cluster analysis results showing means (standard deviations) of avatar distance and engagement index.

	Number of items	**Distance**	**Engagement index**
Cluster 1 (Left)	3153	1.32 (0.40)	0.59 (0.24)
Cluster 2 (Centre)	4731	3.36 (0.47)	0.65 (0.29)
Cluster 3 (Right)	1709	5.27 (0.63)	0.64 (0.34)

5.7 Additional Analysis

5.7.1 Engagement and Different User Activities. We categorized six types of user activities in VR: 1) no interaction, 2) gaze on objects, 3) gaze on information labels, 4) gaze on UGC labels, 5) gaze on virtual avatar, and 6) controller interaction with objects. One-way ANOVA revealed significant differences in engagement and arousal indices among the six types of user activities (see Table 14). Post hoc tests indicated that engagement and arousal indices for gaze on information labels and gaze on UGC labels were significantly higher ($p < .001$) as compared to the other four types of user activities.

Table 14. Analysis results showing means (standard deviations) of engagement, arousal and valence index for six types of user activities.

	Engagement index	**Arousal index**	**Valence index**
No interaction	0.59 (0.27)	1.06 (0.78)	0.10 (1.14)
Gaze on objects	0.59 (0.26)	1.05 (0.75)	0.10 (1.12)
Gaze on information labels	0.68 (0.30)	1.32 (0.77)	0.12 (0.84)
Gaze on UGC labels	0.77 (0.33)	1.49 (0.85)	0.04 (0.68)
Gaze on virtual avatar	0.60 (0.25)	1.18 (0.76)	0.15 (1.24)
Controller interaction with objects	0.61 (0.27)	1.11 (0.8)	0.17 (1.12)
Significance	$F(5, 12876) = 76.56, p < .001$	$F(5, 12876) = 73.85, p < .001$	$F(5, 12876) = 1.80, p = .11$

5.7.2 Engagement, Time of Gaze on Labels, and Object Interactivity. Considering the significant effects of reading labels on user engagement, we calculated the time ratio of gaze on labels for the three object interactivity groups (see Table 15). The results showed that when users were around objects of low interactivity, 74.36% of their time of gaze was on the information and UGC labels, which is much higher than 39.93% and 49.01% for objects of medium and high interactivity respectively.

Table 15. Analysis results showing means (standard deviations) of engagement index and time ratio of gaze on labels for high, medium and low object interactivity.

	Engagement index	**Time ratio of gaze on labels**
High	0.61 (0.27)	49.01%
Medium	0.64 (0.27)	39.93%
Low	0.72 (0.34)	74.36%

5.7.3 Engagement, Time of Gaze on Labels, and Avatar Proximity. The comparison of the time ratio of gaze on labels and avatar proximity showed that users spent more time reading information and UGC labels when they were alone with the objects (50.55%). The time ratio of gaze on labels was lower when they were close to another user, in which case their partners' virtual avatars were shown (see Table 16).

Table 16. Analysis results showing means (standard deviations) of engagement index and time ratio of gaze on labels for close and far avatar proximity.

	Engagement index	**Time ratio of gaze on labels**
Close (<2.34, Cluster 1)	0.59 (0.24)	43.80%
Far (>2.34, Cluster 2 and 3)	0.65 (0.30)	50.55%

6 DISCUSSION

The present study examines factors that have influence on engagement in hybrid virtual and augmented reality environments for object interactivity, user-generated contents, and avatar proximity. We adopted questionnaires in combination with psychophysiological measure with mEEG and user activity monitoring to obtain objective engagement measures. In this section, we summarize our findings by discussing each hypothesis, provide a discussion on the measure of engagement and the interpretation of results, and the limitations and future work.

6.1 Summary of Study Findings

Subjective assessment of engagement is greater when higher presence is perceived in HVAR (**H1**). Overall, the perceived presence is significantly greater for users in VR as compared to AR. Users in VR reported higher levels of spatial presence, engagement and ecological validity as compared to users in AR, although the negative effects, such as disorientation, tiredness, eyestrain and nauseous are also higher. The scales of the UES are also significantly higher in VR as compared to AR. Engagement correlated positively with presence in HVAR.

The total engagement index for each object is greater for those of higher interactivity in HVAR (**H2**). We found that object interactivity and total engagement index are positively correlated, i.e., users tended to spend more time around objects of higher interactivity and engaged more with them. This finding is in line with previous literature that greater interactivity will lead to greater presence and engagement [69]. However, our analysis of real-time engagement have shown contradictory results. Further analysis on user activities indicates that the engagement index is the highest for reading UGC labels, followed by reading object information labels, controller interaction with objects, gaze on virtual avatars, gaze on objects, and no interaction. Users tended to spend a smaller amount of time on reading labels on objects of high interactivity because of more interaction possibilities. The greater real-time engagement index for objects of low interactivity is thus as a result of the limited interaction possibilities and the greater time ratio of gaze spend on reading labels.

Real-time engagement is greater when UGC of virtual objects is presented in HVAR (**H3**). This hypothesis is supported by mEEG data but not the questionnaire data. By comparing HVAR-Info and HVAR-UGC, we see no significant difference in user engagement in either ITC-SOPI or UES questionnaire data. However, the real-time engagement index showed that users do engage more in HVAR-UGC as compared to HVAR-Info. In addition, users are more engaged when reading UGC labels as compared to reading object information labels.

Engagement was not greater when avatars were in close proximity in HVAR and therefore, **H4** is not supported. Although there was greater engagement when avatars were shown in the scene, the correlation between avatar proximity and user engagement was not in line with our expectation. A cluster analysis of avatar distance showed three clusters, and most distances were within the social distance (1-4 meters). Users reported the greatest engagement around 3.36 meters, which is within the social distance defined by Hall [23]. However, user engagement was unexpectedly lower when users moved closer to around 1.32 meters. Further analysis showed that users spent a greater amount of time on reading labels when avatar proximity was greater than 2.34 meters. This can partly account for their greater engagement, as we found that reading labels yielded greater user engagement than other types of user activities.

6.2 Discussion of Results and Findings

6.2.1 Engagement and Presence in HVAR. Presence is a significant indicator of VR and AR experience. Our study demonstrated that both presence and engagement are greater in VR as compared to AR and the results confirm our expected positive correlation between engagement and presence in HVAR. This indicates that the comparison of VR and AR in a shared session is different from comparing them two separate sessions. Here, we discuss three reasons as identified in our observations and interviews.

First, the head-mounted display device for VR supported a greater field of view of situated environments as compared to smartphone-based AR. In our study, users evaluated presence in AR based on the augmented environment where the real world is still in view. We attempted to minimise the effect of the real environment and to support the sense of spatial presence by providing AR users with an augmented exhibition room layout that is almost identical to the VR environment, but without the room (i.e., walls, floor and ceiling). However, the blend of the virtual and the real environment could have resulted in a gap in realism as indicated in prior studies [7]. How do we blend the virtual and the real to improve a greater sense of spatial presence is still a challenge that needs to be explored in future research.

Second, interactions in VR involve both hands and body movements. These natural interactions could have contributed to the greater ecological validity and user engagement in VR [4]. As smartphones are pervasive, we observed that users are more familiar with AR interactions. However, smartphones afforded limited interaction possibilities in AR. The lower ecological validity in AR indicated the need to develop interaction approaches that intersect well between the real and the virtual. More natural control mechanisms are expected for interactions in AR.

Third, the perceptions and expectations of AR users are affected by the VR users in HVAR. Users commented in the interview that although the AR application itself is intriguing, seeing the paired VR user performing controller interactions in VR made them feel limited and in want of more natural interaction. These factors could have caused disengagement during the co-located session.

6.2.2 Object Interactivity and Engagement. Object interactivity in our study was implemented based on the amount of controls over an object, i.e., whether it is static, grabbable, or with additional context-specific interactions. The results implied that users spent a greater length of time and showed greater total engagement index around objects of higher interactivity. We observed that these objects did seem to leave a deeper impression on users. The two objects of high interactivity, *Pottery Unicorn* and *Bronze Music Instrument*, are the most mentioned objects in the interviews. Users reported that objects of high interactivity are more likely to encourage them to have closer views and to obtain detailed information. We also observed that high interactivity objects motivated users to read labels and explore the affordances of these objects. Such an exploratory process is helpful in facilitating user engagement in experiential learning [13].

Although the real-time engagement index is unexpectedly greater for objects of low interactivity, our further analysis found that user engagement was greater for gaze on UGC and object information labels as compared to controller interactions. Users spent a greater amount of time reading labels around objects of low interactivity because they afforded fewer interaction possibilities. This demonstrated that accessing information about virtual objects forms a significant part of user interactions with virtual objects and imposes a significant effect on user engagement. This is partly due to cultural interests that motivated users to read information in order to learn the virtual objects, and also the fact that objects in HVAR are the point of connection between users [36]. In this case, virtual objects mediate user communication and social interactions. This will motivate users to collect and share information so as to build a shared understanding of the hybrid environment.

We summarise that high object interactivity has positive effects on the focused attention and thereby total engagement for these objects. However, the UGC and object information can facilitate user engagement with other users and mediate their communication in a social context. These also add to the interactivity of an object. In addition to the findings on object interactivity and engagement, the lessons learned from label interactions could be generalised to the future design of interactive objects in scenarios other than museums. For example, a virtual book could be designed for a virtual classroom to sustain student engagement in reading activities. Similarly, a manual or a map could bring the focused attention of multiple users to a shared object, engaging them in collaborative design and navigation tasks, such as in architectural design and construction.

6.2.3 User-Generated Contents and Engagement. The effects of user-generated contents on engagement in this study are supported by mEEG data, based on the comparisons between two experimental settings (HVAR-Info and HVAR-UGC), and between reading information labels and UGC labels. In contrast to blocks of texts on the information labels, user commented that the texts on the UGC labels are shorter, more directed, and memorable. Users found that contents generated by others are helpful in filtering information that leads to the understanding of the objects. Quotes from the UGC labels have been mentioned frequently in the interviews. For example, many participants mentioned the comment *'Is it a pig or goat?'* for the *Pottery Unicorn*, because they found this comment amusing and matches with their first impression of the object's appearance. Several participants said that they would not have noticed the *Figure of an Assistant to the Judge of Hell* is exhibited in the British Museum if it had not been identified by others in the UGC label. The UGC labels seemed to be an interface of asynchronous communication which establish a connection between users.

However, some participants have raised concerns about the potential negative effects for children to be misled by unverified information and negative comments, much like those negative feedback on social media [31]. In this context, they deemed museums as places for obtaining knowledge, and therefore, they believe all presented information should be written or moderated by curators or subject domain experts prior to it being presented to users. This finding can be a guideline for designing cultural applications and systems to be cautious about the use of user-generated contents.

During the interview, we also asked participants if they want to leave some comments for the virtual objects. Interestingly, almost all participants said that they are averse to posting meaningless or negative contents. This may suggest that such spaces can possess a solemnity due to the objects presented. Users said they would only share something interesting or leave supplementary information for others. Contents generated by participants included indicators of the objects' high interactivity. Some participants also replied to existing comments on the labels, e.g., a comment the *Pottery Unicorn* says that *'It is not a pig, it is a unicorn! [with "laughing out loud" emoji]'*. The use of emojis was mentioned several times in the interview. Two participants reported that instead of posting texts, they prefer to put stickers and emojis on top of virtual objects. The preference of using visual languages such as emojis and stickers to express sentiments is in accordance with the current trend and culture of social media [15]. Overall, participants are in favour of UGC that are funny, imaginative, with punchlines or helpful information. These are also the types of contents that users are likely to post. Such sharing of subjective interpretations expressed through user-generated contents are also essential in the process of learning and meaning-making [18]. Aside from the museum setting, the findings on UGC can inform future actionable practice of virtual and augmented environments to support social interactions, sharing of thoughts, and also the creation of personalised experiences. For example, users can create a space for their personal collections and notes to host their memories, which can be a private space or a public shared area to socialise with others.

6.2.4 Avatar Proximity and Engagement. Our study demonstrated the positive effect of virtual avatars on user engagement. Users mentioned that the virtual avatars has made the virtual space more real and social as they expected to have visitors in the exhibition type of setting. We have found that close avatar proximity did not support greater user engagement. In fact, the engagement index for the avatar distance at around 1.32 meters was less than the avatar distance at around 3.36 meters. Here, we delve into the reflection of users and identified three reasons for the phenomenon.

First, users had expectations on avatar interactions. Users reported that they expected the avatar to be able to move around and converse with them. However, disengagement occurs when they realised that the avatar was not as interactive as they expected. This is possibly the primary reason for the lower engagement when avatars are in close proximity.

Second, users perceived a significant contrast between the virtual avatar and virtual objects in terms of realism and interactivity. Participants reported that the virtual objects looked realistic and they were allowed interaction possibilities that would not be possible in a real exhibition. In comparison, there were mismatches between the virtual avatar and virtual objects in terms of visual details and interactivity at close distances. Such contrast is likely to result from user mental models of real physical exhibitions where real persons exist and where objects are not allowed to be touched. Therefore, the experience of limited social interaction with avatars could be a reason for the disengagement.

Third, although most users knew that the avatar represented the other participant in AR, the visual appearance of the avatar did influence perception as far as the identity of the user is concerned. Misinterpretations of the virtual avatars include treating the virtual avatar as a non-player character (NPC) or a random visitor. One participant has also perceived the avatar as herself because of the similarity. These misinterpretations could have affected user engagement. This is the challenge of using anthropomorphic avatars because the visual appearance of the avatars can influence perception and as a consequence, engagement [54].

Nevertheless, avatars are essential as the level of engagement is higher when avatar is presented. The average distance of 3.03 meters is to be expected as the acceptable social distance for users embodied in a full-body avatars in VR. This reckons with Hall's proxemic zones in the real world. To the best of our knowledge, there has not been any work that measures avatar distances in VR so future environments could be designed taking this as a benchmark. The highest engagement occurrence around a distance of 3.36 meters indicates that engagement can be better supported when avatars are present within the social distance, within which they can be reminded that there are people in the same space. In the meantime, it is less likely for the limitation in the visual appearance of avatars and interactivity to be an issue, because this situation is similar to being in a public space with strangers whom you are not interacting with. This is a contribution to the future design of multi-user hybrid environments.

6.3 Engagement Measures and the Interpretation of Results

6.3.1 Combined Psychophysiological Measure and User Activity Monitoring. Our study showed that the Muse mEEG works well with the VR HMD: it is minimally invasive and provides an effective objective measure of user engagement based on the frequency signals. It will be difficult to study object interactivity and avatar proximity if engagement is assessed using only retrospective questionnaires. Our hypothesis on UGC is supported by mEEG data, but not questionnaires. We believe the results are related to the quantity of data that we are able to obtain. Questionnaires provide subjective assessment for an overall engagement after each session, whereas mEEG data and user activity data can record data at *t=1* second intervals during each session for real-time engagement and interactions. The combined psychophysiological measure and user activity monitoring approach has significantly increased the quantity and precision of our data in VR. It provided data at a sufficiently large scale and of ecological validity, which have allowed us to obtain an objective understanding of real-time engagement associated with different user activities. With the combined approach, we were able to test our hypotheses on object interactivity and avatar proximity, and to obtain an objective understanding of the effects of UGC on user engagement. Evaluation studies of VR and AR are often hard to manage as they involve complex developments, experimental settings and procedures. Our study has demonstrated that the combined psychophysiological measure and user activity monitoring is an effective means of understanding user engagement in VR with a valid and significant dataset. This approach is helpful for studying user engagement in VR and AR as it can provide objective data recorded at high velocity and that it requires no additional effort from participants.

6.3.2 Objective and Subjective Measures. Our study adopted combined objective (i.e. mEEG and user activity monitoring) and subjective measures (i.e. questionnaires) of engagement in VR, but the AR condition was evaluated primarily based on subjective measures. When interpreting results, it is not to favor one against the other, and it does not make the findings less valid if the mEEG measures are not available. Instead, having additional objective data provides a new perspective for evaluating user experience, and brings the possibility of revealing different findings. It should be interpreted more carefully when conflicting results emerge. For example, our H3 was supported by mEEG data, but not questionnaire data. The questionnaire results indicate that adding UGC labels may not significantly increase the overall user engagement, given that the trivial change to the virtual environment and its effect on the overall experience might not be perceptible and recognizable by participants while filling out the questionnaire post hoc. In the meantime, the continuous mEEG signals provided some markers of engagement and showed that users were indeed engaged when reading UGC labels. Using subjective measures alone may not provide meaningful design guidelines for user engagement, but the mEEG data revealed a ranking of activities in which users were engaged with. This is the main value that lies in the continuous objective data measures.

6.4 Limitations and Future Work

Our study revealed several limitations and challenges for future work in the community. First, the psychophysiological measure and user activity monitoring were only applied in VR, but not AR. Our rationale was to avoid the introduction of additional invasive measures to the AR condition. The need for physical movements in the AR environment also compromised the mEEG data quality. Data loss and poor data quality could be caused by electrode disconnection, as the headband is more prone to movements as compared to EEG headsets. This is one limitation of the mEEG. A practical alternative is the functional near-infrared spectroscopy (fNIRS), which could produce robust data collection results that are less prone to muscular derived noises. Second, the Muse headband has only five dry sensors. While the study of engagement is not affected by the limited number of sensors, more sophisticated study and analysis could not be achieved without the impedance checkers found on full-cap medical-grade EEG systems. Third, some event-related potentials (ERPs) during interactions cannot be assessed with the mEEG due to the missing electrodes around the central area, such as the FCz electrode ERPs associated with haptic feedback [20]. Finally, future research should carefully consider the ethics concerns related to the use of EEG. Participants should be fully informed about the devices, benefits and risks, the data collected, and be allowed to opt out at any time. Companies that choose to attach EEG sensors to a VR headset should also give users the option to switch this feature off.

7 CONCLUSION

We conducted an experiment with 60 participants to understand factors influencing user engagement in Hybrid Virtual and Augmented Reality (HVAR) environments, with the specific focus on object interactivity, user-generated contents (UGC), and avatar proximity. Our analysis confirmed the positive correlations between presence and engagement in HVAR, and that both presence and engagement are greater in VR as compared to AR. Our results supported hypotheses on the positive effects of object interactivity and user-generated contents on user engagement. Furthermore, we demonstrated that UGC engages users more than the labels of information that are attached to objects. We also found that the effects of avatar proximity on user engagement is not according to our expectations. Users tended to engage most when avatars are within the wider range of social distance, i.e., 2-4 meters, but engagement decreases in a closer proximity, i.e. within 1-2 meters. Our study has validated a metric for engagement using combined psychophysiological measure and user activity monitoring at system runtime. The approach can inform the future study of VR and AR systems to obtain an objective understanding of user engagement with an increased scale of data. There is value in future research that extends the objective measure of engagement to multiple users and that can cross-examine real-time engagement index within hybrid VR and AR environments.

ACKNOWLEDGEMENT

We would like to thank our participants for their time and valuable comments during the study at the NVIDIA Joint-Lab on Mixed Reality, the University of Nottingham Ningbo China. This work is partially supported by the National Natural Science Foundation of China (62207022), Natural Science Foundation of the Jiangsu Higher Education Institutions of China (22KJB520038), and Xi'an Jiaotong-Liverpool University (RDF-20-02-47).

REFERENCES

[1] Rebecca L. Acabchuk, Mareyna A. Simon, Spencer Low, Julie M. Brisson, and Blair T. Johnson. 2021. Measuring Meditation Progress with a Consumer-Grade EEG Device: Caution from a Randomized Controlled Trial. *Mindfulness* 12, 1 (2021), 68–81. https://doi.org/10.1007/s12671-020-01497-1

[2] Mafkereseb Kassahun Bekele, Cape Town, Roberto Pierdicca, Emanuele Frontoni, and E V A Savina Malinverni. 2018. A Survey of Augmented, Virtual, and Mixed Reality for Cultural Heritage. *ACM Journal on Computing and Cultural Heritage* 11, 2 (2018), 36. https://doi.org/10.1145/3145534

[3] Chris Berka, Daniel J. Levendowski, Michelle N. Lumicao, Alan Yau, Gene Davis, Vladimir T. Zivkovic, Richard E. Olmstead, Patrice D. Tremoulet, and Patrick L. Craven. 2007. EEG correlates of task engagement and mental workload in vigilance, learning, and memory tasks. *Aviation Space and Environmental Medicine* 78, 5 II (2007).

[4] Nadia Bianchi-Berthouze. 2013. Understanding the Role of Body Movement in Player Engagement. *Human-Computer Interaction* 28, 1 (2013), 40–75. https://doi.org/10.1080/07370024.2012.688468

[5] Benjamin Blankertz, Michael Tangermann, Carmen Vidaurre, Siamac Fazli, Claudia Sannelli, Stefan Haufe, Cecilia Maeder, Lenny Ramsey, Irene Sturm, Gabriel Curio, and Klaus Robert Müller. 2010. The Berlin brain-computer interface: Non-medical uses of BCI technology. https://doi.org/10.3389/fnins.2010.00198

[6] Doug A. Bowman and Ryan P. McMahan. 2007. Virtual Reality: How Much Immersion Is Enough? *Computer* 40, 7 (jul 2007), 36–43. https://doi.org/10.1109/MC.2007.257

[7] Jennifer Brade, Mario Lorenz, Marc Busch, Niels Hammer, Manfred Tscheligi, and Philipp Klimant. 2017. Being There Again – Presence in Real and Virtual Environments and its Relation to Usability and User Experience Using a Mobile Navigation Task. *International Journal of Human-Computer Studies* 101, March 2016 (2017), 76–87. https://doi.org/10.1016/j.ijhcs.2017.01.004

[8] Jeanne H. Brockmyer, Christine M. Fox, Kathleen A. Curtiss, Evan McBroom, Kimberly M. Burkhart, and Jacquelyn N. Pidruzny. 2009. The development of the Game Engagement Questionnaire: A measure of engagement in video game-playing. *Journal of Experimental Social Psychology* 45, 4 (jul 2009), 624–634. https://doi.org/10.1016/j.jesp.2009.02.016

[9] Emily Brown and Paul Cairns. 2004. A grounded investigation of game immersion. In *Conference on Human Factors in Computing Systems - Proceedings*. https://doi.org/10.1145/985921.986048

[10] Marc Busch, Mario Lorenz, Manfred Tscheligi, Christina Hochleitner, and Trenton Schulz. 2014. Being there for real: presence in real and virtual environments and its relation to usability. *Proceedings of the 8th ...* (2014), 117–126. https://doi.org/10.1145/2639189.2639224

[11] Cédric Cannard, Helané Wahbeh, and Arnaud Delorme. 2021. Electroencephalography Correlates of Well-Being Using a Low-Cost Wearable System. *Frontiers in Human Neuroscience* 15, December (2021). https://doi.org/10.3389/fnhum.2021.745135

[12] Eugene Ch'ng and Neil Cooke. 2015. User Study on 3D Multitouch Interaction (3DMi) and Gaze on Surface Computing. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 9187 (2015), 425–433. https://doi.org/10.1007/978-3-319-20898-5_41

[13] Eugene Ch'ng, Yue Li, Shengdan Cai, and Fui Theng Leow. 2020. The effects of VR environments on the acceptance, experience, and expectations of cultural heritage learning. *Journal on Computing and Cultural Heritage* 13, 1 (2020), 1–20. https://doi.org/10.1145/3352933

[14] James J. Cummings and Jeremy N. Bailenson. 2016. How Immersive Is Enough? A Meta-Analysis of the Effect of Immersive Technology on User Presence. *Media Psychology* 19, 2 (2016), 272–309. https://doi.org/10.1080/15213269.2015.1015740

[15] Marcel Danesi. 2016. *The semiotics of emoji: The rise of visual language in the age of the internet.* Bloomsbury Publishing.

[16] Kevin Doherty and Gavin Doherty. 2018. Engagement in HCI: Conception, Theory and Measurement. *ACM Computing Surveys (CSUR)* 51, 5 (2018), 99:1—-99:39. https://doi.org/10.1145/3234149

[17] Steven Dow. 2007. User engagement in physically embodied narrative experiences. *Creativity and Cognition 2007, CC2007 - Seeding Creativity: Tools, Media, and Environments* (2007), 280. https://doi.org/10.1145/1254960.1255016

[18] John Howard Falk and Lynn Diane Dierking. 2000. *Learning from museums: Visitor experiences and the making of meaning.* arXiv:arXiv:1011.1669v3

[19] Frederick G. Freeman, Peter J. Mikulka, Lawrence J. Prinzel, and Mark W. Scerbo. 1999. Evaluation of an adaptive automation system using three EEG indices with a visual tracking task. *Biological Psychology* 50, 1 (1999), 61–76. https://doi.org/10.1016/S0301-0511(99)00002-2

[20] Lukas Gehrke, Sezen Akman, Pedro Lopes, Albert Chen, Avinash Kumar Singh, Hsiang-ting Chen, Chin-Teng Lin, and Klaus Gramann. 2019. Detecting Visuo-Haptic Mismatches in Virtual Reality using the Prediction Error Negativity of Event-Related Brain Potentials. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems.* ACM, New York, NY, USA, 1–11. https://doi.org/10.1145/3290605.3300657

[21] Philippe Gelisse, Pierre Genton, and Arielle Crespel. 2020. Failure to recognize muscular artifacts on the EEG may cause a wrong diagnosis of myoclonic status epilepticus. *Epilepsy and Behavior Reports* 14 (2020), 0–2. https://doi.org/10.1016/j.ebr.2020.100362

[22] Sergio Giraldo and Rafael Ramirez. 2013. Brain-Activity-Driven Real-Time Music Emotive Control. *3rd International Conference on Music & Emotion* June (2013), 11–15. https://doi.org/10.3217/978-3-85125-260-6-68

[23] Edward T. Hall. 1973. The Hidden Dimension. *Leonardo* (1973). https://doi.org/10.2307/1572461

[24] Mariam Hassib, Stefan Schneegass, Philipp Eiglsperger, Niels Henze, Albrecht Schmidt, and Florian Alt. 2017. EngageMeter: A system for implicit audience engagement sensing using electroencephalography. *Conference on Human Factors in Computing Systems - Proceedings* 2017-May (2017), 5114–5119. https://doi.org/10.1145/3025453.3025669

[25] Carrie Heeter. 1992. Being There: The Subjective Experience of Presence. *Presence: Teleoperators and Virtual Environments* 1, 2 (jan 1992), 262–271. https://doi.org/10.1162/pres.1992.1.2.262

[26] Richard M Held and Nathaniel I. Durlach. 1992. Telepresence. *Presence: Teleoperators and Virtual Environments* 1, 1 (jan 1992), 109–112. https://doi.org/10.1162/pres.1992.1.1.109

[27] Yvonne Hellin-Hobbs. 2010. The constructivist museum and the web. https://doi.org/10.14236/ewic/eva2010.13

[28] Patrick Hennig, Philipp Berger, Christoph Meinel, Maria Graber, Jens Hildebrandt, Stefan Lehmann, and Cathleen Ramson. 2013. Tracking visitor engagement in the blogosphere for leveraging rankings. In *Proceedings - SocialCom/PASSAT/BigData/EconCom/BioMedCom 2013.* https://doi.org/10.1109/SocialCom.2013.57

[29] Wijnand Ijsselsteijn and Giuseppe Riva. 2003. Being There : The experience of presence in mediated environments. *Being There: Concepts, effects and measurement of user presence in synthetic environments* (2003), 14. https://doi.org/citeulike-article-id:4444927

[30] Alejandro Jaimes, Mounia Lalmas, and Yana Volkovich. 2011. First international workshop on Social Media Engagement (SoME 2011). In *Proceedings of the 20th International Conference Companion on World Wide Web, WWW 2011.* https://doi.org/10.1145/1963192.1963326

[31] Maria Koutamanis, Helen G.M. Vossen, and Patti M. Valkenburg. 2015. Adolescents' comments in social media: Why do adolescents receive negative feedback and who is most at risk? *Computers in Human Behavior* 53 (2015), 486–494. https://doi.org/10.1016/j.chb.2015.07.016

[32] Olave E. Krigolson, Mathew R. Hammerstrom, Wande Abimbola, Robert Trska, Bruce W. Wright, Kent G. Hecker, and Gordon Binsted. 2021. Using Muse: Rapid Mobile Assessment of Brain Performance. *Frontiers in Neuroscience* 15, January (2021), 1–11. https://doi.org/10.3389/fnins.2021.634147

[33] Olave E. Krigolson, Chad C. Williams, and Francisco L. Colino. 2017. Using Portable EEG to Assess Human Visual Attention. Vol. 10284. 56–65. https://doi.org/10.1007/978-3-319-58628-1_5

[34] Cliff Lampe. 2013. Behavioral trace data for analyzing online communities. In *The SAGE Handbook of Digital Technology Research.* https://doi.org/10.4135/9781446282229.n17

[35] Jane Lessiter, Jonathan Freeman, Edmund Keogh, and Jules Davidoff. 2001. A Cross-Media Presence Questionnaire: The ITC-Sense of Presence Inventory. *Presence* 10, 3 (2001), 282–297. https://doi.org/10.1162/105474601300343612

[36] Yue Li, Eugene Ch'ng, Shengdan Cai, and Simon See. 2018. Multiuser Interaction with Hybrid VR and AR for Cultural Heritage Objects. In *Digital Heritage 2018.* IEEE, San Francisco, USA. https://doi.org/10.1109/DigitalHeritage.2018.8810126

[37] Yue Li, Eugene Ch'ng, Sue Cobb, and Simon See. 2022. Presence and Communication in Hybrid Virtual and Augmented Reality Environments. *PRESENCE: Virtual and Augmented Reality* (2022).

[38] Yue Li, Paul Tennent, and Sue Cobb. 2019. Appropriate Control Methods for Mobile Virtual Exhibitions. In *VRTCH'18*. Springer, Brasov, Romania. https://doi.org/10.1007/978-3-030-05819-7_13

[39] Matthew Lombard and Theresa Ditton. 1997. At the Heart of It All: The Concept of Presence. *Journal of Computer-Mediated Communication* 3, 2 (1997), 0. https://doi.org/10.1111/j.1083-6101.1997.tb00072.xView/save

[40] Matthew Lombard, Theresa B. Ditton, and Lisa Weinstein. 2009. Measuring Presence: The Temple Presence Inventory. In *Proceedings of the 12th Annual International Workshop on Presence*. 1–15. http://www.temple.edu/ispr/prev{_}conferences/proceedings/2009/Lombard{_}et{_}al.pdf

[41] Jack M. Loomis. 1992. Distal Attribution and Presence. *Presence: Teleoperators and Virtual Environments - Premier issue* 1, 1 (1992), 113–119.

[42] Nicolai Marquardt and Saul Greenberg. 2012. Informing the design of proxemic interactions. *IEEE Pervasive Computing* (2012). https://doi.org/10.1109/MPRV.2012.15

[43] Alison McMahan. 2003. Immersion, Engagement, and Presence: A Method for Analyzing 3-D Video Games. In *The Video Game Theory Reader*.

[44] Ryan P. McMahan, Doug A. Bowman, David J. Zielinski, and Rachael B. Brady. 2012. Evaluating display fidelity and interaction fidelity in a virtual reality game. *IEEE Transactions on Visualization and Computer Graphics* 18, 4 (2012), 626–633. https://doi.org/10.1109/TVCG.2012.43

[45] Timothy McMahan, Ian Parberry, and Thomas D. Parsons. 2015. Evaluating Player Task Engagement and Arousal Using Electroencephalography. *Procedia Manufacturing* 3, Ahfe (2015), 2303–2310. https://doi.org/10.1016/j.promfg.2015.07.376

[46] Michael Meehan, Brent Insko, Mary Whitton, and Frederick P. Brooks. 2002. Physiological measures of presence in stressful virtual environments. In *Proceedings of the 29th annual conference on Computer graphics and interactive techniques - SIGGRAPH '02*. ACM Press, New York, New York, USA, 645. https://doi.org/10.1145/566570.566630

[47] Marvin Minsky. 1980. Telepresence. *OMNI magazine* June (1980), 1–6.

[48] Diego Monteiro, Hai Ning Liang, Andrew Abel, Nilufar Bahaei, and Rita De Cassia Monteiro. 2018. Evaluating engagement of virtual reality games based on first and third person perspective using EEG and subjective metrics. *Proceedings - 2018 IEEE International Conference on Artificial Intelligence and Virtual Reality, AIVR 2018* (2018), 53–60. https://doi.org/10.1109/AIVR.2018.00015

[49] Helen Neale and Sarah Nichols. 2001. Theme-based content analysis: A flexible method for virtual environment evaluation. *International Journal of Human Computer Studies* 55, 2 (2001), 167–189. https://doi.org/10.1006/ijhc.2001.0475

[50] Heather L. O'Brien, Paul Cairns, and Mark Hall. 2018. A practical approach to measuring user engagement with the refined user engagement scale (UES) and new UES short form. *International Journal of Human Computer Studies* 112, December 2017 (2018), 28–39. https://doi.org/10.1016/j.ijhcs.2018.01.004

[51] Heather L. O'Brien and Elaine G. Toms. 2008. What is User Engagement? A Conceptual Framework for Defining User Engagement with Technology. *Journal of the American Society for Information Science and Technology* 59, 6 (2008), 938–955. https://doi.org/10.1002/asi20801

[52] Heather L. O'Brien and Elaine G. Toms. 2010. The development and evaluation of a survey to measure user engagement. *Journal of the American Society for Information Science and Technology* 61, 1 (jan 2010), 50–69. https://doi.org/10.1002/asi.21229

[53] Jeeyun Oh and S. Shyam Sundar. 2016. User Engagement with Interactive Media: A Communication Perspective. In *Why Engagement Matters*. Springer International Publishing, Cham, 177–198. https://doi.org/10.1007/978-3-319-27446-1_8

[54] Jorge Peña and Seung-Chul Yoo. 2014. Under Pressure: Avatar Appearance and Cognitive Load Effects on Attitudes, Trustworthiness, Bidding, and Interpersonal Distance in a Virtual Store. *Presence: Teleoperators and Virtual Environments* 23, 1 (feb 2014), 18–32. https://doi.org/10.1162/PRES_a_00166

[55] Matthew Pike and Eugene Ch'ng. 2016. Evaluating virtual reality experience and performance: a brain based approach. In *Proceedings of the 15th ACM SIGGRAPH Conference on Virtual-Reality Continuum and Its Applications in Industry - VRCAI '16*. ACM Press, New York, New York, USA, 469–474. https://doi.org/10.1145/3013971.3014012

[56] Alan T. Pope, Edward H. Bogart, and Debbie S. Bartolome. 1995. Biocybernetic system evaluates indices of operator engagement in automated task. *Biological Psychology* 40, 1-2 (1995), 187–195. https://doi.org/10.1016/0301-0511(95)05116-3

[57] Lawrence J. Prinzel, Frederick G. Freeman, Mark W. Scerbo, Peter J. Mikulka, and Alan T. Pope. 2000. A closed-loop system for examining psychophysiological measures for adaptive task allocation. *International Journal of Aviation Psychology* 10, 4 (2000), 393–410. https://doi.org/10.1207/S15327108IJAP1004_6

[58] Holger Regenbrecht. 2002. Measuring presence in augmented reality environments: design and a first test of a questionnaire. *Annual International Workshop Presence* (2002), 1–4. https://www.google.com/{%}5Cnpapers3://publication/uuid/BE89B550-B84B-4411-9464-71837C29609F

[59] Marco C. Rozendaal, David V. Keyson, Huib de Ridder, and Peter O. Craig. 2009. Game feature and expertise effects on experienced richness, control and engagement in game play. *AI and Society* (2009). https://doi.org/10.1007/s00146-009-0188-3

[60] Thomas Schubert, Frank Friedmann, and Holger Regenbrecht. 2001. The experience of presence: Factor analytic insights. *Presence: Teleoperators and Virtual Environments* 10, 3 (2001), 266–281. https://doi.org/10.1162/105474601300343603

[61] Martijn J. Schuemie, Peter van der Straaten, Merel Krijn, and Charles A.P.G. van der Mast. 2001. Research on Presence in Virtual Reality: A Survey. *CyberPsychology & Behavior* 4, 2 (2001), 183–201. https://doi.org/10.1089/109493101300117884

[62] Neta Shaby, Orit Ben Zvi Assaraf, and Tali Tal. 2017. The Particular Aspects of Science Museum Exhibits That Encourage Students' Engagement. *Journal of Science Education and Technology* (2017). https://doi.org/10.1007/s10956-016-9676-7

[63] Thomas B Sheridan. 1992. Musings on telepresence and virtual presence. *Presence: Teleoperators & Virtual Environments* 1, 1 (1992), 120–126.

[64] Richard Skarbez, Frederick P Brooks, and Mary C Whitton. 2017. A Survey of Presence and Related Concepts. *ACM Comput. Surv. Article* 50, 96 (2017). https://doi.org/10.1145/3134301

[65] Mel Slater. 2009. Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society B: Biological Sciences* 364, 1535 (2009), 3549–3557. https://doi.org/10.1098/rstb.2009.0138

[66] Mel Slater, Christoph Guger, Guenter Edlinger, Robert Leeb, Gert Pfurtscheller, Angus Antley, Maia Garau, Andrea Brogni, and Doron Friedman. 2006. Analysis of Physiological Responses to a Social Situation in an Immersive Virtual Environment. *Presence: Teleoperators and Virtual Environments* 15, 5 (2006), 553–569. https://doi.org/10.1162/pres.15.5.553

[67] Mel Slater, Martin Usoh, and Anthony Steed. 1994. Depth of Presence in Virtual Environments. *Presence: Teleoperators and Virtual Environments* 3 (1994), 130–144.

[68] Mel Slater and Sylvia Wilbur. 1997. A Framework for Immersive Virtual Environments (FIVE): Speculations on the Role of Presence in Virtual Environments. *Presence: Teleoperators and Virtual Environments* 6, 6 (dec 1997), 603–616. https://doi.org/10.1162/pres.1997.6.6.603

[69] Jonathan Steuer. 1992. Defining Virtual Reality: Dimensions Determining Telepresence. *Journal of Communication* 42, 4 (1992), 73–93. https://doi.org/10.1111/j.1460-2466.1992.tb00812.x

[70] Stella Sylaiou, Katerina Mania, Athanasis Karoulis, and Martin White. 2010. Exploring the Relationship Between Presence and Enjoyment in a Virtual Museum. *International Journal of Human Computer Studies* 68, 5 (2010), 243–253. https://doi.org/10.1016/j.ijhcs.2009.11.002

[71] Arthur Tang, Frank Biocca, and Lynette Lim. 2004. Comparing Differences in Presence during Social Interaction in Augmented Reality versus Virtual Reality Environments : An Exploratory Study. *7th Annual International Workshop on Presence* (2004), 204–208. papers3://publication/uuid/81D1A395-AA28-4DC5-9C50-3FD2A207F6B6

[72] Paul Tennent, Sarah Martindale, Steve Benford, and Dimitrios Darzentas. 2020. Thresholds : Embedding Virtual Reality in the Museum. *Journal on Computing and Cultural Heritage* 13, 2 (2020), 12:1–12:35.

[73] Daniel Vogel and Ravin Balakrishnan. 2004. Interactive public ambient displays: Transitioning from implicit to explicit, public to personal, interaction with multiple users. *UIST: Proceedings of the Annual ACM Symposium on User Interface Softaware and Technology* 6, 2 (2004), 137–146.

[74] Cassandra M. Wilkinson, Jennifer I. Burrell, Jonathan W.P. Kuziek, Sibi Thirunavukkarasu, Brian H. Buck, and Kyle E. Mathewson. 2020. Predicting stroke severity with a 3-min recording from the Muse portable EEG system for rapid diagnosis of stroke. *Scientific Reports* 10, 1 (2020), 1–11. https://doi.org/10.1038/s41598-020-75379-w

[75] Bob G. Witmer, Christian J. Jerome, and Michael J. Singer. 2005. The factor structure of the Presence Questionnaire. *Presence: Teleoperators and Virtual Environments* 14, 3 (2005), 298–312. https://doi.org/10.1162/105474605323384654

[76] Bob G. Witmer and Michael J. Singer. 1998. Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence: Teleoperators and Virtual Environments* (aug 1998), 225–240. http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.2233447