



OPEN ACCESS

EDITED BY

Ekaterina Prasolova-Førland,
Norwegian University of Science and
Technology, Norway

REVIEWED BY

Judith Molka-Danielsen,
Molde University College, Norway
Jose Garcia,
University of Applied Sciences
Technikum Wien, Austria

*CORRESPONDENCE

Anthony Steed,
✉ a.steed@ucl.ac.uk

SPECIALTY SECTION

This article was submitted to Virtual
Reality and Human Behaviour,
a section of the journal
Frontiers in Virtual Reality

RECEIVED 21 December 2022

ACCEPTED 01 March 2023

PUBLISHED 23 March 2023

CITATION

Steed A, Archer D, Izzouzi L, Numan N,
Shapiro K, Swapp D, Lammiman D and
Lindeman RW (2023), Immersive
competence and immersive literacy:
Exploring how users learn about
immersive experiences.
Front. Virtual Real. 4:1129242.
doi: 10.3389/frvir.2023.1129242

COPYRIGHT

© 2023 Steed, Archer, Izzouzi, Numan,
Shapiro, Swapp, Lammiman and
Lindeman. This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is
permitted, provided the original author(s)
and the copyright owner(s) are credited
and that the original publication in this
journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Immersive competence and immersive literacy: Exploring how users learn about immersive experiences

Anthony Steed^{1*}, Dan Archer¹, Lisa Izzouzi¹, Nels Numan¹,
Kalila Shapiro¹, David Swapp¹, Dinah Lammiman² and
Robert W. Lindeman³

¹Department of Computer Science, University College London, London, United Kingdom, ²Department of Anthropology, University College London, London, United Kingdom, ³Human Interface Technology Lab, University of Canterbury, Christchurch, New Zealand

While immersive experiences mediated through near-eye displays are still a relatively immature medium, there are millions of consumer devices in use. The level of awareness of the forms of the interface and media will vary enormously across the potential audience. Users might own personal devices or might encounter immersive systems in various venues. We introduce the term *immersive competence* to refer to the general practical knowledge and skills that users accumulate about how typical immersive interfaces work—the ways in which buttons are used, main locomotion techniques, etc. We then introduce the term *immersive literacy* to refer to awareness of how immersive interfaces are unique, when they might be appropriate, typical forms of media, etc. We sketch out how users develop competence and literacy with immersive media, and then highlight various open questions that are raised.

KEYWORDS

immersive, virtual reality, mixed reality, user experience, literacy, competence

1 Introduction

Millions of units of consumer virtual reality (VR) systems have now been sold. There is a developing ecosystem of content that is targeted at these devices. Much of this content is games or game-like social experiences, but there has been a very diverse range of content produced, from music experiences to documentaries from conflict zones. As this new medium becomes more widespread and mainstream, there is a need to better understand how users encounter immersive systems so that research and development can adapt to the users' growing awareness of the technology. In particular, we note that in the population there is a very broad range of levels of expertise in the use of immersive technology.

Prior research back to the 1990s has highlighted that the user experience of immersive content has some distinctive characteristics: the user is usually immersed and thus separated from their physical environment; the view of the content is first person and slaved to user head movements and thus there is a visuo-proprioceptive match; the user can use their own body to interact with the content. This has led to a very active research field on presence and immersion, which has recently intersected with themes of embodied interaction and the neuroscience of embodiment e.g., see surveys of (Kilteni et al., 2012; Skarbez et al., 2017). Much has been made of the ability of immersive technologies to generate reactions to virtual

simulations that are similar to analogous situations in the real world (Slater, 2009). These range from looming responses (Heeter, 1992) through to socially-conditioned responses to avatars (Heeter, 1992). Further, because the user of head-mounted display (HMD) systems can be represented by an avatar inside the system, it has been shown that the form of that avatar can change biases or engender different behaviours [e.g., (Kiltner et al., 2013; Pan and Steed, 2019)]. These results give great motivation to study immersive systems as an interface type with unique capabilities.

What this body of research fails to address head-on is the evolving expectations of users of consumer applications. In particular, it is unclear what more-experienced users of these systems are learning and expecting from the interfaces, how they acquire that experience and how their behaviour changes over time. For example, some phenomena such as reactions to virtual cliffs might only have their dramatic impact a small number of times, due to habituation. We might expect experienced users to be familiar with the options available to them in different applications to represent themselves as an avatar. We might also expect them to have strong preferences for specific games or social experiences, because they fit their interests for activity or self-expression. In addition, we might expect them to have preferences for certain platforms because of their investment in certain hardware and app platforms.

We should expect naïve users to react very differently to immersive content compared with experienced users. But we should also expect naïve users to have their own expectations about the content because of an increasing amount of media on other platforms that is about immersive content.

The goal of this paper is thus to outline an area of study that spans concepts that we have termed *immersive competence* and *immersive literacy*. We define these as:

Immersive competence is a user's familiarity and skill with the interface and controls of immersive systems.

and:

Immersive literacy is a user's familiarity, awareness and experience with the theory and motivations behind immersive system design.

Thus we consider immersive competence to be more related to practical skills and immersive literacy to be more critical and reflective knowledge. We acknowledge that these are not orthogonal concepts and any experience of an immersive system gives you both some literacy and competence, but we argue that there are differences we can identify between users, both naïve and expert. In particular, we note that more-experienced users might be familiar with specific concepts, such as presence, that might bias their behaviours. We will discuss the parallels and differences between these in more detail in Section 3.

Throughout the paper, we will refer to immersive technologies as a term that primarily refers to VR, but we would consider certain head-mounted augmented reality (AR) systems as partly covered by the same concerns. There are already systems that can switch between the two such as the Varjo XR systems¹.

In this paper, we cover a broad range of prior research work that has reflected on the form of immersive interfaces. We also draw on the authors' observations of developing and demonstrating immersive experience to thousands of users, in both academic and commercial contexts. We argue that the main issues around immersive competence are about ensuring that new interfaces are learnable and discoverable by a broad range of users with different levels of immersive competence and immersive literacy: from users on their first experience, through to highly competent veteran users who have tried lots of different systems.

The first contribution of this paper is the introduction of the terms immersive competence and immersive literacy, and an initial development of their scope. The second contribution is the posing of some future research directions that will hopefully motivate more researchers to focus on research and development of immersive systems that support the broad range of users' competence and literacy. The final contribution is that the paper can act as a reference to crystallise further interdisciplinary reflection on the uniqueness of immersive media.

2 Background

2.1 Immersive experiences

Immersive VR and AR systems evolved from earlier non-immersive simulation technologies. While VR has more recently come of age as a viable consumer technology, many key concepts and observations about the potential power of immersive systems date back to the 1990s and earlier. Ivan Sutherland's tracked AR HMD (Sutherland, 1968) is usually cited as the first example of an immersive system driven by computer-generated graphics. Kalawsky documents the developments in the 1980s that led to commercial VR systems, with significant developments in computer graphics hardware being necessary to support the needs of the field (Kalawsky, 1993). Turn-key VR systems such as the "Reality Built for Two" from VPL (Blanchard et al., 1990) and the "ProVision 100" from Divison (Grimsdale, 1991) were available for relatively well-funded labs to purchase and experiment with. An active development community evolved [e.g., see (Rheingold, 1992; Delaney, 2017)]. Early research highlighted the sense of "presence" as an interesting phenomenon reported by users of these systems (Held and Durlach, 1992; Heeter, 1992).

However, the high costs and low quality of output meant that systems were too expensive for consumer use. The introduction of the CAVE (Cruz-Neira et al., 1993) brought high resolution and very wide field of view, and such systems were successfully exploited by a number of industries [e.g., see (Brooks, 1999)]. Textbooks such as Sherman and Craig (2003) document the rapid expansion of the scope of research on systems in the 1990s.

It is in the 2000s that developments in computer graphics hardware started to converge with the latent demand for immersive applications. While the explosion of Oculus onto the scene is perhaps the key event, there was already a small but active industry and a range of academic labs looking at novel systems. Jerald's (2015) book is a good reference of where the technologies had got to just as the consumer VR market flourished. There is an associated rich literature of critique of the technologies, descriptions

¹ <https://varjo.com/products/xr-3/>.

of their applications in various domains and speculation on their potential [e.g., (Grau et al., 2004; Laurel, 2014; Lanier, 2017; Bailenson, 2018)].

Thus, while it might be claimed that modern consumer VR and AR systems are in some ways not so different in capability from lab-based systems going back to the 1990s (Steed et al., 2021), certainly in the past decade a very large number of new commercial systems have become available, many millions more people have experienced VR or AR and there has been a very large increase in the amount of content. This new diversity of users, systems and content motivates a re-evaluation of what it means to be competent and literate in this medium.

2.2 Competence and literacy

The term *literacy* has traditionally referred to the written word—one's ability to read and write. With the advent of new forms of media, the term evolved to consider not only production and reception skills, but also critical skills in contextualising and questioning the information that we receive through these media. Computer technology has rapidly expanded the range of media and this has to some degree further fractured distinctions between what are referred to as literacy and competence. For example, the term computer literacy is used to refer to technical skills in using information technology. Much of the theoretical work around literacy and competence is in the area of education research. Two recent review papers both cite some of the difficulties in terminology, in particular the conflation of literacy and competence. Zhao et al. (2021) note that the distinction between the terms is blurred, and further complicated by regional variations (across languages and countries) in definitions or common usage. Spante et al. (2018) note a variation in the use of the terms literacy and competence depending on whether the concepts are defined by policy, research or both, and whether they focus on technical skills or social practices.

2.3 Competence and literacy in other domains

Some of the most common uses of the term literacy are in the overlapping domains of computer literacy, information literacy and digital literacy. These have a very broad scope, but this is well documented and dissected in the relevant communities [e.g., (Horton, 1983; Bawden, 2001)]. These literacies have become increasingly important to the average person as more and more services and information sources have gone online. These types of literacy are taught in schools. The European Union Digital Competence Framework (Vuorikari et al., 2022) attempts to document the broad range of digital skills and knowledge that are required in modern society.

Among the related literacy concepts discussed by Bawden (2001) is media literacy, which likewise has varying definitions depending on the context and medium of use (Silverblatt et al., 2015). Potter (2010) provided an overview of the various schools of thought, with some scholars viewing it as developing skill and ability (the way this

paper describes competence) or knowledge, while others see it as a hands-on learning activity requiring engagement with the content.

Physical literacy describes one's ability to move, but also the ability to read the environment and to respond effectively (Whitehead, 2001). Immersive systems certainly require some amount of physical literacy, and we expect there will be some cross-overs to develop. Another overlapping domain is that of game literacy [e.g., Buckingham and Burn (2007)], and we expect that immersive games will be critiqued within that domain. We argue in later sections that there are many unique aspects that distinguish immersive systems from other interactive systems. Sherman and Craig (1995) were the first to use the term immersive literacy in an early paper discussing the flexibility of this new medium. They do not provide a definitive definition but propose the challenge that as VR is a new medium, one can be literate in that medium, and literacy would encompass being familiar with the capabilities of that medium. One of the key points in that paper was that literacy would be hard to achieve as the medium was relatively unexplored. Finally, we note recent work in the HCI community developing literacy about artificial intelligence systems as users need to develop competencies to interact with them (Long and Magerko, 2020).

3 Comparing immersive competence and immersive literacy

As noted in the introduction, we do not claim that immersive competence and immersive literacy are orthogonal. As we discuss in Section 4, each covers a very broad range of knowledge of users. But a key contention is that users will have different levels of competence and of literacy, and that these will not develop in parallel. We now consider two users how have different trajectories in the levels of competence and literacy.

User 1 has had some gaming background on traditional gaming platforms (consoles, PCs, etc.). They buy a Meta Quest for home use. In their first experience with an immersive system, they thus start with a higher level of immersive competence because they can transfer some skills (e.g., joystick use) to the immersive situation. They are more likely to be predisposed to learning interactive systems more quickly so they rapidly learn how to use the immersive system. However, they focus on a narrow range of games and do not explore some of the more broad types of experiences, so they develop some literacy about immersive technologies, but are constrained to the gaming-focused styles of content.

User 2 does not own any immersive equipment, nor is much of a games player, but they are engaged in VR training at work, and thus they are very interested in applications of VR. When they first try an immersive system, they have already read about the technology and know a bit about how VR has been used to train people in other companies. They continue to try immersive systems when they encounter them in public installations. They thus get to try various different systems. They also read more about the technology, and so become exposed to more critique about the technology and come across concepts such as presence. They develop a broad understanding of the technology and become moderately

competent in using different systems and adapting to new systems as they encounter them.

These two user stories highlight that the population that encounters and uses immersive technology cannot simply be described as naïve or expert. Both naïve users and experts could have quite different competencies and literacies. This will be one of our major topics in the discussion in [Section 4](#), in particular when we discuss how users develop competence and literacy in [Section 4.3](#).

3.1 Immersive competence

Following our definition in [Section 1](#), in this section, we elaborate immersive competence as being more about the practical skills and competences with the controllers and controls of immersive systems. This is analogous to competence with other physical and virtual elements of computer user interfaces. Thus, analogies might be the familiarity with mouse and keyboard interfaces, mobile and touch-screen input, and the desktop metaphor.

It is clear that despite some definitions of VR claiming the lack of an “interface,” or emphasising the perceptual illusion of non-mediation ([Lombard and Ditton, 1997](#)), as encountered today and in the near term, VR and AR are usually experienced through multiple interface devices that impinge on the user experience. That is, HMDs are obviously heavy to wear and are visible (and sometimes audible if fan-cooled) to the user. The user also often carries controller devices, though hand tracking can be used on a number of devices. For example, on the Microsoft HoloLens 2, hand gestures are required for interaction, whereas on Meta Quest this is an option that some applications use. In either case, the user needs to learn some conventions for control.

In the remainder of this section we start to outline some of the components of immersive competence. These should be considered an initial survey of the components, with the aim to start a community discussion about what is felt to be the key skills and knowledge. These key skills and knowledge could then be the subject of assessment or the target of training and/or onboarding materials.

3.1.1 Hardware

We are focused here on the majority of current consumer VR and AR systems that comprise an HMD and either controller-based input or hand tracking-based input. These typically comprise an HMD and, usually, one or more hand-held controllers.

3.1.1.1 HMD

Taking the HMD first, users need to be instructed on how to put on the HMD. The user needs to learn and understand:

- How to grossly adjust the HMD straps or fit the optics in a good position;
- How to finely adjust the HMD to optimise the position of the HMD for the clearest view, including potentially shifting the HMD in and out;
- How to confirm that the screens are (both) on and showing the correct aspect ratio;
- How to confirm whether there are interpupillary distance (IPD) adjustments, potentially even their personal IPD measurements, and how to adjust them;
- Whether there are any other optical adjustments or configurations to make (e.g., diopter adjustment, alternative lenses);
- How to calibrate any eye-tracking sub-systems present in the HMD;
- How to position any audio headphones that are either a separate component or attached to the HMD;
- How to adjust the audio levels.

Our experience with demonstrating to the public or in laboratory situations is that we need to confirm these aspects with users, because users might be naïve about what to expect, including simply whether the screens start on or off. A more-experienced user might understand that the screens are supposed to have a specific sharpness, and so might adjust them, but a less-experienced user may not even know what should be considered sharp.

The most obvious aspect, which still needs to be explained to some naïve users, is that the HMD is tracked [three or six degrees of freedom (DOF)] and that graphics are slaved to this in a first-person egocentric manner. Thus the user can look around naturally and explore the environment (in VR) or find additional objects or media (in AR). This is often the first “Wow!” moment for the naïve user. In subsequent uses of VR or AR, this will not have the same impact. In our experience, some users might turn their bodies or heads but are hesitant to move very far. If the experience supports six-DOF movements less experienced users might need to be told that they can bend down, or even crouch, crawl, etc. It may not be clear that these options are possible, and there might be a lingering worry that they will not work (of course some movements are not possible due to the bulk of the HMD itself).

Users pick up these immersive competence skills quickly. After experiencing a few different HMDs, users will probably understand the main differences between them. But some adjustments need to be experienced to be known about. For example, not all HMDs have IPD adjustment, and eye tracking is only available in some HMDs available at the time of writing.

3.1.1.2 Hand-held controllers

An idealised controller would support everything that we can do with our bodies in the real world: walking, running, reaching and grasping, throwing and a whole range of other actions, such as catching, writing, typing and fine motor skills. This is far beyond current technology, and so contemporary hand controllers have evolved from gaming controllers, borrowing and extending interface metaphors from these, as well as from 2D mouse interfaces.

For controller-based systems, the user needs to learn:

- How many buttons there are, their locations and the easiest grasp to reach them;
- Whether there is a joystick or other 2D locator device, how to use this, and whether it acts as a button;
- Which buttons, if any, are analogue distance sensitive or pressure sensitive;
- Whether the buttons can detect proximity;

- Whether the controller is tracked and in what cases this fails (e.g., when held out of sight);
- What the mappings are of the various inputs (buttons/joystick) to control the experience.

The controller is usually tracked in six DOF. There are devices which are only three DOF (rotation only), and some that only act as button input. However, the main consumer VR systems at least are converging on a standard set of buttons and joysticks (Steed et al., 2021). These control systems typically have two buttons and a joystick roughly on the top, for operation with the thumb, one trigger button (or two) under the first (and second) fingers, and options/menu buttons. Of course, when the user puts their hand into the field of view of the HMD, most systems will draw a controller or virtual hand. This is often the second “Wow!” moment in most HMDs, and a naïve user will often explore this experience of seeing their hand, and seeing what the button presses do. The user may learn that some controllers can only be tracked in certain regions, such as in front of them for systems such as the Meta Quest that use cameras on the HMD to track controllers.

A naïve user might not expect much from any particular hardware. A more-experienced user might be expected to understand the differences between HMDs and controllers, and potentially be aware of the constraints and opportunities that certain installation types have. For example, a more-experienced user might recognise that Google Cardboard or Oculus Go only support three-DOF experiences with minimal interaction. We discuss this further in Section 4.2 when we discuss types of immersive installations.

3.1.2 Controller embodiment

By *controller embodiment*, we refer to how the controls are manifest inside the virtual environment. In VR, the system has options. The default in SteamVR-based systems is to present models of the controllers that show the manipulations that the user makes to the controllers (e.g., pushed buttons are shown as pushed in the virtual environment). This gives a subtle steer to the user that the controller is present and might act as a tool. The design space of options here is massive, with some systems endowing the virtual tool with interesting virtual capabilities or configurations (e.g., the game *Budget Cuts*²). Another option is to draw a hand, either in a fixed pose that might match the default grip of the controller, or dynamically changing with the hand movements of the user. This hand may animate to show that the fingers and thumb are interacting with the buttons and joystick, or it may mimic the pose of the hand as detected by proximity.

Controller embodiment is something that even experienced users need to assess: in the first few seconds of using most immersive systems, experienced users will usually test the controllers and operate the buttons to assess what works. This is because applications are very different in what they display. At a basic level, looking at the controllers ensures that they are turned on, the tracking is working as expected, and the program is operating as expected because the control buttons function.

3.1.3 Avatar embodiment

By *avatar embodiment* we refer to systems that go beyond simple tracking to infer a partial or whole body of the user (a “self-avatar”). While many single-user systems do not support self-avatars, in social applications it is common to see avatars that represent other users, and thus it makes sense to draw a self-avatar for the user so that they can see how others see them. This also affords opportunities for avatar customisation.

Avatar embodiment is one area that all users need in order to express themselves. While a naïve user might need help understanding that the avatar they see is themselves in a mirror, more-experienced users will be interested in how the avatar moves, any animations applied to it, gestures it can make, how to control facial expressions, etc. We distinguish avatar embodiment from the more general term embodiment as used in the VR literature because that is focused on the relation to the concepts of body schema and body image as explored on the boundaries of immersive media and neuroscience (see Section 3.2.5).

In AR, the user almost always sees their own hands, so it might seem that avatar embodiment is less interesting. However, the systems might include representations of tools held by the user or adornments such as jewellery worn by the user.

Some systems, including the Meta Quest, include hand-tracking capabilities. A representation of the hand as it is tracked is thus almost necessary in the system to show that the hands are being tracked and their momentary poses. The hand shape can be used to drive physical interaction, or interpreted as symbolic gestures.

3.1.4 Guardian systems

HMDs can obscure some or most of the real world, so there is a risk that users will bump into physical objects. To address this problem, most consumer systems include some sort of guardian system that represents the boundaries of the activity space. If the user gets too close to the boundary, a visual barrier appears on the display. Users need to be aware that these systems exist, the importance of defining and using them correctly, and also what their weaknesses are. For example, they currently use a static capture of the surrounding environment, and so will not detect moving objects or objects that have been introduced to the space since the last description of the boundary.

These issues are not so relevant for exhibition spaces or other installation spaces, where users can expect to be safe from collisions. Not only is the space likely to be properly calibrated, but content will likely be chosen or customised to fit the space. Additionally, there is less likelihood of other users or household pets entering the space (see Section 4.2).

3.1.5 Basic actions

Most VR and AR systems focus on visual and audio presentation and thus lack the ability to fully simulate environments as they cannot provide all the sensory cues that we expect when interacting with real objects. The most obvious example is that virtual objects are not solid and do not have weight, so grasping needs to be simulated. This will seem unnatural to naïve users, and will also hamper natural interaction, since many common cues will be missing.

Thus, in implementing interactions, the designer must make a distinction between realistic (e.g., direct manipulation) and

² https://store.steampowered.com/app/400940/Budget_Cuts/.

supernatural interaction metaphors. In the former case, as the user moves and manipulates the (real) controller, its virtual representation moves linearly in accordance with its real motion. However, some interaction metaphors allow extended reach [e.g., (Poupyrev et al. 1996)] by mapping the virtual controller non-linearly over a much longer range.

The most prevalent use of non-realistic interaction is for virtual navigation. Real walking by the user is restricted by the physical space available. Real walking is thus normally limited to a handful of steps. Applications typically use the joystick on the controllers for locomotion, or implement a teleportation technique. A very broad survey of locomotion techniques is given by Di Luca et al. (2021).

3.1.6 Simulator sickness

While knowledge about why simulator sickness occurs might be considered more as part of immersive literacy under our definitions, we nevertheless acknowledge that tactics to avoid simulator sickness are skills, and thus fall under immersive competence. Simulator sickness still occurs (Kennedy et al., 1993; Saredakis et al., 2020), though there is now a lot of knowledge about how to diminish its effects, such as the technique of narrowing the field of view when turning (Bolas et al., 2014) as demonstrated in the game *Eagle Flight*³, amongst others. Development guidelines for consumer systems often mention interaction techniques that should be avoided to reduce the prevalence of simulator sickness (Meta Inc., 2022).

Thus, we note that users may develop skills for avoiding simulator sickness (e.g., not turning the head too quickly), or develop behaviours for how to minimise the likelihood of suffering symptoms (e.g., choosing to use teleportation rather than joystick locomotion). We note that current content distribution platforms normally have some sort of comfort rating system for VR games (e.g., Meta classifies content into Comfortable, Moderate or Extreme).

3.2 Immersive literacy

In this section we discuss some elements of knowledge about immersive systems and content that might be considered elements of immersive literacy. As defined in Section 1, we consider literacy to be the knowledge about the expected types of experience that one might have on immersive systems. Of course, the design space of virtual environments is massive. It ranges from re-creations of other media (e.g., simulations of comics, watching of TV) through architectural explorations (e.g., Museum of Other Realities⁴, reconstructions of famous monuments) to the psychedelic and fantastic (e.g., Visionarium⁵). Visually it ranges from cartoony to near photorealistic.

This section will not try to be comprehensive about what we consider comprises immersive literacy. Instead, we highlight some key features of forms of immersive media, theories about its impact,

and some of the emerging key decisions that inform the design of virtual experiences. The key point is that a user's experiences with immersive media, both the content and situation of use (see Section 4.2), will lead them to develop their own literacies, and keener users may engage with the rapidly expanding literature that critiques immersive media. This increased immersive literacy may then impact their attitudes and behaviours.

3.2.1 Diegetic versus non-diegetic interaction

While most immersive applications with controllers or hand-tracked input exploit direct manipulation of some environmental objects with the hands, there are limitations to this. A realistic interaction metaphor only goes so far, as objects are not solid and thus cannot constrain the hand. For example, pulling a virtual drawer will open the drawer. But once the drawer is fully extended, the drawer will not stop the real hand from moving, so either the virtual hand detaches from the drawer, or the virtual hand location will no longer have a one-to-one mapping with the real hand.

While some of this might reasonably be understood as a skill or competence that one needs to develop to interact with different applications, it informs a pervasive design choice: whether to adopt a real-world metaphor for all controls of the experiences and underlying application [referred to as diegetic interaction, see e.g., Steed et al. (2021)], or whether to adopt a metaphor for interactions based on transforming 2D user interface components directly into 3D (e.g., virtual menus) or *via* simulation on virtual tablets or computers. For example, the game *I Expect You to Die*⁶ takes the diegetic metaphor to its limit. It is a game based around different levels set in room-scale scenes. All interactions are done by direct manipulation of objects. More interestingly, the inter-level scene, where the next game level is selected, also uses a direct manipulation metaphor of placing a film reel in a projector to switch levels. In contrast, the game *Beat Saber*⁷ is primarily a rhythm action game that encourages player movement, but its control system (level selection, game configuration, etc.) uses a menu-style system that is laid out in 2D on surfaces facing the user. It is effective, but relies on a different set of skills to the game (e.g., accurate pointing at a distance).

3.2.2 Egocentric action

Another key feature of immersive technologies is that the media is presented from an egocentric point of view, and (mostly) surrounds the user. This is a natural by product of the display devices: they are typically unconstrained, so the user's turning around needs to be supported. However, a related aspect is whether or not actions are directed at the user or not. One could view the user as a passive, potentially mobile observer, or one can be present and embodied in the scene in the sense that the participant both sees themselves and other characters, and that objects in the scene react to them. In 360-degree video this is often a design decision: do actors face the camera or not, and do they address the future participant (Bevan et al., 2019)? In a study of a 3D animation

3 <https://www.ubisoft.com/en-us/game/eagle-flight>.

4 <https://www.museumor.com/>.

5 <https://store.steampowered.com/app/928300/Visionarium/>.

6 <https://iexpectyoutodie.schellgames.com/>.

7 <https://beatsaber.com/>.

created by the BBC and Aardman Animation, Steed et al. (2018) varied both the embodiment of the user and the control of the animated characters to address their monologues towards the participant or to an empty space. Both had an impact on user-reported presence.

Another interesting aspect that has been emphasised by some content developers is that because the media is all-enveloping, the user might not be able to watch everything that goes on in the scene (e.g., the 360-degree video, *Inside the Box of KURIOS*⁸). Thus, an interesting question is how much the user understands about the scene, and the potential to experience a “fear of missing out” (FOMO) (MacQuarrie and Steed, 2017).

3.2.3 Platform issues

An aspect of media literacy that partly inspired this paper is the recognition that a user’s immersive competence and literacy will develop over time in terms of how content is shaped by the control of production and distribution methods. One interesting angle on this with respect to immersive content, is that the system capabilities are developing rapidly by the platform holders. This is perhaps more profound than the older debate about platform competition in video games. For video games, it was a relatively straightforward debate about which platform would have the highest power and thus afford the richest visual media [e.g., see Kent (2010)]. With immersive technology, even researchers such as ourselves might be aware that there are many new features coming (e.g., face tracking, multi-focal displays), but we cannot predict when they will be available, what their capabilities and limitations are, how they will be used by content creators, and how they will be received by users. This causes us some anxiety about where to focus more exploratory user interface research and development.

Otherwise we note that, in a similar manner to the games and mobile app industries, the platform holders are also the gatekeepers for distribution of content. Thus they censor some content, and maintain quality standards. Notably, the Meta Store suggests that Meta Quest devices are suitable for ages 13+.

3.2.4 Presence

The topic of presence has motivated researchers since the earliest days of the technology. Most researchers have a common understanding that there is something about being immersed in the system that leads people to feeling as if they are transported elsewhere (Held and Durlach, 1992) or have a subjective feeling of “being there” (Heeter, 1992). There have been hundreds of papers written about what presence is, how it emerges from conscious perception, and what impact it has. Any short introduction cannot do justice to the topic, but a recent survey addresses some of the current debate (Skarbez et al., 2017).

Some of the key early understandings about presence depending on self-representation inside the system (Slater et al., 1998) are now underpinned by research on perception and neuroscience especially an understanding of the role of embodiment (see next section).

What has been interesting in the past couple of years is that presence as a concept has been publicised as the selling point of the

technology [e.g., Mark Zuckerberg using it to motivate their work in the Metaverse (Zuckerberg, 2021)]. Thus, feeling present with other users and in the places constructed by simulations is itself a platform objective. Therefore, we can expect more literate users to be aware of this objective and it will be interesting to tease out how this might bias them.

3.2.5 Embodiment and body-centred interaction

A key aspect of presence is how users will act, behave, and thus take ownership of the representation of their own body in VR. Otherwise referred to as embodiment (Kiltner et al., 2012), this notion is rooted in a psychological illusion known as the rubber hand illusion (Botvinick and Cohen, 1998). In VR, the embodiment illusion allows us to believe, or have agency over, a virtual hand that we see in a virtual environment and that moves according to our real hand’s movements [e.g., see Yuan and Steed (2010)]. While embodiment is recognised as an important aspect underpinning a user’s sense of presence in VR, the virtual representations of the user’s body have become recognised as a complex issue, and a key topic of VR research. A realistic representation of the user, along with accurate tracking, such that the virtual body moves according to the user’s movements (Slater and Steed, 2000), provide the ideal conditions for a strong sense of user embodiment. A comprehensive overview of the tensions between effective interaction and user embodiment is provided in (Dewez et al., 2021).

Once the embodiment illusion was proven effective in VR, researchers investigated how different types of avatars would impact users, including varying the body types, gender or age (Banakou et al., 2013; Peck et al., 2013; Maister et al., 2015). Researchers have also investigated embodiment in bodies with altered layouts such as additional limbs (Stephoe et al., 2013; Won et al., 2015).

3.2.6 Avatars and the uncanny valley

First-time users might have expectations about human-like avatars, such as expecting realistic appearance and behaviour. When these expectations are violated, users may feel disappointed or disengaged from the virtual experience. Mori et al. (2012) introduced the term *uncanny valley* to describe user reactions to human representations that approach but fail to fully meet the expectations of human likeness. Users may experience an eerie, unsettling or repulsive feeling in response to avatars that do not look quite right.

However, as the user becomes more experienced with the system, or more literate about the technology, they might understand that these types of non-photorealistic representation are not unusual. Thus, their negative feelings might reduce over time as interactions with the human representation are repeated, as was found in a study by Zlotowski et al. (2015). It might also be expected that a user’s prior experience could have an effect on interactions with avatars. For example, strong familiarity with games or cartoons might make users more accepting of the non-realistic appearance of characters.

3.2.7 Breaks in presence and the dualistic nature of media

While users are immersed in the technology, they still have to deal with the real world. Sometimes the real world will remind them of its presence: the user may collide with an object, or an external sound might occur. Sometimes the user will be reminded of their situation because of an inconsistency in the environment or

⁸ https://fb.watch/fz__u1uqoE/.

something implausible. The notion of “breaks in presence” has been proposed to highlight that users switch attention and engagement between the real and virtual (Slater and Steed, 2000). Users might also be able to interact with both environments at once. One example is talking to an external person while engaging in a task. More experienced users might be able to maintain some awareness of the likely physical space boundaries or cabling. Of course, a user might also spontaneously stop believing that the scene in front of them is actually happening.

Hartmann and Hofer (2022) related the experience of immersive systems to the idea that experiences of various media are dualistic in nature. Within communication sciences, this dualism is described as comprising an “involved reception mode” and an “analytical reception mode.” In the involved reception mode, the user suspends disbelief, thinking, and operating within the constructed narrative/world of a mediated experience, whereas in the analytical reception mode, the user is able to think about the medium itself. In the context of VR, the reception mode might be attractive to storytellers that wish to convey particular content, but our observations of users is that while some are engrossed in the experience and following along having suspended any disbelief, others switch at some point to a more active experimental mode, where the story content is background. This might be a more analytical mode and it might involve thinking about how the underlying technology is supporting the immersive experience.

4 Discussion and challenges

The large variability of previous experience that different users have while using immersive technologies poses several challenges to researchers and designers. As discussed, users might have tried very different types of immersive technologies, different types of content with different levels of interactivity and more or less guidance. Further, they may be more or less familiar with some of the theories and reflections on immersive content. They might have expectations that come from reading or watching material that describes immersive experiences, while they themselves have relatively little experience. They might have tried poor immersive systems from earlier decades. All of this means that researchers and designers need to be very aware of this variance in user competence and literacy. In this section we discuss some implications, with the expectation that each needs further development.

4.1 Prior knowledge

The term VR has been in usage for three decades, with certain aspects of the interface such as HMDs and gloves covered in many popular media from the 1990s onwards (Rheingold, 1992; Delaney, 2017). Many users will have seen depictions of immersive systems in a diverse range of TV and films (e.g., *Disclosure*⁹ or *Ready Player One*¹⁰). This means that most users encounter immersive

technologies with at least some notion of the capabilities (i.e., very few people are immersive illiterate).

In our early immersive experiments in the 1990s, we could usually assume that participants were naïve to the premise of the immersive system. While VR was sometimes covered in the media, we would encounter users who would state that they had tried VR when they had only tried something such as a screen-based experience that was marketed as VR. Today, in lab-based experiences or exhibition situations, we still encounter people who have no experience using immersive systems. However, there are now many more media publicising VR. For example, at the time of writing, in the United Kingdom, Meta is running an advertising campaign across various media, publicising the possibilities of VR and the metaverse, and electronic billboards in market squares and subways in Germany and Japan show visual advertisements for Pico VR headsets.

4.2 Encountering immersive systems

One of the interesting aspects of immersive technologies at the moment is the range of ways that a user might encounter the technology. We list several of these and note salient details of how they differ in support of development of competence and literacy.

4.2.1 Home-use VR

The majority of experiences with VR today are probably due to home use of consumer VR systems such as the Meta Quest or PlayStation VR, which have each sold millions of units. In addition, there have been a large number of three-DOF devices such as Oculus Go and devices such as Google Cardboard that hold smartphones. At home the user is self-guided, but recent devices have sophisticated introductory experiences that help the user through basic interactions. The user might be watched by people (e.g., family members) while playing, but these are well-known to the user and probably have a similar level of competence, and so can assist the user with tasks such as donning the equipment or advising on what to do and where to go. Of course, only the keenest consumers will have more than one HMD, so competence and literacy will develop within the constraints of the single platform.

4.2.2 Location-based VR

Location-based experiences such as *The Void*¹¹ are somewhat unique in that they are based on consumer equipment, or variations of it, but support large-scale motion. They tend to be group experiences, and have tutorials and instruction phases. There are staff to aid the users in getting started and to ensure the equipment is working. The experiences often exploit large-scale motion to move users around, so they experience travel through an environment and encounter physical props. There is the opportunity for other environmental effects such as fans or heat lamps to be used. If a user has already developed some competence and literacy about immersive technologies, then the large motion, props and

⁹ <https://www.imdb.com/title/tt0109635/>.

¹⁰ <https://www.imdb.com/title/tt1677720/>.

¹¹ <https://www.thevoid.com/> or *Zero Latency*¹².

environmental effects might generate new “Wow!” effects. While such experiences will expand the user’s literacy about the long-term potential of immersive technology, they will need to develop more competence when they encounter immersive technologies in other contexts.

4.2.3 Themed installations

We separated this category to cover installations that support high throughput over long periods, such as theme parks or other types of out-of-home activities. Examples include the *Derren Brown Ghost Train*¹², a VR installation at Thorpe Park in the UK, or VR rollercoaster rides from the German company VR Coaster GmbH¹³, that have been installed in various theme parks since 2016. The need for high traffic means that the experiences tend to use HMDs only, and any interaction would thus be limited to being based on head gaze. As discussed by Mine in the context of some of Disney’s early attempts at themed installations (Mine, 2003), a key part of such installations is the guidance visitors receive about the experiences, such as the type of experience and visuals from the experience, that can be given as part of the themed installation, perhaps as part of a queue area. Any specific instruction on devices is very quick, often as part of an onboarding to a ride vehicle. Thus while these might reach a lot of people and be compelling experiences, they are a very narrow subset of the potential of immersive technologies and will not develop much competence with the broader range of technology.

4.2.4 Site installations

The availability of cheap consumer systems has led to a diverse range of installations in museums and other public places. A site installation would be characterised by not being strongly themed to the site, but showing content related to the spatial context. Key here is that these installations are also staffed, so that visitors can be shown the equipment and controls for the experience. Installations of this type generally do not have such high throughput needs, so experiences may be longer (5–10 min is common) in small numbers at a time (1–4 is common). The types of content and interactions can be very diverse. The experience of such a system might give one a broader understanding than, say, a themed installation, but typically only one experience is available. A key part is that visitors can usually see others in the experience, so they get some insights into how the experience will be. A site installation might react to the architecture and theme of the space [e.g., (Tennent et al., 2020; Ioannidis et al., 2021)]. This can be done in order to make the immersive content not so unexpected and thus perhaps more plausible (see Section 4.8).

4.2.5 Hosted installations

This type of installation is distinct from location-based VR as it typically involves consumer equipment with limited range, but installed in a public site, such as a bar or arcade. These sites might have a set of themed content with timed experiences, or might effectively be offering access to a large library of games that

users select from. These spaces facilitate group access to immersive experiences out of the home. They can offer quite diverse experiences, and equipment might be shared between members of a group in an *ad hoc* manner, supporting sharing of particular content or scenes, or daring each other to take part in certain scenes (e.g., a horror scene or one that induces vertigo).

4.2.6 Exhibition installations

A final type of installation is rather niche but interesting. It is the demonstration of immersive content in an immersive exhibition or festival (e.g., *Laval Virtual*)¹⁴. The audience for these is probably already competent in immersive technologies, but we mention them because they are a venue for developing broad literacy with the forms of content, as they are often the first venues to show original, exploratory content.

A final note is that there is an emerging literature on how to best support demonstrations to the public (Emerson et al., 2019; BBC Virtual Reality, 2019; Watson, 2019; Oakes, 2018). These include best practices for equipment, expectations of users of different kinds and reflections on the best types of content.

4.3 Developing competence and literacy

How do competence and literacy build over multiple encounters? In one use case in Section 3 we suggested that users might stick with one platform and a relatively narrow genre of content, or they might be open to trying and have access to different systems. Some of the competencies we identified in Section 3.1 transfer: basic knowledge of embodiment, the need to use the controllers, the need to adjust the HMD, the need to locomote and the ways to do this. Others need to be adapted, such as control mappings, moving to hand tracking from controllers or *vice versa*, and adjustments on the HMD. Literacy is more cumulative as it concerns the types of content, types of experiences and knowledge about the capabilities of systems.

One specific threshold that might be crossed as the user gains experience is purchasing their own immersive system. There have been a number of recent papers developing technology-acceptance models for consumer hardware (Manis and Choi, 2019; Sagnier et al., 2020; Lee et al., 2019). The models bring in aspects of knowledge that are out of scope in this paper, such as perceived utility, but of course, utility would also be a motivation for continuing to engage with the systems and would also grow both competence and literacy.

These types of issues present researchers and designers with the problem of how to instruct the broad range of users in the specifics of their technology and content. The naïve user might still need to be told how to adjust the HMD, look around, raise their hands, reach out, etc., so maybe this should be “baked in” to our introductory procedures. For the purchaser of a new HMD, there is usually an introductory experience that shows off some of the features of the system, some common content types and simple interactions. These can be reasonably long (> 5 min), so it would not be possible to have

¹² <https://www.thorpepark.com/explore/theme-park/rides/derren-browns-ghost-train/>.

¹³ <https://www.vrcoaster.com/>.

¹⁴ <https://laval-virtual.com/>.

participants do this first in many situations where throughput of users is an issue. Installations often have introductory videos and instructors who convey much of the necessary material in a “pre-show.” Exhibitions and location-based VR facilities also often have helpers who can explain the technology, help with fitting the equipment, etc.

A first suggestion is that there might be value in building a common tutorial or set of tutorials that are shorter than the device introductory demonstrations, but introduce and focus on the main skills. This could be valuable in some professional or experimental usages of immersive technology where it is important that users have a basic level of competence with the technology. An alternative might be explanatory videos that users could be referred to prior to their initial exposure session. While there might not be common agreement about what such a tutorial entails, we feel that an open source effort around a sandbox-style environment that facilitates instructional experiences would be useful.

Another observation is around effective coaching or instruction of users once they have donned the equipment. Sony’s PlayStation VR allows developers to construct an explanatory view of the scene for the default TV and monitor so that other players or viewers have some understanding of what the immersed user is experiencing. Some installations support wireless streaming from mobile devices or second screens from PC-based devices, but this is not common. Enhanced platform support for open streaming or other types of environment sharing with other devices would help. Appropriate standards for this would be beneficial for bringing this to fruition.

A final observation is that while there are guidelines for giving good demonstrations [e.g. Oakes (2018)], it might be desirable to develop further guidelines to help people who are demonstrating immersive systems. This would include how to support users at different competence and literacy levels.

4.4 Levels of competence and literacy

In some situations, participants have the luxury of time to spend with an immersive system that is novel to them. However, in many situations, we might want to assess the level of competence and literacy a person has with immersive systems before they engage. The most obvious case is when throughput must be kept high. A demonstrator or operator must ensure both a quick change-over between attendees, as well as deep engagement by each attendee with the experience. Participants might be asked if they can see their hands and confirm placement of the hands on controllers and operation on buttons. The experiences will need to be fairly obvious to operate, and interactions would need to be staged and signposted. Especially in theme parks, the experiences can be mostly linear with little control over the interactions.

In other situations, we might need to support users who engage with more complex applications for longer periods of time, or regularly, such as an experiment in a lab or some form of professional training application. Here, it is important to avoid frustration with the basics of interactions and actions in the world. One suggestion, aside from providing good tutorials as suggested in the previous section, is that developers provide some tools to assess some types of competence inside the virtual world. Some of our experiments include training environments that require simple tasks

to complete. This is primarily a compulsory tutorial, but it acts as a test of competence: the task can usually be completed quickly (e.g., navigate to a series of hotspots and move objects), but otherwise it can be used to coach the user through the experience. Potentially this could be developed into a standard approach, or even a standard environment and toolset to have users show competence before proceeding, or default to a longer tutorial. This would need careful consideration, especially in situations such as experiments where training effects are important and thus a floor of competence is expected.

One measure that is increasingly used as an explanatory variable in experiments on immersive systems is some form of prior experience with immersive systems. Typically questions that have been used in the past ask users to specify how many times they have used VR or AR (as applicable) in the past month, with answers on a scale such as “Never” through “Once” to “2–5 times” to “More Frequently.” One might group users into high and low expertise and then explain performance differences between groups, or explain an interaction as taking place in one group or another, or use expertise in a regression. However, even someone who frequently uses a VR system might be competent in a narrow domain. Perhaps they play a particular game on Meta Quest. They still might be unfamiliar with other controllers or different types of content. Thus, we suggest that the community needs to develop a standardised questionnaire that explores competence and literacy.

4.5 Transfer of competence and literacy

In one of the examples in Section 3 we mentioned that a user might have experience with gaming and thus be broadly familiar with the buttons and controls on the typical consumer VR controllers. One interesting discussion is what other competences and literacies might impact a user’s ability in immersive systems. Certainly, some level of computer literacy would help the user understand how non-diegetic menus (see Section 3.2.1) might work. Digital or games literacy would help the user understand certain media forms, such as level selection, goal prioritisation, skill trees, etc. Users with higher physical literacy might be more experienced with exploring the environment and the function of their own body and avatar representation within the environment.

A second area to explore is the transfer of domain task knowledge into the immersive experience. While it is a relatively common assumption that training inside an immersive system can transfer to the real world [e.g., (Michalski et al., 2019) study table tennis training], users can presumably use knowledge of real tasks and bring this into the immersive system. Of course, if a task mostly involves procedural knowledge, then even if the immersive simulation is not an exact match in physical actions, the user should be able to interpret the steps required and then find the equivalent actions in the immersive system. Again we note the tension between diegetic and non-diegetic styles of interfaces. With diegetic, there is a limitation of how true to real-world the simulation can be. With non-diegetic immersive interfaces, standard issues of usability engineering come into play. We note that usability engineering for users with expert domain knowledge is a known problem. For example, Chilana et al. (2010) note that experienced users tend to find more serious faults in applications. Thus, while

domain experts might transfer skills, this would still require careful design of the application. Subtle differences might lead to quite different behaviours [e.g., see (Ioannidis et al., 2021) for a discussion of the performance of expert surgeons on a surgical simulator].

4.6 Reactions as if real

One common demonstration of the power of VR is the reaction to a visual cliff, or what is sometimes known as “The Pit” demo [e.g., (Slater et al., 1995; Usoh et al., 1999)]. Here, the user sees themselves on the edge of a virtual drop. User reactions can be quite strong including exclamations, visible nervousness, physiological changes, and immediate retreat from the edge. However, either prior exposure to a variant of The Pit, or prior knowledge about Pit-like effects could diminish the response. Thus, we could ask questions about whether behavioural presence responses diminish over time, what is the trajectory of this change, and what implications this has for the design of experiences.

One might expect that a person who sees The Pit illusion for the first time and is naïve, might react strongly and chalk this up as an interesting experience, but be more prepared for it the next time. But what happens over time when they realise that the drop is not real and thus they do not need to react as if it is real?

The authors’ personal experience is that The Pit can still have an impact on experienced users. Despite seeing dozens of variants of The Pit over the years, one of the authors still has a strong reaction to the drop, partly because he suffers from mild vertigo. We expect that there would still be a measurable response in biosignals that are often used to measure responses to stressful situations [e.g., Meehan et al. were the first to measure physiological responses to The Pit (Meehan et al., 2002)]. However the behavioural response might be more subtle. The same author recounts that if they take part in a demonstration with a visual cliff, they are still wary of the visual cliff because in a number of previous demonstrations they have virtually fallen (i.e., the experience has simulated plummeting downwards). They thus assume that if they want to navigate quickly around the environment, they might as well avoid the pit. If forced to cross a pit, then they will do so without hesitation, because this must be the way forward in the experience.

Thus, literacy with the ability of immersive systems to generate this response allows users to deal with them in ways that are different to what might be expected from a similar or analogous situation in the real world. The perceptual illusion might persist, but its effects are controlled by other cognitive processes. There is a lot of interesting work to do on how responses change over time, whether literacy makes users more or less likely to respond to novel situations, and whether there is a baseline of physiological response that can be relied upon.

Finally, the authors have encountered colleagues who are particularly susceptible to presence illusion, for better or worse. Most of these colleagues would be competent with the technology, so perhaps they are good at suspending disbelief. Perhaps they have knowledge that explains small implausibilities so they can focus on the overall story. This seems amenable to investigation with questionnaires and interviews and we note that the Immersive Tendencies Questionnaire (Witmer and Singer, 1998) probes somewhat in this direction, though it predates these new models of presence.

4.7 Ecological validity

A strongly related concept is the use of immersive technologies to explore user behaviour. A number of papers have had success with, or have advocated for use of, immersive technologies in a variety of sciences, including social science (Pan and Hamilton, 2018) and neuroscience (de Gelder et al., 2018). We do not want to just caution that more literate users might not treat the scenario as real, however it is worth noting the analogy to one author’s reaction to The Pit: users might not believe the scenario, but they understand it and thus play along, as a sort of make-believe. If we imagine a training scenario, we have no doubt that being immersed would likely be more powerful than, say, reading or watching a movie of a similar scenario. This might be because it affords natural interactions with the objects and other participants in the scenario. But can we tell the difference between a user who has suspended their disbelief and one who is playing along because that is what is expected?

4.8 Plausibility

This directly leads into a very recent discussion about the relationship between coherence, congruence, plausibility and the sense of presence. Plausibility Illusion (Psi) was introduced by Slater as being distinct from Place Illusion (PI) (Slater, 2009; Slater et al., 2022). PI was more related to the perception of being in a place, whereas Psi was related to believing that the scenario was really happening. Thus, someone who was experiencing Psi would react in an appropriate way, given a common or expected understanding about what appropriate means based on prior experience (i.e., mostly real-world experiences).

In the past few years there has been a lot of debate about how to support Psi. Under the control of the developers are the level of visual complexity, animation complexity, scripting, etc. It would be useful to know if these require certain quality levels to support users believing the scenario or not. Most current immersive simulations are not photorealistic, but still support Psi. A compelling demonstration of this is Pan et al. (2016)’s demonstration of a doctor’s responses to antibiotic requests from simulated patients. Certainly, any application needs to have some sort of internal coherence (Skarbez et al., 2021). Latoschik and Wienrich (2022) recently argued that there should be congruence between processed and expected information on the sensory, perceptual, and cognitive layers. These authors note that coherence and congruence are partly due to properties of the environment, but also due to individual differences in users.

We observe that plausibility will depend on literacy with different types of content. If one has seen wild, futuristic environments, or even simple, realistic environments without simulated gravity, neither is immediately implausible the next time one encounters them. Thus, the preconditions of plausibility for a user will depend on their literacy about immersive systems.

4.9 Training competence and literacy

A final set of discussions revolves around how we can support the development of competence and literacy. We note that

immersive technologies are evolving rapidly, and no doubt even experienced users will need to learn new system capabilities as they develop. For example, at the time of writing, face tracking is available as an attachment for some HMDs (e.g., the Vive Facial Tracker¹⁵) for the committed hobbyist, but there is relatively little software support. Face tracking is also now available as a feature on the Meta Quest Pro released in late 2022, however, this HMD is priced and marketed a level above consumer-level devices. We expect that there will be some skill involved in using face tracking effectively and ensuring it is producing the desired face shapes.

We believe that competence and literacy deserve a much fuller discussion within the community. For example, we have not established key dimensions of competence, as prior works in other domains have done (see Section 4.5). It is perhaps too early for this, but it is a discussion worth launching. Literacy will perhaps require more work to develop. There is significant work to do to curate and distil key genres and forms from the very broad range of experiences and reflections on experiences. We highlight general work on very general types of user experience such as Pine and Gilmore (1998) that classifies experience across active/passive and absorbing/immersive (with the latter term used in a different way to us), as starting points for classification.

There are, however, some promising analyses of sub-domains. The survey of non-fiction 360-degree video by Bevan et al. (2019) highlights several emerging features of such media. The essay of Steed et al. (2021) reflects on how modern immersive VR systems are a strict subset of the types of interfaces found in the broader 3D user interface field. Di Luca et al. (2021) do a very broad analysis of the existing types of locomotion techniques, knowledge of which would help build literacy and preparedness to better approach new experiences (see also Section 3.2).

One activity that would be interesting and useful, though undoubtedly controversial, would be to curate lists or multiple lists of content that we think are particularly insightful to build competence and literacy. For example, one of the authors is very keen that his students try not just the top-ten VR games of the time, but content such as *Virtual Virtual Reality*¹⁶ and *The Under Presents*¹⁷ as these mix interesting interaction and different types of storytelling. They also both poke fun at the medium. There is certainly a role for compelling and interesting content in motivating users to engage further with the media, perhaps first by developing competence in exploring the media, and then in literacy by exploring the potential of the medium.

Finally, we note that the first and early experiences of users are very important for their engagement. In some of the situations that users first encounter immersive experiences, the hosts are motivated to give a good experience to attract customers, but the length of time is limited (see Section 4.2). We would note that as designers of experiments and installations, we have a responsibility to naïve users that we do not give them a bad experience with the technology. Not that we can necessarily make our experiences compelling (experiments are often

repetitive and relatively dull), but those experiences should not be discomforting or off-putting.

5 Conclusion

In this paper we have started to outline immersive competence and immersive literacy as distinct concepts. Immersive competence describes the practical ability to operate immersive systems and understand the range of ways interfaces operate in immersive systems. Immersive literacy encapsulates knowledge not only about the range of systems and content, but the potentially unique properties of user experiences within such systems, such as knowledge about presence and embodiment. We have argued that, while immersive displays might be considered as falling under the scope of digital literacy or media literacy, they deserve to be studied as independent topics.

As noted in the discussion, we believe the descriptions of the scope of immersive competence and immersive literacy that we have given in Section 3 and Section 4 need further development. We look forward to engaging in a community discussion about how to flesh these out and turn them into descriptive and training materials. We are also very interested in how we as a community can support individuals in becoming more competent and literate, and how we can assess competence in ways that support research and design processes.

Finally, our discussion has drawn out some challenges for researchers and developers in this field. Developing immersive systems is not just challenging because we need to support users with very different levels of competence, but also we can increasingly expect users to be aware of some of the areas of immersive literacy. If the user is very aware of concepts such as presence or the uncanny valley, do they engage in the same way as a more naïve user? As users become more literate, researchers need to understand the impact of this, so that we can enable the medium to flourish.

Author contributions

AS coordinated the paper. All authors contributed to discussion and writing. The discussion originated in a reading group at UCL. RL was visiting UCL for some of that discussion. All authors contributed to the article and approved the submitted version.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

15 <https://www.vive.com/us/accessory/facial-tracker/>.

16 <https://tenderclaws.com/vvr>.

17 <https://tenderclaws.com/theunderpresents>.

References

- Bailenson, J. (2018). *Experience on demand: What virtual reality is, how it works, and what it can do*. W. W. Norton and Company. Google-Books-ID: 2fkqDwAAQBAJ.
- Banakou, D., Groten, R., and Slater, M. (2013). Illusory ownership of a virtual child body causes overestimation of object sizes and implicit attitude changes. *Proc. Natl. Acad. Sci.* 110, 12846–12851. doi:10.1073/pnas.1306779110
- Bawden, D. (2001). Information and digital literacies: A review of concepts. *J. Document.* 57, 218–259. doi:10.1108/EUM0000000007083
- BBC Virtual Reality (2019). Making VR a reality: Storytelling and audience insights 2019. Tech. rep. Available at: http://downloads.bbc.co.uk/aboutthebbc/internetblog/bbc_vr_storytelling.pdf.
- Bevan, C., Green, D. P., Farmer, H., Rose, M., Cater, K., Stanton Fraser, D., et al. (2019). “Behind the curtain of the “ultimate empathy machine”: On the composition of virtual reality nonfiction experiences,” in Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (New York, NY, USA: Association for Computing Machinery), 1–12. CHI '19. doi:10.1145/3290605.3300736
- Blanchard, C., Burgess, S., Harvill, Y., Lanier, J., Lasko, A., Oberman, M., et al. (1990). Reality built for two: A virtual reality tool. *ACM SIGGRAPH Comput. Graph.* 24, 35–36. doi:10.1145/91394.91409
- Bolas, M., Jones, J. A., McDowall, I., and Suma, E. (2014). *Dynamic field of view throttling as a means of improving user experience in head mounted virtual environments*. University of Southern California USC, US9645395B2.
- Botvinick, M., and Cohen, J. (1998). Rubber hands ‘feel’ touch that eyes see. *Nature* 391, 756. doi:10.1038/35784
- Brooks, F. (1999). What’s real about virtual reality? *IEEE Comput. Graph. Appl.* 19, 16–27. doi:10.1109/38.799723
- Buckingham, D., and Burn, A. (2007). Game literacy in theory and practice. *J. Educ. Multimedia Hypermedia* 16, 323–349.
- Chilana, P. K., Wobbrock, J. O., and Ko, A. J. (2010). “Understanding usability practices in complex domains,” in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (New York, NY, USA: Association for Computing Machinery), CHI '10, 2337–2346. doi:10.1145/1753326.1753678
- Cruz-Neira, C., Sandin, D., and DeFanti, T. (1993). “Surround-screen projection-based virtual reality: The design and implementation of the CAVE,” in SIGGRAPH 93 conference proceedings, *Computer graphics Annual conference series* (New York: ACM)
- de Gelder, B., Kätysyri, J., and de Borst, A. W. (2018). Virtual reality and the new psychophysics. *Br. J. Psychol.* 109, 421–426. doi:10.1111/bjop.12308
- Delaney, B. (2017). *Virtual reality 1.0 – the 90’s: The birth of VR in the pages of CyberEdge journal*. OCLC: 985344386. CyberEdge Information Services. Place of publication not identified.
- Dewez, D., Hoyet, L., Lécuyer, A., and Argelaguet Sanz, F. (2021). “Towards “avatar-friendly” 3D manipulation techniques: Bridging the gap between sense of embodiment and interaction in virtual reality,” in Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (Yokohama Japan: ACM), 1–14. doi:10.1145/3411764.3445379
- Di Luca, M., Seifi, H., Egan, S., and Gonzalez-Franco, M. (2021). “Locomotion vault: The extra mile in analyzing VR locomotion techniques,” in Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (New York, NY, USA: Association for Computing Machinery), CHI '21, 1–10. doi:10.1145/3411764.3445319
- Emerson, D., Tucker, A., and Badu, A. C. (2019). *Taking VR stories to UK audiences: Common ground, a touring report*. Available at: <https://www.bfi.org.uk/industry-data-insights/reports/taking-vr-stories-uk-a-udiences-common-ground-uk-touring-report>.
- Grau, O., Malina, R. F., and Cubitt, S. (2004). *Virtual art: From illusion to immersion* (Cambridge, Mass, London: MIT Press), first edition edn.
- Grimsdale, C. (1991). “PROvision-parallel processing meets virtual reality” in IEE Colloquium on Real World Visualisation - Virtual World - Virtual Reality. 11/1–11/2.
- Hartmann, T., and Hofer, M. (2022). I know it is not real (and that matters) media awareness vs. Presence in a parallel processing account of the VR experience. *Front. Virtual Real.* 3. doi:10.3389/frvir.2022.694048
- Heeter, C. (1992). *Being there: The subjective experience of presence*. *Presence Teleoperators Virtual Environ.* 1, 262–271. doi:10.1162/pres.1992.1.2.262
- Held, R. M., and Durlach, N. I. (1992). Telepresence. *Presence Teleoperators Virtual Environ.* 1, 109–112. doi:10.1162/pres.1992.1.1.109
- Horton, F. W. (1983). Information literacy vs. Computer literacy. *Bull. Am. Soc. Inf. Sci.* 9, 14–16.
- Meta Inc. (2022). *VR locomotion design guide*. Oculus Developers.
- Ioannidis, P., Eklund, L., and Lovlie, A. S. (2021). We dare you: A lifecycle study of a substitutable reality installation in a museum space. *J. Comput. Cult. Herit.* 14, 34:1–34:21. doi:10.1145/3439862
- Jerald, J. (2015). *The VR book*. Association for Computing Machinery. doi:10.1145/2792790
- Kalawsky, R. (1993). *The science of virtual reality and virtual environments*. 1st edn. USA: Addison-Wesley Longman Publishing Co., Inc.
- Kennedy, R. S., Lane, N. E., Berbaum, K. S., and Lilienthal, M. G. (1993). Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *Int. J. Aviat. Psychol.* 3, 203–220. doi:10.1207/s15327108ijap0303_3
- Kent, S. (2010). *The ultimate history of video games*. Crown/Archetype.
- Kilteni, K., Groten, R., and Slater, M. (2012). The sense of embodiment in virtual reality. *Presence Teleoperators Virtual Environ.* 21, 373–387. doi:10.1162/PRES_a_00124
- Kilteni, K., Bergstrom, I., and Slater, M. (2013). Drumming in immersive virtual reality: The body shapes the way we play. *Vis. Comput. Graph. IEEE Trans.* 19, 597–605. doi:10.1109/tvcg.2013.29
- Lanier, J. (2017). *Dawn of the new everything: A journey through virtual reality*. Random House. Google-Books-ID: k8fIDAAQBAJ.
- Latoschik, M. E., and Wienrich, C. (2022). Congruence and plausibility, not presence: Pivotal conditions for XR experiences and effects, a novel approach. *Front. Virtual Real.* 3. doi:10.3389/frvir.2022.694433
- Laurel, B. (2014). *Computers as theatre*. 2nd edn. Addison-Wesley. Google-Books-ID: kfEEAQAQBAJ.
- Lee, J., Kim, J., and Choi, J. Y. (2019). The adoption of virtual reality devices: The technology acceptance model integrating enjoyment, social interaction, and strength of the social ties. *Telematics Inf.* 39, 37–48. doi:10.1016/j.tele.2018.12.006
- Lombard, M., and Ditton, T. (1997). At the heart of it all: The concept of presence. *J. Comput. Mediated Commun.* 3, 0. doi:10.1111/j.1083-6101.1997.tb00072.x
- Long, D., and Magerko, B. (2020). “What is AI literacy? Competencies and design considerations,” in Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (New York, NY, USA: Association for Computing Machinery), CHI '20, 1–16. doi:10.1145/3313831.3376727
- MacQuarrie, A., and Steed, A. (2017). “Cinematic virtual reality: Evaluating the effect of display type on the viewing experience for panoramic video,” in 2017 IEEE Virtual Reality (VR), 45–54. doi:10.1109/VR.2017.7892230
- Maister, L., Slater, M., Sanchez-Vives, M. V., and Tsakiris, M. (2015). Changing bodies changes minds: Owning another body affects social cognition. *Trends Cognit. Sci.* 19, 6–12. doi:10.1016/j.tics.2014.11.001
- Manis, K. T., and Choi, D. (2019). The virtual reality hardware acceptance model (VR-HAM): Extending and individualizing the technology acceptance model (TAM) for virtual reality hardware. *J. Bus. Res.* 100, 503–513. doi:10.1016/j.jbusres.2018.10.021
- Meehan, M., Insko, B., Whitton, M., and Brooks, F. P., Jr. (2002). Physiological measures of presence in stressful virtual environments. *Acm Trans. Graph.* 21, 645–652. doi:10.1145/566654.566630
- Michalski, S. C., Szpalk, A., Saredakis, D., Ross, T. J., Billingham, M., and Loetscher, T. (2019). Getting your game on: Using virtual reality to improve real table tennis skills. *PLOS ONE* 14, e0222351. doi:10.1371/journal.pone.0222351
- Mine, M. (2003). “Towards virtual reality for the masses: 10 years of research at Disney’s VR studio,” in Proceedings of the workshop on Virtual environments 2003 (New York, NY, USA: Association for Computing Machinery), EGVE '03, 11–17. doi:10.1145/769953.769955
- Mori, M., MacDorman, K. F., and Kageki, N. (2012). The uncanny valley [from the field]. *IEEE Rob. Autom. Mag.* 19, 98–100. doi:10.1109/MRA.2012.2192811
- Oakes, J. (2018). “VR demos: Best practices so people won’t hate you,” in XR Developer Conference 2018. Available at: <https://www.gdcvault.com/play/1025620/VR-Demos-Best-Practices-So>.
- Pan, X., and Hamilton, A. F. d. C. (2018). Why and how to use virtual reality to study human social interaction: The challenges of exploring a new research landscape. *Br. J. Psychol.* 109, 395–417. doi:10.1111/bjop.12290
- Pan, Y., and Steed, A. (2019). “Avatar type affects performance of cognitive tasks in virtual reality,” in 25th ACM Symposium on Virtual Reality Software and Technology (Parramatta NSW Australia: ACM), 1–4. doi:10.1145/3359996.3364270
- Pan, X., Slater, M., Beacco, A., Navarro, X., Rivas, A. I. B., Swapp, D., et al. (2016). The responses of medical general practitioners to unreasonable patient demand for antibiotics - a study of medical ethics using immersive virtual reality. *PLOS ONE* 11, e0146837. doi:10.1371/journal.pone.0146837
- Peck, T. C., Seinfeld, S., Aglioti, S. M., and Slater, M. (2013). Putting yourself in the skin of a black avatar reduces implicit racial bias. *Conscious. Cognition* 22, 779–787. doi:10.1016/j.concog.2013.04.016
- Pine, B. J., and Gilmore, J. H. (1998). Welcome to the experience economy. *Harv. Bus. Rev.* 76, 97–105.
- Potter, W. J. (2010). The state of media literacy. *J. Broadcast. Electron. Media* 54, 675–696. doi:10.1080/08838151.2011.521462

- Poupyrev, I., Billinghurst, M., Weghorst, S., and Ichikawa, T. (1996). "The go-go interaction technique: Non-linear mapping for direct manipulation in VR," in Proceedings of the 9th annual ACM symposium on User interface software and technology - UIST '96 (Seattle, Washington, United States: ACM Press), 79–80. doi:10.1145/237091.237102
- Rheingold, H. (1992). *Virtual reality: The revolutionary technology of computer-generated artificial worlds-and how it promises to transform society*. Simon and Schuster. Google-Books-ID: 6ZkoAQAAAMAJ.
- Sagnier, C., Loup-Escande, E., Lourdeaux, D., Thouvenin, I., and Valléry, G. (2020). User acceptance of virtual reality: An extended technology acceptance model. *Int. J. Human-Computer Interact.* 36, 993–1007. doi:10.1080/10447318.2019.1708612
- Saredakis, D., Szpak, A., Birkhead, B., Keage, H. A. D., Rizzo, A., and Loetscher, T. (2020). Factors associated with virtual reality sickness in head-mounted displays: A systematic review and meta-analysis. *Front. Hum. Neurosci.* 14, 96. doi:10.3389/fnhum.2020.00096
- Sherman, W. R., and Craig, A. B. (1995). Literacy in virtual reality: A new medium. *ACM SIGGRAPH Comput. Graph.* 29, 37–42. doi:10.1145/216876.216887
- Sherman, W. R., and Craig, A. B. (2003). "Understanding virtual reality: Interface, application, and design," in *Morgan Kaufmann series in computer graphics and geometric modeling*. 1st edn (Amsterdam ; Boston: Morgan Kaufmann Publishers).
- Silverblatt, A., Ferry, J., and Finan, B. (2015). *Approaches to media literacy: A handbook*. 2 edn. New York: Routledge. doi:10.4324/9781315706344
- Skarbez, R., Brooks, F. P., Jr., and Whitton, M. C. (2017). A survey of presence and related concepts. *ACM Comput. Surv.* 50, 1:1–96:39. doi:10.1145/3134301
- Skarbez, R., Brooks, F. P., and Whitton, M. C. (2021). Immersion and coherence: Research agenda and early results. *IEEE Trans. Vis. Comput. Graph.* 27, 3839–3850. doi:10.1109/TVCG.2020.2983701
- Slater, M., and Steed, A. (2000). A virtual presence counter. *Presence* 9, 413–434. doi:10.1162/105474600566925
- Slater, M., Usoh, M., and Steed, A. (1995). Taking steps: The influence of a walking technique on presence in virtual reality. *ACM Trans. Comput. Human Interact.* 2, 201–219. doi:10.1145/210079.210084
- Slater, M., Steed, A., McCarthy, J., and Maringelli, F. (1998). The influence of body movement on subjective presence in virtual environments. *Hum. Factors J. Hum. Factors Ergonomics Soc.* 40, 469–477. doi:10.1518/001872098779591368
- Slater, M., Banakou, D., Beacco, A., Gallego, J., Macia-Varela, F., and Oliva, R. (2022). A separate reality: An update on place illusion and plausibility in virtual reality. *Front. Virtual Real.* 3, 914392. doi:10.3389/frvir.2022.914392
- Slater, M. (2009). Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philos. Trans. R. Soc. B Biol. Sci.* 364, 3549–3557. doi:10.1098/rstb.2009.0138
- Spante, M., Hashemi, S. S., Lundin, M., and Algers, A. (2018). Digital competence and digital literacy in higher education research: Systematic review of concept use. *Cogent Educ.* 5, 1519143. doi:10.1080/2331186X.2018.1519143
- Steed, A., Pan, Y., Watson, Z., and Slater, M. (2018). "We Wait"—the impact of character responsiveness and self-embodiment on presence and interest in an immersive news experience. *Front. Robotics AI* 5, 112. doi:10.3389/frobt.2018.00112
- Steed, A., Takala, T. M., Archer, D., Lages, W., and Lindeman, R. W. (2021). Directions for 3D user interface research from consumer VR games. *IEEE Trans. Vis. Comput. Graph.* 27, 4171–4182. doi:10.1109/TVCG.2021.3106431
- Stephoe, W., Steed, A., and Slater, M. (2013). Human tails: Ownership and control of extended humanoid avatars. *Vis. Comput. Graph. IEEE Trans.* 19, 583–590. doi:10.1109/tvcg.2013.32
- Sutherland, I. (1968). "A head-mounted three dimensional display" in Proceedings of the December 9-11, 1968, fall joint computer conference, part I. 757–764.
- Tennent, P., Martindale, S., Benford, S., Darzentas, D., Brundell, P., and Collishaw, M. (2020). Thresholds: Embedding virtual reality in the museum. *J. Comput. Cult. Herit.* 13, 1:1–12:35. doi:10.1145/3369394
- Usoh, M., Arthur, K., Whitton, M. C., Bastos, R., Steed, A., Slater, M., et al. (1999). "Walking > walking-in-place > flying, in virtual environments," in Proceedings of the 26th annual conference on Computer graphics and interactive techniques - SIGGRAPH '99 (ACM Press), 359–364. doi:10.1145/311535.311589
- Vuorikari, R., Kluzer, S., and Punie, Y. (2022). *DigComp 2.2: The Digital Competence Framework for Citizens-With new examples of knowledge, skills and attitudes*. Tech. Rep. EUR 31006 EN. Luxembourg: Publications Office of the European Union.
- Watson, Z. (2019). *BBC shares insights into great virtual reality storytelling*. Available at: <https://medium.com/bbc-product-technology/bbc-shares-insights-into-great-virtual-reality-storytelling-e5e721dae08e>.
- Whitehead, M. (2001). The concept of physical literacy. *Eur. J. Phys. Educ.* 6, 127–138. doi:10.1080/1740898010060205
- Witmer, B. G., and Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence Teleoperators Virtual Environ.* 7, 225–240. doi:10.1162/105474698565686
- Won, A. S., Bailenson, J., Lee, J., and Lanier, J. (2015). Homuncular flexibility in virtual reality. *J. Computer-Mediated Commun.* 20, 241–259. doi:10.1111/jcc4.12107
- Yuan, Y., and Steed, A. (2010). "Is the rubber hand illusion induced by immersive virtual reality?" in 2010 IEEE Virtual Reality Conference (VR) (IEEE), 95–102.
- Zhao, Y., Pinto Llorente, A. M., and Sánchez Gómez, M. C. (2021). Digital competence in higher education research: A systematic literature review. *Comput. Educ.* 168, 104212. doi:10.1016/j.compedu.2021.104212
- Zlotowski, J., Sumioka, H., Nishio, S., Glas, D., Bartneck, C., and Ishiguro, H. (2015). Persistence of the uncanny valley: The influence of repeated interactions and a robot's attitude on its perception. *Front. Psychol.* 6, 883. doi:10.3389/fpsyg.2015.00883
- Zuckerberg, M. (2021). *Founder's letter*. Available at: <https://about.fb.com/news/2021/10/founders-letter/>.