This is the accepted manuscript of the following article: Li, C., & Ip, H. H. S. (2022). Defining virtual reality enabled learning. International Journal of Innovation and Learning, 31(3), 291-306, which has been published in final form at https://doi.org/10.1504/IJIL.2022.122128.

Defining Virtual Reality enabled Learning

Chen Li*

Department of Computing,
The Hong Kong Polytechnic University,
Hung Hom, Kowloon
Hong Kong SAR
E-mail: richard-chen.li@polyu.edu.hk
*Corresponding author

Horace Ho Shing Ip

Centre for Innovative Applications of Internet and Multimedia Technologies (AIMtech Centre) & the Department of Computer Science
City University of Hong Kong,
Kowloon Tong, Kowloon
Hong Kong SAR
E-mail: horace.ip@cityu.edu.hk

Abstract: Virtual reality (VR), as a set of Human-computer Interaction (HCI) technologies, allows the creation of computer-generated virtual environments, in which users are free to interact with the environments, virtual objects, agents, or even other users. Previous empirical studies have demonstrated the feasibility and effectiveness of using VR for education. However, the ambiguity in the definition has caused both theoretical and practical obstacles in this research area. This article aims to address this issue by considering the interdisciplinary nature of this research area. Specifically, the term VR-enabled learning is coined here, and the definition of VR-enabled learning is decomposed into five dimensions of concerns, namely immersion, presence, pedagogy, intended learning outcomes, and learner specifics. These five dimensions of concerns should not only help to address the ambiguity in the definition but also serve as a framework to guide the planning and practising of future research studies in VR-enabled learning.

Keywords: virtual reality, immersion, presence, pedagogy, learning outcome, learner specifics.

Reference to this paper should be made as follows: [author info]. (2021) 'Defining Virtual Reality enabled Learning', *Int. J. Innovation and Learning*,

Biographical notes: Chen Li received his BSc in computer science and technology from Nanjing University in 2008, MSc and PhD in computer science from City University of Hong Kong in 2011 and 2018, respectively. He is currently a research assistant professor of the Department of Computing, The Hong Kong Polytechnic University. He has developed many virtual reality applications for education, visual art, architecture design, and cultural heritage preservation. His research interests include virtual reality, augmented reality, human-computer interaction, learning analytics, and innovative technologies

[author info]

for education.

Horace Ho Shing Ip is Lee Shau Kee Professor of e-Learning, Chair Professor of Computer Science, and the founding director of the AIMtech Centre (Centre for Innovative Applications of Internet and Multimedia Technologies) at City University of Hong Kong. He graduated from University College London (UCL), UK, with a BSc(Hons) and PhD degrees in 1980 and 1983 respectively. His research interests include interactive multimedia, pattern recognition, virtual environment and e-Learning technologies. He has received regional and international awards for his work including a Gold Medal from the Salon International Des Inventions, Geneva; Prix Ars Electronica, Austria; and the Asia-Pacific Information and Communications Technology Award. In 2007, Prof. Ip was recognized as one of the world's distinguished innovators by TIME magazine. He is a Fellow of the Hong Kong Institution of Engineers (HKIE), a Fellow of the Institution of Engineering and Technology (IET) in the United Kingdom, a Fellow of the British Computer Society. Prof. Ip was awarded in 2004 the prestigious Fellow of the International Association for Pattern Recognition (IAPR) for contributions to the field of Pattern Recognition. He has published over 300 papers in international journals and conference proceedings.

1 Introduction

Ever since Jaron Lanier coined the terminology "virtual reality (VR)" to describe a computer-generated virtual world in the late 1980s, people have dreamed about the infinite number of application domains that could benefit from VR. However, in the early days of VR, the applications were mainly restricted to training pilots, astronauts, and soldiers. This is mainly due to the complexity of the software and hardware system causing high cost in both development and maintenance, as well as the concerns over system usability and system robustness (Sutcliffe & Kaur, 2000; Virvou & Katsionis, 2008). The Virtual Interface Environment Workstation (VIEW) developed and used by the National Aeronautics and Space Administration (1990) at that time closely resembles the VR systems as we know today but costed a fortune to develop and maintain. The VIEW consists of a head-mounted stereoscopic display (HMD) which displays computer-rendered images or video streams from remote cameras, a pair of data gloves that track user's hand and finger movements allowing real-time interaction, and other peripherals that support the functioning of the system and applications running on top of the system.

As the enabling technologies getting more and more accessible and the concerns regarding usability being gradually addressed, the use of VR for educational purposes has become a new trend in both K-12 and higher education (Virvou & Katsionis, 2008). Discussed in prior reviews, there are a few advantages of using VR for educational purposes over conventional means of learning. First, VR provides more freedom to implement pedagogical approaches that might be difficult to implement in the physical world. Luo et al. (2021) observed and reported in their recent systematic review covering 157 empirical studies that more and more recent studies employed the pedagogical approaches, which were dedicated to VR-enabled learning. The authors also suggested that with these pedagogical approaches being implemented, VR could be more suitable for teaching abstract concepts or procedural knowledge via authentic problem solving, as well as changing learners' affect and attitude. Merchant et al. (2014) systematically reviewed 67 articles reporting empirical studies that used VR for K-12 and higher education. The meta-analysis showed the potential of using VR for education and presented its numerous advantages, including the ease of implementing certain pedagogical approaches, but also called for future studies to *"test more design variables and interesting interaction effects of design features"* (Merchant et al., 2014, p. 37). Second, the use of VR for educational purposes provides new ways for assessment. During VR exposure, a lot of interactive data can be captured and such data, with proper handling and processing, can be quite valuable for learning assessment. Besides, Chang et al. (2020) also investigated the integration of the peer assessment approach with a VR design system of learning earth science. This can also be a new direction where VR can contribute to learning assessment. Lastly, the nature of VR opens new possibilities in distance learning. VR as a multimodal interface is expected to be able to better connect learners with their peers and with educators. The use of connected and collaborative virtual worlds, such as SecondLife, for distance learning in both K-12 and higher education has demonstrated its feasibility and effectiveness (Warburton, 2009; Inman et al., 2010; Martín-Gutiérrez et al., 2017), although educators might find the function-packed platform difficult to use given its interface design (Inman et al., 2010). The above-mentioned advantages of using VR for educational purposes, compared to conventional means of learning, should be more prominent, given the current global pandemic and the urgent need for technology-assisted learning. However, previous review articles on this area of research often used different definitions of using VR for educational purposes. For example, in early review articles, the learning experience delivered by desktop computers was included in the analysis and discussion, such as in the work of Mikropoulos & Natsis (2011) and Merchant et al. (2014). The findings may no longer applicable as discussed by Luo et al. (2021), because the more recent experience is often delivered by HMDs or Cave Automatic Virtual Environment (CAVE) (Cruz-Neira et al., 1992); the advancements in relevant technologies can significantly affect the experience (Schubert et al., 2001), thus may impact the learning outcomes. Moreover, the use of video games for educational purposes and the gamification design in educational technology system generated even more confusion in this area of research, which will be discussed in section 2.

In this paper, the use of VR for educational purposes is first coined as the term - *"virtual reality enabled learning"*. This term is believed to be more precise and concise for describing the relation between technology and learning - VR and related technologies enable learning. Based on this term, five major dimensions of concerns regarding the definition of VR-enabled learning are discussed. The five major dimensions of concerns are immersion, presence, pedagogy, intended learning outcomes, and learner specifics; the first two dimensions of concerns focus on the technological and psychological aspects of VR experience, while the rest three focus on learning. Corresponding guidelines are proposed under each dimension, which can be summaries as a more practical, more structured, and less ambiguous definition of VR-enabled learning.

2 Defining Virtual Reality enabled Learning

To define VR-enabled learning, the first step is to clearly define VR itself. Unfortunately, the definition of VR is quite diverse as well. This might be due to the interdisciplinary nature of VR research, since many of the terminologies were borrowed from other domains of research, such as psychology and cognitive science (Skarbez et al., 2017). The two most accepted definitions of VR come from two different but interconnected perspectives - the enabling technology and the human experience.

Brooks (1999) defined VR from the perspective of the enabling technology. The author listed four crucial and four auxiliary technologies that enabled the experience of being *"immersed in a responsive virtual world"* (Brooks, 1999, p. 16). Specifically, the four crucial technologies concerned with the delivery of the visual stimuli according to the position and orientation of the user's head and limbs, while the four auxiliary technologies mostly concerned with the delivery of the auditory stimuli and the enabling of user-VR interactions. This definition is mainly based on the capability of the enabling technologies, which is closely related to the concept of immersion as defined by Slater (1998). Most of the off-the-shelf HMDs and highly customised CAVE-like system nowadays are equipped with both crucial and auxiliary technologies. However, based on this definition, desktop VR (a.k.a., non-immersive VR) (Robertson et al., 1993), which displays interactive computer-rendered images on a conventional monitor, should be out of the discussion in this paper, although its name clearly contains VR.

On the other hand, Steuer (1992) defined VR from the perspective of human experience. The author suggested that "the key to defining VR in terms of human experience rather than technological hardware is the concept of presence" and defining VR solely by the technological capabilities was lacking in providing insights on the "processes or effects" of using VR. In the contexts of using VR for educational purposes, Steuer's concern is exceptionally prominent - learning itself is a process and the effects of learning, such as gaining new knowledge or practising new skills, are the ultimate objective of learning. The two definitions are not contradictory but in fact complementary. We should admit that human experience is subjective thus difficult to measure. The four crucial and four auxiliary technologies listed in Brooks (1999) are short in considering the human involvement of the experience but are more objective and much easier to measure. Moreover, to deliver a better human experience requires the technologies to reach a certain capability threshold, such as the responsiveness of the system, which is well known to be highly correlated with the feeling of presence and cybersickness (Frank et al., 1988; Kennedy et al., 1993; Dinh et al., 1999; Cummings & Bailenson, 2016). Hence, defining VR-enabled learning relies on these two complementary perspectives. In this paper, immersion and presence are the two dimensions of concerns when defining VR-enabled learning; the former corresponds to the technical perspective of VR or virtual environment (VE) while the latter covers the fundamentals of human experience in VR.

On top of a clear definition of VR, a well-thought and practical definition of VR-enabled learning also should be able to distinguish this particular domain of application from others, such as the use of VR for entertainment. Although there is a lack of in-depth research on using commercial VR games directly for educational purposes, we can take video games as an example of discussion here; previous studies have found that commercial video games could be used for educational purposes, with the prerequisite of providing proper facilitation and support to learners (Coyne, 2003; Prensky, 2003; Charsky & Mims, 2008). Moreover, in the education research domain, gamification, a means that originated from the game design being adapted in non-game applications (Deterding et al., 2011), is widely used in educational technology system design to motivate learners and enhance their learning engagement (Simões et al., 2013). From this point of view, the boundary between commercial video games and the so-called interactive multimedia contents for education is blurry. However, just like learning cannot be simply defined by the characteristics of textbooks or the venue it takes place, VR-enabled learning cannot be simply defined by the nature of learning materials or the mediated learning environment to be used. This paper suggests defining VR-enabled learning as a unique type of learning experience. The definition suggested in this paper, besides the constructs related to VR, also concerns the pedagogy, the intended learning outcome, and the targeted learning specifics.

3 Dimensions of Concerns

In this section, five dimensions of concerns will be discussed. Based on the discussion, corresponding guidelines will be proposed to help formalise a more practical, more structured, and less ambiguous definition of VR-enabled learning.

3.1 Immersion

Immersion nowadays is often referred to as a set of objective and usually technological characteristics of a VE (Slater, 1999). In the early days of VR application and research, the use of the term immersion was quite ambiguous since the term was originally borrowed from media research. Witmer & Singer (1998) first defined immersion as *"a psychological state characterized by perceiving oneself to be enveloped by, included in, and interacting with an environment that provides a continuous stream of stimuli and experiences"* (Witmer & Singer, 1998, p. 227). This definition sometimes is also referred to as psychological immersion in later literatures, such as Lombard et al. (2000), to distinguish the definition of immersion by Witmer & Singer (1998) from the definition by Slater (1999). As discussed previously in section 2, the definition of immersion by Witmer & Singer (1998) is more subjective, more challenging to measure, and has a significant overlap with the concept of presence. Hence, this paper will follow the definition by Slater (1999) on the concept of immersion, which is used to objectively describe the technological capability of VEs. By doing so, an interesting outcome is that the immersion of some common VEs can now be compared objectively. For example, it is obvious that the immersion of HMD is higher than that of CAVE-like VEs, because CAVE-like VEs can be simulated in HMD but not the other way around. Similarly, L-Shape projection-based VEs can be simulated in CAVE-like VEs by simply disabling a few projection surfaces of the CAVE. If a cut-off threshold regarding the immersion of VEs could be set, the ambiguity in the definition of VR-enable learning could be greatly reduced, addressing the issues raised by Luo et al. (2021). However, the above comparison and the followed comparative definition overlooked the multimodal nature of VEs. The previous examples only take the VEs' capability in delivering visual stimulations into consideration; and the capabilities of VEs are apparently more than delivering stimulations and should also facilitate user-VE interactions or even multi-user interactions in VEs to deliver a fully mediated experience (Riva et al., 2003). Hence, considering the multimodal nature of VEs, in this paper, it is suggested to clearly report the technological capabilities of VEs used in future research studies by covering the details regarding the delivery of sensory stimulations, as well as the facilitations of interactions. To better guide practice and future systematic analysis, based on the work of Brooks (1999) and the latest development of the enabling technologies, a three-level taxonomy concerning the immersion of VR-enabled learning is proposed in Table 1 for referencing.

Table 1. A three-level taxonomy concerning the immersion of VR-enabled learning.

		Immersion		
		Low	Medium	High
Stimulation Fidelity	Visual	The experience is delivered through a non-stereoscopic display or a stereoscopic display covering a limited field of view (less than 120 degrees at normal experiencing position).	The experience is delivered through a stereoscopic display covering a large field of view (more than 120 degrees at normal experiencing position).	The experience is delivered through a stereoscopic display covering a large field of view and supporting viewing project adjustments based on real-time motion capturing data.
	Auditory	The experience is delivered through a standard stereo audio system with no spatial effects.	The experience is delivered through a standard stereo or multi-channel audio system with spatial effects.	The experience is delivered through a standard stereo or multi-channel audio system with spatial effects and other advanced auditory effects simulating the physics of the virtual world.
	Haptic	No haptic stimulation is delivered, or generic vibration is delivered.	Textural and/or force feedback haptics are delivered covering limited locations of the learner’s body (most likely, the two hands of the learner).	Textural and/or force feedback haptics are delivered covering much of the learner’s body.

Interactivity	Learner-VE	Learners can only interact with the VE using a conventional user interface or an enhanced user interface designed for the VE's input devices.	Learners can either interact with the VE via a conventional user interface or Natural User Interface (NUI).	Multimodal interactions between the learner and the VE are supported.
	Among peer learners	The experience is individual or nonsynchronous, so there is no or very limited real-time interactions among peer learners.	Peer learners can interact with each other using text.	Multimodal interactions among peer learners are supported.

3.2 Presence

Although it is widely accepted that presence is a complex construct concerning human experience during VR exposure, there is no widespread agreement on how to define and, more importantly, operationalise the construct (Skarbez et al., 2017). The most common interpretation of presence is the sense or feeling of *"being there"*, which was originally discussed by Steuer (1992). Slater (2009) further proposed that presence is composed of two components, Place Illusion (PI) and Plausibility Illusion (Psi); PI is defined as the illusion of *"being in a place in spite of the sure knowledge that you are not there"* (Slater, 2009, p. 3551), while Psi is defined as the illusion that *"what is apparently happening is really happening (even though you know for sure that it is not)"* (Slater, 2009, p. 3553). The former corresponds to the feeling of *"being there"* as discussed by Steuer (1992) and other conventional views of presence, while the latter sees presence from a new perspective; that is believing what you are experiencing in the VEs. This expansion to the conventional concept of presence is well accepted by the VR research communities, because of its two major advantages. First, PI is referred only to the illusion of being in a place and not to other illusions that were conventionally attributed to presence. Second, the introduction of Psi emphasizes the overall VR experience is not only about the illusion of being in a place (i.e., PI), but also how this multimodal interface behaves in a way that is genuinely plausible to users.

Presence can also be comparable, although it is not as objective as immersion. One can always ask a user to compare two VR experiences in terms of how present he/she feels. However, such a comparison might be challenging in practice. First, different users may simply interpret "how present" in various ways. We should admit that this question is extremely difficult to answer, given the discussion above and the nature of presence - it is an internal and subjective feeling. Second, the comparison between two VR experiences is only valid when the same user can access both experiences. In practice, this is an exceedingly rare case, because, for most of the time, the comparison happens between a newly developed experience and a previously studied experience. The latter may not be available in the same VEs where the former can be deployed. Luckily, a few instruments have been developed to measure the construct of presence (Usoh et al., 2000). Table 2 shows a list of widely administrated instruments for measuring presence. Note that some of the instruments were developed by following a slightly different definition of presence from what has been discussed in this paper.

Whether more presence is a good thing is still questionable nowadays. Previous research suggested that the linking between presence and task-specific performance in VR is not as strong as we thought (Witmer & Singer, 1998; Snow, 1998; Modjeska & Waterworth, 2000). Given the VR-enabled learning context of discussion, it is particularly important to investigate whether more presence can always yield better learning effectiveness. Currently, there are very few in-depth research studies on this question. Because presence is the core construct of VR, it is suggested to measure this construct in future research practices on VR-enabled learning. This will also help us better understand the complex interaction between presence and learning.

Table 2. A list of widely administrated instruments for measuring presence

Reference (Abbreviation)	Publication Year	Number of Item	Usage
Slater et al. (SUS)	1994	6	VE
Barfield & Hendrix	1995	6	VE
Baños et al.	1998	77	VE
Kim & Biocca	1997	8	VE
Witmer & Singer (WS)	1998	32	VE
Dinh et al.	1999	14	VE
Nichols et al.	2000	9	VE
Lessiter et al. (ITC-SOPI)	2001	44	Cross-Media
Schubert et al. (IPQ)	2001	14	VE
Nowak & Frank	2003	9	Shared VE
Lombard & Weinstein (TPI)	2009	4-8	Cross-Media

3.3 Pedagogy

Pedagogical approaches are often overlooked in this interdisciplinary research area. Conventional approaches to learning often focus on the knowledge transfer strategies that are centred on text- or dialogue-based engagements with learners. As discussed in section 3.1, the distinctive nature of VEs is multimodality, which provides more freedom to implement pedagogical approaches that might be difficult to implement in conventional means of learning (see Fig. 1). Although there is a new trend of implement pedagogical design dedicated to VR-enabled learning as observed and reported by Luo et al. (2021), these newly developed pedagogical designs are adapted from a few conventional approaches; they can better support the learning activities from happening inside VR or a mixture of inside and outside VR. Most of such conventional approaches are constructivism approaches, such as inquiry-based learning (Rutherford, 1964; Papert, 1980), problem-based learning (Savery et al., 2015), and experiential learning (Kolb, 2014). Jonassen (1994) first summarised the implications of constructivism for instructional design considering using VR for educational purposes. According to Janssen (1994), the knowledge construction process may be facilitated in learning environments with a few key characteristics, such as the ability to represent the natural complexity of the real world, present authentic tasks rather than simply provide abstract instructions, and support the fostering of reflective practice. In response to these characteristics, VR is a perfect tool to facilitate constructivism approaches. Despite the practical limitations such as the limited computational power of the computer systems, theoretically, the natural complexity of the real world can be simulated with an exceptionally high level of fidelity in the virtual world. On top of this, authentic tasks can be provided via the multimodal interface, allowing the reflective practice.

Although constructivism approaches have dominated practices, VR-enabled learning is not immune from the criticism of constructivism being practised in conventional learning environments (Kirschner et al., 2016). On the contrary, the challenges of providing appropriate and effective guidance and facilitation can be more prominent in VR-enabled learning. If the learning activities take place in certain VEs, such as HMDs, learners cannot see and sometimes can barely hear from the educators. Although presence can be enhanced in this case, it also makes the provision of appropriate and effective guidance and facilitations in time extremely challenging. Li et al. (2019) explored the possibilities of adapting Kolb's experiential learning model for enhancing the social and emotional skills of learners with autism spectrum disorder (ASD). In this work, pre-programmed agents were placed in different places of the virtual world, where the educators think the learners might need facilitations or guidance. These agents could fit into the virtual world and the social context natural; and they were programmed to detect if the learners might have encountered challenging situations using the interactive data as the inputs. Another possible solution to better employ constructivism approaches is to connect learners with educators in the virtual world. Both educators and learners can choose avatars representing themselves, join the same virtual world, and interact with each other through voice- or text-based chatting. Many of the early research studies in VR-enabled learning adopted this approach using networked virtual worlds such as SecondLife (Warburton, 2009; Inman et al., 2010). The utilisation of networked virtual worlds also enables collaborative learning among peer-learners. In short, the design of the learning content and how the design can support the chosen pedagogical approaches is crucial to the success of VR-enabled learning. It is suggested to report the underlying rationale of the learning content design choices given the chosen pedagogical approaches in future research studies.

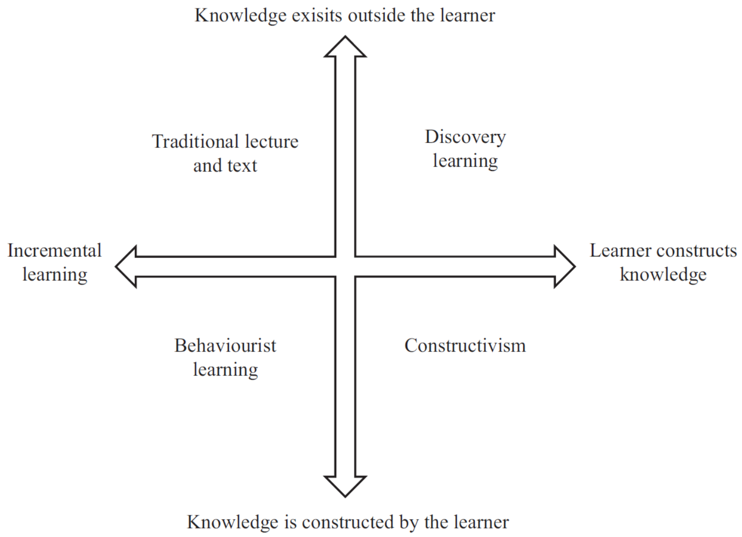

Figure 1. Constructivism in the quadrants of learning theory and knowledge theory.

3.4 Intended Learning Outcome

The clear definitions of intended learning outcomes can greatly help distinguish VR-enabled learning from other applications of VR, such as the use of VR for entertainment. Followings Bloom’s Taxonomy, like learning in conventional means, the intended learning outcomes of VR-enabled learning can also be cognitive (Bloom et al., 1956; Krathwohl, 2002), affective (Krathwohl et al., 1964), or psychomotor (Harrow, 1972). According to previous surveys, VR-enabled learning being practised in K-12 education settings often cover the cognitive domain of learning alone. For example, among the analysed studies, using VR to learn basic science dominated the disciplinary fields in Luo et al. (2021). The learning of basic science naturally emphasises the understanding and acquiring of factual and conceptual knowledge. On the contrary, in the higher education setting, VR-enabled learning has been practised in medical and engineering education a lot. Learning in these disciplinary fields clearly involves both cognitive and psychomotor domain of learning (Abulrub et al., 2011; Merchant et al., 2014; Radianti et al., 2020).

Recently, there is a new trend that focuses on the affective domain in VR-enabled learning. The affective domain of learning mainly involves learners’ feelings, emotions, and attitudes. It is about how leaners deal with things emotionally. According to Krathwohl et al. (1964), this domain is categorized into five hierarchical subdomains, namely receiving, responding, valuing, organization, and characterization, among which characterization is the highest of the domain while receiving is the lowest. Kwok et al. (2010) empirically studied the use of VR for affective domain of learning in the higher education setting. Based on the affective learning theory (Kort et al., 2001), this study explored the complex relationship between learners’ affective experience, learning engagement, and creativity during VR-enabled learning. The results showed that VR could facilitate effective learning both cognitively and affectively. Li et al. (2020) also empirically studied the effectiveness of using VR in classroom settings for enhancing youth's intercultural sensitivity by adopting the affective learning theory. The study called for more in-depth research regarding the learning effectiveness, the motivation of instructional material delivered in VEs, and the enjoyment perceived by learners. Compared to other domains of learning, using VR to enabled affective domain of learning is heavily underexplored and more empirical evidence is urgently needed, especially evidence on the retained effects in the affective domain of learning.

3.5 Learner Specifics

The fifth dimension of concerns focuses on the attributes of learners, which is often summarised as and called learner specifics. Learner specifics may include the basic demographics of targeted learners, such as age and gender, as well as the specific attributes or traits describing how they learn. For example, learning style is believed to be an especially important component that describes how learners learn. Learning styles (Claxton & Murrell, 1987) and corresponding learning style models (e.g., Gregorc & Ward (1977), Fleming (2011), and Kolb (2014)) offered descriptive typologies that brought the possibility to the provision of modifiable preferences for learning. Although VR is believed to be quite flexible in catering modifiable preferences for learning (de Freitas & Levene, 2004; Mikropoulos & Natsis, 2011; Ip & Li, 2015; Merchant et al., 2014; Luo et al., 2021), we should admit its limitations as well. For example, the experience is often full of graphical and auditory contents, but displaying text-based information, which is key to conventional learning settings, is supposed to be tricky in VEs (Dingler et al., 2018; Kojić et al., 2020). How learners with strong reading and/or writing preference (Fleming, 2011) can also benefit from VR-enable learning is yet to be investigated. In fact, most of the empirical studies did not report on the learning styles of the participants, making it difficult to draw any conclusion regarding this research question. Hence, although this is the fifth dimension of concerns covered here in this paper, learner specifics should be first considered when designing VR-enabled learning. Learner specifics also should be carefully collected and reported in research papers to foster the generalisation of the findings.

Recently, there is also a new trend of exploring the use of VR for the training of learners with special learning needs, which brings even more challenges in catering highly diverse learners specifics. Cromby et al. (1996) first listed three major advantages of VR in its application for people with learning difficulties. First, VEs are safe but authentic environments, in which learners can learn through trials and errors but without the risk of being exposed to real, humiliating, or dangerous consequences. Second, the virtual world can be manipulated in ways the real world cannot be under the control of educators. This allows the employment of many pedagogical approaches previously mentioned in section 3.3. And finally, VR allows the convey of educational information with limited or without the use of language or other symbol systems, which are believed to be challenging for people with learning difficulties. These three major advantages have been mentioned in almost all later studies in this research area and were hypothesized to be extremely useful in supporting learning among people with ASD. Lorenzo et al. (2019) reviewed empirical studies on using VR for learners with ASD that were published from

1990 to 2017. The authors suggested that besides these three major advantages, for learners with ASD, VR can also cater repetition in learning with manipulable deviations to enhance the generalization of learned knowledge and skills. The same advantage of VR-enabled learning for learners with ASD was also mentioned by Ip & Li (2015). In a short conclusion, it is suggested that besides conventional learning styles and other special attributes or traits of learning, the special learning needs should be carefully analysed and considered when designing and developing VR-enabled learning, so that the advantages can benefit learning on a broader spectrum of learners.

4 Conclusion and Future Work

The use of VR for educational purposes has brought both opportunities and challenges. One of the distinctive characteristics of this research area is interdisciplinarity, which also brings arguments towards the scope and boundary of the area. This paper first coined the term VR-enabled learning to better depict the relation between technology and learning, which are the two major components of this research area. We then defined VR-enabled learning as a learning experience; VR-enabled learning greatly benefits from the affordances of presence, which is enabled and delivered through relevant immersive technologies. The learning experience should be based on solid pedagogical approaches and should also be designed with well-defined intended learning outcomes while considering the learner specifics in the given educational contexts. The contributions of this paper also include the five dimensions of concerns of this research area; they are immersion, presence, pedagogy, intended learning outcome, and learner specifics. The first two concerns are more about VR, while the rest three are under the umbrella of educational research. Specifically, the development of the two constructs, saying immersion and presence, are introduced and compared. Based on previous research studies, it is advocated to report immersion systematically and measure presence using appropriate instruments in future studies. This can greatly help to address the issues regarding the scope and boundary of this research area mentioned in Luo et al. (2021) and Merchant et al. (2014). More importantly, by doing so, future studies can be reviewed and analysed more systematically to yield more concrete evidence on the effectiveness of VR-enabled learning. We also discussed the pedagogical approaches that can be employed for VR-enabled learning, which is often overlooked in this interdisciplinary research area. Although constructivism approaches have dominated practices, VR-enabled learning is not immune from the criticism of constructivism being practised in conventional learning environments (Kirschner et al., 2016). Thus, the development of pedagogical approaches dedicated to VR-enabled learning is still needed due to the nature of the medium. The importance of intended learning outcome and learner specifics are also discussed in this paper; they both are the key dimensions of concerns to help distinguish VR-enabled learning from adapting entertainment-oriented interactive media contents for learning.

Based on previous research studies and this paper, the followings are a few possible directions for future research. First, the cost-effectiveness of VR-enabled learning needs to be further studied. As mentioned by Luo et al. (2021), making high-quality VR-enabled learning is costly and whether the effect can justify the high cost needs to be investigated. In many empirical studies in this research area, the authors are often overly optimistic about the development of the enabling technology and believe in the improvement of accessibility as the enabling technology develops. However, the cost of the enabling technology is only the tip of the iceberg; the development of learning contents is far costly than the procurement of the enabling technology, and the learning contents often need to be updated regularly. Hence, the cost-effectiveness of VR-enabled learning is largely an underexplored area of the research. Second, the long-term effectiveness of VR-enabled learning is yet to be fully studied. Almost all empirical studies measured the immediate effects of learning but ignored the retained effects. This is mainly due to the difficulty in assessment and the design of the experiment. Since more and more research studies now cover the affective domain of learning, it is particularly important to measure the retained learning effects and compare them with such effects in other means of learning. Third, the transferring of knowledge and/or skills from the virtual environment to real-life situations needs to be further investigated. When using VR in the special education setting, the transferring of knowledge is often facilitated through learning activities that happen outside the virtual environment (Ip & Li, 2021). However, in the mainstream education setting, this is often overlooked. Lastly, the underlying affordance of VR needs to be further explored and studied empirically. Despite HMDs getting more and more accessible, desktop VR (a.k.a., non-immersive VR) still dominates the practices. As such primitive enabling technologies cannot ensure a high level of presence during exposure, which is one of the major affordances of VR-enabled learning, the complex relationship between the affordances and learning effectiveness needs to be further investigated (Coulter et al., 2007); and such investigation may help us better justify the high cost of VR-enable learning as well. We hope this paper can serve as an anchor for future studies and intrigue more in-depth research in VR-enabled learning and its associated challenges.

Acknowledgement

[acknowledgement]

References

Abulrub, A. H. G., Attridge, A. N., & Williams, M. A. (2011). Virtual reality in engineering education: The future of creative learning. In *2011 IEEE global engineering education conference (EDUCON)* (pp. 751-757). IEEE.

Baños, R. M., Botella, C., Garcia-Palacios, A., Villa, H., Perpiñá, C., & Alcaniz, M. (2000). Presence and reality judgment in virtual environments: a unitary construct?. *CyberPsychology & Behavior, 3*(3), 327-335.

Barfield, W., & Hendrix, C. (1995). The effect of update rate on the sense of presence within virtual environments. *Virtual Reality, 1*(1), 3-15.

Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook 1: Cognitive domain*. New York: David McKay.

Brooks, F. P. (1999). What's real about virtual reality?. *IEEE Computer graphics and applications, 19*(6), 16-27.

Chang, S. C., Hsu, T. C., & Jong, M. S. Y. (2020). Integration of the peer assessment approach with a virtual reality design system for learning earth science. *Computers & Education, 146*, 103758.

Charsky, D., & Mims, C. (2008). Integrating commercial off-the-shelf video games into school curriculums. *TechTrends, 52*(5), 38-44.

Claxton, C. S., & Murrell, P. H. (1987). *Learning styles*. Washington, DC: George Washington University (ERIC).

Coulter, R., Saland, L., Caudell, T., Goldsmith, T. E., & Alverson, D. (2007). The effect of degree of immersion upon learning performance in virtual reality simulations for medical education. *InMedicine Meets Virtual Reality*, 15, 155.

Coyne, R. (2003). Mindless repetition: Learning from computer games. *Design studies, 24*(3), 199-212.

Cromby, J. J., Standen, P. J., & Brown, D. J. (1996). The potentials of virtual environments in the education and training of people with learning disabilities. *Journal of Intellectual Disability Research, 40*(6), 489-501.

Cruz-Neira, C., Sandin, D. J., DeFanti, T. A., Kenyon, R. V., & Hart, J. C. (1992). The CAVE: audio visual experience automatic virtual environment. *Communications of the ACM, 35*(6), 64-73.

Cummings, J. J., & Bailenson, J. N. (2016). How immersive is enough? A meta-analysis of the effect of immersive technology on user presence. *Media Psychology, 19*(2), 272-309.

de Freitas, S., & Levene, M. (2004). An investigation of the use of simulations and video gaming for supporting exploratory learning and developing higher-order cognitive skills. In *Proceedings of the IADIS Cognition and Exploratory Learning in the Digital Age Conferenc*e, Lisbon, Portugal, 12–15 December.

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness: defining" gamification". In *Proceedings of the 15th international academic MindTrek conference: Envisioning future media environments* (pp. 9-15).

Dingler, T., Kunze, K., & Outram, B. (2018). Vr reading uis: Assessing text parameters for reading in vr. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems* (pp. 1-6).

Dinh, H. Q., Walker, N., Hodges, L. F., Song, C., & Kobayashi, A. (1999). Evaluating the importance of multi-sensory input on memory and the sense of presence in virtual environments. In *Proceedings IEEE Virtual Reality (Cat. No. 99CB36316)* (pp. 222-228). IEEE.

Fleming, N. D. (2011). *Teaching and learning styles: VARK strategies*. IGI global.

Frank, L. H., Casali, J. G., & Wierwille, W. W. (1988). Effects of visual display and motion system delays on operator performance and uneasiness in a driving simulator. *Human factors, 30*(2), 201-217.

Gregorc, A. F., & Ward, H. B. (1977). *A new definition for individual: implications for learning and teaching*. NASSP Bulletin, 401(6), 20–23.

Harrow, A. J. (1972). *A taxonomy of the psychomotor domain: A guide for developing behavioral objectives*. New York: Longman.

Inman, C., Wright, V. H., & Hartman, J. A. (2010). Use of Second Life in K-12 and higher education: A review of research. *Journal of Interactive Online Learning, 9*(1).

Ip, H. H., & Li, C. (2015). Virtual reality-based learning environments: recent developments and ongoing challenges. In *International Conference on Hybrid Learning and Continuing Education* (pp. 3-14). Springer, Cham.

Ip, H. H. S., & Li, C. (2021). Introducing Immersive Learning into Special Education Settings: A Comparative Review of Two Studies. *Creative and Collaborative Learning through Immersion*, 135-150.

Jonassen, D. H. (1994). Thinking technology: Toward a constructivist design model. *Educational technology, 34*(4), 34-37.

Kennedy, R. S., Lane, N. E., Berbaum, K. S., & Lilienthal, M. G. (1993). Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The international journal of aviation psychology, 3*(3), 203-220.

Kim, T., & Biocca, F. (1997). Telepresence via television: Two dimensions of telepresence may have different connections to memory and persuasion. *Journal of computer-mediated communication, 3*(2), JCMC325.

Kirschner, P., Sweller, J., & Clark, R. E. (2006). Why unguided learning does not work: An analysis of the failure of discovery learning, problem-based learning, experiential learning and inquiry-based learning. *Educational Psychologist, 41*(2), 75-86.

Kojić, T., Ali, D., Greinacher, R., Möller, S., & Voigt-Antons, J. N. (2020). User Experience of Reading in Virtual Reality—Finding Values for Text Distance, Size and Contrast. In *2020 Twelfth International Conference on Quality of Multimedia Experience (QoMEX)* (pp. 1-6). IEEE.

Kolb, D. A. (2014). *Experiential learning: Experience as the source of learning and development*. FT press.

Kort, B., Reilly, R., & Picard, R. W. (2001, August). An affective model of interplay between emotions and learning: Reengineering educational pedagogy-building a learning companion. In *Proceedings IEEE International Conference on Advanced Learning Technologies* (pp. 43-46). IEEE.

Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory into practice, 41*(4), 212-218.

Krathwohl, D. R., Bloom, B. S., & Masia, B. B. (1964). *Taxonomy of educational objectives: the classification of educational goals. Handbook 2: Affective domain*. New York: David McKay Company.

Kwok, R. C. W., Cheng, S. H., Ip, H. H. S., & Kong, J. S. L. (2010). Design of affectively evocative smart ambient media for learning. *Computers & Education, 56*(1), 101-111.

Lessiter, J., Freeman, J., Keogh, E., & Davidoff, J. (2001). A cross-media presence questionnaire: The ITC-Sense of Presence Inventory. *Presence: Teleoperators & Virtual Environments, 10*(3), 282-297.

Li, C., Ip, H. H. S., & Ma, P. K. (2019). A Design Framework of Virtual Reality Enabled Experiential Learning for Children with Autism Spectrum Disorder. In *International Conference on Blended Learning* (pp. 93-102). Springer, Cham.

Li, C., Ip, H. H., Wong, Y. M., & Lam, W. S. (2020). An empirical study on using virtual reality for enhancing the youth's intercultural sensitivity in Hong Kong. *Journal of Computer Assisted Learning, 36*(5), 625-635.

Lombard, M., Ditton, T. B., Crane, D., Davis, B., Gil-Egui, G., Horvath, K., ... & Park, S. (2000). Measuring presence: A literature-based approach to the development of a standardized paper-and-pencil instrument. In *Third international workshop on presence*, delft, the Netherlands (Vol. 240, pp. 2-4).

Lombard, M., Ditton, T. B., & Weinstein, L. (2009). Measuring presence: the temple presence inventory. In *Proceedings of the 12th annual international workshop on presence* (pp. 1-15).

Lorenzo, G., Lledó, A., Arráez-Vera, G., & Lorenzo-Lledó, A. (2019). The application of immersive virtual reality for students with ASD: A review between 1990–2017. *Education and Information Technologies, 24*(1), 127-151.

Luo, H., Li, G., Feng, Q., Yang, Y., & Zuo, M. (2021). Virtual reality in K-12 and higher education: A systematic review of the literature from 2000 to 2019. *Journal of Computer Assisted Learning*.

Martín-Gutiérrez, J., Mora, C. E., Añorbe-Díaz, B., & González-Marrero, A. (2017). Virtual technologies trends in education. *EURASIA Journal of Mathematics, Science and Technology Education, 13*(2), 469-486.

Merchant, Z., Goetz, E. T., Cifuentes, L., Keeney-Kennicutt, W., & Davis, T. J. (2014). Effectiveness of virtual reality-based instruction on students' learning outcomes in K-12 and higher education: A meta-analysis. *Computers & Education, 70*, 29-40.

Mikropoulos, T. A., & Natsis, A. (2011). Educational virtual environments: A ten-year review of empirical research (1999–2009). *Computers & Education, 56*(3), 769-780.

Nichols, S., Haldane, C., & Wilson, J. R. (2000). Measurement of presence and its consequences in virtual environments. *International Journal of Human-Computer Studies, 52*(3), 471-491.

Modjeska, D., & Waterworth, J. (2000). Effects of desktop 3D world design on user navigation and search performance. In *2000 IEEE Conference on Information Visualization. An International Conference on Computer Visualization and Graphics* (pp. 215-220). IEEE.

Nowak, K. L., & Biocca, F. (2003). The effect of the agency and anthropomorphism on users' sense of telepresence, copresence, and social presence in virtual environments. *Presence: Teleoperators & Virtual Environments, 12*(5), 481-494.

Papert, S. (1980). *Mindstorms: Children, computers, and powerful ideas*. New York: Basic Books, Inc.

Prensky, M. (2003). Digital game-based learning. *Computers in Entertainment (CIE), 1*(1), 21-21.

Radianti, J., Majchrzak, T. A., Fromm, J., & Wohlgenannt, I. (2020). A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. *Computers & Education, 147*, 103778.

Riva, G., Davide, F., & IJsselsteijn, W. A. (2003). Being there: The experience of presence in mediated environments. *Being there: Concepts, effects and measurement of user presence in synthetic environments, 5*.

Robertson, G. G., Card, S. K., & Mackinlay, J. D. (1993). Three views of virtual reality: nonimmersive virtual reality. *Computer, 26*(2), 81.

Rutherford, F. J. (1964). The role of inquiry in science teaching. *Journal of Research in Science Teaching, 2*(2), 80-84.

Savery, J. R. (2015). Overview of problem-based learning: Definitions and distinctions. *Essential readings in problem-based learning: Exploring and extending the legacy of Howard S. Barrows*, 9, 5-15.

Schubert, T., Friedmann, F., & Regenbrecht, H. (2001). The experience of presence: Factor analytic insights. *Presence: Teleoperators & Virtual Environments, 10*(3), 266-281.

Simões, J., Redondo, R. D., & Vilas, A. F. (2013). A social gamification framework for a K-6 learning platform. *Computers in Human Behavior, 29*(2), 345-353.

Sutcliffe, A. G., & Kaur, K. D. (2000). Evaluating the usability of virtual reality user interfaces. *Behaviour & Information Technology, 19*(6), 415-426.

Slater, M. (1999). Measuring presence: A response to the Witmer and Singer presence questionnaire. *Presence, 8*(5), 560-565.

Slater, M. (2009). Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*(1535), 3549-3557.

Slater, M., Usoh, M., & Steed, A. (1994). Depth of presence in virtual environments. *Presence: Teleoperators & Virtual Environments, 3*(2), 130-144.

Snow, M. P. (1998). *Charting presence in virtual environments and its effects on performance*. Virginia Polytechnic Institute and State University, Blacksburg, VA.

National Aeronautics and Space Administration. (1990). The Virtual Interface Environment Workstation (VIEW). Retrieved from: https://www.nasa.gov/ames/spinoff/new_continent_of_ideas/

Skarbez, R., Brooks, Jr, F. P., & Whitton, M. C. (2017). A survey of presence and related concepts. *ACM Computing Surveys (CSUR), 50*(6), 1-39.

Usoh, M., Catena, E., Arman, S., & Slater, M. (2000). Using presence questionnaires in reality. *Presence: Teleoperators & Virtual Environments, 9*(5), 497-503.

Virvou, M., & Katsionis, G. (2008). On the usability and likeability of virtual reality games for education: The case of VR-ENGAGE. *Computers & Education, 50*(1), 154-178.

Warburton, S. (2009). Second Life in higher education: Assessing the potential for and the barriers to deploying virtual worlds in learning and teaching. *British journal of educational technology, 40*(3), 414-426.

Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence, 7*(3), 225-240.