

Immersing learners in stories: A systematic literature review of educational narratives in virtual reality

James Calvert, Margee Hume

Torrens University Australia

The aim of this study, via a systematic review, was to investigate the addition of narratives in immersive virtual reality (IVR) and the associated impacts on learning. Narratives in IVR put the learner in the story, which, until recent developments in IVR head-mounted display technology, was out of reach in most classrooms. The review found that added immersion afforded by VR is particularly important in relation to learning where situational context is desirable. Importantly, IVR experiences with a narrative may have the potential to increase affective outcomes for learners, without reducing cognitive gains. Additionally, data concerning learning theories and design methodologies was extracted from the studies. The systematic review yielded 12 relevant and applied papers with this demonstrating there are still significant gaps in the research concerning the impact on learning in narrative IVR. This review highlights that the inclusion of narratives in educational IVR offers many potential benefits, however is yet to be fully explored. There are endless opportunities to tell stories in IVR. How they can be used and their impact on education is a rich ground for further research and development.

Implications for practice or policy

- Educators and VR developers should consider including a narrative in an educational IVR experience because of the potential to increase affective outcomes such as motivation and engagement, without reducing cognitive gains.
- Learners can benefit from educational IVR experiences with narratives because they can provide situational context, elicit an emotional response and scaffold complex learning.

Keywords: virtual reality, review, narrative, immersive, education

Introduction

Since the launch of the Oculus Rift in 2013, consumer virtual reality (VR) head-mounted displays (HMDs) have become increasingly affordable and capable, attracting attention as education technologies supporting learning outcomes (Kavanagh et al., 2017). Highly immersive HMDs are now readily available for schools, universities and organisations wanting to improve learner outcomes, yet research to date is rudimentary and disorganised. Spurred on by the increasing availability and excitement around this technology, research has increased, with studies examining the benefits for learners having varied and diverse objectives. However, the benefits from research lack consensus and are not being fully realised. One significant oversight is that the increased immersion afforded by modern HMDs, unlike their less immersive predecessors such as desktop-based VR, can create narrative experiences that were not previously possible in the digital realm (Calvert & Abadia, 2020; Misak, 2018). This paper contributes to a larger research agenda by collating current research focused on immersive VR experiences (IVR) where the learner is immersed in a narrative. This is distinct from when learners are provided narrative experiences in non-immersive media such as desktop VR or multimedia materials. It is also distinct from educational material delivered to learners in IVR, but where the experience is devoid of a narrative.

We use the terminology immersive narrative VR (INVR), which is described as a narrative experience delivered via a suitably capable IVR HMD, affording heightened immersion in the narrative via a combination of increased sense of presence, embodiment, agency, realism and other factors (Jeunet et al., 2018; Shin, 2017; Slater, 2018). Put simply, INVR lets the learner be in the narrative, rather than reading or watching the narrative on a screen. Via systematic review, our paper demonstrates that to date researcher and practitioner understanding of the impact of INVR on learning and the design of the experience and learner outcomes are rudimentary, and studies focused on this system of relationships are very limited.

IVR affords educators the chance to immerse learners in an experience which takes place in a crafted reality. Educators can create experiences in IVR that are not bound by physical, geographic or temporal limitations.

It is an ideal medium for constructing rich narrative environments using realistic representations and cognitive artefacts, extending the range of our experience to other social and physical environments (Biocca, 2002). Afforded by the sense of presence and the embodied experience, learners can be immersed in a narrative (Shin, 2018). As highlighted by Misak (2018, p. 40), “VR mimics the work of the readers’ imagination” and can convey narrative elements more realistically than its traditional counterparts of texts or screens, the application of which can be applied to any narrative scenario required for enhancing learning, extending its potential use to all areas of education. In the context of IVR, the addition of the extra dimensions of immersion afforded by modern HMDs, specifically the sense of embodiment and presence (Caserman et al., 2019; Slater, 2018) warrants a closer investigation. With little known how INVR experiences impact learning, further research is warranted to advance this practice. This creates a compelling argument for the function of INVR for learning.

This practice of utilising INVR in education is poorly defined and little explored. It is rarely included in the broader agenda, yet as a practice it is anecdotally proposed to extend the learner experience (Breien & Wasson, 2020). Narrative stories and the passing of knowledge through storytelling have been applied and studied in digital learning applications prior to the commercialisation of VR (Dickey, 2005, 2006). Narratives in digital learning applications include games and interactive learning material, and for the purposes of this review differ from narratives in IVR, in that they are not delivered via an immersive HMD (Adams et al., 2012; Dettori & Paiva, 2009; Jemmali et al., 2018). However, the findings from these previous studies serve as suitable background for this systematic review and can inform the potential impacts of INVR. Several studies exist on utilising game narrative techniques for interactive educational experiences, and how these narrative devices allow for increased engagement, presence, self-efficacy, interest and meaning construction (Garneli et al., 2017; Jemmali et al., 2018). There are studies which have demonstrated that adding a narrative to an interactive media learning experience results in no or even reduced learning gains when compared to traditional learning materials (Adams et al., 2012; Martey et al., 2017). This seems incongruous to what would be expected of increased engagement, presence, self-efficacy, interest and meaning construction, but to date this has not been explicitly investigated. Hence, the current evidence regarding narratives in digital learning experiences points towards a mixture of positive and negative outcomes for learners.

Because IVR for learning is relatively new, we must look to other digital-based learning experiences such as video games and virtual learning environments to see if the integration of a narrative in IVR will aid learning. If there are complex instructional elements, for example, including a narrative in the digital-based learning experience, these can lead to reduced learning (Pilegard & Mayer, 2016). However, if a narrative is added, it can support learning by increasing presence and generating an emotional response (Gorini et al., 2011). The narrative may also support learning by helping learners organise knowledge, increase motivation and make sense of their experience (Dettori & Paiva, 2009). The problem with such discrepancies in outcomes can be traced back to differences in learning experience design and study set-up. Narrative is additional content in the experience, and it may or may not be related to the learning outcomes. How effectively the narrative is integrated can play a decisive part in the outcomes for the learners experiencing it (Jemmali et al., 2018). There appears in the current research little attention to the design process and the quality of the design and how that might impact the learning outcomes. Through systematic review we investigated this in more depth, with the aim of advancing the role of design and alignment with the learning outcomes.

Educational VR research agenda

There is a growing body of research examining the potential benefits of VR for learning across a range of educational domains and disciplines. Studies are seeking evidence to support the implementation of IVR in curricula across various fields trying to leverage what the technology might offer. As highlighted in the review by Jensen and Konradsen (2018), some of the 21 studies reported the positive effect IVR can have on cognitive, affective or psychomotor skill acquisition in specific scenarios. However, it was found that study quality was below average, and that outside those scenarios, IVR was not more effective than less immersive material. One of the review’s conclusions was that learning in IVR can be beneficial when the depicted simulation is correctly utilised. In the review by Hamilton et al. (2020), 29 studies examined the learning outcomes of IVR when compared to less immersive methods. In the review, cognitive learning was the most commonly assessed learning domain (87%), over procedural and affective. The same review examined learning outcomes and found that most studies concluded the IVR intervention was significantly

better for education. Only two studies in the review reported no difference, and two studies found detrimental effects to learning when using IVR. The findings from these reviews highlight the mixed results occurring in the field of study. Further work is warranted to better categorise the types of IVR experience used in education and the impact on learning each category has.

A significant gap emerging from previous reviews is an inquiry into IVR for learners through the lens of learning theories. Radiani et al. (2020) are the sole researchers in the discovery that included an examination of learning theories in their review of 38 studies. Even though the review identified the growing number of application domains where IVR is beneficial, they found that learning theories were not often considered as an assist or guide towards measuring learning outcomes. Considering this information, we hypothesised that a review of INVR for learners will yield a similar lack of learning theory application. This lack of learning theory application in educational IVR studies further demonstrates the immaturity of the field.

Learning outcomes in IVR

This systematic review examined the literature focused on INVR for learning. As such, we can compare and analyse the learning outcomes of implementing INVR in education. An overview of the possible learning outcomes and how they align with learning theory is therefore important. Several studies have demonstrated that IVR can offer cognitive benefits such as knowledge mastery (Abadia et al., 2019), knowledge transfer (Bhargava et al., 2018), knowledge retention (Krokos et al., 2019; Meyer et al., 2019) and task engagement (Bhargava et al., 2018), while others have focused on the affective benefits of IVR, demonstrating that IVR can increase engagement and elicit emotions (Allcoat & von Mühlenen, 2018; Schutte & Stilinović, 2017).

Designing effective instructional material that maximises cognitive outcomes requires knowledge of cognition and the mechanisms of learning and problem solving (Plass et al., 2010). Cognitive load theory (CLT) can explain learning outcomes by considering the limitations and capabilities of “human cognitive architecture” (Plass et al., 2010). Linked to CLT is the cognitive theory of multimedia learning (CTML), which aims to improve instructional materials to make best use of a learner’s cognitive limits (Mayer, 2009). CTML aims to make the most efficient use of the auditory and visual channels by which we process words and pictures, the two main methods by which we receive instructional material. An IVR experience engages these auditory and visual senses, and effecting learning in this environment can be explained by CLT and CTML.

In addition to cognitive outcomes, there is research that supports affective outcomes as being important to the learning process as well (Um et al., 2012). Affective outcomes include learner satisfaction, motivation, perceived difficulty and positive perceptions (Um et al., 2012). Together, these terms contribute to the cognitive-effective theory of learning with media (Moreno, 2006) and expand on CTML (Mayer, 2009) by adding extra dimensions of learning found in cognitive research. Of particular interest is affective mediation, which facilitates learning by increasing or decreasing cognitive engagement through motivational factors. Further supporting affective mediation is interest theory, which states that learners pay more attention when they are interested (Dewey, 1913). The need to explore the additional affective benefits INVR learning experiences can offer the learner is considered one of the main categories for data collection in this systematic review.

Motivated by the current survey of relevant literature presenting as diverse and varied, the research field requires further investigation to gain a better understanding of IVR in education, and more specifically, the role of INVR and associated learning outcomes. Moreover, the context of recent research in training, simulation, game-based learning and general education further compounds the inconsistency in the application of findings in research and practice. We undertook a systematic review to organise the research and build on our understanding of INVR to assist learners and contribute to the call for research underpinning the functions of VR for learning (Bower & Jong, 2020).

To investigate the role of design of INVR for education and to ascertain associated cognitive and affective learning outcomes, we utilised a systematic review to interrogate the problem, research the question and construct the inclusion and methods. We specifically focused the systematic review on IVR learning experiences that feature narratives and those papers aiming on achieving learning outcomes. This review

will support a better understanding of immersing learners in stories and the benefits of educational INVR, while advancing the organisation of this research agenda and field of enquiry. Currently, the research has inconsistent quality and a clear agenda is not evident, which has resulted in a small subset of relevant papers. INVR may have other benefits, but this is unknown due to the small subset. This small subset shows an increase in affective learning outcomes and engagement with no reduction in cognitive learning gains, thereby warranting further empirical research.

Emerging from the current literature reviewed, this systematic review addressed the following research questions (RQs):

RQ1 To what educational domains are INVR being applied?

The inclusion of this domain analysis focuses on the area of application to ascertain where INVR are most effective. Looking across all the studies will assist in identifying the key motivation for including a narrative in the educational VR experience. Comparisons can be made with the educational domains reported in other systematic reviews to determine if INVR experiences are better suited or more commonly employed in specific areas.

RQ2 How does an INVR experience impact learning?

VR is a new form of media with new opportunities for immersive learning. As research begins to explore the affordances of INVR, more needs to be known about how it benefits learning and if the results from these studies contribute to learning theories.

RQ3 How are narratives being designed and structured in the context of educational IVR?

The purpose of surveying the narrative structures used is to ascertain what narrative elements there are and what they do for the learner. Are there appropriate design methodologies applied during development to ensure the narrative is fit for purpose? Pulling data from the selected studies will provide useful information for future inclusion of narratives in IVR learning experiences.

Method

For this review, we used the preferred reporting items for systematic reviews and meta-analysis (PRISMA; Moher et al., 2009). A complete outline of all PRISMA phases follows below.

Search strategy

To begin with, we determined that both journal articles and conference proceedings would be included in the review. After an initial scoping test, we chose three key databases for the literature search: IEEE Xplore, Scopus and EBSCO. We conducted searching on these digital libraries during August and September of 2020. In addition to the search results, we included papers found via researcher checking and journal hand searching.

Search terms

The following search string was used for all databases:

- “virtual reality” OR VR
- AND immers*
- AND educat* OR learn*
- AND story* OR stories OR narrative* OR scenario*

This search string was designed to capture all relevant articles to the systematic review. Immers*, which covers immersive, immersion and immersed, was important because it will return results for immersive VR as opposed to desktop VR. The terms educat* and learn* were included to ensure the results would meet the criteria of the VR experience being in an education context or one where learning was measured. The terms story*, stories, narrative* and scenario* were included to ensure the results fit the criteria of the study which is focused on the use of narratives in educational VR. In this instance, story*, stories* and scenario*

were included to ensure that articles where narratives were described or featured without the specific term of *narrative* applied to them, were captured in the search results.

In addition, only articles published in English were included in the search. The final step in the search was to restrict the year of publication to between 2013 and 2020. The first consumer headset capable of high levels of immersion, in the form of the Oculus Rift Development Kit 1, became available in early 2013 (Avila & Bailey, 2014). Articles published prior to this would not meet the eligibility criteria.

Selection and screening

The search returned the following results from each database: IEEE Xplore ($n = 108$), Scopus ($n = 397$) and EBSCO ($n = 117$). To this, we added articles sourced from researcher checking and journal hand searching ($n = 6$) to the total pool. Once duplicates had been removed, a final pool of $n = 453$ articles remained. These articles were then screened by reading the title and abstract to remove those which unmistakably did not fit within this review. This resulted in $n = 79$ articles for full-text review. Each article was then assessed against the following criteria to determine the final set of articles included in this review. To be included, an article must meet all of the eligibility criteria. Figure 1 shows the PRISMA flow diagram for the process of article search and refinement.

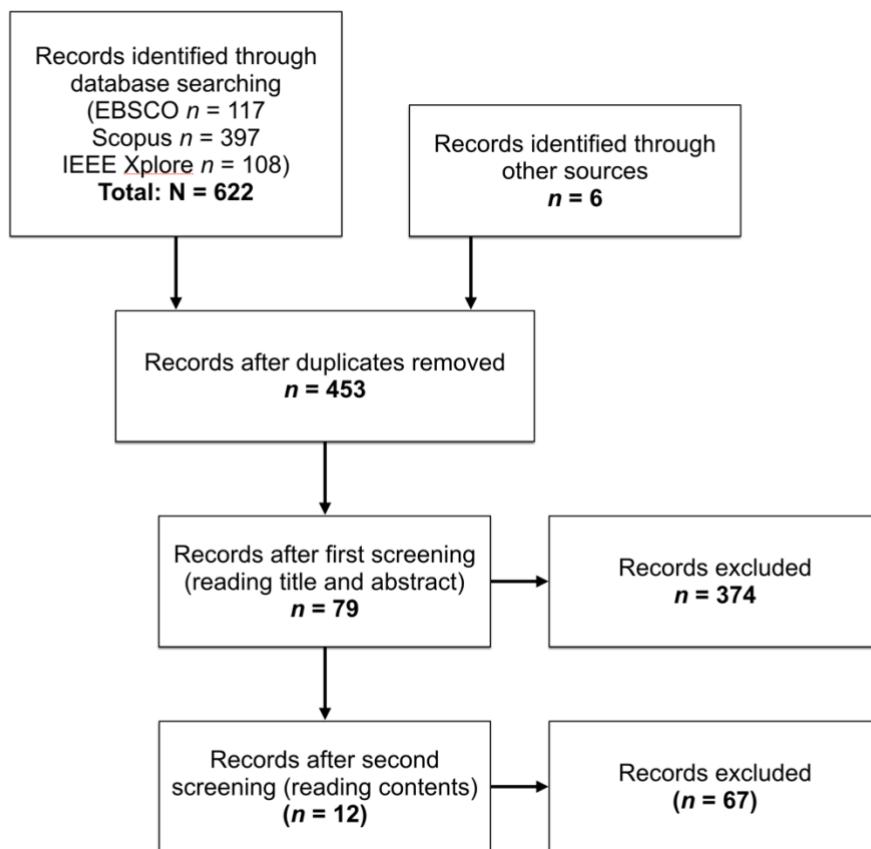


Figure 1. PRISMA flow diagram

Eligibility criteria

The first phase required a full-text read through of a sample ($n = 20$) of the shortlisted papers, which was conducted by both of the authors. The articles were reviewed against the following criteria:

- (1) A narrative being presented to the learner in IVR via an immersive HMD. This needs to be presented to the viewer and not be created by the viewer. Multiple story branches in a choose-your-own-adventure format are acceptable. The narrative must be a significant feature of the experience.
- (2) The INVR experience had to be for learners. However, the learners can be either in a school, higher education, professional or museum setting.
- (3) The aim of the study is to improve the learner experience in some capacity. The narrative need not necessarily be the exclusive intervention for achieving this.
- (4) The improvements to the learning experience must be measured and/or assessed.
- (5) The design of the INVR experience must be described in satisfactory detail.

A specific set of criteria just for narrative assessment was required in order to determine which of the IVR experiences featured a true and significant narrative and which did not. After reviewing the initial subset of 20 papers, it was decided to override the narrative definition offered within each paper, with a pre-defined narrative definition. The following quote from Wertsch (1998) offers a succinct and suitable definition of a narrative that papers can be assessed by:

Narrative is organized around temporality, it has a central subject, a plot with a beginning, middle and end, and an identifiable narrative voice; it makes connections between events; it achieves a closure, a conclusion, a resolution. (p. 80)

Based on this, the final list of required items that must be present in order to be determined a narrative is outlined in Table 1.

Table 1
Narrative features

Narrative criteria	Description
Character	<ul style="list-style-type: none"> • Player character that is adequately described or represented (virtually) that may or may not have agency • And/or character(s) that inhabit the virtual world which are adequately described or represented
Setting	<ul style="list-style-type: none"> • Time and location where the narrative can take place
Plot	<ul style="list-style-type: none"> • An event or sequence of events that involves/impacts the character(s) and takes place in the setting • A plot sets in motion a sequence of events from which the conflict or crisis will emerge
Conflict or crisis	<ul style="list-style-type: none"> • A conflict which impacts the character(s) and drives the story forward
Resolution	<ul style="list-style-type: none"> • The conflict is resolved, either successfully or unsuccessfully

Once there was unanimous agreement between reviewers on results of the sample pool assessment, the reviewers moved on to the final stage and the remaining articles were then subject to a full-text review. This resulted in $n = 12$ articles for inclusion in this review.

Data collection process

Categories for data collection were first piloted and then discussed to determine inter-rater reliability. This ensured that data gathered would be relevant to the overall aims of the study and that the data extracted would be consistent amongst the researchers. Data were gathered in the following categories: application domain, learning theory, narrative structure or theory, measures, and outcomes.

The last step was to obtain a valid quality score for each paper by administering the Medical Education Research Study Quality Instrument (MERSQI). The MERSQI test was originally developed to assess medical research (Reed et al., 2007), but was adapted here to include education specific terminology for each item. The revised MERSQI guide can be found in the appendix. Papers can score a maximum of 18 points on the MERSQI, and higher scores indicate a higher quality study. The intention behind obtaining a MERSQI score for each article was to gather data and make an assessment on the quality of research performed in this field.

Results

The following section outlines the results of the systematic review, with analysis conducted to address the research questions. Table 2 provides a summary of the main characteristics of the articles, covering participants, IVR device, learner type and MERSQI score.

Table 2
Summary table

Paper	Participants	IVR HMD	Learner type	MERSQI score
(Alrehaili & Al Osman, 2019)	31	Oculus Rift	Student (secondary)	14
(Calvert & Abadia, 2020)	79	HTC Vive	Student (secondary and university)	15.5
(Caserman et al., 2019)	20	HTC Vive	Professional	9
(Donald & Scott-Brown, 2019)	53	Oculus Rift	Student (secondary)	7.5
(Buttussi & Chittaro, 2018)	96	Oculus Rift	General public	14.5
(Feng, González, Amor, et al., 2020)	93 (87 completed)	Oculus Rift	General public	12
(Feng, González, Mutch, et al., 2020)	191	Oculus Rift	Students (secondary and university) and professionals	10.5
(Ferguson et al., 2020)	42	PlayStation VR	Student (secondary)	13.5
(Nowak et al., 2020)	180 (171 completed)	HTC Vive	General public	13
(Pagano et al., 2020)	70 & 75	HTC Vive	General public	6
(Parong & Mayer, 2018)	55 & 57	HTC Vive	Student (university)	16
(Zhang et al., 2019)	90 (approximately)	HTC Vive	Students (secondary and university) and the general public	6

Analysis from the MERSQI scores

Overall, the data from the MERSQI test returned a mean of 11.4 ($SD = 3.6$), which is 1.3 points lower than the mean scores from the MERSQI test in the review by Hamilton et al. (2020). The variance in definition of the individual MERSQI items as outlined above may have contributed to the lower mean score.

There were four papers that scored below 10, which is considered low based on previous applications of the MERSQI (Hamilton et al., 2020). For all these low scoring papers, the focus was on design innovations with a limited lead user study, which was conducted to validate IVR design methodologies. Any learning

outcomes which were measured in the studies centred around the affective domains of learner satisfaction, attitudes and perceptions. Their contributions are valuable in the rapidly evolving field of INVR by increasing our understanding of the learning domains to which INVR can be applied and how they might be designed. However, these findings suggest that more studies with robust study designs are required.

Application domain

RQ1 To what educational domains are INVR being applied?

In total, only four application domains were present across the 12 studies: history ($n = 4$), training serious games ($n = 4$), science ($n = 3$) and public health ($n = 1$).

The reasons given for including a narrative in the IVR history lessons were to immerse learners in a significant historical event (Calvert & Abadia, 2020; Pagano et al., 2020); experience personal stories (Donald & Scott-Brown, 2019); elicit emotional response (Calvert & Abadia, 2020); and to improve information recall and recall of spatial information (Ferguson et al., 2020).

The reasons given for including a narrative in the training serious games were to mimic real events (Buttussi & Chittaro, 2018), or to utilise an interactive story that changes based on how the player reacts, mimicking potential real-life scenarios (Caserman et al., 2019; Feng, González, Amor, et al., 2020; Feng, González, Mutch, et al., 2020). Essentially, all four training serious games aimed for a narrative experience that gave learners a chance to experience a real-life scenario. The only difference was whether the virtual scenario was fictionalised or based on real events.

The reasons given for including a narrative in the science lessons were to scaffold learning of complex concepts (Zhang et al., 2019) and to support learning through interactivity (Alrehaili & Al Osman, 2019). The reasons given for including a narrative in the public health lesson were to let users experience events and feel empathy in an immersive environment (Nowak et al., 2020).

Looking across all the studies, four common reasons emerged as the key motivation for including a narrative in the educational IVR experience, which were to immerse learners in real-life scenarios, experience personal stories, elicit emotional response and scaffold complex learning.

Learning theories and outcomes

Data relating to learning theory, design methodology and narrative structure was extracted from the articles, which is summarised in Table 3. This data was then used to answer RQs 2 and 3.

RQ2 How does an INVR experience impact learning?

When extracting data relating to learning theories, only explicitly mentioned learning theories which are clearly linked to the study design, measures or outcomes were included. It was found that 42% ($n = 5$) of studies did not mention a learning theory. Mayer's (2009) CTML was mentioned in two studies. These two studies were also the most thorough when discussing how these learning theories relate to learning in IVR and discussed the impact their results have on the theory. The theory was also embedded in the study design.

The following learning theories were mentioned once across all the studies. Self-efficacy theory is used to explain how well a student performs when engaging in learning activities (Chemers et al., 2001). It is defined as the effort and persistence on a task that is applied when the learner believes they are competent in the task (Bandura, 1977). Self-efficacy is an important measure when assessing the effectiveness of educational IVR materials. Generative learning theory describes one of the key aspects of cognitive processing in learning. The basis of the theory is that by selecting, organising and integrating information with prior knowledge, meaningful learning can occur (Fiorella & Mayer, 2015). When audio visual material is effectually presented to the learner in IVR, generative learning can take place. The zone of proximal development is defined as the distance between what the learner is currently capable of as an independent problem solver and what they are potentially capable of under adult guidance (Vygotsky, 1978). It is used as a basis for developing educational experiences in IVR that lead to optimal learning. Social cognitive theory explains that learning occurs not just through direct experiences, but also through vicarious experiences such as social interactions (Bandura, 1977, 2001). In the context of IVR, for example, learners can be involved in social situations where learning can occur as explained by social cognitive theory. The

media-enables-method hypothesis is used to explain why IVR may afford enhanced methods for learning by drawing upon the unique affordances of the technology, such as enhanced immersion, presence and embodiment (Moreno, 2006).

Table 3
Learning, design and narrative data extraction

Paper	Learning theory	Design methodology	Narrative structure
(Alrehaili & Al Osman, 2019)	Multimedia learning theory	Analysis, design, development, implementation, and evaluation (ADDIE model)	Role playing game (RPG)
(Buttussi & Chittaro, 2018)	Social cognitive theory	N/A	Scenario inspired by real incident
(Calvert & Abadia, 2020)	Media-enables-method hypothesis, mastery approach	N/A	Historical accuracy, scenarios, student at the centre of the action
(Caserman et al., 2019)	N/A	N/A	Multiple paths with alternate endings
(Donald & Scott-Brown, 2019)	N/A	Co-design, design council's double diamond	Interviews, lived experience, oral history
(Feng, González, Amor, et al., 2020)	N/A	N/A	Action-driven narrative method
(Feng, González, Mutch, et al., 2020)	Problem-based gaming framework	Action-driven narrative method	Story modules, explicit narrative, action-driven narrative method
(Ferguson et al., 2020)	Zone of proximal development	N/A	Non-fiction, educational environmental narrative games, active navigation, passive navigation, explicit story structure, implicit story structure
(Nowak et al., 2020)	N/A	N/A	Users experience events
(Pagano et al., 2020)	Dale's cone of experience, Bloom's taxonomy	Learning mechanics-game mechanics model, activity theory-based model for serious games	Ancient history, help, measuring goals, scoring, interaction
(Parong & Mayer, 2018)	Cognitive theory of multimedia learning, cognitive load theory, interest theory, self-efficacy theory, generative learning theory	N/A	Interactive tour with narration
(Zhang et al., 2019)	N/A	Iterative human centered design, design make learn cycle	Interactivity, gameplay, guided, trial and error, non-linear, three act structure

Figure 2 shows that knowledge gain was the most popular measure when assessing the outcomes of narrative VR for education, featuring in 67% of articles. Of the studies that measured knowledge gain, only four did so in comparison with other non-IVR learning interventions (Alrehaili & Al Osman, 2019; Buttussi & Chittaro, 2018; Calvert & Abadia, 2020; Parong & Mayer, 2018). Of these four, only one reported a significant gain in knowledge from the IVR intervention, while the other three reported no significant gain. In the study by Calvert and Abadia (2020), learners scored higher on knowledge tests after experiencing an historical story in IVR versus those who viewed the same experience as desktop 360° video. The study by Alrehaili and Al Osman (2019) tested for knowledge gain in an immersive science lesson and found that

there was no significant difference between IVR versus desktop-based simulation. They did, however, find a significant increase in knowledge retention for the IVR group. The two other studies found no significant knowledge gain for learners in IVR in a training serious game (Buttussi & Chittaro, 2018) and science lesson (Parong & Mayer, 2018). More studies are required that measure learning gains in INVR experiences compared to non-IVR learning materials, to better report on this measure.

The following measures belong to the affective outcomes of INVR. The measure of engagement featured in 57% of articles, where it was considered a key objective for deploying an educational IVR experience. Of these articles, only three compared IVR against a non-IVR intervention, all of which reported that engagement in IVR was significantly higher. The measure of the sense of presence featured in 33% of articles. The increased sense of presence is common motivator for developing IVR (Jensen & Konradsen, 2018) and in the context of learners, can be used to give context to learning material. Immersing learners in historical scenarios is one such example. Other measures that reported gains in IVR over non-IVR learning materials include empathy, immersion, realism, motivation, interest, affect and content-specific items. There is growing support for the affective benefits of IVR for education, with 83% ($n = 10$) of papers measuring affective outcomes. In 75% ($n = 9$) of papers, cognitive and affective measures were both reported on.

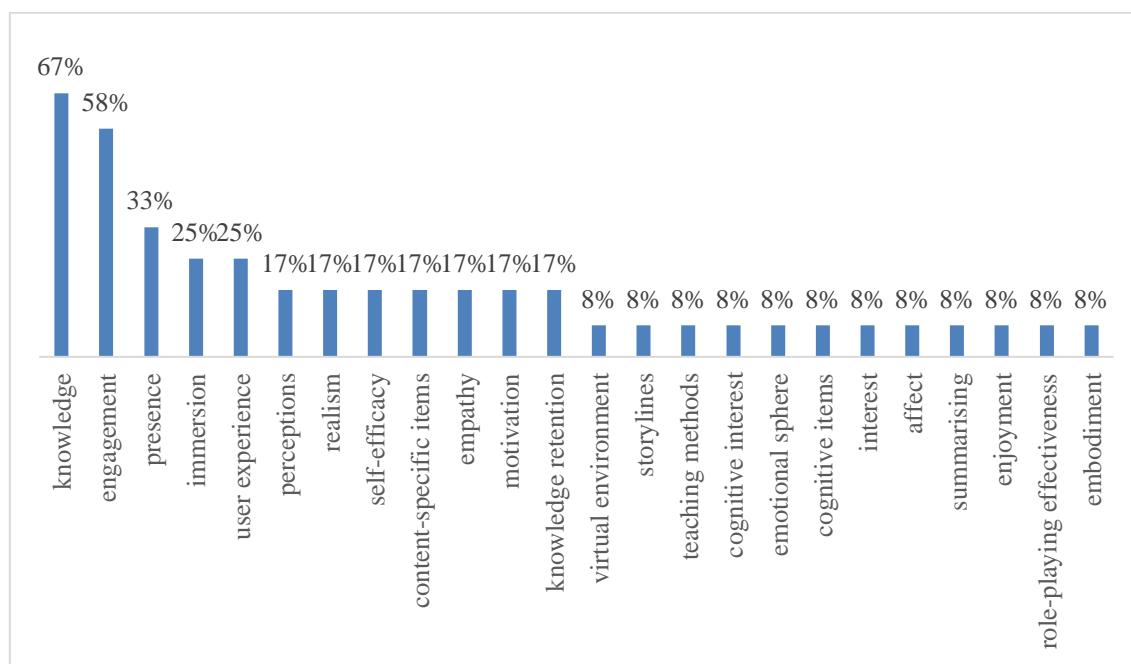


Figure 2. Assessment measures

Narratives in VR

RQ3 How are narratives being designed and structured in the context of educational IVR?

When extracting data relating to design methodology, only explicitly mentioned applications of design methodologies were included. Only 42% ($n = 5$) of studies utilised a design methodology in the process of developing the VR experience. In all instances where a design methodology was mentioned, it was explained in detail and followed with the aim of improving learning outcomes. Out of the 58% ($n = 7$) articles that did not specify a design methodology, two used a commercially available IVR experience (Ferguson et al., 2020; Parong & Mayer, 2018).

When extracting data relating to narrative structure, only items explicitly mentioned in the article were included. The terminology used within the articles to describe narrative structure varied greatly. This could be due to the varying nature of the application domains, design methodology and goals of each study. INVR experiences for education are also a newly developing area, with many different narrative structures being employed in the design of the experience. In response to this, the narrative structures mentioned in Table 3

were grouped into themes, which were then assigned a narrative term, which is displayed in Table 4. This list of terms can be used to better describe the structural features of a narrative VR experience for education.

Table 4
Structural features of INVR

Narrative structure	Mentions
Non-fiction	4
Fiction	8
Passive participant	2
Active participant	10
Implicit story structure	3
Explicit story structure	3
Scenarios	2
Goal driven	1
Branching storylines	1
Guided	3

Storylines are added in IVR to help scaffold complex learning or to provide a situational context to the material. The added immersion afforded by IVR is particularly important in relation to experiences where situational context is desirable. This applies to either fiction or non-fiction educational IVR narratives. A fiction or non-fiction storyline in IVR provides this situational context in a highly engaging manner.

More studies favoured active participants (83%) than passive participants (17%). Active participation includes features such as gameplay mechanics and interactivity. Passive participation includes passive navigation and learners being in the centre of the action. Both active and passive participation were effective for learning gains.

In the papers that mentioned implicit or explicit story structure, neither structure had comprehensive advantages for learning over the other. However, with minimal studies from which to draw information about implicit versus explicit story structure in IVR, there is no comprehensive guide as to which is more effective in an INVR learning scenario. It is also worth noting that an IVR learning experience can feature moments of both implicit and explicit story structure. Implicit versus explicit story structure is a very interesting area, and more studies need to specify which of these they will employ, as both methods have their unique benefits.

The remaining narrative structures employed in INVR for education were scenarios, goal driven, branching storylines and guided. Unlike the features discussed above, these features are optional and are not a choice of either one or the other, such as fiction or non-fiction.

Discussion

We conclude that current research has inconsistent quality, and a clear agenda is not evident, resulting in a small subset of relevant papers. This small subset shows an increase in affective learning outcomes and engagement with no reduction in cognitive learning gains. This highlights the significance of further research required in this area. INVR may have other benefits yet to be acknowledged and investigated but this is unknown due to the small subset.

For all the articles included for review, the following trends emerged. Firstly, in the validity of instrument domain, only two papers reported relationships to other variables. This is a major omission from the studies, and for there to be significant gains in the research of INVR, correlation between domains require further investigation. Secondly, only two papers reported a contribution to educational theory. These are notable absences and demonstrate that this emerging field is still maturing.

Research foci and future agenda

The small subset of results suggests an emergent field with limited critical mass in any constructs and domains. The systematic review process exhaustively searched and interrogated papers yet failed to reveal

a strong consistent approach to the research conducted. The systematic reviews on educational IVR by Hamilton et al. (2020) and Radiani et al. (2020) found that science was the most common application domain, with no mention of history, suggesting IVR in a history context as a relatively recent development, and it is particularly interesting as it is the equal most popular application domain in this review. There are two possible reasons for this: most articles included in both of those studies were published in 2018 and some early access in 2019. The history-focused IVR studies in this review were published in 2019 and 2020. It is also possible that as IVR development and research become more accessible, the application domains for learning expanded to include more general education material, such as history. As more and more narrative-based learning experiences for IVR are developed, this list of educational domains will likely increase. The library of INVR experiences is likely to continue growing to cater for a wide range of pedagogical applications.

From the analysis of educational domains, four main benefits for including a narrative in an IVR learning experience emerged. These benefits are immerse learners in real-life scenarios; experience personal stories; elicit emotional response; and to scaffold complex learning. Each of these benefits require further investigation to better understand the full impact of INVR. The added immersion afforded by IVR is particularly important in relation to experiences where situational context is desirable.

Lack of foci on learning outcomes

The survey of articles found that only 68% included a learning theory. Of those, only two papers reported a contribution to an existing educational theory. This low number aligns with the reviews conducted by Radiani et al. (2020) and Hamilton et al. (2020). Therefore, it is evident that some studies in this field are making conclusions that IVR will improve learning without analysing their findings against existing theory. Although it is understandable that some studies will focus more on design innovations and classroom application, empirical research aligned with learning theory is still required to better understand the role INVR plays in learning.

Cognitive outcomes in the two studies with CMTL were supported by the theory and positive outcomes for learners achieved. With the high number of studies featuring affective measures, it is surprising that there were no learning theories focused on affective features. Considering this, the cognitive-affective theory of learning with media (Moreno, 2006) would be a suitable theory for analysis of INVR for education because of the inclusion of affective outcomes in conjunction with cognitive outcomes.

Regarding learning outcomes, there was one example of knowledge gain and one example of knowledge retention in the studies. There was no other evidence to suggest that immersive narrative VR can improve cognitive factors over other forms of media, or of IVR experience without narratives. Similar to research from other interactive educational media, the cognitive benefits are yet to be fully supported in INVR for education. What was consistent was the growing support for the affective benefits of INVR for education.

INVR design and learning outcomes

The papers in this review all included a narrative and the results did not demonstrate negative learning impacts. This sits in contrast to the findings of Pilegard and Mayer (2016), but more importantly, gives support to the argument by Jemmal et al. (2018) that there are many variables in the design of the narrative experience and the associated studies that means a comprehensive analysis is difficult. The application of appropriate design methodologies was evident in the literature; however, this was from fewer than half the surveyed studies. Considering that how an IVR experience is designed can have significant impact on both cognitive and affective outcomes, this is an important finding. However, the outcome of this review is that the inclusion of a narrative in an IVR experience for learning should be considered when the educational goal is to immerse learners in real-life scenarios; experience personal stories; elicit an emotional response; or to scaffold complex learning.

The lack of learning theory and application of design methodology can be improved by bringing together more interdisciplinary teams of researchers and active practitioners in the fields of education, psychology, design, computer games and creative writing. When a team tries to develop an IVR experience and then undertake empirical research, cost time and design aspects are in conflict – this may be solved by cross-functional teams. Writing, designing and developing an immersive narrative that engages students and

achieves learning goals is a complex undertaking. This is supported by Jemmali et al. (2018), who suggested that adding narratives is further complicated by the variation in duration, relationship to learning content and overall structure. To assist, structural features of INVR found in the articles of this systematic review (Table 4) can act as a guide for future research and development. These assistive narrative structures are learners being active or passive in the narrative; implicit or explicit story structure; scenarios; goals; branching storylines; and guidance. Importantly, these structures apply to either fiction or non-fiction narratives. Indeed, each of these narrative structures can be further explored in future research to determine the unique characteristics and impacts on learning of each.

Conclusion and future directions

The results from this systematic review reveal that the existing research is returning varying results and offers little critical mass in any of the focal areas. Of the small subset of papers found in the review, the results regarding cognitive outcomes from INVR lack consensus. This variation in findings is a result of the diverse quality and study design of the papers, in addition to the variation in narrative design structures within each study. Some of the studies found a positive impact on knowledge gain; however, there was no clear narrative structure that leads to this. Compounding the lack of consistency, several studies showed no increase in cognitive outcomes compared to other forms of media, suggesting this variability warrants further research. Conversely, affective outcomes resulting from the INVR showed consistent gains in learner engagement, empathy, interest, self-efficacy, motivation and presence over other forms of educational media, with these benefits applying to all narrative structures present in the articles. Therefore, the INVR narrative design structures found in this study may assist in adding clarity to future research and in building a strong research agenda advancing IVR and learning outcomes. This study highlights the small number of comprehensive studies linking learning theory, design methodology, narrative structure and learning outcomes, and as such, confirm further research is still required for a better understanding of the learning outcomes and role of INVR in education. Regardless, there is still compelling evidence that INVR offers affective benefits in a wide range of educational applications and should be considered a valuable educational tool.

Statement on open data, ethics and conflict of interest

The authors report that there is no conflict of interest in the reporting of this research and that no human participants were used in this study. Data will be made available by submitting an individual request to the first author.

References

- Abadia, R., Calvert, J., & Dasika, R. (2019). Effectiveness of using an immersive and interactive virtual reality learning environment to empower students in strengthening empathy and mastery learning. In M. Chang, H.-J. So, L.-H. Wong, F.-Y. Yu, J.-L. Shih, I. Boticki, M.-P. Chen, A. Dewan, S. Haklev, E. Koh, T. Kojiri, K.-C. Li, D. Sun, & Y. Wen (Eds.), *Proceedings of the 27th International Conference on Computers in Education, Taiwan* (pp. 495–504). Asia-Pacific Society for Computers in Education. https://apsce.net/icce/icce2019/proceedings/paper_108.pdf
- Adams, D. M., Mayer, R. E., MacNamara, A., Koenig, A., & Wainess, R. (2012). Narrative games for learning: Testing the discovery and narrative hypotheses. *Journal of Educational Psychology*, 104(1), 235–249. <https://doi.org/10.1037/a0025595>
- Allcoat, D., & von Mühlenen, A. (2018). Learning in virtual reality: Effects on performance, emotion and engagement. *Research in Learning Technology*, 26. <https://doi.org/10.25304/rlt.v26.2140>
- Alrehaili, E. A., & Al Osman, H. (2019). A virtual reality role-playing serious game for experiential learning. *Interactive Learning Environments*, 1–14. <https://doi.org/10.1080/10494820.2019.1703008>
- Avila, L., & Bailey, M. (2014). Virtual reality for the masses. *IEEE Computer Graphics and Applications*, 34(5), 103–104. <https://doi.ieeecomputersociety.org/10.1109/MCG.2014.103>
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84(2), 191–215. <https://doi.org/10.1037/0033-295X.84.2.191>
- Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology*, 52, 1–26. <https://doi.org/10.1146/annurev.psych.52.1.1>

- Bhargava, A., Bertrand, J. W., Gramopadhye, A. K., Madathil, K. C., & Babu, S. V. (2018). Evaluating multiple levels of an interaction fidelity continuum on performance and learning in near-field training simulations. *IEEE Transactions on Visualization and Computer Graphics*, 24(4), 1418–1427. <https://doi.org/10.1109/TVCG.2018.2794639>
- Biocca, F. (2002). The evolution of interactive media: Toward “being there” in nonlinear narrative worlds. In M. C. Green, J. J. Strange, & T. C. Brock (Eds.), *Narrative impact: Social and cognitive foundations* (pp. 97–130). Lawrence Erlbaum Associates Publishers.
- Bower, M., & Jong, M. S. (2020). Immersive virtual reality in education. *British Journal of Educational Technology*, 51(6), 1981–1990. <https://doi.org/10.1111/bjet.13038>
- Breien, F. S., & Wasson, B. (2020). Narrative categorization in digital game-based learning: Engagement, motivation & learning. *British Journal of Educational Technology*, 52(1), 91–111. <https://doi.org/10.1111/bjet.13004>
- Buttussi, F., & Chittaro, L. (2018). Effects of different types of virtual reality display on presence and learning in a safety training scenario. *IEEE Transactions on Visualization and Computer Graphics*, 24(2), 1063–1076. <https://doi.org/10.1109/TVCG.2017.2653117>
- Calvert, J., & Abadia, R. (2020). Impact of immersing university and high school students in educational linear narratives using virtual reality technology. *Computers & Education*, 159, Article 104005. <https://doi.org/10.1016/j.compedu.2020.104005>
- Caserman, P., Garcia-Agundez, A., & Goebel, S. (2019). A survey of full-body motion reconstruction in immersive virtual reality applications. *IEEE Transactions on Visualization and Computer Graphics*, 1–1. <https://doi.org/10.1109/TVCG.2019.2912607>
- Chemers, M. M., Hu, L., & Garcia, B. F. (2001). Academic self-efficacy and first year college student performance and adjustment. *Journal of Educational Psychology*, 93(1), 55–64. <https://doi.org/10.1037/0022-0663.93.1.55>
- Dettori, G., & Paiva, A. (2009). Narrative learning in technology-enhanced environments. In N. Balacheff, S. Ludvigsen, T. de Jong, A. Lazonder, & S. Barnes (Eds.), *Technology-Enhanced Learning* (pp. 55–69). Springer Netherlands. https://doi.org/10.1007/978-1-4020-9827-7_4
- Dewey, J. (1913). *Interest and effort in education*. Houghton Mifflin.
- Dickey, M. D. (2005). Engaging by design: How engagement strategies in popular computer and video games can inform instructional design. *Educational Technology Research and Development*, 53(2), 67–83. <https://doi.org/10.1007/BF02504866>
- Dickey, M. D. (2006). Game design narrative for learning: Appropriating adventure game design narrative devices and techniques for the design of interactive learning environments. *Educational Technology Research and Development*, 54(3), 245–263. <https://doi.org/10.1007/s11423-006-8806-y>
- Donald, I., & Scott-Brown, K. (2019). Their memory: Exploring veterans’ voices, virtual reality and collective memory. *Journal For Virtual Worlds Research*, 12. <https://doi.org/10.4101/jvwr.v12i2.7360>
- Feng, Z., González, V. A., Amor, R., Spearpoint, M., Thomas, J., Sacks, R., Lovreglio, R., & Cabrera-Guerrero, G. (2020). An immersive virtual reality serious game to enhance earthquake behavioral responses and post-earthquake evacuation preparedness in buildings. *Advanced Engineering Informatics*, 45, Article 101118. <https://doi.org/10.1016/j.aei.2020.101118>
- Feng, Z., González, V. A., Mutch, C., Amor, R., Rahouti, A., Baghouz, A., Li, N., & Cabrera-Guerrero, G. (2020). Towards a customizable immersive virtual reality serious game for earthquake emergency training. *Advanced Engineering Informatics*, 46, Article 101134. <https://doi.org/10.1016/j.aei.2020.101134>
- Ferguson, C., van den Broek, E. L., & van Oostendorp, H. (2020). On the role of interaction mode and story structure in virtual reality serious games. *Computers & Education*, 143, Article 103671. <https://doi.org/10.1016/j.compedu.2019.103671>
- Fiorella, L., & Mayer, R. E. (2015). *Learning as a generative activity: Eight learning strategies that promote understanding*. Cambridge University Press; Cambridge Core. <https://doi.org/10.1017/CBO9781107707085>
- Garneli, V., Giannakos, M., & Chorianopoulos, K. (2017). Serious games as a malleable learning medium: The effects of narrative, gameplay, and making on students’ performance and attitudes: Serious games as a malleable learning medium. *British Journal of Educational Technology*, 48(3), 842–859. <https://doi.org/10.1111/bjet.12455>
- Gorini, A., Capideville, C. S., De Leo, G., Mantovani, F., & Riva, G. (2011). The role of immersion and narrative in mediated presence: The virtual hospital experience. *Cyberpsychology, Behavior, and Social Networking*, 14(3), 99–105. <https://doi.org/10.1089/cyber.2010.0100>

- Hamilton, D., McKechnie, J., Edgerton, E., & Wilson, C. (2020). Immersive virtual reality as a pedagogical tool in education: A systematic literature review of quantitative learning outcomes and experimental design. *Journal of Computers in Education*. <https://doi.org/10.1007/s40692-020-00169-2>
- Jemmali, C., Bunian, S., Mambretti, A., & El-Nasr, M. S. (2018). Educational game design: An empirical study of the effects of narrative. *Proceedings of the 13th International Conference on the Foundations of Digital Games*, 1–10. <https://doi.org/10.1145/3235765.3235783>
- Jensen, L., & Konradsen, F. (2018). A review of the use of virtual reality head-mounted displays in education and training. *Education and Information Technologies*, 23(4), 1515–1529. <https://doi.org/10.1007/s10639-017-9676-0>
- Jeunet, C., Albert, L., Argelaguet, F., & Lecuyer, A. (2018). “Do you feel in control?”: Towards novel approaches to characterise, manipulate and measure the sense of agency in virtual environments. *IEEE Transactions on Visualization and Computer Graphics*, 24(4), 1486–1495. <https://doi.org/10.1109/TVCG.2018.2794598>
- Kavanagh, S., Luxton-Reilly, A., Wuensche, B., & Plimmer, B. (2017). A systematic review of virtual reality in education. *Themes in Science & Technology Education*, 10(2), 85–119.
- Krokos, E., Plaisant, C., & Varshney, A. (2019). Virtual memory palaces: Immersion aids recall. *Virtual Reality*, 23(1), 1–15. <https://doi.org/10.1007/s10055-018-0346-3>
- Martey, R. M., Shaw, A., Stromer-Galley, J., Kenski, K., Clegg, B., Folkestad, J., Saulnier, T., & Strzalkowski, T. (2017). Testing the power of game lessons: The effects of art style and narrative complexity on reducing cognitive bias. *International Journal of Communication*, 11. <https://ijoc.org/index.php/ijoc/article/view/5032>
- Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). Cambridge University Press. <https://doi.org/10.1017/CBO9780511811678>
- Meyer, O. A., Omdahl, M. K., & Makransky, G. (2019). Investigating the effect of pre-training when learning through immersive virtual reality and video: A media and methods experiment. *Computers & Education*, 140, Article 103603. <https://doi.org/10.1016/j.compedu.2019.103603>
- Misak, J. (2018). A (virtual) bridge not too far: Teaching narrative sense of place with virtual reality. *Computers and Composition*, 50, 39–52. <https://doi.org/10.1016/j.compcom.2018.07.007>
- Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *BMJ*, 339, Article b2535. <https://doi.org/10.1136/bmj.b2535>
- Moreno, R. (2006). Does the modality principle hold for different media? A test of the method-affects-learning hypothesis. *Journal of Computer Assisted Learning*, 22(3), 149–158. <https://doi.org/10.1111/j.1365-2729.2006.00170.x>
- Nowak, G. J., Evans, N. J., Wojdynski, B. W., Ahn, S. J., Len-Rios, M. E., Carera, K., Hale, S., & McFalls, D. (2020). Using immersive virtual reality to improve the beliefs and intentions of influenza vaccine avoidant 18-to-49-year-olds: Considerations, effects, and lessons learned. *Vaccine*, 38(5), 1225–1233. <https://doi.org/10.1016/j.vaccine.2019.11.009>
- Pagano, A., Palombini, A., Bozzelli, G., De Nino, M., Cerato, I., & Ricciardi, S. (2020). ArkaeVision VR game: User experience research between real and virtual paestum. *Applied Sciences*, 10(9), Article 3182. <https://doi.org/10.3390/app10093182>
- Parong, J., & Mayer, R. (2018). Learning science in immersive virtual reality. *Journal of Educational Psychology*, 110(6), 785–797. <https://doi.org/10.1037/edu0000241>
- Pilegard, C., & Mayer, R. E. (2016). Improving academic learning from computer-based narrative games. *Contemporary Educational Psychology*, 44–45, 12–20. <https://doi.org/10.1016/j.cedpsych.2015.12.002>
- Plass, J. L., Moreno, R., & Brünken, R. (2010). *Cognitive load theory*. Cambridge University Press.
- Radiani, J., Majchrzak, T. A., Fromm, J., & Wohlgemant, I. (2020). A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. *Computers & Education*, 147, Article 103778. <https://doi.org/10.1016/j.compedu.2019.103778>
- Reed, D. A., Cook, D. A., Beckman, T. J., Levine, R. B., Kern, D. E., & Wright, S. M. (2007). Association between funding and quality of published medical education research. *JAMA*, 298(9), 1002–1009. <https://doi.org/10.1001/jama.298.9.1002>
- Schutte, N. S., & Stilinović, E. J. (2017). Facilitating empathy through virtual reality. *Motivation and Emotion*, 41(6), 708–712. <https://doi.org/10.1007/s11031-017-9641-7>

- Shin, D. (2017). The role of affordance in the experience of virtual reality learning: Technological and affective affordances in virtual reality. *Telematics and Informatics*, 34(8), 1826–1836. <https://doi.org/10.1016/j.tele.2017.05.013>
- Shin, D. (2018). Empathy and embodied experience in virtual environment: To what extent can virtual reality stimulate empathy and embodied experience? *Computers in Human Behavior*, 78, 64–73. <https://doi.org/10.1016/j.chb.2017.09.012>
- Slater, M. (2018). Immersion and the illusion of presence in virtual reality. *British Journal of Psychology*, 109(3), 431–433. <https://doi.org/10.1111/bjop.12305>
- Um, E., Plass, J., Hayward, E., & Homer, B. (2012). Emotional design in multimedia learning. *Journal of Educational Psychology*, 104, 485–498. <https://doi.org/10.1037/a0026609>
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press. <https://doi.org/10.2307/j.ctvjf9vz4>
- Wertsch, J. V. (1998). *Mind as action*. Oxford University Press.
- Zhang, L., Bowman, D. A., & Jones, C. N. (2019). Enabling immunology learning in virtual reality through storytelling and interactivity. In J. Y. C. Chen & G. Fragomeni (Eds.), *Lecture notes in computer science: Vol. 11575. Virtual, augmented and mixed reality: Applications and case studies* (pp. 410–425). Springer International Publishing. https://doi.org/10.1007/978-3-030-21565-1_28

Corresponding author: James Calvert, jcalvert@torrens.edu.au

Copyright: Articles published in the *Australasian Journal of Educational Technology* (AJET) are available under Creative Commons Attribution Non-Commercial No Derivatives Licence ([CC BY-NC-ND 4.0](https://creativecommons.org/licenses/by-nd/4.0/)). Authors retain copyright in their work and grant AJET right of first publication under CC BY-NC-ND 4.0.

Please cite as: Calvert, J., & Hume, M. (2022). Immersing learners in stories: A systematic literature review of educational narratives in virtual reality. *Australasian Journal of Educational Technology*, 38(5), 45-61. <https://doi.org/10.14742/ajet.##>

Appendix

Original MERSQI by Reed et al. (2007, p. 1004), adapted for education

Domain	MERSQI Item	Original subscale	Educational technology specific subscale
Study design	Study design	Single group cross-sectional or single group post-test only Single group pre-test and post-test Nonrandomised, 2 group Randomised controlled trial	<i>no change</i> <i>no change</i> <i>no change</i> Experimental trials with participant randomisation
Sampling	Institutions	1 institution sampled 2 institutions sampled >2 institutions sampled	1 learner type sampled (primary, secondary, university, professional or general public) 2 learner types sampled >2 learner types sampled
	Response rate	N/A <50% or not reported 50%-74% response rate >75% response rate	<i>no change</i> <i>no change</i> <i>no change</i> <i>no change</i>
Type of data	Type of data	Assessment by study participant Objective measurement	Assessment by study participant (user evaluation, surveys, interviews or focus groups) Objective measurement (knowledge tests or physiological measures)
Validity of evaluation instrument	Internal structure	N/A Not reported Reported	<i>no change</i> <i>no change</i> <i>no change</i>
	Content	N/A Not reported Reported	<i>no change</i> <i>no change</i> <i>no change</i>
	Relationships to other variables	N/A Not reported Reported	<i>no change</i> <i>no change</i> <i>no change</i>
Data analysis	Appropriateness of analysis	Data analysis inappropriate for study design or type of data Data analysis appropriate for study design or type of data	<i>no change</i> <i>no change</i>
	Complexity of analysis	Descriptive analysis only Beyond descriptive analysis	<i>no change</i> <i>no change</i>
Outcomes	Outcomes	Satisfaction, attitudes, perceptions, opinions, general facts Knowledge, skills Behaviours Patient/health care outcome	Learner satisfaction (usability), learner attitudes (engagement and motivation), learner perceptions (in the case of VR this can be presence, embodiment or immersion), opinions or general facts Knowledge and/or skills acquisition Learner behaviours Education theoretical contribution