



A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda

Jaziar Radianti ^a, Tim A. Majchrzak ^{a,*}, Jennifer Fromm ^b, Isabell Wohlgenannt ^c

^a University of Agder, Kristiansand, Norway

^b University of Duisburg–Essen, Duisburg, Germany

^c University of Liechtenstein, Vaduz, Liechtenstein

ARTICLE INFO

Keywords:

Augmented and virtual reality
Cooperative/collaborative learning
Distance education and online learning
Human–computer interface
Media in education

ABSTRACT

Researchers have explored the benefits and applications of virtual reality (VR) in different scenarios. VR possesses much potential and its application in education has seen much research interest lately. However, little systematic work currently exists on how researchers have applied immersive VR for higher education purposes that considers the usage of both high-end and budget head-mounted displays (HMDs). Hence, we propose using systematic mapping to identify design elements of existing research dedicated to the application of VR in higher education. The reviewed articles were acquired by extracting key information from documents indexed in four scientific digital libraries, which were filtered systematically using exclusion, inclusion, semi-automatic, and manual methods. Our review emphasizes three key points: the current domain structure in terms of the learning contents, the VR design elements, and the learning theories, as a foundation for successful VR-based learning. The mapping was conducted between application domains and learning contents and between design elements and learning contents. Our analysis has uncovered several gaps in the application of VR in the higher education sphere—for instance, learning theories were not often considered in VR application development to assist and guide toward learning outcomes. Furthermore, the evaluation of educational VR applications has primarily focused on usability of the VR apps instead of learning outcomes and immersive VR has mostly been a part of experimental and development work rather than being applied regularly in actual teaching. Nevertheless, VR seems to be a promising sphere as this study identifies 18 application domains, indicating a better reception of this technology in many disciplines. The identified gaps point toward unexplored regions of VR design for education, which could motivate future work in the field.

1. Introduction

Digital devices are being increasingly adopted for learning and education purposes (Zawacki-Richter & Latchem, 2018). This can particularly be observed in the 1997–2006 period, when *networked computers for collaborative learning* were intensively used, and in the 2007–2016 period, when so-called *online digital learning* became widespread. During these two periods, people examined the

* Corresponding author.

E-mail addresses: jaziar.radianti@uia.no (J. Radianti), timam@uia.no (T.A. Majchrzak), jennifer.fromm@uni-due.de (J. Fromm), isabell.wohlgenannt@uni.li (I. Wohlgenannt).

<https://doi.org/10.1016/j.compedu.2019.103778>

Received 22 July 2019; Received in revised form 21 November 2019; Accepted 25 November 2019

Available online 9 December 2019

0360-1315/© 2019 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license

(<http://creativecommons.org/licenses/by/4.0/>).

exploitation potential of emerging technologies, such as virtual learning environments (Boulton, Kent, & Williams, 2018; Muñoz-Cristóbal, Gallego-Lema, Arribas-Cubero, Martínez-Monés, & Asensio-Pérez, 2017), mobile devices (Wang, Wiesemes, & Gibbons, 2012), and computer-mediated communication (Mason & Bacsich, 1998). Recently, *virtual reality (VR) technologies* are actively being incorporated into education, teaching, and training in various application domains.

Although VR is not new, the recent developments in immersive technologies – in terms of visualization and interactions – have made VR increasingly attractive to scholars. The latest VR head-mounted displays (HMDs), such as HTC Vive or Oculus Rift, allow users to experience a high degree of *immersion*. Immersion describes the involvement of a user in a virtual environment during which his or her awareness of time and the real world often becomes disconnected, thus providing a sense of “being” in the task environment instead. Freina and Ott (2015) define this term as “a perception of being physically present in a non-physical world by surrounding the user of the VR system created with images, sound, or other stimuli” so that a participant feels he or she is actually “there”.

Likewise, low-budget HMDs for mobile devices, such as Samsung Gear VR and Google Cardboard, enable *everyone* to experience immersive virtual environments. In addition, the current devices also offer interaction capabilities. While high-end VR equipment comes with dedicated controllers that support game play (e.g., Oculus Touch), low-budget HMDs, such as Google Cardboard, support gaze control or enable the user to interact with the virtual environment by providing a magnetic switch.

The HMD market is expected to be valued at USD 25 billion by 2022, growing at a Compound Annual Growth Rate of 39.52% between 2019 and 2025 (Global Augmented Reality, 2019). Therefore, the time is ripe to study immersive VR, primarily due to the increased capability of VR technology and reduced costs. For example, the recently released Oculus Quest comes in the form of a cordless HMD. While the design enables the user to move more freely, the cost is set at approximately 400 USD—pretty much the same price as the previous generation Rift with cables.

Moreover, VR has been described as the learning aid of the 21st century (Rogers, 2019). A study suggests that students retain more information and can better apply what they had learned after participating in VR exercises (Krokos, Plaisant, & Varshney, 2019). Considering the potential learning enhancement through VR use, it is understandable why researchers, organizations, and educators nowadays scrutinize this technology intensively, looking to add an extra dimension to the classroom with respect to both teaching and learning.

Due to the increasing scholarly attention paid to VR technologies, there are already a few comprehensive overviews and systematic mappings of VR applications for education (Jensen & Konradsen, 2018; Merchant, Goetz, Cifuentes, Keeney-Kennicutt, & Davis, 2014). For example, Merchant et al. (2014) focus on desktop VR in education while Jensen and Konradsen (2018) emphasize the application of HMD technologies. Although Jensen and Konradsen (2018) have a similar focus to this review, which is in terms of learning experience and learning outcome, they are not specific about their target audience (such as K12, higher education, vocational training, or job training). Besides, these two studies do not examine underlying learning theories governing the VR application development and the design elements of the VR applications.

Hence, in this paper, we focus on the use of immersive VR in higher education and answer the following questions:

- How are immersive VR technologies embedded into higher education?
- What design elements, learning contents, and immersive technologies have been laid out to support VR-based learning?
- What learning theories have been applied to guide VR application design and development for higher education?
- What evaluation methods have been employed to evaluate the learning outcomes?
- What higher education domains have applied VR to teaching and learning?

In this article, we carry out a systematic mapping of the existing VR literature and highlight the immersive aspect of VR to answer these questions, which are presented in detail in Table 2. Our aim is to contribute to the extant body of knowledge on the application of digital devices for educational purposes.

This work is organized as follows: In Section 2, related works on systematic reviews of VR in an educational context are presented. The section concludes with a gap analysis and an explanation of how our work addresses the noted gaps. Then, our research design is described in Section 4, including our literature identification search procedure, our semi-automatic filtering process, and analysis methods. In Section 5, four analysis frameworks (1. *research design*, 2. *research method and data analysis*, 3. *learning*, and 4. *design element*) and their sub-frameworks are formulated as an overarching category foundation for the consolidation and analysis of the results. Drawing on this classification, we identify patterns of VR design element use for specific learning contents in higher education. The results are discussed in Section 6 to highlight the implications, opportunities for future research, recommendations for lecturers, and limitations of our work. Finally, we draw a conclusion in Section 7. Three appendices (from p. 32) provide more detailed information about certain aspects of our research design as well as a list of articles included in our literature study.

2. Theoretical background

In this section, we draw the background of the two main topics combined in this paper: immersive VR and learning theories. Bringing these topics together allows us to analyze the usage of VR in the higher education context.

2.1. VR and immersion

VR can be defined as “the sum of the hardware and software systems that seek to perfect an all-inclusive, sensory illusion of being present in another environment” (Biocca & Delaney, 1995). Immersion, presence, and interactivity are regarded as the core

characteristics of VR technologies (Ryan, 2015; Walsh & Pawlowski, 2002). The term *interactivity* can be described as the degree to which a user can modify the VR environment in real-time (Steuer, 1995). *Presence* is considered as “the subjective experience of being in one place or environment, even when one is physically situated in another” (Witmer & Singer, 1998). While researchers largely agree on the definitions of interactivity and presence, differing views exist on the concept of *immersion*. One branch of researchers suggests that immersion should be viewed as a *technological* attribute that can be assessed objectively (Slater & Wilbur, 1997), whereas others describe immersion as a subjective, individual belief, i.e., a *psychological* phenomenon (Witmer & Singer, 1998).

Jensen and Konradson (2018) suggest an additional perspective concerning the positive effects of immersion and presence on learning outcomes. The results of the reviewed studies in their work show that learners who used an immersive HMD were more engaged, spent more time on the learning tasks, and acquired better cognitive, psychomotor, and affective skills. However, this study also identifies many factors that can be reinforcers or barriers to immersion and presence. Both graphical quality of VR and the awareness when using VR, for instance, can reduce the sense of presence. Individual personality traits may also be associated with a limited acquisition of skills from using VR technologies.

According to the technological view, the term immersion means the “extent to which the computer displays are capable of delivering an inclusive, extensive, surrounding, and vivid illusion of reality” (Slater & Wilbur, 1997). More precisely, this includes the degree to which the physical reality is excluded, the range of sensory modalities, the width of the surrounding environment as well as the resolution and accuracy of the display (Slater & Wilbur, 1997). The technological attributes of a VR technology – such as the frame rate or the display resolution – consequently determine the degree of immersion that a user experiences (Bowman & McMahan, 2007). In contrast, the psychological point of view considers immersion to be a psychological state in which the user perceives an isolation of the senses from the real world (Witmer & Singer, 1998). According to this view, the perceived degree of immersion differs from person to person and the technological attributes barely influence it (Mütterlein, 2018).

With our systematic mapping study, we aim to identify the current applications of immersive VR in higher education. For this reason, we set specific inclusion and exclusion criteria to distinguish papers that describe immersive VR applications from papers that describe non-immersive VR applications. As the subjective experience of immersion could hardly be used as a selection criterion, we defined certain types of VR technologies as either immersive or non-immersive based on their technological attributes. These specific technology types were then used as either the inclusion or exclusion criteria. For this paper, we considered *mobile* VR (e.g., Google Cardboard, Samsung Gear), *high-end HMDs* (e.g., Oculus Rift, HTC Vive), and *enhanced* VR (e.g., a combination of HMDs with data gloves or bodysuits) as immersive.¹ These VR technologies allowed the user to fully immerse into the virtual environment (Khalifa & Shen, 2004; Martín-Gutiérrez, Mora, Añorbe Díaz, & González-Marrero, 2017). Desktop VR and Cave Automatic Virtual Environment (CAVE) systems, however, were considered as *non-immersive* because the user can still recognize the screen or a conventional graphics workstation (Biocca & Delaney, 1995; Robertson, Czerwinski, & Van Dantzich, 1997).

2.2. Learning paradigms

An understanding of the existing learning paradigms is essential for performing an analysis of the current state of VR applications in higher education. Thus, we introduce the main ideas behind the existing learning paradigms. Literature distinguishes between behaviorism, cognitivism, and constructivism (Schunk, 2012). Other scholars also include experiential learning (Kolb & Kolb, 2012) to this list and, recently, connectivism has been introduced as a new learning paradigm (Kathleen Dunaway, 2011; Siemens, 2014). Each learning paradigm has developed various theories about educational goals and outcomes (Schunk, 2012). Each of these theories also offers a different perspective on the learning goals, motivational process, learning performance, transfer of knowledge process, the role of emotions, and implications for the teaching methods.

Behaviorism assumes that knowledge is a repertoire of behavioral responses to environmental stimuli (Shuell, 1986; Skinner, 1989). Thus, learning is considered to be a passive absorption of a predefined body of knowledge by the learner. According to this paradigm, learning requires repetition and learning motivation is extrinsic, involving positive and negative reinforcement. The teacher serves as a role model who transfers the correct behavioral response.

Cognitivism understands the acquisition of knowledge systems as actively constructed by learners based on pre-existing prior knowledge structures. Hence, the proponents of cognitivism view learning as an active, constructive, and goal-oriented process, which involves active assimilation and accommodation of new information to an existing body of knowledge. The learning motivation is intrinsic and learners should be capable of defining their own goals and motivating themselves to learn. Learning is supported by providing an environment that encourages discovery and assimilation or accommodation of knowledge (Shuell, 1986). Cognitivism views learning as more complex cognitive processes such as thinking, problem-solving, verbal information, concept formation, and information processing. It addresses the issues of how information is received, organized, stored, and retrieved by the mind. Knowledge acquisition is a mental activity consisting of internal coding and structuring by the learner. Digital media, including VR-based learning can strengthen cognitivist learning design (Dede, 2008). Cognitive strategies such as schematic organization, analogical reasoning, and algorithmic problem solving will fit learning tasks requiring an increased level of processing, e.g. classifications, rule or procedural executions (Ertmer & Newby, 1993) and be supported by digital media (Dede, 2008).

¹ It should be noted that immersion apparently has a sociocultural aspect the notion of which changes over time. When the first 3D games were introduced in the early 1990s, players would consider these to be *immersive* despite seeing software-rendered graphics in a resolution of 320X160 pixels with 2D sprites, often displayed on a small CRT monitor.

Constructivism posits that learning is an active, constructive process. Learners serve as information constructors who actively construct their subjective representations and comprehensions of reality. New information is linked to the prior knowledge of each learner and, thus, mental representations are subjective (Fosnot, 2013; Fosnot & Perry, 1996). Therefore, constructivists argue that the instructional learning design has to provide macro and micro support to assist the learners in constructing their knowledge and engaging them for meaningful learning. The macro support tools include related cases, information resources, cognitive tools, conversation, and collaboration tools, and social or contextual support. A micro strategy makes use of multimedia and principles such as the spatial contiguity principle, coherence principle, modality principle, and redundancy principle to strengthen the learning process. VR-based learning fits the constructivist learning design (Lee & Wong, 2008; Sharma, Agada, & Ruffin, 2013). Constructivist strategies such as situated learning, cognitive apprenticeships, and social negotiation are appropriate for learning tasks demanding high levels of processing, for instance, heuristic problem solving, personal selection, and monitoring of cognitive strategies (Ertmer & Newby, 1993).

Experientialism describes learning as following a cycle of experiential stages, from concrete experience, observation and reflection, and abstract conceptualization to testing concepts in new situations. Experientialism adopts the constructivist's point of view to some extent—e.g., that learning should be drawn from a learner's personal experience. The teacher takes on the role of a facilitator to motivate learners to address the various stages of the learning cycle (Kolb & Kolb, 2012).

Connectivism takes into account the digital-age by assuming that people process information by forming connections. This newly introduced paradigm suggests that people do not stop learning after completing their formal education. They continue to search for and gain knowledge outside of traditional education channels, such as job skills, networking, experience, and access to information, by making use of new technology tools (Siemens, 2014).

Of course, there are many varieties of learning theories in addition to the main paradigms listed here, such as those developed based on the information processing theory or the social cognitive theory. Regardless of which learning theories under each paradigm are used by VR researchers, it is crucial that the development of VR applications for higher education is firmly grounded on existing learning theories because learning theories offer guidelines on the motivations, learning process and learning outcomes for the learners.

3. Related work

In the following section, we first present an overview of literature reviews before discussing the gaps found in existing works.

3.1. Literature reviews

In our search for existing systematic reviews and mapping studies on VR in an educational context, the Scopus digital library returned 59 peer-reviewed articles published between 2009–2018, which included search terms “virtual reality” and “systematic review” in their titles and were linked to *education*, *training*, *teaching*, or *learning*. The most popular application domains covered in these systematic reviews were medicine (78%), social science (15%), neuroscience (11%), and psychology (11%). We filtered the search results further by checking the abstracts of these studies for whether the immersion aspect was also included in them.

We found 18 potentially relevant articles, of which four were very relevant for our work and reported the results of a systematic mapping study. In addition, we found two other thematically relevant articles, one of which described a review and one a meta analysis. Hence, we examine six review articles in total and introduce them shortly to illustrate the research gap we address with our systematic mapping study.

Feng, González, Amor, Lovreglio, and Cabrera-Guerrero (2018) conducted a systematic review of immersive VR serious games for evacuation training and research. The application area was quite specific, i.e., building evacuation and indoor emergencies. The authors proposed the use of serious games, as these have increasingly been adopted for pedagogic purposes. Feng et al. (2018) focused on examining immersion through a virtual environment, where the participants could feel that they are physically inside the artificially simulated environment. The authors identified the pedagogical and behavioral impacts, the participants' experience as well as the hardware and software systems that were used to combine serious games and immersive VR-based training. They reported several pedagogical and behavioral outcomes, which included knowledge of evacuation best practices, self-protection skills, and spatial knowledge. The typical methods used for measuring the learning outcomes in this study were questionnaires, open-ended question interviews, paper-based tests, and logged game data (e.g., evacuation time, damage received). Regarding behavioral impacts, the following points were specified:

- Evacuation facility validation (tested and validated different evacuation facility designs and installations),
- Behavioral compliance (investigated whether participants followed the evacuation instructions),
- Hazard awareness (investigated whether participants could notice hazards in the environment),
- Behavioral validation (validated the hypothetical behavior model),
- Social influence (examined the social influence on evacuation behavior),
- Behavior recognition (identified different behaviors under different evacuation conditions), and
- Way-finding behavior (studied evacuation way-finding behavior).

Participants were observed as to whether they showed fear, engagement, stress, and high mental workload during the training. To track these experiences, the researchers employed an electrodermal activity sensor to evaluate fear and anxiety, a photoplethysmography sensor to obtain blood volume pulse amplitude, and a multichannel physiological recorder to measure the

emotional responses of participants. [Feng et al. \(2018\)](#) also mentioned the characteristics of the game environments identified during the study, i.e., the teaching methods through direct or post-training feedback, navigation, static and dynamic hazard simulation, narratives – such as action, performance or instruction-driven, interactivity/non-interactivity with non-playable characters – as well as audio-visual and motion sense stimulation.

[Wang, Wu, Wang, Chi, and Wang \(2018\)](#) specifically surveyed the use of VR technologies for education and training in the construction engineering field. The authors focused on VR technologies, applications, and future research directions. The study found five major technology categories used in the construction engineering context—desktop VR, immersive VR, 3D game-based VR, building information modeling VR, and augmented reality. The authors noted that desktop VR was used to improve the students' motivation and comprehension. Under immersive VR, the authors subsumed HMDs combined with sensor gloves and suits, the so-called virtual structural analysis program and CAVE systems, where the immersive virtual environment is formed around the user's location by using a 3D immersive VR power wall. [Wang et al. \(2018\)](#) concluded that immersive VR was important for improving concentration and giving trainees a measure of control over the environment. Building information modeling VR has been applied in construction engineering to visualize schedule information and construction work on site, enabling students to interact with building elements in a VR environment, and a system that includes a question-and-answer game to enhance the learning experience. The authors found that VR applications are mostly used in architectural visualization and design education, construction safety training, equipment and operational task training, and structural analysis education. In addition, [Wang et al. \(2018\)](#) revealed five future directions for VR-related education in construction engineering:

- Integrations with emerging education paradigms,
- Improvement of VR-related educational kits,
- VR-enhanced online education,
- Hybrid visualization approaches for ubiquitous learning activities, and
- Rapid as-built scene generation for virtual training.

[Chavez and Bayona \(2018\)](#) looked at VR in the learning context and examined the characteristics that determine successful implementation of this technology as well as its positive effects on learning outcomes. The authors defined 24 characteristics of VR—e.g., interactive capability, immersion interfaces, animation routines, movement, and simulated virtual environment. They discovered that, in certain subjects such as medicine, the “movement” feature is vital for learning about the reaction of a person's body, whereas, in general education, the “immersion interfaces” are often used as a way to learn through “live experience” that is closer to reality. The authors further specified 17 positive effects of VR, including improving learning outcomes, living experiences that are closer to reality, intrinsic motivation, increasing level of interest in learning, and improved skills, although in some fields, such as psychology, no meaningful learning effects were observed.

To some extent, there are similarities between this work and our work; however, clear differences can be identified. First, we specifically focus on immersive HMDs as recent technology developments, while [Chavez and Bayona \(2018\)](#) did not distinguish between immersive and non-immersive VR. Second, we use a systematic mapping method to reveal what VR design elements have been used for teaching different types of learning content, such as declarative knowledge or procedural and practical knowledge. In contrast, [Chavez and Bayona \(2018\)](#) analyzed the characteristics of VR applications on a more abstract level and mapped these using broad application domains such as medicine or psychology.

[Suh and Prophet \(2018\)](#) discussed the state of immersive VR research in their systematic study and, in particular, named current research trends, major theoretical foundations, and research methods used in previous immersive technology research. Regarding research trends, [Suh and Prophet \(2018\)](#) found four popular domains that use immersive technologies: education, entertainment, healthcare, and marketing. They also identified two main research streams. First, studies that examine the user experience and the effects of unique system features of immersive technology. Second, research that scrutinizes how the use of immersive technologies enhances user performance through, for instance, learning and teaching effectiveness, task performance, and pain management. With respect to the theoretical foundations employed in existing immersive technology studies, these authors further identified the flow theory, conceptual blending theory, cognitive load theory, constructive learning theory, experiential theory, motivation theory, presence theory, situated cognitive theory, media richness theory, stimulus–organism–response model, and the technology acceptance model. In terms of the research method, experiments, surveys, and multi-method approaches were frequently used. [Suh and Prophet \(2018\)](#) provided the following classification framework for immersive technology use:

- Stimuli aspect (i.e., sensory, perceptual, and content),
- Organism aspect (i.e., cognitive and affective reactions),
- Response aspect (i.e., positive and negative outcomes), and
- Individual differences in VR use (i.e., gender, age, sensation-seeking tendency, and personal innovativeness).

[Jensen and Konradsen \(2018\)](#) reviewed the use of HMDs in education and training for skill acquisition. The authors examined factors influencing immersion and presence for applying VR for education, the influence of immersion and presence on learning, and situations where HMDs are useful for cognitive, psychomotor and affective skills acquisition. They also studied physical discomfort due to HMD usage and learners' attitudes toward HMDs. The authors stressed the importance of the content, e.g. simulations as learning enablers rather than the HMD itself. Some barriers of using HMDs in education and training were identified, i.e. lack of content and the designs of the HMDs which are more entertainment-oriented than education-oriented. To be relevant for teachers, HMDs should possess user content editing capability.

Table 1
Summary of systematic reviews on VR technologies in an educational context.

Authors	Database source	Keywords search	# Articles and timespan	Focus
Feng et al. (2018)	Scopus, Engineering Village	Virtual reality, VR, virtual environment, virtual simulation	15 out of 567, no time constraint	Indoor building evacuation, pedagogical and behavioral impacts, gaming environment development and outcome, participant experience measures
Wang et al. (2018)	Scopus, Web of Science	Virtual reality, virtual environment, augmented reality, 3D, game	66 out of 347, years 1997–2017	VR and VR-related technologies, their applications, implementation areas, and future directions for construction engineering education and training
Chavez and Bayona (2018)	IEEE Xplore Digital, Science Direct, PsycINFO, Wiley Online Library, ACM Digital Library, Blackwell Publishing, and Google Scholar	Virtual reality, VR, serious game, virtual environment, VE, decision making, influence, education, higher education, professional education, not augmented reality	30 out of 4060, years 1999–2017	The characteristics of VR, teaching subjects and outcomes of VR implementation into the learning process
Suh and Prophet (2018)	Scopus, Social Sciences Citation Index (SSCI)	Immersive technology, augmented reality, virtual reality, mixed reality	54 out of 926, years 2010–2017	Trends, theoretical foundations and research methods in immersive technology research
Jensen and Konradsen (2018)	Scopus, Web of Science, EBSCOhost, PubMed, IEEE Xplore, ERIC, PsycINFO, IBSS	Virtual reality, head-mounted display, education, training, learning.	21 out of 8185, 2013–2017	Skill acquisition (psychomotor, cognitive and affective); documentation of experimental or quasi-experimental studies.
Merchant et al. (2014)	PsycINFO (EBSCO), Medline (Pub Med), Dissertation and Thesis, Eric (EBSCO), Education Full Text, PaperFirst, CINHAL, Manual search in some journals, Google Scholars	Virtual reality, virtual worlds, virtual learning environments, computer assisted learning, AI, mixed reality, synthetic environment, virtual classrooms, augmented reality, immersive learning environment, computer games, game-based learning environment, serious games, simulations; education, learning, instruction, and instructional design	67 out of 7078, all studies until 2011	Desktop-based virtual reality technologies as an assessment, diagnostic, or therapeutic tool; games, simulations or virtual worlds

Despite the focus on immersive VR technologies, the study of Jensen and Konradsen (2018) is still different from ours in several points: First, the work does not analyze immersive VR applications in terms of learning content, design elements, underlying learning theories, and the VR application domains. Second, the findings are limited to declarative knowledge, gesture skills, and emotional control while our study considers other types of learning outcomes as well.

Merchant et al. (2014) conducted a meta-analysis to address the impact of instructional design principles for VR-based instructions. They pointed out improved learning outcomes through games compared to simulations and virtual worlds, but there was an inverse relationship between the number of treatment sessions and learning outcome gains. On the contrary, virtual worlds deteriorate students' learning outcome gains. In simulation studies, the elaborated explanation feedback type is more suitable for declarative tasks whereas knowledge of the correct response is more appropriate for procedural tasks.

We identified distinct contributions compared to our study. First, the review looks at both K-12 and higher education settings. Second, it examines desktop VR technologies that are excluded in our study. Third, the authors examined the suitability of particular VR instruction designs for certain learning outcomes. In other words, our study will be complimentary with the findings of the work of Merchant et al. (2014).

The aspects covered in systematic reviews are summarized in Table 1.

3.2. Gaps in the systematic review of VR for education literature

In order to make a contribution to theory, our study aims to fill gaps in existing literature.

Immersion: Suh and Prophet (2018) focused on immersive VR in their systematic review. However, they also examined several different application areas besides education, such as health care and marketing. Thus, in this article, immersive VR applications in the education field were not discussed in-depth. Likewise, the review by Merchant et al. (2014) stresses the learning outcomes. The review of Feng et al. (2018) only considered serious games as a specific type of immersive VR applications. In contrast, Chavez and Bayona (2018) did not look at immersive VR in particular but rather examined the characteristics of VR in general. Wang et al. (2018) discussed immersive VR among other related technologies, such as non-immersive VR (i.e., desktop VR) and augmented reality. Jensen and Konradsen (2018) do not analyze immersive VR from the perspective of learning content, design elements, underlying learning theories, and the VR application domains.

Table 2
Research Questions of this study.

#	Research Question (RQ)
RQ1	What types of immersive VR technologies are used in higher education?
RQ2	Which research designs, data collection methods, and data analysis methods are applied to examine the use of immersive VR in higher education?
RQ3	What learning theories are applied to examine the use of immersive VR in higher education?
RQ4	Which research methods and techniques are applied to evaluate the learning outcomes of immersive VR usage in higher education?
RQ5	In what higher education application domains are immersive VR applications used?
RQ6	For which learning contents in higher education are immersive VR applications used?
RQ7	What design elements are included in immersive VR applications for higher education?
RQ8	What is the relationship between application domains and learning contents of immersive VR applications for higher education?
RQ9	What is the relationship between learning contents and design elements of immersive VR applications for higher education?

Application areas of VR-based education: While some existing systematic reviews focused on a very specific application area such as construction engineering (Wang et al., 2018) or evacuation training (Feng et al., 2018), others considered education only as one application area among many (Suh & Prophet, 2018). The work of Chavez and Bayona (2018) provided an overview of VR application areas in education but can be distinguished from our work through other aspects (i.e., the focus on immersive VR). Merchant et al. (2014) analyze the literature that is based on the desktop VR technologies.

Overview of teaching content: In addition to the gaps described earlier, all six reviews considered only broad VR application domains in education (i.e., medicine or psychology) but did not shed light on specific types of learning content that can be taught using VR applications (e.g., declarative knowledge or procedural and practical knowledge).

Design elements: While Chavez and Bayona (2018) analyzed the characteristics of VR in an educational context on a more abstract level, the other five reviews did not focus on the design elements underlying the content of HMD-based teaching and education. Furthermore, a mapping of VR design elements that are used for teaching specific types of learning content is missing thus far. Merchant et al. (2014) address a single design principle, i.e. instruction design.

Methods and theories: Suh and Prophet (2018) provided an overview of research methods and theories applied in immersive technologies research but did not particularly focus on education. The other existing reviews focused on entirely different aspects such as application domains. Thus, an overview of the learning theories that are used as a theoretical foundation for studies on VR-based learning is still missing.

Evaluation: Little knowledge has been accumulated on how to evaluate learning outcomes when using VR in teaching activities. While Feng et al. (2018) collected methods for measuring the learning outcomes in the specific application area of evacuation training, the other existing reviews did not consider learning outcome evaluation methods. Even though the technology being reviewed was desktop VR, Merchant et al. (2014) discusses the number of treatment sessions and feedback mechanism as factors affecting the learning outcome. However, the applicability of these treatments for immersive VR needs further examination.

4. Research design

In the following section, we describe our research design, which consists of the method, review process, and classification framework.

4.1. Research method and research questions

Consensus about the design space of VR-based teaching would significantly aid future developments. To this end, we apply a systematic mapping approach to the literature by extracting key information from documents indexed in four scientific digital libraries. Our research aims to obtain an overview of the relationship between application domains and learning contents and also between design elements and learning contents. Based on our results, we propose an agenda for future research and first recommendations for the future development of VR applications for higher education.

Systematic literature reviews, as proposed by Kitchenham et al. (2009) or Webster and Watson (2002), have been widely used as an approach to obtain comprehensive insights into a specific research domain. Furthermore, to answer questions about the structure of a broad field, relevant topics within this field, as well as research trends, Kitchenham et al. (2009) recommend using a mapping study, which is a specific form of a systematic literature review. In contrast to a standard systematic literature review, which is driven by a particular research question, Kitchenham et al. (2009) point out that a mapping study reviews a broader topic and classifies the primary research papers within the specific domain under study. The research questions suggested in such a study have a high level of abstraction and include issues such as: what sub-topics have been discussed, what empirical methods have been used, and what sub-topics have adequate empirical studies to support a more detailed systematic review. An example of applying the mapping study approach to a literature review is the work of Wendler (2012).

To accomplish the objective of this study, we propose several research questions that focus on systematizing and structuring the research on VR applications for higher education. They are listed in Table 2.

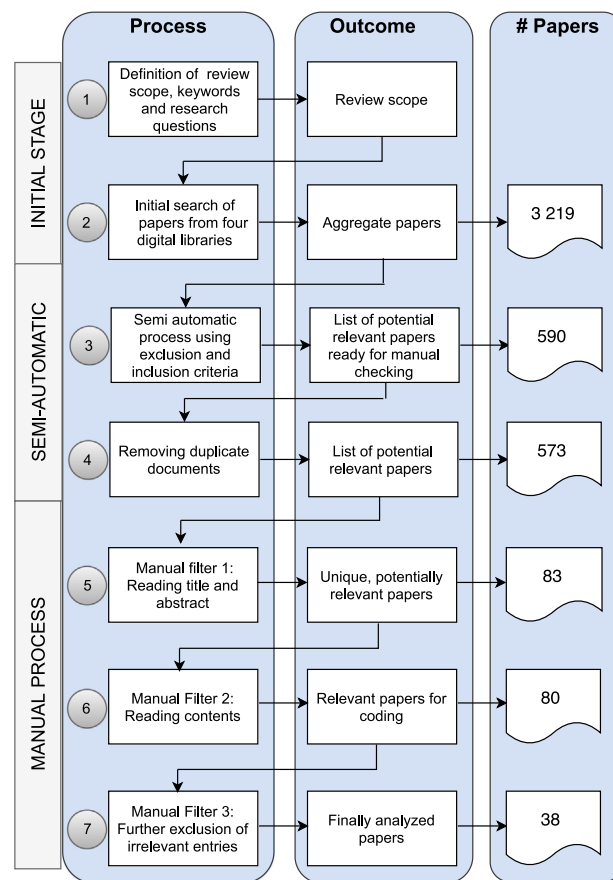


Fig. 1. Review process.

4.2. Review process and literature search method

Our review process included procedures, considerations, and decisions that lead to a consolidated list of articles to be reviewed in-depth. We conducted a systematic mapping study, as suggested by Wendler (2012), to obtain an overview of the field. The overall review process, from defining the review scope to identifying a final selection of articles for analysis, is illustrated in Fig. 1. In total, the article review process consisted of seven steps.

4.2.1. Definition of the review scope, keywords, and research questions (step 1)

While the research questions have been described in Table 2, defining the scope and keywords was quite challenging because, in fact, VR research is extensive and the number of publications in this area is abundant. We followed the procedures, as suggested by Webster and Watson (2002), starting by selecting keyword search strategies in relevant digital libraries. These relevant libraries for our search were the IEEE Xplore Digital Library, ProQuest, Scopus, and Web of Science. IEEE Xplore is a rich repository that covers the domains of computer science, information technology, engineering, multimedia, and other software-related publications. ProQuest comprises articles in the areas of medicine, surgery, and nursing sciences. Scopus provides a wide range of publication domains, covering the fields of technology, natural sciences, information technologies, social sciences, and medicine. The Web of Science indexes social science, arts, and humanities spheres. We were aware of other databases, such as the ACM digital library, Science Direct, JSTOR, EBSCO, and Taylor & Francis, but we expected most of the articles they contain to already be a part of the databases we selected. This expectation was confirmed by exemplary cross-checks.

4.2.2. Initial paper search in four digital libraries (step 2)

For our database search, we defined the following search string²:

² Appendix A (p. 32) gives the search strings used for each database.

```
"virtual reality" OR VR
AND educat* OR learn* OR train* OR teach*
AND "higher education" OR university OR college
AND NOT "machine learning" OR "deep learning" OR "artificial intelligence" OR "neural network"
AND NOT rehabilitation OR therapy
```

Specifying ‘‘higher education’’ OR university OR college was crucial in order to minimize irrelevant application areas, such as VR for primary, secondary, or vocational education. By adding these keywords, we reduced the results of one database from more than 3000 articles to approximately 800 articles. The term NOT (‘‘machine learning’’ OR ‘‘deep learning’’ OR ‘‘artificial intelligence’’ OR ‘‘neural network’’) was added to avoid articles that reported on artificial intelligence without the (human) learning context. In addition, the keywords *rehabilitation* and *therapy* were often associated with physical training, which was also out of scope. The exclusion keywords were obtained after conducting several search tests, followed by a thorough examination of the results.

The search results covered peer-reviewed scientific journal articles and conference papers written in English and published between 2016 and 2018. A Google Trends search revealed an increasing interest in the topic of VR since 2016, when the immersive HTC Vive headset was released. Thus, starting the search from 2016 increased the likelihood of obtaining immersive VR-based learning articles. Due to the novelty of immersive HMDs, the inclusion of conference papers was necessary, as the majority of innovative research and development using HMDs was being documented in conference papers instead of journal articles. The aggregated results of the initial search were 3219 articles. Since we dealt with a considerable number of results, we implemented a two-stage filtering process: (1) semi-automatic filters for the exclusion and inclusion criteria and (2) manual filters to identify potential papers.

4.3. Semi-automatic process

Steps 3 and 4 in Fig. 1 were semi-automatic processes created to exclude and include articles by checking and extracting a list of the most critical words and word clusters from the abstracts. We made use of the list to manually select representative exclusion and inclusion keywords to narrow down the list of articles to review. The goal of this process was to ensure that the articles about immersive VR technologies were appropriately captured.

4.3.1. Exclusion and inclusion method (step 3)

First, we performed a content analysis of all databases using KH Coder 3 that can be used for quantitative content analysis, text mining, and computational linguistic purposes (Coder, 2017). To conduct this analysis, pre-processing was performed by removing punctuation marks, such as periods, commas, and question marks. Stop words (e.g., and, or, of), which provide no additional meaning to a sentence, were also removed. The words with conjugated or inflected forms, such as verbs or adjectives, were reduced to their word stems. For instance, *buy*, *bought*, and *buying*, in a given text, would all be extracted as *buy*. We conducted further pre-processing by only including nouns, proper-nouns, and verbs in the analysis, while ignoring prepositions, adjectives, and adverbs, before conducting word clustering.

KH Coder 3 usage was intended for extracting the collections of articles’ abstracts into word clusters and not in order to conduct complete computational linguistics-based content analysis. The tool provides *TermExtract*, which is a feature that automatically extracts clusters of words that often occur together—e.g., “virtual spatial navigation”. It could happen that the same word is clustered into two clusters, such as “virtual spatial” and “navigation”, or that unintended cluster of words emerge. However, each word cluster was scored and, therefore, the highly scored clusters were deemed to be reliable (see Higuchi (2016) for technicalities of KH Coder 3). Thus, this process is entirely different from simply extracting a frequency word list, where each extracted word is ranked based on the frequency of its occurrence.

This process returned a list of word clusters with scores and we used the 1000 most important words and clusters that resulted from this process to manually select the *good enough* keywords for additional exclusions and inclusions. This process led to a reduced selection of potentially relevant articles, which had to be marked as “relevant” and “not relevant” during the subsequent manual process.

It should be noted that we could have applied more straightforward filtering criteria, such as the number of citations, e.g., to only review papers that received at least twenty citations. However, this would not have been an ideal approach because we might have overlooked or excluded interesting, relevant, recent papers due to a small number of citations. It should be recalled that most collected papers were relatively newly published (between 2016 and 2018).

Second, we went through the list of extracted words and clusters to identify the terms that were useful as additional exclusion and inclusion criteria. Exemplary exclusion words were “augmented reality” and “desktop virtual reality” because we only focused on immersive VR and not on augmented or mixed reality. Certain terms, such as “primary education”, “secondary education”, and “vocational education”, were also noticeable exclusion terms because we only focused on higher education. An exhaustive list of exclusion words is documented in Appendix B (p. 32).

The following inclusion keywords were selected:

```
vr application, headset, glasses, goggles, immersive, immersion, immers, head-mounted display,
head mounted display, oculus, vive, samsung gear, google cardboard, playstation vr,
playstation virtual reality, pimax, google daydream, samsung odyssey.
```

Table 3
Semi-automatic filtering results and manual checking.

Source	Original	Semi-automatic		Manual checking	
		Exclusion	Inclusion	Duplicate	Abstract
IEEE Explore	1 656 articles	899 articles	345 articles	340 articles	35 articles
ProQuest	877 articles	494 articles	168 articles	167 articles	10 articles
Scopus	574 articles	374 articles	60 articles	54 articles	30 articles
Web of Science	112 articles	97 articles	17 articles	12 articles	5 articles
Total	3 219 articles	1 864 articles	590 articles	573 articles	80 articles

These keywords were selected to ensure that our search would return papers that dealt with immersive VR technologies such as HMDs. It should be noted here that, even though we applied a semi-automatic approach at this stage, we performed quality checks by skimming the abstracts to check whether the excluded papers were truly irrelevant.

The word extraction and selection was conducted for each database separately because the results from each database varied significantly. For example, the IEEE database often returned computer science-related articles, while the papers indexed in ProQuest were mostly related to medicine, health, and nursing sciences. These keyword selections were performed by all authors in order to agree on the choices of the inclusion and exclusion terms and to clarify if disagreements occurred.

4.3.2. Removing duplicate documents (step 4)

Duplication check was a rather straightforward process. ProQuest, for instance, consists of multiple databases and sometimes returns two identical articles. By activating ProQuest's automatic removal feature, redundant results were reduced. In this manner, 28 out of 905 articles were purged from ProQuest so that the number of identified articles was reduced to 877. The duplication check of the articles across the databases was done using title-based sorting in an Excel spreadsheet. The aggregated results, after the implementation of the semi-automatic process and duplication removal, were 590 articles, as summarized in [Table 3](#).

4.4. Manual selection process

This process was comprised of three steps—reading the titles and abstracts, reading the contents, and further exclusion of irrelevant articles.

4.4.1. Manual filter 1: Reading the titles and abstracts (step 5)

In step 5, all authors read through the 590 abstracts obtained from the semi-automatic process, marking the articles as either relevant or not relevant. To increase the judgment reliability at this stage, each abstract was read by at least two authors. If there were disagreements, the remaining two authors would also judge the abstracts. This process resulted in 83 articles ready for further processing. The summary of the papers, from the semi-automatic process to the list of articles ready for thorough reading, can be seen in the rightmost column of [Table 3](#).

4.4.2. Manual filter 2: Reading the contents (step 6)

This second manual stage not only served as a filtering process but also as a pre-coding stage. As the coding process involved all authors (four persons), there was a risk that we would code and judge the articles differently. To increase the intercoder reliability of the coding process and to assure that all paper reviewers worked uniformly, all four authors took part in a coding test with twenty selected papers. At this stage, we compared the results and discussed discrepancies in the way we coded the articles until we reached consensus, instead of calculating the discrepancies quantitatively as suggested by [Krippendorff \(2004\)](#) or [Holsti \(1969\)](#). To help the reviewers judge the category of the paper content in a similar way, we added definitions and explanations for each concept that were likely to be a subject of multiple interpretations. At this cross-validation stage, the list of “ready-to-code” papers was already reduced to 80 articles.

4.4.3. Manual filter 3: Further exclusion of irrelevant entries (step 7)

During this stage, we started the actual coding process and continuous reading through all papers allowed irrelevant articles to be further discarded. We continued practicing the intercoder reliability process in which each paper was coded by two persons. Any discrepancies were discussed until we had a set of articles with agreed coding categories ready. In the end, we included 38 relevant articles into our systematic mapping study. They are listed in [Appendix C](#) (p. 32).

4.5. Classification framework for analysis

Before beginning the coding process, we developed a *concept matrix*, as suggested by [Webster and Watson \(2002\)](#), to allow us to identify the learning contents, application domains, and VR design elements in relation to VR-based education. The analysis followed five steps, as can be seen in [Fig. 2](#). A slight overlap with the third stage of the manual filtering papers exists, as manual filtering and coding were performed simultaneously. The concept matrix was a prerequisite to conducting a third manual filtering.

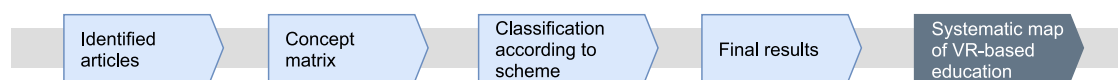


Fig. 2. Classification stages.

Table 4

Definition of the categories.

Categories	Explanation
Empirical, qualitative research	A study that adopts well-established qualitative methodology (Creswell & Creswell, 2017), such as narrative research, phenomenology, grounded theory, ethnography, and case studies.
Empirical, quantitative research	A study that includes elements such as true experiments, with random assignment of subjects to treatment conditions, and less rigorous experiments such as quasi-experimental and correlational approaches. This type of study can also consist of a survey that includes cross-sectional and longitudinal studies using questionnaires (Creswell & Creswell, 2017).
Conceptual	A study that is designed with a specific focus on theoretical advancements (Stolterman & Wiberg, 2010).
Design-oriented	A study that is intended to reveal new knowledge as its primary objective. This is particularly the case if this knowledge is such that it would not have been attainable if the design – the bringing forth of an artifact (e.g., a research prototype) – had not been a crucial element of the research process (Fallman, 2003).
No method explained	A study does not have a recognizable method at all.

4.5.1. Coding

The identified articles, depicted in Fig. 2, were all 80 papers obtained from the second manual filtering. We adapted the concept matrix used by Webster and Watson (2002), who illustrates it as a logical approach that defines several *concepts* (e.g., variables, theories, topics, or methods) that serve as a classification scheme for grouping all relevant articles. Based on existing literature, we developed an initial concept matrix and added new concepts to it during the classification process.

Traditionally, the systematic literature review process is often done directly on a spreadsheet, including the list of identified articles and theoretical concepts for analysis. While this method was also considered for this study, we wanted to avoid handling a large matrix when conducting the coding with several authors at the same time, thus increasing the risk of making errors. Admittedly, working directly on the spreadsheet has some advantages as well; for example, it would have been easier to make changes when new concepts were introduced during the coding process.

However, at the same time, working with a spreadsheet can be overwhelming—in our case, 134 concepts were extracted from the literature. Consequently, as an alternative, we used an online questionnaire that was designed in a way that was consistent with the concept matrix we developed. Each page of the online questionnaire consisted of one analysis framework with corresponding concepts. The questionnaire form also enabled us to include the definitions and explanations of each concept, which was helpful to ensure a unified coding process. We further added an open text field for each framework. In this manner, the coders could propose new concepts or add notes about their coding decisions. In summary, the method ensured more consistent coding results and made it easier to work simultaneously during the coding process. Eventually, the results were stored in a single spreadsheet for analysis. For consolidating the results, we used the classification frameworks as described in Section 4.5.2 to 4.5.5.

4.5.2. Research design framework

We used existing research design categories without proposing new ones. Different authors have different ways of categorizing research designs, depending on the research perspective being used. Kumar (2019), for instance, divides research based on the study results' applications, the objective of the study, and the mode of enquiry used. According to Kumar (2019), research can either be pure or applied from an application perspective and can be descriptive, exploratory, correlational, or explanatory from a research objective perspective. From the mode of enquiry perspective, research can either be structured or unstructured. Creswell and Creswell (2017) divides research designs into qualitative, quantitative, and mixed methods, while Spector and Spector (1981) distinguishes between experimental and non-experimental research designs. In this study, we adapted the work of Wendler (2012), who classifies research design into the following categories: empirical qualitative research, empirical quantitative research, conceptual research, and design-oriented research. The definition of each category is given in Table 4.

4.5.3. Data collection methods and data analysis framework

We combined existing, established data collection methods both in qualitative and quantitative studies with our findings during the review process. The categories are compiled in Table 5.

The data analysis framework emerged through the coding process, as we extracted only the methods that were explicitly mentioned by authors. We further added the category of descriptive statistics, which was chosen when authors of the papers from our search presented their results by reporting frequencies, percentages, means, and standard deviations. The following data analysis methods were mentioned by the authors: descriptive statistics, t-test, correlation, analysis of variance (ANOVA), chi-square test, Fisher's exact test, Mann–Whitney U test, Mc Nemar's Test, multilevel linear modeling, qualitative analysis technique, and analysis of co-variance (ANCOVA). If the study did not have recognizable data analysis at all, then “No Method Applied” was chosen.

Table 5
Data collection methods.

Categories	Explanation
Development	A design- or development-oriented study that documents the overall development process.
Experimental design, usability, user testing	A study that is designed as an experimental study—e.g., a study that compares the performance of two groups or a study that includes usability or user testing.
Survey	A study that collects data from questionnaires, either paper-based or in the form of an online survey.
Interview, focus group discussion	A study that collects or explores the attitudes, opinions, or perceptions toward an issue, a service, a technology, or an application that allows for an open discussion between members of a group. Studies that collect opinions from individual subjects are also included here.
Observation	A study that collects and records information descriptively by observing the behaviors or social interactions of a subject or a group, either in an obtrusive or non-obtrusive way.
Case study, action research	A study that intends to improve and advance practice and is conducted in an iterative way, through identifying areas of concern, developing and testing alternatives, and experimenting with a new approach. This category also includes case studies in which the information is collected from a bounded system, a population, or a specific entity.
Literature Review	A study that collects information from existing literature and uses a systematic method to synthesize the results.
Mobile sensing	A study that collects information from mobile sensors.
Interaction log in VR app	A study that collects information from a developed VR app (e.g., user activities) and uses the resulting interaction log for analysis.
No method applied	A study that does not have a recognizable data collection method at all.

Table 6
Learning theory framework.

Categories	Explanation
Behavioral learning	When students receive either rewards or punishments for correct or incorrect answers and can thus learn the consequences of certain behavior (Shuell, 1986; Skinner, 1989). This applies to VR applications that include a system that allows the students to learn—e.g., responses resulting in satisfying (rewarding) consequences or responses producing annoying (punishing) consequences or learning what the consequences are for following or not following the rules. The students learn when their responses produce certain outcomes and allow them to adapt to their environments.
Experiential learning	When students learn through hands-on experience and using analytical skills to reflect on their experience (Kolb & Kolb, 2012). These reflections lead to changes in judgment, feelings, or skills of the student.
Generative learning	When students engage in cognitive processing during learning, including selecting (i.e., paying attention to relevant incoming information), organizing (i.e., mentally arranging the information into a coherent structure), and integrating (i.e., connecting the verbal and pictorial representations with one another and with relevant prior knowledge activated from long-term memory) (Parong & Mayer, 2018).
Operational learning	When students learn how to construct or assemble an object—as they do in Zhou, Ji, Xu, and Wang (2018)), where the students can interact, select, grasp, move, point, and place objects to learn computer assembly.
Game-based learning	When students learn through a gamification process, i.e., the use of game design elements and mechanics—such as points, levels, and badges—and game dynamics—such as rewards, statuses, and competition—in the learning process. An example of this is found in the work of Bryan, Campbell, and Mangina (2018), where the authors include gamification in the VR application, allowing the students to travel to different countries around the world, explore these locations, learn facts, and answer questions.
Contextual learning	When students learn by emphasizing the context, i.e., the set of circumstances that are relevant for the learners to build their knowledge. Hence, the learning content can assist in guiding students toward developing insights through balanced, organic, and successful environments and strategies. An example of this is a setting in which students use VR that encourages a more complex and higher level of thinking in order to improve their phonological, morphological, grammar, and syntax knowledge, as applied by Chen (2016).
Jeffries simulation theory	When students learn through a simulation process and experience in a trusted environment that is incorporated in the VR design (Jeffries, Rodgers, & Adamson, 2015).
Cone of learning theory	When students learn through both active and passive learning that involves direct, purposeful learning experiences, such as hands-on or field experience. According to this theory, students learn best when they go through a real experience or that experience is simulated in a realistic way. This theory is also known as Dale's Cone theory (Dale, 1969).

4.5.4. Learning framework: Learning theory and learning content

Similar to the data analysis framework, the learning theory framework emerged through the coding process and we only extracted those learning theories that were explicitly mentioned by authors. Details are given in Table 6. In contrast, Table 7 was first populated with initial concepts but, then, new categories were gradually established and incorporated into the framework during the coding process. The first four types of learning content emerged from the literature (cf. with the works of Anderson (1982) and Crebert, Bates, Bell, Patrick, and Cragolini (2004)).

Table 7
Learning content framework.

Categories	Explanation
Analytical and problem-solving	Whether the use of VR can encourage students to improve their analytical skills, such as collecting and analyzing data, writing computer programs, or making complex decisions.
Communication, collaboration, soft skills	Whether the use of VR is intended to strengthen the students' ability to work in a team or whether students can improve their communication skills (e.g., presenting in front of an audience). This category also includes soft skills, such as management and leadership competencies.
Procedural-practical knowledge	Where the use of VR aims to assist students with internalizing procedures, such as knowing how to perform a surgery or how to perform firefighting procedures.
Declarative knowledge	Where the use of VR is intended to help students memorize factual knowledge (e.g., theoretical concepts and scientific principles). This includes, for example, learning the names of planets in our solar system.
Learning a language	Where the use of VR aims to improve students' foreign language capabilities, such as reading, listening, writing, and speaking.
Behavioral impacts	Where the use of VR aims to change the behavior of students by, for example, improving their learning habits, awareness of mobbing, and compliance to rules.
Others	Articles that could not be classified into the alternative concepts above.
Not specified	Where there is no statement or implicit information about the expected learning outcome of VR usage.

4.5.5. Design element framework

As in the learning framework, the design elements of VR for education also emerged from articles during the review process. Some of the categories had previously been proposed by [Wohlgenannt, Fromm, Stieglitz, Radianti, and Majchrzak \(2019\)](#) but our work, as listed in [Table 8](#), provides an extensive update.

5. Results and analysis

In the following section, the results of the systematic mapping study are described according to the research questions (see [Table 2](#), p. 9). Of the 38 articles included in our analysis, 68% originated from conferences, whereas 32% were published in journals (cf. [Fig. 3](#)). Our search window was 2016–2018. However, we also included one paper that was issued in 2015 because the acceptance for the online publication of this paper was 2015 but the actual journal publication date was in 2016. The number of journal publications in 2018 was remarkable, indicating an increasing scholarly interest in VR for higher education.

5.1. VR technologies (RQ1: What types of immersive VR technologies are used in higher education?)

As illustrated in [Fig. 4](#), our review shows that 76% of the studies used high-end HMDs, such as Oculus Rift or HTC Vive. Many of these high-end VR systems use various supporting tools, such as controllers, touchpads, and haptic feedback. Out of 41 VR technology counts, eight used low-budget mobile VR, for instance, [Ye, Hu, Zhou, Lei, and Guan \(2018\)](#) used a smartphone and Google Cardboard for the VR environment. However, interactive manipulation was performed through a desktop monitor connected to the mobile app.

Only a few of the articles used enhanced VR, for instance, [dela Cruz and Mendoza \(2018\)](#), [Veronez, Gonzaga, Bordin, Kupssinsku, Kannenberg, Duarte, et al. \(2018\)](#), and [Pena and Ragan \(2017\)](#). For example, [Veronez et al. \(2018\)](#) used an additional G27 Racing Wheel to control the VR environment. 2% of the articles did not specifically mention the VR technology employed. It should also be noted that, in some experiments, two technologies were used, as in [Bujdosó, Novac, and Szimkovics \(2017\)](#), [Webster and Dues \(2017\)](#) and [Buñ, Trojanowska, Ivanov, and Pavlenko \(2018\)](#), leading to a higher count than that of the number of papers. Overall, high-end HMDs were the most commonly used immersive VR technology.

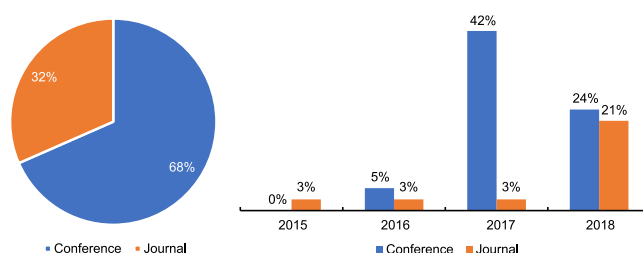


Fig. 3. Publication outlets and publication years of the 38 selected papers.

Table 8
Design element definition.

Categories	Explanation
Realistic surroundings	The virtual environment is of high graphic quality and has been designed to replicate a specific environment in the real world. For example, this applies to medical students who develop their surgery skills in an authentic-looking operation room.
Passive observation	Students can look around the virtual environment. This design element also applies to applications in which users can travel along a predefined path and look around while doing so. However, they are neither able to move around on their own nor to interact with virtual objects or other users.
Moving around	Students can explore the virtual environment on their own by teleporting or flying around.
Basic interaction with objects	Students can select virtual objects and interact with them in different ways. This includes retrieving additional information about an object in written or spoken form, taking and rotating it, zooming in on objects to see more details, and changing an object's color or shape.
Assembling objects	Students can select virtual objects and put them together, including the creation of new objects by assembling several individual objects.
Interaction with other users	Students can interact with other students or teachers. The interaction can take place in form of an avatar and via communication tools such as instant messaging or voice chat. This design element also includes the possibility of students visiting each other's virtual learning spaces.
Role management	The VR application offers different functionalities for different roles. A distinction is made between the role of a student and the role of a teacher. For a teacher, the VR application offers extended functionalities, such as assigning and evaluating learning tasks or viewing the learning progress of students.
Screen sharing	The VR application allows students and teachers to stream applications and files from their local desktop onto virtual screens. This allows them to share and edit content from their local desktops with other users in the virtual environment (e.g., PowerPoint, Google Drive, and Google Docs).
User-generated content	Students can create new content, such as 3D models, and upload this new content to the virtual environment. This design element also applies when the user-generated content can be shared with other users so that they can use it in their virtual environment as well. This design element does not apply when students can only access virtual objects that were created by developers and provided by a library in the virtual environment.
Instructions	Students have access to a tutorial or to instructions on how to use the VR application and how to perform the learning tasks. The instructions can be given by text, audio, or a virtual agent. This design element does not apply when students have to discover how to use the virtual environment or how to perform learning tasks on their own.
Immediate feedback	Students receive immediate textual, auditory, or haptic feedback. The feedback informs students about whether they have solved the learning tasks correctly and whether interactions with virtual objects were successful. In some cases, feedback may also be provided by simulating the results of an interaction with virtual objects, for example, when the corresponding chemical reaction is simulated after chemicals have been mixed in a virtual laboratory.
Knowledge test	Students can check their learning progress through knowledge tests, quizzes, or challenges.
Virtual rewards	Students can receive virtual rewards for successfully completing learning tasks. Students can be rewarded virtually by receiving achievements, badges, higher ranks on a leader board, and by unlocking exclusive content, such as hidden rooms or additional learning content.
Making meaningful choices	Students learn in the virtual environment through participating in a scenario (role-playing) that can end in different ways. In this scenario, they have to make decisions that affect the outcome of the scenario. This design element does not apply when the students' decisions have no influence on the outcome of the scenario.

5.2. Research designs (RQ2: Which research designs, data collection methods, and data analysis methods are applied to examine the use of immersive VR in higher education?)

Concerning the research method, *development* research was popular in 26 articles, as seen in the upper-left bar chart of Fig. 5, followed by experimental design, usability, and user testing (18 articles) as well as survey (16 articles). The bar chart in the upper-right of Fig. 5 shows the data analysis method. Both qualitative and quantitative data analysis methods were applied. The t-test (16) was the most commonly applied quantitative data analysis method. Conversely, other methods were rarely used, in only one to three articles on average, except for the correlation method (5 articles). However, it was a bit counterintuitive that most papers had no explicit or recognizable data analysis method (18 articles). Only four papers used qualitative data analysis methods, which included observations and focus group discussions.

We looked deeper into the relationship between the research design, research method, and data analysis method, as presented in two bubble charts in the lower part of Fig. 5. It should be noted here that the number of articles in the bubble charts adds up to a higher number than shown in the bar charts because one article may contain more than one category of research design. The research design categories in the middle can be read in combination with the bubble charts on the left and the right sides. The bubble charts depict the concentration of the studies. The combination of the data collection method and the research design is shown in the bottom left bubble chart of Fig. 5.

It shows that design-oriented, empirical quantitative, and empirical qualitative research were the dominant research contents. Most of the design-oriented works had combined their studies with development (24 articles), experimental design, usability, and user testing (11 articles), or had used interviews and focus group discussions to collect data (11 articles). This pattern is similar to other empirical qualitative and empirical quantitative studies. We discovered very few studies that were designed as conceptual

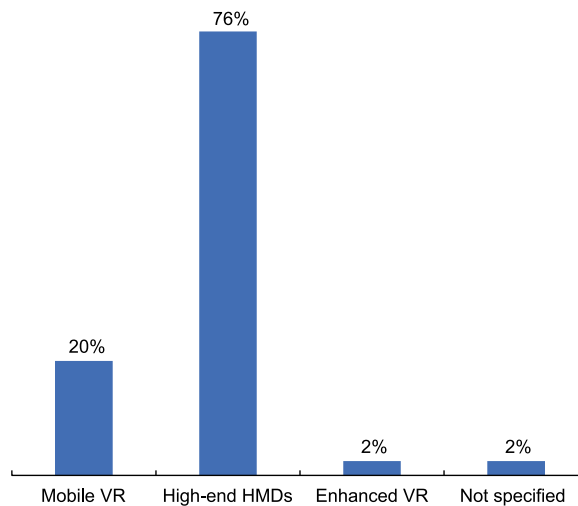


Fig. 4. VR technologies being used (n=41).

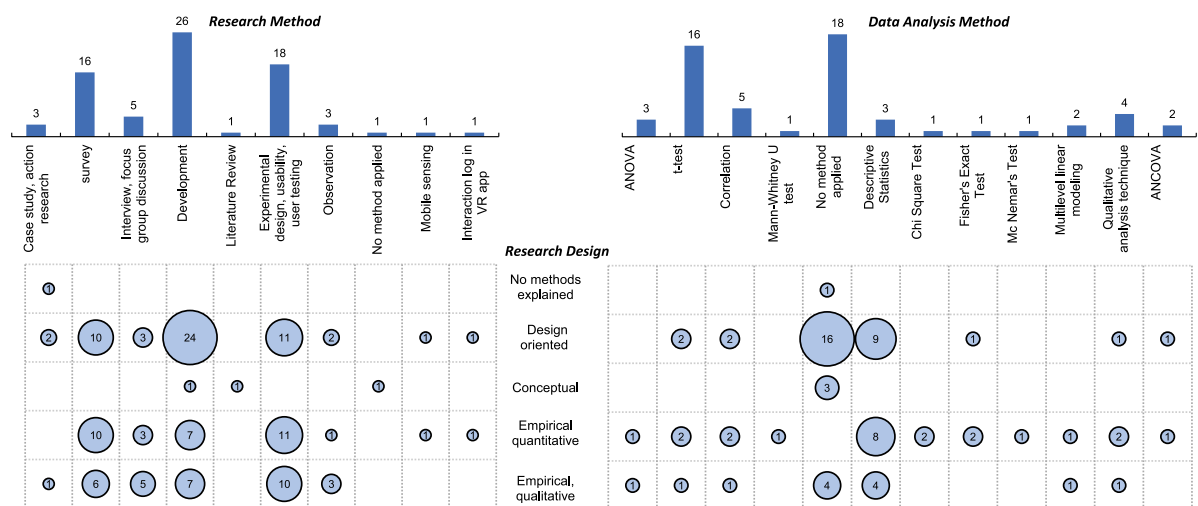


Fig. 5. Research design, data collection method, and data analysis method.

and these used a development approach (Chin et al., 2017), a literature review (Ekkelenkamp, Koch, de Man, & Kuipers, 2016), or did not mention any data collection method.

The bubble chart on the opposite side shows the combination of research design and data analysis method. Interestingly, most studies that reported no data analysis methods were design-oriented (16 articles). The majority of studies that employed a design-oriented research approach primarily reported only descriptive statistics (9 articles). Apart from this, empirical quantitative research designs employed a wide range of identified data analysis methods, with less visible concentrations. In addition, several of the empirical qualitative studies reported descriptive statistics as well (4 articles). Furthermore, we identified that this category had also employed ANOVA, t-test, correlation, multilevel linear modeling, and other quantitative data analysis methods. Four articles mentioned no explicit data analysis method. Likewise, of the conceptual research design articles, three had not applied any method.

5.3. Learning theories (RQ3: What learning theories are applied to examine the use of immersive VR in higher education?)

Fig. 6 indicates that the literature which claims to create VR content or design VR applications for higher education surprisingly lacked reference to explicit learning theories. Thus, the “not mentioned” category consists of 68% of all articles. It should also be noted that, during the coding, we avoided “reading between the lines” and only extracted the learning theories that were explicitly mentioned by study authors as their theoretical foundation. Among the articles that had a theoretical foundation, experiential learning accounts for 11% of all articles, while each of the remaining theories accounts for 3%. The other category consists of an article that explicitly mentions different learning theories, such as the constructivism learning theory, flow theory, gamification

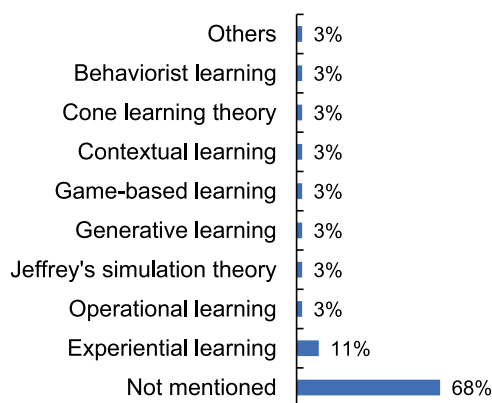


Fig. 6. Applied learning theories (n=38).

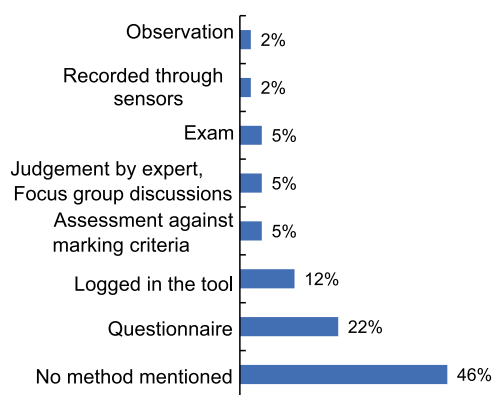


Fig. 7. Evaluation methods.

learning, transfer of learning in the literature review section, but its authors did not really claim which learning theory they prefer to use (e.g. [Chen, 2016](#)). The cone of learning theory was kept as a separate category, given the fact that authors explicitly used this theory in all development VR processes. Recent literature has grouped the cone of learning theory with that of experiential learning, (e.g., ([Garrett, 1997](#))). However, many recent authors are still using this theory on its own ([Davis & Summers, 2015](#)) and, therefore, we placed it in a separate category.

5.4. Learning outcome evaluation (RQ4: Which research methods and techniques are applied to evaluate the learning outcomes of immersive VR usage in higher education?)

Almost half of the reviewed articles did not specify a learning outcome evaluation method (see [Fig. 7](#)). A few articles used questionnaires (22%) or user activities while logged into the VR application (12%). Exams, expert judgments, or focus group discussions accounted for a mere 5%, respectively, while the remaining articles used observations or sensor data. This fact is intriguing because many articles described evaluations of developed VR applications, however, the focus was mainly placed on usability or user experience. Nevertheless, there were also several articles that actually measured or evaluated how much the students' knowledge or skills progressed after the use of immersive VR, for example, [Farra, Smith, and Ulrich \(2018\)](#) and [Zhang, Suo, Chen, Liu, and Gao \(2017\)](#).

5.5. Application domain (RQ5: In what higher education application domains are immersive VR applications used?)

Engineering was the most popular application area identified in 24% of the articles (see [Fig. 8](#)). The next most popular application domains were computer science (10%) and astronomy (7%). The total number adds up to more than 38 articles because we found several articles that fell into more than one category. Furthermore, some articles were very generic and did not mention a specific domain (12%). This applied to the works by [Hu, Su, and He \(2016\)](#), [Webster and Dues \(2017\)](#), [Yang, Cheng, and Yang \(2016\)](#), [Zizza et al. \(2019\)](#), and [Misbhauddin \(2018\)](#).

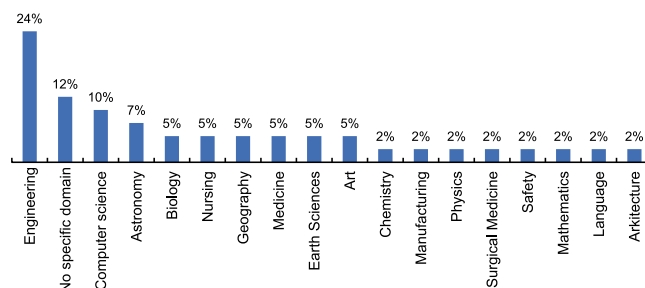


Fig. 8. Application domain (n=42).



Fig. 9. Learning content.

5.6. Learning content (RQ6: For which learning contents in higher education are immersive VR applications used?)

Fig. 9 shows that VR applications for higher education were most frequently used to teach: procedural-practical knowledge (33%), such as filing a report (Pena & Ragan, 2017) or extinguishing fires (Zhang et al., 2017); declarative knowledge (25%), such as learning planet names (Papachristos, Vrellis, & Mikropoulos, 2017) or theoretical concepts in pneumatics (de la Cruz & Mendoza, 2018); and analytical and problem-solving skills (12%), such as diagnosing patients (Harrington et al., 2018) or learning how to code (Román-Ibáñez, Pujol-López, Mora-Mora, Pertegal-Felices, & Jimeno-Morenilla, 2018). The rest of the learning content categories found in the literature were communication, collaboration, and soft skills (10%), behavioral impact (6%), and learning a language (2%). The categories others and not specified accounted for 6% and 4%.

5.7. Design elements (RQ7: What design elements are included in immersive VR applications for higher education?)

Basic interaction and realistic surroundings were the most frequently used design elements, accounting for 24% and 17% of the articles, respectively (see Fig. 10). Immediate feedback and instructions share an equal percentage, i.e., 10%. The next popular element is interaction with other users, which accounted for 9% of all category counts. Passive observation and assembly also shared an equal percentage of representation at 8% each. The rest of the design elements were included in 1%–4% of all papers, i.e., moving around (4%), user-generated content (3%), virtual rewards (2%), role management (2%), and knowledge test (2%). Both screen sharing and making meaningful choices accounted for 1% of all papers.

5.8. Mapping 1 (RQ8: What is the relationship between application domains and learning contents of immersive VR applications for higher education?)

The bubble chart in Fig. 11 links learning contents and application domains. The magnitude of the bubble corresponds to the number of articles found for a given combination. The number of articles was too small to generate clusters and such clustering would also be limited by the fact that application domains are not necessarily independent. Nevertheless, the mapping allowed for observations, although we could not generalize the relationships between the learning content and the application domain because

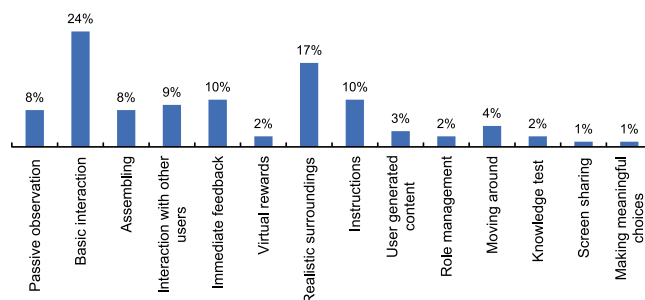


Fig. 10. Identified design elements.

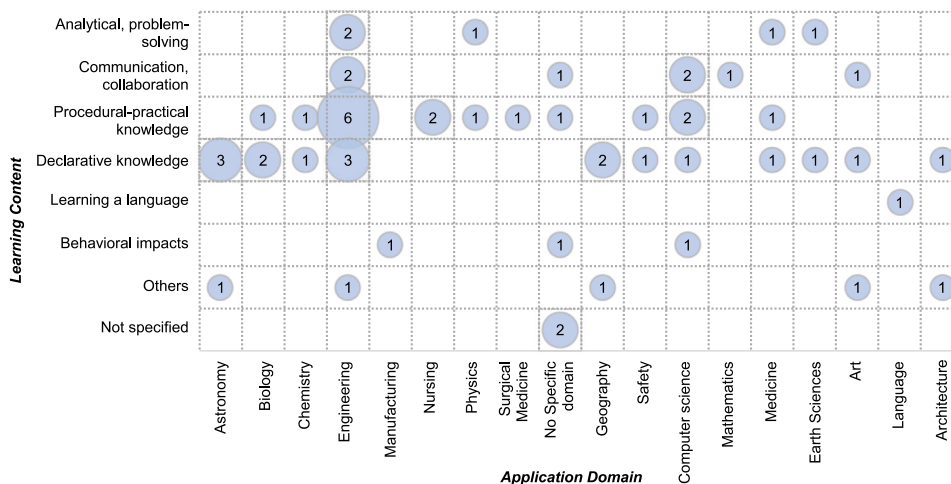


Fig. 11. Relationship between learning content and application domain.

too few relevant papers would be included. At the same time, learning contents were spread among application domains, indicating a wide variety of learning contents across disciplines.

The matrix in Fig. 11 is relatively sparse. While this, undoubtedly, can partly be explained by our attempt at making clear distinctions, it also hints at the experimental nature of the field. Moreover, we observed that no reasonable grouping of application domains – for example, biology, chemistry, and physics as natural sciences – would yield additional insights.

From the mapping, we observed that astronomy used either declarative knowledge (Němec, Fasuga, Trubač, & Kratochvíl, 2017; Rosenfield et al., 2018; Veronez et al., 2018) or “others”, which included learning content for visualization and motivation. Both biology and chemistry paid attention to declarative knowledge and procedural or practical knowledge (e.g., the articles by AlAwadhi et al. (2017), Parong and Mayer (2018)). Since engineering is an application domain frequently described in the literature, it allowed for a diverse assessment. Engineering used the following learning contents: analytical and problem solving (Harrington et al., 2018; Parong & Mayer, 2018), communication and collaboration or soft skills (Hickman & Akdere, 2018; Ye et al., 2018), procedural and practical knowledge (AlAwadhi et al., 2017; Buñ et al., 2018; dela Cruz & Mendoza, 2018; Román-Ibáñez et al., 2018; Schroeder, Bailey, Johnson, & Gonzalez-Holland, 2017), and declarative knowledge (AlAwadhi et al., 2017; Němec et al., 2017; Ye et al., 2018). The only manufacturing article we identified described behavioral impacts (Carruth, 2017).

Nursing (Farra et al., 2018; Smith et al., 2018), physics (Pirker, Lesjak, & Guetl, 2017), and surgical medicine (Ekkelenkamp et al., 2016) mainly focused on procedural knowledge. The articles on geography proposed an application to improve declarative knowledge (Bryan et al., 2018; Parong & Mayer, 2018). Similarly, for safety (Zhang et al., 2017), surgical medicine (Harrington et al., 2018), earth science (Chin et al., 2017; Gerloni et al., 2018), and architecture (Němec et al., 2017), articles describing declarative knowledge as the learning content could be identified, among diverse other aims. Not surprisingly, the only article we identified in the domain of language focused on teaching the skill of language learning (Chen, 2016).

Diverse pictures can be drawn for computer science and art. Art papers (Cortiz & Silva, 2017; Song & Li, 2018) varied in terms of learning content, as did the computer science articles. However, they emphasized collaboration and communication (Bujdosó et al., 2017; Hickman & Akdere, 2018) as well as procedural and practical knowledge (Parmar, Isaac, Babu, D'Souza, Leonard, Jörg, Gundersen, & Daily, 2016; Zhou et al., 2018).

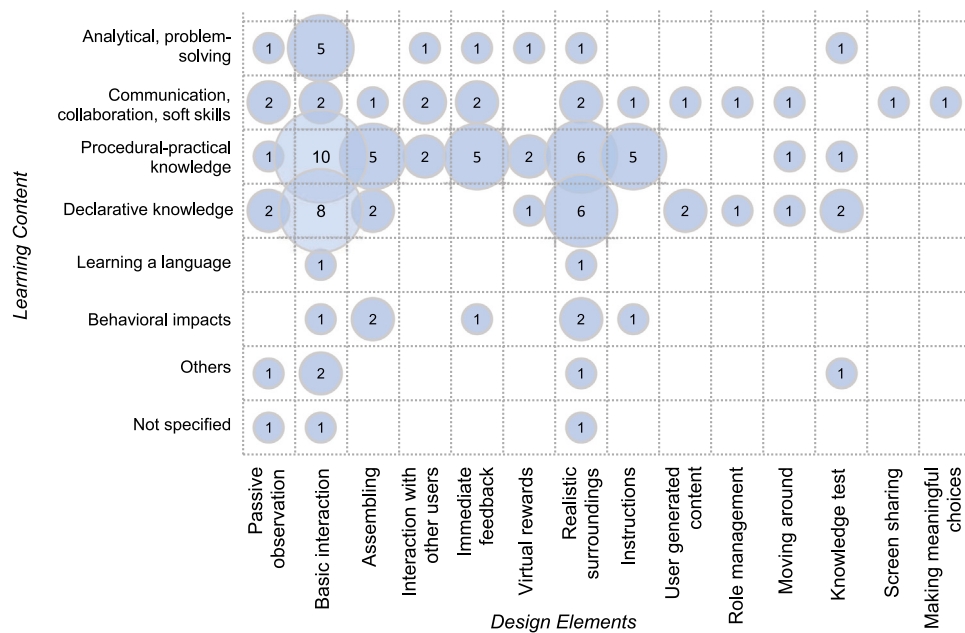


Fig. 12. Relationship between learning content and design elements.

5.9. Mapping 2 (RQ9: What is the relationship between learning contents and design elements of immersive VR applications for higher education?)

Fig. 12 illustrates the overall mapping results of design elements and learning contents. In this way, we analyzed whether there were patterns of certain design elements that were used in VR applications intended to teach specific learning contents. From Fig. 12, we can conclude that basic interaction and realistic surroundings were used for each type of learning content. These tendencies can be observed in the distribution of the bubbles spreading across all learning content categories. However, this observation is reasonable because most VR applications would need to have at least basic interaction elements and to be realistic enough to increase the immersive experience of users.

From the design element perspective, the basic interaction element was mostly applied in the procedural or practical knowledge-oriented VR applications (10 articles) (AlAwadhi et al., 2017; Buñ et al., 2018; dela Cruz & Mendoza, 2018; Farra et al., 2018; Harrington et al., 2018; Ibanez, Kloos, Leony, Rueda, & Maroto, 2011; Pena & Ragan, 2017; Pirker et al., 2017; Smith et al., 2018; Zhang et al., 2017) and in the declarative knowledge-oriented VR applications (8 articles) (AlAwadhi et al., 2017; Bryan et al., 2018; Dolezal, Chmelik, & Liarakapis, 2017; Hu et al., 2016; Im, An, Kwon, & Kim, 2017; Némec et al., 2017; Rosenfield et al., 2018; Zhang et al., 2017). Likewise, the realistic-surroundings element was adopted for teaching procedural or practical knowledge and declarative knowledge contents. There were six articles in the former category (AlAwadhi et al., 2017; Buñ et al., 2018; Ekkelenkamp et al., 2016; Im et al., 2017; Smith et al., 2018; Zhang et al., 2017) and six articles the latter category (AlAwadhi et al., 2017; Chin et al., 2017; Némec et al., 2017; Rosenfield et al., 2018; Ye et al., 2018; Zhang et al., 2017).

Furthermore, from the learning content perspective, seven or more design elements were applied by VR applications intending to teach the following learning contents: communication, collaboration, and soft skills (11 elements), procedural or practical knowledge (10 elements), declarative knowledge (9 elements), and analytical and problem solving skills (7 elements). The remaining categories used four or fewer design elements. On average, an article addressed one to five design elements. Hence, the total number of design elements per category could be higher than the number of articles. For instance, basic interaction and immediate feedback elements were identified in the study of Román-Ibáñez et al. (2018). Six design elements were recognized in the article by Dolezal et al. (2017)—basic interaction, interaction with users, immediate feedback, instructions, role management, and moving around. Zhang et al. (2017) adopted five design elements—basic interaction, immediate feedback, virtual rewards, realistic surrounding, and knowledge test. The majority of the articles in each category described more than two design elements.

Only a few articles used the following design elements: user-generated content, role management, moving around, knowledge test, screen sharing, and making a meaningful choice; these are mostly found in articles that contain multiple design elements.

There is no single rule about what design elements are valid for specific learning contents. Basic interaction, in particular, was found in all, regardless of the learning contents being addressed. AlAwadhi et al. (2017), Rosenfield et al. (2018), Zhang et al. (2017), and Némec et al. (2017) are examples of articles that aimed for declarative knowledge. Rosenfield et al. (2018) proposed the Worldwide Telescope 3D application for exploring the planetary surface, elevation maps, orbital path, ephemerides, and solar systems in a realistic manner. Zhang et al. (2017) built a VR environment for learning related to fire safety – fire type, fire hazard,

fire alarm – and for providing knowledge about various safety evacuation signs in a virtual campus. AlAwadhi et al. (2017) suggested a VR design in order to conduct experiments, watch pre-recorded lectures, perform a campus tour, and recognize the components of informative labs. Both papers also contained practical–procedural content, i.e., how to conduct experiments (AlAwadhi et al., 2017) and how to prevent fire (Zhang et al., 2017). Némec et al. (2017) presented multiple realistic VR environment objects, such as human skeletons, solar system planets, architecture, machinery rooms, and power plants.

The articles by Smith et al. (2018) and Buñ et al. (2018) are examples of studies that aimed at practical–procedural knowledge, containing both basic interaction and realistic surrounding design elements. Buñ et al. (2018) used VR for improving the users' procedural knowledge regarding assembly operation in a manufacturing activity. The basic interactions could be recognized in the possibilities given to users to manipulate objects in a realistic manufacturing environment. Smith et al. (2018) applied VR to teach decontamination skills—e.g., how to pick up a patient's freshly cut off, contaminated clothing, and place them in a nearby biohazard bin. It was tailored with a realistic emergency room environment (Smith et al., 2018).

There were several design elements found in the analytical- and problem solving-oriented literature. Román-Ibáñez et al. (2018) described learning about robotic arm trajectories and collision detection with immediate feedback. Pena and Ragan (2017) suggested improving the analytical skill of recognizing accidents in an industrial setting by allowing users to move around in a virtual environment. Harrington et al. (2018) exploited gesture detection to increase the analytical skills used for decision-making when working with critically injured patients. Gerloni et al. (2018) proposed the ARGO3D platform for learning about geological information and geohazards, such as the crater of an active volcano, in a realistic environment. The platform allows students to navigate inside the environment, fly around, and take pictures. Veronez et al. (2018) developed a VR system to teach the analytical skills necessary for assessing road safety, while Pirker et al. (2017) provided an example of how VR could improve analytical skills through experimentation in a virtual physics lab.

To improve communication, collaboration, and soft skills, several design elements were implemented in the literature. Bujdosó et al. (2017) suggested MaxWhere as a VR collaborative arena (ViRCA) for encouraging interactions with other students. In this example, it is possible to assign tasks to others, generate new content, and work in shared documents. Zizza et al. (2019) suggested the VRLE platform, involving one instructor and a group of students, to improve social competencies and social interaction. The users interact with the environment, as well as with instructors and peers, using virtual hands and voices. Ye et al. (2018) used assembling elements as part of a soft skill to comprehend the control parameters in the automatic control system and to detect abnormalities of the system.

Cortiz and Silva (2017) used passive observation to learn art history through VR in which students can add more content into the learning material, such as artwork from a well-known artist. Hickman and Akdere (2018) reported a conceptual work that describes a possible VR platform for improving intercultural leadership, which included design elements such as passive observation, immediate feedback, and realistic surroundings.

There were 17 papers identified for improving procedural and practical knowledge. For example, Muller, Panzoli, Galaup, Lagarrigue, and Jessel (2017) adopted assembling as the main design element to teach the procedure of setting up a tailstock, palpating, and operating machine. An information and instruction set appears in the VR menu as part of its design elements. Im et al. (2017) used VR for technical training. Misbhauddin (2018) suggested VREdu for providing an immersive lecture, where the platform incorporates role management elements that allow different access for students and instructors. It was streamed to users using the chunking method, i.e., a segment of the video was saved, uploaded to the server, and broadcast.

There were 13 papers that sought to improve declarative knowledge. Dolezal et al. (2017) developed a collaborative virtual environment for geovisualization, where a virtual world is shared by users with avatars through a computer network. This virtual environment supports role management by providing different user rights to teachers and students. The tutorial was important in this work, as part of the instruction design element. The immediate feedback element could be recognized in the possibility provided to the teacher to press a button and reveal the correct and incorrect answers. Zhang et al. (2017) developed fire safety VR with mostly basic interaction elements. Bryan et al. (2018) introduced Scenic Spheres, which allows students to travel around the world, learn facts, and get a feeling about local culture.

Only a few articles focused on learning a language, pedagogical impacts, behavioral impacts, other areas (not falling into an existing category), and areas not specified (having no recognizable learning content). Typically, such works had multiple design elements. Only one paper proposed learning language content (Chen, 2016) and used the realistic surroundings element that captures a situation to be used to learn morphology, phonology, grammar, and syntax. It also used some basic interaction elements in VR such as selecting an object.

The articles aiming to change perceptions, attitude, and actions toward certain issues were categorized as having behavioral impacts learning content. Parmar et al. (2016) introduced Virtual Environment Interaction (VENVi) for teaching programming, logic, and computational thinking in combination with dancing. Users can select dance movements and generate pseudo-code as visual programming. Additionally, they can select drag and drop options for different movements or steps, such as moving forward, backward, squatting, and jumping, and can also simultaneously see a co-located virtual self-avatar that mimics their own movements, look around in the virtual world, and experience a high degree of presence. The behavioral impact happened is found in terms of changing perception toward computer science and computer scientists by highlighting that learning programming could be fun.

Hu et al. (2016) used VR for safety education and environmental protection. The users can collect and classify virtual objects as a basic interaction element and can receive rewards when they perform correctly as a feedback mechanism. The behavioral impact here is with respect to improving actions toward environmental protection. Carruth (2017) developed two VR showcases for industrial applications, such as familiarizing oneself with industrial workspaces, tools, and safety, in a realistic VR environment. Users can interact with this VR environment by picking up tools, manipulating them, and exploring a full range of physical interactions that

represent the assembling element. The behavior changes in this work regard improved awareness of industrial safety and increased knowledge on the usage of tools in an industrial setting.

The visualization-oriented content (Němec et al., 2017) and motivation content papers (Song & Li, 2018) were merged into the “others” category. In addition, a few of the papers did not easily allow us to derive any design elements and learning content (Webster & Dues, 2017; Yang et al., 2016). Hence, these were classified as papers with “no specific elements”.

In a final remark to RQ9, we can summarize the common design elements found in the literature, such as basic interaction and immediate feedback. Typical basic interactions found in the literature were achieved through two levels: first, through user interactions with the VR hardware (e.g., controllers) and, second, through the user interactions inside the VR environment.

Concerning the user interactions with the VR hardware, the following eight types of design were primarily found in the literature. First, head movement detection through sensors embedded in the headset was a basic interaction used to obtain an approximation of the head orientation (Román-Ibáñez et al., 2018) or the addition of a 360° camera (AlAwadhi et al., 2017). Second, motion and infrared sensors were used to transcribe head movements of pitch (up–down), yaw (side to side), and roll (rotational) in order to allow for orientation within the virtual space (Harrington et al., 2018) and track the user’s position (Gerloni et al., 2018; Im et al., 2017). Third, pressing-down a touchpad on the controller or clicking on a map to teleport to the desired location (Pena & Ragan, 2017) was used in order to interact with objects, instruments, or text within the VR platform (Harrington et al., 2018). Fourth, handheld input with thumbsticks (Gerloni et al., 2018), hand-held controllers, and base-stations were used for accurate localization and tracking of the controllers (Zizza et al., 2019). Fifth, attaching a racing wheel system was used to enable realistic driving, such as changing gears manually (Veronez et al., 2018). Sixth, integrated realistic haptic of the VR controllers was sometimes accompanied by a dual-stage trigger with 24 sensors (Pirker et al., 2017; Zhang et al., 2017)—this can be used to track the natural behavior of users, such as moving forward and backward, squatting, grabbing, and releasing. Seventh, sensor technologies – such as galvanic skin response, and facial emotion detection – were used to detect emotions (Hickman & Akdere, 2018). Eighth, interacting through one button only was also used (Muller et al., 2017).

With respect to the interaction inside the VR environment, nine capabilities were revealed in the literature. First, the ability to manipulate objects in various ways—rotation, placement, or moving the objects using a natural virtual hand metaphor (Zizza et al., 2019). Second, the ability to grab objects, such as stickers that can be put on a map (Dolezal et al., 2017). Third, basic pause and repeat, customizable voice notes, and recall abilities (Im et al., 2017). Fourth, the ability to select menu, drag, drop, and look around using a built-in head tracking functionality (Bryan et al., 2018; Song & Li, 2018). Fifth, the ability to interact with language modules. Sixth, the ability to observe objects, such as numerous art exhibitions or artistic design works (Song & Li, 2018). Seventh, the ability to perform experiments, watch pre-recorded lectures, move around the campus, and touch objects (AlAwadhi et al., 2017). Eighth, the ability to collect and classify virtual objects (Hu et al., 2016), pick up tools, manipulate them, and explore a full range of physical interactions (Carruth, 2017). Ninth, the ability to look around in the environment (Bryan et al., 2018).

Typical immediate feedback often used in the literature was as follows: alert systems (Román-Ibáñez et al., 2018); sounds to signal an incorrect gear change (Veronez et al., 2018) or the success and failure when conducting technical works (Im et al., 2017); virtual hands and voices (Zizza et al., 2019); display of the data stream that updates the status of an object (Ye et al., 2018); provision of immediate feedback to students through controller vibration when selecting an object in a VR environment (Muller et al., 2017); revealing correct and incorrect answers by pressing a button (Dolezal et al., 2017); receiving scores after answering a quiz that appears in the VR environment (AlAwadhi et al., 2017) or when a user performs a correct action (Hu et al., 2016); and multisensory, visual feedback, such as highlights, signs, and haptic feedback, that allow users to “feel” the virtual objects (Carruth, 2017).

6. Discussion

In the following section, we discuss the results by describing implications, suggesting a research agenda, deriving recommendations, and identifying limitations.

6.1. Implications

The use of the term *immersion* in connection to VR technology usage has been understood differently. We had to exclude many papers during our evaluation process due to their incompatible use of this term, which was often applied to non-immersive technologies. We included a high number of inclusion terms related to immersive technologies, such as Oculus, Vive, Samsung Gear, Google Cardboard, and Samsung Odyssey, and excluded as many non-immersive technologies as possible, such as 360-degree videos, Desktop VR, CAVE, and panoramic videos. Unfortunately, there is still ambiguity and non-homogeneous understanding of the equipment that can be considered as “immersive technology”.

Regarding *learning theories*, three common types of articles were found in the literature. First, those articles that described VR applications for higher education in detail often did not mention explicit learning theories as their theoretical foundation (68%). In these articles, the development of the VR applications was described thoroughly but the evaluation focused mainly on usability. Such described works, therefore, had an experimental character and were reproducible to some degree only, let alone generalizable. Second, the articles that described VR development in-depth and also mentioned learning theories were often disconnected. For instance, the authors only evaluated app features or usability but not the learning outcomes. Third, there were articles that highlighted the underlying theories of educational VR design but did not report on the technical development in detail. As a result, it was often hard to extract design elements from these papers.

In the 38 reviewed articles, we could identify 18 *application domains*, indicating that there is an interest in the use of immersive VR technologies in many different fields, especially in engineering and computer science. However, this impression must be treated cautiously because most articles did not report experiences with or lessons learned from implementing VR in real university courses. The majority reported on the development process or explored potential uses of VR-based learning.

In some areas, VR seemed to be *mature* enough to be used for teaching procedural, practical knowledge and declarative knowledge. Examples included fire safety, surgery, nursing, and astronomy. In these cases, more professional VR applications were used and were proven to be appropriate for learning in higher education. However, the majority of articles indicated that VR for education is still in its experimental stages—prototyping and testing with students.

Few papers *evaluated* the learning outcomes after applying VR in a specific domain and most of the evaluations that were made consisted of usability-oriented tests. This is also another indicator of the VR maturity level, which still remains a barrier for its adoption in regular teaching activities.

As the only *design element*, basic interactions apparently appear in all types of VR learning content. However, upon a closer look, authors highlight two different levels of user interactions: (1) interactions that occur inside the VR environment and (2) interactions with the hardware, such as the exploitation of haptic and sensors in the headset or other physical objects that connect users with the VR objects. Many papers claim to have created VR for specific purposes in a *realistic* environment and the element “realistic surrounding” was found in almost all learning contents. The terminology of realism is still not uniform, however, as some papers understood realistic surroundings to be a high-fidelity VR environment with complex, high-quality graphics, while for others it could also mean “realistic enough”, in that the user can recognize the environments or the objects, without the creation of objects with real-world details.

6.2. Research agenda and recommendations

Our literature analysis indicated that the interest in applying immersive VR for educational purposes has increased. At the same time, a very low maturity of the field needs to be ascertained. Based on the lessons learned, the following research agenda can be proposed.

First, the theory on VR for educational applications is apparently not advanced enough to allow for a homogeneous usage of related terms, such as immersion or realism. To mitigate the retarding effect of ambiguity and unclarity, more work is required that seeks to contribute toward a common understanding. On the basis of articles such as ours, which summarize categorization frameworks from theory and which seek to classify other works based on sound criteria, a common understanding can be built. Ultimately, proposing a taxonomy of learning theories and other framing factors for educational VR applications is a future research task.

Second, future VR development for higher education needs to build on existing experiments (rather than being exploratory from scratch) and to provide results that allow for generalization. There is no general issue with design-oriented or even mostly developmental works. However, to make thorough contributions, such works need to take a holistic standpoint. Our work clearly indicates that a high contribution and impact can be expected from articles that have a sound theoretical foundation (e.g., learning theories) and technological foundation (e.g., careful selection of design elements) but are also explicit in describing the design and development process. Admittedly, such interdisciplinary work can be hard to develop. Thus, providing the necessary theoretical background to build novel educational VR applications is a future research task. Several aspects of generalization will be supportive. Technology generalization is driven by work that is outside of the educational context but can make the design of such applications much easier. Generalization of learning methods and design elements would help to better the sharing of good practices and to design new course content more rapidly even outside of the fields in which VR has already worked well. As part of this future research task, researchers will also need a comprehensive market overview of existing VR applications that support education. Industrial applications might lead the way to theory building by scientists as well as to the creation of new applications in the educational context.

Third, such works will also allow for a better adoption of VR in higher education. There is currently little exchange of best practices related to education-oriented VR—either within or across disciplines. For example, VR applications in the natural sciences might be more similar to one another than, for example, VR applications in the arts. Thus, the research aim again needs to be holistic and to include well-evaluated works that intend to extend the body of knowledge rather than to simply report anecdotal findings. This will also help derive best practices from the few applications in which VR is already being used with very good learning outcomes.

Fourth, to fulfill the aim of deriving best practices and of describing useful application cases, better evaluation procedures are needed. It is typical that experimental works place the main focus on usability. However, future educational VR applications should be more thoroughly evaluated by employing quantitative and qualitative research methods to assess the students' increase of knowledge and skills as well as the students' learning experience. Evaluations of educational VR applications need to be conducted both in terms of technical feasibility (i.e., from a software engineering standpoint) and of the learning outcomes (i.e., from a pedagogical standpoint). We also suggest that future evaluations assess whether developed applications reflect the users' needs, from the perspective of both teachers and students. Thus, future research needs to include workshops, surveys, and focus group discussions in order to extract the necessary learning content and the expected learning outcomes as well as the usability requirements for VR applications from teachers and students.

Fifth, technological progress is needed to create environments that are perceived as realistic, thus providing *real* immersion. This research aim is not bound to the realm of education. In fact, education can build on high-paced developments, e.g., with respect to gaming, which leads to better development frameworks and fosters the understanding of *how* to create immersion.

Sixth, the eventual aim might be an inclusion of VR into higher education curricula. Such an inclusion could be twofold: (1) VR could be utilized as a teaching tool in real university courses to improve learning outcomes and (2) VR could become the teaching subject itself. The latter would cater to VR becoming more important for application in more and more professional domains. For the inclusion of VR in curricula, much more work is needed on the role of the design elements and the design of learning contents for VR.

6.3. Recommendations for lecturers

Our research shows that the “realistic surroundings” and “basic interaction” design elements occur in all types of VR applications in our sample. Thus, these can be seen as the basic design requirements for educational VR applications. VR applications that aim to improve declarative knowledge can be recommended for initiating VR in the courses. In our mapping of design elements and learning outcomes, we can see that most applications for declarative knowledge use only these two basic design elements. Therefore, these elements can be a soft start for VR development, easy to use by lecturers and students, and might not require any changes in the curriculum. In many study programs, students are mainly taught declarative knowledge in lectures and are expected to memorize what they have learned for exams. The use of VR applications to impart declarative knowledge could support this learning method to make lectures more exciting.

From our own experience as lecturers, we recognize the students’ wishes on practice-oriented learning contents rather than memorizing facts. If lecturers have already had good experiences with VR in their lectures, they could use advanced applications to make teaching more practice-oriented. In our sample, most articles describe VR applications for teaching procedural knowledge. At the same time, our mapping of design elements and learning outcomes shows that these applications have the largest number of design elements. This means that the development of such applications is more complex and that teachers and students might need more VR experience to apply them. Furthermore, it might require complex changes of the curriculum to shift the focus from teaching declarative knowledge to more practice-oriented content. Nevertheless, we believe that the true potential of VR lies not in better teaching of declarative knowledge, but in offering opportunities to “learn by doing” which is often very difficult to implement in traditional lectures.

We have limited results concerning VR aims at improving communication, collaboration and soft skills, behavioral impacts and analytical skills to be able to derive recommendations for the most useful design elements that will meet specific learning goals.

Revisiting our findings and the research agenda, we conclude that the state of the art does not allow for the provision of an exhaustive set of recommendations or even a catalogue of best practices. Nonetheless, it is possible to propose an initial list of ideas for teachers in higher education. It should help them decide when to use VR, for which purposes to use it, and which technological steps to take.

Undoubtedly, VR is a hyped topic that does not only tend to make technological enthusiasts excited. This momentum can be used by educators and chances are good for someone to be an early adopter and to make a contribution. Thus, our recommendation to interested educators is not to be too shy about applying VR to their curricula. They should also be empowered to become capable of deciding on buying apps, cooperating with the industry, and developing requirements—it does not seem realistic that non-technologists will be able to develop arbitrary virtual worlds from scratch in the near future.

Since VR applications exist in some fields, it makes much sense to learn from existing *good* cases. Even without a body of best practices, *solid cases* can be used as an indicator of what works. Solid in this sense could mean that a VR application in higher education is built with sound technology, is explicit about learning aims and design elements, and has been evaluated thoroughly. Thus, we recommend that such works are built on and that attempts are made to transfer positive experiences as far as possible to their own application domain.

There are many technological hurdles. The price tag of VR can be very high—maybe too high to be convincing, as long as a positive impact on learning is not proven. Thus, we recommend starting with a realistically small scope and tailoring the application case to what seems feasible and achievable in a timespan of a few months. Setting out to virtualize a whole course with all its assets is bound to fail even if plentiful resources (e.g., funding and VR-proficient personnel) are available. However, the integration of low-budget mobile HMDs or the use of VR applications in a few selected lecture units over the entire semester could represent first steps toward a broad adoption of VR in higher education.

6.4. Limitations

Due to the nature of the review, selection, and filtering process, our work has several limitations. First, we only looked at immersive VR applications for education. There are various VR technologies, other than HMDs, that have been used for educational purposes, such as 360-degree videos, desktop VR, and mixed reality. These technologies might already have achieved a higher level of maturity and have been successfully applied for higher education purposes. It is worth acknowledging that as the focus of this study has been on learning through HMDs VR engagement, the paper does not consider emerging applications and studies applying collaborative and group engagement in VR through VR domes/other multiple person VR environments (Grogorick, Ueberheide, Tauscher, Bittner, & Magnor, 2019; Liu, Liu, & Jin, 2019). Future work might need to seek drawing the *whole picture*.

Second, we also limited ourselves to publications appearing between 2016 and 2018, assuming this to be a timeline when HMDs gained popularity and may have changed the way VR is used in the education context. If the types of VR technologies used in education are not a limiting factor, there are some innovative VR-based teaching and learning approaches that have been

Table A.9
Search strings.

Database	Results	Keyword search and other applied filters
Scopus	574	(TITLE-ABS-KEY (“virtual reality”) OR TITLE-ABS-KEY (“VR”) AND TITLE-ABS-KEY (educat*) OR TITLE-ABS-KEY (teach*) OR TITLE-ABS-KEY (learn*) OR TITLE-ABS-KEY (train*) AND TITLE-ABS-KEY (university) OR TITLE-ABS-KEY (“higher education”) OR TITLE-ABS-KEY (college) AND NOT TITLE-ABS-KEY (“machine learning”) OR TITLE-ABS-KEY (“deep learning”) AND NOT TITLE-ABS-KEY (“artificial intelligence”) OR TITLE-ABS-KEY (“neural network”) AND NOT TITLE-ABS-KEY (rehabilitation) OR TITLE-ABS-KEY (therapy)) AND DOCTYPE (ar OR cp) AND PUBYEAR > 2016 AND PUBYEAR < 2018 AND (LIMIT-TO (LANGUAGE , “English”)))
IEEEExplore Digital Library	1 656	(“Abstract”:“virtual reality” OR “Abstract”:“VR” AND “Abstract”:“educat*” OR “Abstract” OR “Abstract”:“teach” OR “Abstract”:“train” AND “Abstract”:“university” OR “Abstract”:“higher education” OR “Abstract”:“college” NOT “Abstract”:“artificial intelligence” OR “Abstract”:“neural network” NOT “Abstract”:“rehabilitation” NOT “Abstract”:“machine learning” OR “Abstract”:“deep learning”). Filters Applied: Journals & Magazines Conferences virtual reality 2016–2018
ProQuest	877 from 905	noft(“virtual reality”) OR noft(VR) AND noft(teach*) OR noft(learn*) OR noft(educat*) AND noft(“higher education”) OR noft(“university”) OR noft(college) NOT noft(“machine learning”) NOT noft(“rehabilitation”) Limits applied Databases: All databases searched; Limited by: Full text, Peer reviewed, Date: From January 01 2016 to December 31 2018. Source type: Conference Papers & Proceedings, Scholarly Journals. Document type: Article, Conference Paper, Conference Proceeding Language: English. Automatic removal of duplications are applied.
Web of Science	112	((TS=(“virtual reality” OR VR)) AND (TS=(educat* OR learn* OR teach* OR train*)) AND (TS=(university OR “higher education” OR college)) NOT (ALL=(“artificial intelligence” OR “neural network” OR “deep learning” OR “machine learning” OR therapy OR rehabilitation)))) AND LANGUAGE: (English) AND DOCUMENT TYPES: (Article)

Table B.10
List of keywords.

2d visual virtual reality system	Disease	Network security	Sport training
2d virtual reality interface	Dynamic balance control	Non-immersive virtual reality	Stroke
3d computer	Employee training	Omnidirectional video	Temporal bone
360-degree virtual reality	Gait ability	Panoramic video	Three dimensional computer graphics
Accident prevention	Gait speed	Parkinson	Treadmill training
Alzheimer’s disease	Gait training	Patient data	Upper limb
Animal health	Game performance	Personnel training	Upper limb motor training
Arm function	High school	Postural balance	Veterinary public health
Augmented reality	Mixed reality	Postural control	Video
Augmented reality exposure	Mental health	Postural stability	Virtual patient
Autism	Mixed virtual reality simulator	Primary school	Virtual reality film
Balance training	Mobile devices	Primary education	Virtual reality video
Brain injury	Mobile game	Primary education	Virtual walking
CAVE	Mobile immersive virtual reality	Radiotherapy	Virtual world
Cinematic virtual reality	Mobile learning	Rehabilitation	Virtual limb
Cognitive impairment	Motor rehabilitation	Safety engineering	Virtual reality-based rehabilitation
Cognitive deficit	Mobile robot	Secondary education	Virtual reality paradigm
Dental treatment	Mobile virtual reality	Second life	Visual speed
Desktop virtual reality	Motor disturbances	Serious game	Voice rehabilitation
Desktop vr	Muscle training	Social anxiety disorder	Vocational
Dichoptic training	National deaf-blind equipment distribution	Social stress	

documented by researchers. In other words, the validity of our conclusions is within the scope of the aforementioned research boundary—despite indications that we have chosen the timescale wisely.

Third, most papers included in this work originated from conference papers instead of journal papers, as part of our definition of immersive VR was the use of HMDs, regardless of costs. Consequently, some of the papers have not yet reported thorough solutions or lack supporting theories.

Fourth, this paper does not examine the barriers to the usage of VR for learning, as many papers have reported common issues for users, such as nausea, dizziness, and some other physical symptoms. Some barriers are related to the technology itself. For example, the display resolution and cables disturb the immersion in and control of the devices (e.g., grabbing vs. teleporting), making it difficult to switch functions. In case of mobile VR, the visual system is deemed to be demanding. Sometimes, computational issues occur during VR experiments that could cause barriers in terms of learning speed. However, technological limitations will gradually be reduced, along with the emergence of new VR technologies. The 2019 Oculus Quest, for example, comes as a cordless HMD, which will help to overcome the cable issues. Nonetheless, our work *cannot* answer the question whether a widespread use of VR in education is reasonable. None of the named limitations impairs the value of our work; in fact, they provide the opportunity to continue advancing the field.

7. Conclusion and future research

In this article, a systematic mapping study was conducted, focusing on the employment of immersive VR technologies for higher education purposes. Immersive VR technologies, application domains, learning contents, and design elements being used in recent

Table C.11
Articles included in the study.

#	Code	Authors	Title
1	P4	Veronez, Gonzaga, Bordin, Kupssinsku, Kannenberg, Duarte, et al. (2018)	RIDERS: Road inspection driver simulation
2	P5	Webster and Dues (2017)	System usability scale (SUS): Oculus Rift®DK2 and Samsung Gear VR®
3	P12	Papachristos, Vrellis, and Mikropoulos (2017)	A comparison between Oculus Rift and a low-cost smartphone VR headset: Immersive user experience and learning
4	P13	Román-Ibáñez, Pujol-López, Mora-Mora, Pertegal-Felices, and Jimeno-Morenila (2018)	A low-cost immersive virtual reality system for teaching robotic manipulators programming
5	P15	Kwon et al. (2017)	A virtual reality based engine training system: A prototype development & evaluation
6	P16	Rosenfield et al. (2018)	AAS worldwide telescope: A seamless, cross-platform data visualization engine for astronomy research, education, and democratizing data
7	P17	Dolezal, Chmelik, and Liarokapis (2017)	An immersive virtual environment for collaborative geovisualization
8	P22	Pena and Ragan (2017)	Contextualizing construction accident reports in virtual environments for safety education
9	P23	dela Cruz and Mendoza (2018)	Design and development of virtual laboratory: A solution to the problem of laboratory setup and management of pneumatic courses in Bulacan State University College of Engineering
10	P24	Zhang, Suo, Chen, Liu, and Gao (2017)	Design and implementation of fire safety education system on campus based on virtual reality technology
11	P25	Bujdosó, Novac, and Szimkovics (2017)	Developing cognitive processes for improving inventive thinking in system development using a collaborative virtual reality system
12	P26	Hickman and Akdere (2018)	Developing intercultural competencies through virtual reality: Internet of Things applications in education and learning
13	P27	Harrington et al. (2018)	Development and evaluation of a trauma decision-making simulator in Oculus virtual reality
14	P29	Smith et al. (2018)	Effectiveness of two varying levels of virtual reality simulation
15	P38	Gerloni et al. (2018)	Immersive virtual reality for earth sciences
16	P46	Muller, Panzoli, Galaup, Lagarrigue, and Jessel (2017)	Learning mechanical engineering in a virtual workshop: A preliminary study on utilizability, utility and acceptability
17	P47	Parong and Mayer (2018)	Learning science in immersive virtual reality
18	P48	Pirker, Lesjak, and Guetl (2017)	Maroon VR: A room-scale physics laboratory experience
19	P50	Shattuck (2018)	Multiuser virtual reality environment for visualizing neuroimaging data
20	P53	Schroeder, Bailey, Johnson, and Gonzalez-Holland (2017)	Presence and usability do not directly predict procedural recall in virtual reality training
21	P54	Parmar, Isaac, Babu, D'Souza, Leonard, Jörg, Gundersen, and Daily (2016)	Programming moves: Design and evaluation of applying embodied interaction in virtual environments to enhance computational thinking in middle school students
22	P55	Zhou, Ji, Xu, and Wang (2018)	Promoting knowledge construction: A model for using Virtual Reality interaction to enhance learning
23	P56	Song and Li (2018)	Research on application of VR technology in art design teaching
24	P57	Bryan, Campbell, and Mangina (2018)	Scenic spheres—An AR/VR educational game
25	P59	Hu, Su, and He (2016)	The design and implementation of the 3D educational game based on VR headsets
26	P61	Chen (2016)	The effects of virtual reality learning environment on student cognitive and linguistic development
27	P62	Yang, Cheng, and Yang (2016)	The impact of three types of virtual reality scene on learning
28	P64	Farra, Smith, and Ulrich (2018)	The student experience with varying immersion levels of virtual reality simulation

(continued on next page)

Table C.11 (continued).

#	Code	Authors	Title
29	P66	Buñ, Trojanowska, Ivanov, and Pavlenko (2018)	The use of virtual reality training application to increase the effectiveness of workshops in the field of lean manufacturing
30	P67	Zizza et al. (2019)	Toward a social virtual reality learning environment in high fidelity
31	P68	Ekkelenkamp, Koch, de Man, and Kuipers (2016)	Training and competence assessment in GI endoscopy: A systematic review
32	P71	Chin et al. (2017)	Using virtual reality for an immersive experience in the water cycle
33	P72	Němec, Fasuga, Trubač, and Kratochvíl (2017)	Using virtual reality in education
34	P73	AlAwadhi et al. (2017)	Virtual reality application for interactive and informative learning
35	P75	Carruth (2017)	Virtual reality for education and workforce training
36	P78	Ye, Hu, Zhou, Lei, and Guan (2018)	VR interactive feature of HTML5-based WebVR control laboratory by using head-mounted display
37	P79	Misbhaudhin (2018)	VREdu: A framework for Interactive Immersive Lectures using virtual reality
38	P80	Cortiz and Silva (2017)	Web and virtual reality as platforms to improve online education experiences

literature on educational VR applications were examined. The review results show that the interest in immersive VR technologies for educational purposes seems to be quite high, which is indicated by the variety of the research domains that have applied this technology in teaching. The majority of authors treated VR as a promising learning tool for higher education, however, the maturity of the use of VR in higher education is still questionable. Technologies described in most of the reviewed articles remained in an experimental state and were mostly tested in terms of their performance and usability. This article also reveals that very few design-oriented studies constructed their VR applications based on a specific learning theory, which serves as technical development guidance. Moreover, few papers thoroughly describe how VR-based teaching can be adopted in the teaching curriculum.

These facts can hinder the rapid adoption of immersive VR technologies into teaching on a regular basis. We acknowledge that, in some domains such as engineering and computer science, certain VR applications have been used on a regular basis to teach certain skills, especially those that require declarative knowledge and procedural-practical knowledge. However, in most domains, VR is still experimental and its usage is not systematic or based on best practices. This paper pinpoints key gaps that serve to provide insights for future improvements, especially for VR application developers and teachers in higher education.

Our work will continue with a market analysis of VR technologies that could be employed in higher education as well with a survey of educators. We aim to continue advancing the field now that we have understood its low-maturity but nevertheless promising nature.

Acknowledgment

This project has been funded with support from the European Commission [Erasmus+ grant number 2018-1-LI01-KA203-000107]. This publication reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

Appendix A. Keyword search in each database

The employed search strings are compiled in [Table A.9](#).

Appendix B. Keywords

An exhaustive list of keywords for deeper exclusion is given in [Table B.10](#).

Appendix C. Selected articles

The articles included in our study are compiled in [Table C.11](#).

References

- AlAwadhi, S., AlHabib, N., Murad, D., AlDeei, F., AlHouti, M., Beyrouthy, T., et al. (2017). Virtual reality application for interactive and informative learning. In *2017 2nd international conference on bio-engineering for smart technologies* (pp. 1–4). IEEE.
- Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review*, 89(4), 369.
- Biocca, F., & Delaney, B. (1995). Immersive virtual reality technology. In *Communication in the age of virtual reality*, (pp. 57–124). Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.

- Boulton, C. A., Kent, C., & Williams, H. T. (2018). Virtual learning environment engagement and learning outcomes at a 'bricks-and-mortar' university. *Computers & Education*, 126, 129–142.
- Bowman, D. A., & McMahan, R. P. (2007). Virtual reality: how much immersion is enough? *Computer*, 40(7), 36–43.
- Bryan, S. J., Campbell, A., & Mangina, E. (2018). Scenic spheres-an AR/VR educational game. In *2018 IEEE games, entertainment, media conference* (pp. 1–9). IEEE.
- Bujdosó, G., Novac, O. C., & Szimkovics, T. (2017). Developing cognitive processes for improving inventive thinking in system development using a collaborative virtual reality system. In *2017 8th IEEE international conference on cognitive infocommunications* (pp. 000079–000084). IEEE.
- Buñ, P., Trojanowska, J., Ivanov, V., & Pavlenko, I. (2018). The use of virtual reality training application to increase the effectiveness of workshops in the field of lean manufacturing. In *4th international conference of the virtual and augmented reality in education* (pp. 65–71).
- Carruth, D. W. (2017). Virtual reality for education and workforce training. In *2017 15th international conference on emerging elearning technologies and applications* (pp. 1–6). IEEE.
- Chavez, B., & Bayona, S. (2018). Virtual reality in the learning process. In *Trends and advances in information systems and technologies* (pp. 1345–1356). Springer International Publishing.
- Chen, Y.-L. (2016). The effects of virtual reality learning environment on student cognitive and linguistic development. *The Asia-Pacific Education Researcher*, 25(4), 637–646.
- Chin, N., Gupta, A., Nguyen, J., Sukhin, S., Wang, G., & Mirizio, J. (2017). Using virtual reality for an immersive experience in the water cycle. In *2017 IEEE MIT undergraduate research technology conference* (pp. 1–4). IEEE.
- Coder, K. (2017). *Open source software, Higuchi, Koichi*. Ritsumeikan University.
- Cortiz, D., & Silva, J. O. (2017). Web and virtual reality as platforms to improve online education experiences. In *2017 10th international conference on human system interactions* (pp. 83–87). IEEE.
- Crebert, G., Bates, M., Bell, B., Patrick, C.-J., & Cragnolini, V. (2004). Developing generic skills at university, during work placement and in employment: graduates' perceptions. *Higher Education Research & Development*, 23(2), 147–165.
- Creswell, J. W., & Creswell, J. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications.
- Dale, E. (1969). *Audiovisual methods in teaching*. The Dryden Press.
- Davis, B., & Summers, M. (2015). Applying Dale's cone of experience to increase learning and retention: A study of student learning in a foundational leadership course. *Qscience Proceedings*, 6.
- Dede, C. (2008). Theoretical perspectives influencing the use of information technology in teaching and learning. In J. Voogt, & G. Knezek (Eds.), *International handbook of information technology in primary and secondary education* (pp. 43–62). Boston, MA: Springer US, http://dx.doi.org/10.1007/978-0-387-73315-9_3.
- dela Cruz, D. R., & Mendoza, D. M. M. (2018). Design and development of virtual laboratory: A solution to the problem of laboratory setup and management of pneumatic courses in bulacan state university college of engineering. In *2018 IEEE games, entertainment, media conference* (pp. 1–23). IEEE.
- Dolezal, M., Chmelik, J., & Liarokapis, F. (2017). An immersive virtual environment for collaborative geovisualization. In *VS-GAMES* (pp. 272–275).
- Ekkelenkamp, V. E., Koch, A. D., de Man, R. A., & Kuipers, E. J. (2016). Training and competence assessment in GI endoscopy: a systematic review. *Gut*, 65(4), 607–615.
- Ertmer, P. A., & Newby, T. J. (1993). Behaviorism, cognitivism, constructivism: Comparing critical features from an instructional design perspective. *Performance Improvement Quarterly*, 6(4), 50–72. <http://dx.doi.org/10.1111/j.1937-8327.1993.tb00605.x>.
- Fallman, D. (2003). Design-oriented human-computer interaction. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 225–232). ACM.
- Farra, S. L., Smith, S. J., & Ulrich, D. L. (2018). The student experience with varying immersion levels of virtual reality simulation. *Nursing Education Perspectives*, 39(2), 99–101.
- Feng, Z., González, V. A., Amor, R., Lovreglio, R., & Cabrera-Guerrero, G. (2018). Immersive virtual reality serious games for evacuation training and research: A systematic literature review. *Computers & Education*, 127, 252–266.
- Fosnot, C. T. (2013). *Constructivism: Theory, perspectives, and practice*. Teachers College Press.
- Fosnot, C. T., & Perry, R. S. (1996). Constructivism: A psychological theory of learning. *Constructivism: Theory, Perspectives, and Practice*, 2, 8–33.
- Freina, L., & Ott, M. (2015). A literature review on immersive virtual reality in education: state of the art and perspectives. In *The international scientific conference elearning and software for education (Vol. 1)* (p. 133). "Carol I" National Defence University.
- Garrett, L. (1997). Dewey, Dale, and Bruner: Educational philosophy, experiential learning, and library school cataloging instruction. *Journal of Education for Library and Information Science*, 38(2), 129–136.
- Gerloni, I. G., Carchiolo, V., Vitello, F. R., Sciacca, E., Becciani, U., Costa, A., et al. (2018). Immersive virtual reality for earth sciences. In *2018 federated conference on computer science and information system* (pp. 527–534). IEEE.
- Global Augmented Reality (2019). <https://www.prnewswire.com/news-releases/global-augmented-reality-ar-and-virtual-reality-vr-market-in-healthcare-market-to-reach-11-14-billion-by-2025--300831306.html>.
- Grogier, S., Ueberheide, M., Tauscher, J., Bittner, P. M., & Magnor, M. A. (2019). Gaze and motion-aware real-time dome projection system. In *IEEE conference on virtual reality and 3D user interfaces, VR 2019, Osaka, Japan, March 23–27, 2019* (pp. 1780–1783). IEEE, <http://dx.doi.org/10.1109/VR.2019.8797902>, URL: <https://ieeexplore.ieee.org/xpl/conhome/8787730/proceeding>.
- Harrington, C. M., Kavanagh, D. O., Quinlan, J. F., Ryan, D., Dicker, P., O'Keeffe, D., et al. (2018). Development and evaluation of a trauma decision-making simulator in oculus virtual reality. *The American Journal of Surgery*, 215(1), 42–47.
- Hickman, L., & Akdere, M. (2018). Developing intercultural competencies through virtual reality: Internet of things applications in education and learning. In *2018 15th learning and technology conference (L&T)* (pp. 24–28). IEEE.
- Higuchi, K. (2016). *KH Coder 3 reference manual*. Kyoto (Japan): Ritsumeikan University.
- Holsti, O. R. (1969). *Content analysis for the social sciences and humanities*. Addison-Wesley.
- Hu, X., Su, R., & He, L. (2016). The design and implementation of the 3D educational game based on VR headsets. In *2016 international symposium on educational technology* (pp. 53–56). <http://dx.doi.org/10.1109/ISSET.2016.15>.
- Ibanez, M., Kloos, C. D., Leony, D., Rueda, J. J. G., & Maroto, D. (2011). Learning a foreign language in a mixed-reality environment. *IEEE Internet Computing*, 15(6), 44–47. <http://dx.doi.org/10.1109/MIC.2011.78>.
- Im, T., An, D., Kwon, O. Y., & Kim, S. Y. (2017). A virtual reality based engine training system: A prototype development & evaluation. In *CSEDU 2017 - proceedings of the 9th international conference on computer supported education (Vol. 1)* (pp. 262–267).
- Jeffries, P. R., Rodgers, B., & Adamson, K. (2015). NLN Jeffries simulation theory: Brief narrative description. *Nursing Education Perspectives*, 36(5), 292–293.
- 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. <http://dx.doi.org/10.1007/s10639-017-9676-0>.
- Kathleen Dunaway, M. (2011). Connectivism: Learning theory and pedagogical practice for networked information landscapes. *Reference Services Review*, 39(4), 675–685.
- Khalifa, M., & Shen, N. (2004). System design effects on social presence and telepresence in virtual communities. In *ICIS 2004 proceedings* (p. 43).
- Kitchenham, B., Brereton, O. P., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009). Systematic literature reviews in software engineering—a systematic literature review. *Information and Software Technology*, 51(1), 7–15.
- Kolb, A. Y., & Kolb, D. A. (2012). Experiential learning theory. In *Encyclopedia of the sciences of learning* (pp. 1215–1219). Springer.

- Krippendorff, K. (2004). Reliability in content analysis. *Human Communication Research*, 30(3), 411–433.
- Krokos, E., Plaisant, C., & Varshney, A. (2019). Virtual memory palaces: immersion aids recall. *Virtual Reality*, 23(1), 1–15.
- Kumar, R. (2019). *Research methodology: A step-by-step guide for beginners*. Sage Publications Limited.
- Kwon, B., Kim, J., Lee, K., Lee, Y. K., Park, S., & Lee, S. (2017). Implementation of a virtual training simulator based on 360° multi-view human action recognition. *IEEE Access*, 5, 12496–12511. <http://dx.doi.org/10.1109/ACCESS.2017.2723039>.
- Lee, E. A.-L., & Wong, K. W. (2008). A review of using virtual reality for learning. In Z. Pan, A. D. Cheok, W. Müller, & A. El Rhalibi (Eds.), *Transactions on edutainment I* (pp. 231–241). Berlin, Heidelberg: Springer Berlin Heidelberg, http://dx.doi.org/10.1007/978-3-540-69744-2_18.
- Liu, Y., Liu, Z., & Jin, Y. (2019). Virtual dome system using HMDs: An alternative to the expensive and less accessible physical domes. In J. Y. Chen, & G. Fragomeni (Eds.), *Virtual, augmented and mixed reality. Applications and case studies* (pp. 312–328). Cham: Springer International Publishing.
- 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.
- Mason, R., & Bacsich, P. (1998). Embedding computer conferencing into university teaching. *Computers & Education*, 30(3), 249–258. [http://dx.doi.org/10.1016/S0360-1315\(97\)00068-7](http://dx.doi.org/10.1016/S0360-1315(97)00068-7), URL: <http://www.sciencedirect.com/science/article/pii/S0360131597000687>.
- 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. <http://dx.doi.org/10.1016/j.compedu.2013.07.033>.
- Misbhaudhin, M. (2018). VREdu: A framework for interactive immersive lectures using virtual reality. In *2018 21st Saudi computer society national computer conference* (pp. 1–6). IEEE.
- Muller, N., Panzoli, D., Galaup, M., Lagarrigue, P., & Jessel, J.-P. (2017). Learning mechanical engineering in a virtual workshop: A preliminary study on utilisability, utility and acceptability. In *2017 9th international conference on virtual worlds and games for serious applications (VS-Games)* (pp. 55–62). IEEE.
- Muñoz-Cristóbal, J. A., Gallego-Lema, V., Arribas-Cubero, H. F., Martínez-Monés, A., & Asensio-Pérez, J. I. (2017). Using virtual learning environments in bricolage mode for orchestrating learning situations across physical and virtual spaces. *Computers & Education*, 109, 233–252.
- Mütterlein, J. (2018). The three pillars of virtual reality? Investigating the roles of immersion, presence, and interactivity. In *Proceedings of the 51st Hawaii international conference on system sciences*.
- Némec, M., Fasuga, R., Trubač, J., & Kratochvíl, J. (2017). Using virtual reality in education. In *2017 15th international conference on emerging elearning technologies and applications* (pp. 1–6). IEEE.
- Papachristos, N. M., Vrellis, I., & Mikropoulos, T. A. (2017). A comparison between oculus rift and a low-cost smartphone VR headset: immersive user experience and learning. In *2017 IEEE 17th international conference on advanced learning technologies* (pp. 477–481). <http://dx.doi.org/10.1109/ICALT.2017.145>.
- Parmar, D., Isaac, J., Babu, S. V., D'Souza, N., Leonard, A. E., Jörg, S., Gundersen, K., & Daily, S. B. (2016). Programming moves: Design and evaluation of applying embodied interaction in virtual environments to enhance computational thinking in middle school students. In *2016 IEEE virtual reality* (pp. 131–140). IEEE.
- Parong, J., & Mayer, R. E. (2018). Learning science in immersive virtual reality. *Journal of Educational Psychology*, 110(6), 785–797.
- Pena, A. M., & Ragan, E. D. (2017). Contextualizing construction accident reports in virtual environments for safety education. In *2017 IEEE virtual reality* (pp. 389–390). IEEE.
- Pirker, J., Lesjak, I., & Guetl, C. (2017). Maroon VR: A room-scale physics laboratory experience. In *2017 IEEE 17th international conference on advanced learning technologies* (pp. 482–484). IEEE.
- Robertson, G., Czerwinski, M., & Van Dantzich, M. (1997). Immersion in desktop virtual reality. In *Proceedings of the 10th annual ACM symposium on user interface software and technology* (pp. 11–19). Citeseer.
- Rogers, S. (2019). Virtual reality: The learning aid of the 21st century. *Forbes*.
- Román-Ibáñez, V., Pujol-López, A. F., Mora-Mora, H., Pertegal-Felices, L. M., & Jimeno-Morenilla, A. (2018). A low-cost immersive virtual reality system for teaching robotic manipulators programming. *Sustainability*, 10(4), <http://dx.doi.org/10.3390/su10041102>.
- Rosenfield, P., Fay, J., Gilchrist, R. K., Cui, C., Weigel, A. D., Robitaille, T., et al. (2018). AAS worldwide telescope: A seamless, cross-platform data visualization engine for astronomy research, education, and democratizing data. *The Astrophysical Journal Supplement Series*, 236(1), 22.
- Ryan, M.-L. (2015). *2, Narrative as virtual reality 2: Revisiting immersion and interactivity in literature and electronic media* (Vol. 2). JHU Press.
- Schroeder, B. L., Bailey, S. K., Johnson, C. I., & Gonzalez-Holland, E. (2017). Presence and usability do not directly predict procedural recall in virtual reality training. In *International conference on human-computer interaction* (pp. 54–61). Springer.
- Schunk, D. H. (2012). *Learning theories an educational perspective*. Pearson.
- Sharma, S., Agada, R., & Ruffin, J. (2013). Virtual reality classroom as an constructivist approach. In *2013 proceedings of IEEE southeastcon* (pp. 1–5). <http://dx.doi.org/10.1109/SECON.2013.6567441>.
- Shattuck, D. W. (2018). Multiuser virtual reality environment for visualising neuroimaging data. *Healthcare Technology Letters*, 5(5), 183–188.
- Shuell, T. J. (1986). Cognitive conceptions of learning. *Review of Educational Research*, 56(4), 411–436.
- Siemens, G. (2014). Connectivism: A learning theory for the digital age. *International Journal of Instructional Technology & Distance Learning*, 2.
- Skinner, B. F. (1989). The origins of cognitive thought. *American Psychologist*, 44(1), 13.
- Slater, M., & Wilbur, S. (1997). A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments. *Presence: Teleoperators & Virtual Environments*, 6(6), 603–616.
- Smith, S. J., Farra, S. L., Ulrich, D. L., Hodgson, E., Nicely, S., & Mickle, A. (2018). Effectiveness of two varying levels of virtual reality simulation. *Nursing Education Perspectives*, 39(6), E10–E15.
- Song, Y., & Li, L. (2018). Research on application of VR technology in art design teaching. In *2018 international conference on engineering simulation and intelligent control* (pp. 343–345). IEEE.
- Spector, P. E., & Spector, P. F. (1981). *Research designs* (Vol. 23). Sage.
- Steuer, J. (1995). Defining virtual reality: Dimensions determining presence. In *Communication in the age of virtual reality* (pp. 33–56). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Stolterman, E., & Wiberg, M. (2010). Concept-driven interaction design research. *Human-Computer Interaction*, 25(2), 95–118.
- Suh, A., & Prophet, J. (2018). The state of immersive technology research: A literature analysis. *Computers in Human Behavior*, 86, 77–90.
- Veronez, M. R., Gonzaga, L., Bordin, F., Kupssinsku, L., Kannenberg, G. L., Duarte, T., et al. (2018). RIDERS: Road Inspection Driver Simulation. In *25th IEEE conference on virtual reality and 3d user interfaces, VR 2018 - proceedings* (pp. 715–716).
- Walsh, K. R., & Pawlowski, S. D. (2002). Virtual reality: A technology in need of IS research. *Communications of the Association for Information Systems*, 8(1), 20.
- Wang, R., Wiese, R., & Gibbons, C. (2012). Developing digital fluency through ubiquitous mobile devices: Findings from a small-scale study. *Computers & Education*, 58(1), 570–578.
- Wang, P., Wu, P., Wang, J., Chi, H.-L., & Wang, X. (2018). A critical review of the use of virtual reality in construction engineering education and training. *International Journal of Environmental Research and Public Health*, 15(6), 1204.
- Webster, R., & Dues, J. F. (2017). System usability scale (SUS): Oculus Rift® DK2 and Samsung Gear VR®. In *ASEE annual conference and exposition, conference proceedings*, (Vol. 2017-June).
- Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review. *MIS Quarterly*, xiii–xxiii.
- Wendler, R. (2012). The maturity of maturity model research: A systematic mapping study. *Information and Software Technology*, 54(12), 1317–1339.
- Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence*, 7(3), 225–240.

- Wohlgenannt, I., Fromm, J., Stieglitz, S., Radianti, J., & Majchrzak, T. A. (2019). Virtual reality in higher education: Preliminary results from a design-science-research project. In *28th international conference on information systems development. AIS*.
- Yang, X., Cheng, P.-Y., & Yang, X. (2016). The impact of three types of virtual reality scene on learning. In *2017 international conference of educational innovation through technology* (pp. 322–324). IEEE.
- Ye, Q., Hu, W., Zhou, H., Lei, Z., & Guan, S. (2018). VR interactive feature of html5-based webvr control laboratory by using head-mounted display. *International Journal of Online Engineering*, 14(3).
- Zawacki-Richter, O., & Latchem, C. (2018). Exploring four decades of research in computers & education. *Computers & Education*, 122, 136–152.
- Zhang, K., Suo, J., Chen, J., Liu, X., & Gao, L. (2017). Design and implementation of fire safety education system on campus based on virtual reality technology. In *2017 federated conference on computer science and information systems* (pp. 1297–1300). IEEE.
- Zhou, Y., Ji, S., Xu, T., & Wang, Z. (2018). Promoting knowledge construction: a model for using virtual reality interaction to enhance learning. *Procedia computer science*, 130, 239–246.
- Zizza, C., Starr, A., Hudson, D., Nuguri, S. S., Callyam, P., & He, Z. (2019). Towards a social virtual reality learning environment in high fidelity. In *2018 15th IEEE annual consumer communications & networking conference* (pp. 1–4). IEEE.