Interactive Learning Environments

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/nile20

Learning in immersed collaborative virtual environments: design and implementation

Tycho T. De Back, Angelica M. Tinga & Max M. Louwerse

To cite this article: Tycho T. De Back, Angelica M. Tinga & Max M. Louwerse (2023) Learning in immersed collaborative virtual environments: design and implementation, Interactive Learning Environments, 31:8, 5364-5382, DOI: 10.1080/10494820.2021.2006238

To link to this article: https://doi.org/10.1080/10494820.2021.2006238

Published online: 12 Dec 2021.

INTERACTIVE LEARNING ENVIRONMENTS
2023, VOL. 31, NO. 8, 5364–5382
https://doi.org/10.1080/10494820.2021.2006238

Routledge
Taylor & Francis Group

Learning in immersed collaborative virtual environments: design and implementation

Tycho T. De Back, Angelica M. Tinga and Max M. Louwerse

Department of Cognitive Science & Artificial Intelligence, Tilburg University, Tilburg, The Netherlands

ABSTRACT

Immersive virtual environments hold unexplored potential to scaffold and stimulate learning in multiple ways for the purpose of increasing potential learning gains. Yet, the number of implementations in educational settings remains very limited. One reason for limited implementation of immersive virtual environment applications may be a lack of recommendations for their effective design. Building on an extensive theoretical framework, this paper provides such recommendations, specific to immersive educational settings. The recommendations are divided into strategies to optimize cognitive load, foster collaborative learning, leverage platform-specific affordances, mitigate platform-specific limitations and to obtain additional benefits. We illustrate the implementation of these recommendations using a novel collaborative virtual reality environment shown to yield learning gains in two prior studies. Moreover, we detail how a non-invasive and cost-effective feature of automated performance analysis can monitor learning gains in collaborative virtual reality environments. The recommendations based on the example of a collaborative virtual reality environment pave the way for more such implementations to maximize benefits for learners and educators alike.

ARTICLE HISTORY

Received 20 March 2021
Accepted 8 November 2021

KEYWORDS

Learning design; learning environments; immersive virtual reality; cognitive load; collaborative learning

Introduction

Due to the recent proliferation of virtual reality (VR), educational institutions increasingly look towards immersive technologies as a means to support learning and motivation. VR may either be of a non-immersive or an immersive type. In non-immersive VR, the virtual environment is often observed from a distance through the screen of a desktop monitor and interacted with using a conventional mouse and keyboard. By contrast, in immersive VR (IVR) the virtual environment is experienced more directly as it surrounds and immerses the viewer, such as when viewed using head-mounted displays or Cave Automatic Environments (CAVEs, projecting a virtual environment on the walls of a room; Cruz-Neira et al., 1992). In addition, in IVR the environment is often interacted with in an embodied fashion using devices that translate natural body motions into the virtual experience. In the context of education, the virtual environments of Second Life are a prominent example of a platform most commonly experienced using non-immersive desktop VR. For IVR educational use is currently explored for multiple immersive online social platforms which generally convey more social cues and allow participants to interact in ways that more closely resemble real life situations.

CONTACT Tycho T. De Back t.t.deback@uvt.nl

The promise of VR to be beneficial for learning is supported by theoretical frameworks such as the one proposed by Dalgarno and Lee (2010), suggesting the use of VR may yield a range of learning benefits. Fowler (2015) extended the framework of Dalgarno and Lee (2010) to incorporate a stronger pedagogical approach matching potential learning benefits of virtual learning environments (VLEs) to the learning requirements. The frameworks of Dalgarno and Lee (2010) and Fowler (2015) are informative as a general directive for the design of pedagogically sound VLEs. The cognitive theory of multimedia learning (Mayer, 2009, 2014) in turn is informative in making design decisions for optimizing cognitive load in virtual environments to the benefit of learning.

Even though these frameworks provide a theoretical foundation, they are not specific to immersive VLEs and contain few practical recommendations for their actual design. This is unfortunate because immersive VLEs hold unique possibilities for stimulus delivery and embodied interaction which can benefit from informed design decisions to be leveraged effectively.

Moreover, even if recommendations specific to immersive VLEs were to be made available, the design process may remain daunting given the large number of design decisions involved. Concrete examples of the implementation of immersive VLEs can assist this process, yet currently, few such examples exist. This is especially the case for collaborative settings (but see Šašinka et al., 2018 for a notable exception), which have the advantage of providing immersive education to groups of learners simultaneously while reducing required lab time.

The aim of the current paper is to facilitate the informed design of immersive VLEs. To this end, we first revisit the theoretical frameworks of Dalgarno and Lee (2010), Fowler (2015) and Mayer (2009, 2014) and highlight aspects mostly absent from these frameworks, yet which may be of critical importance for fostering learning in immersive VLEs. Based on these frameworks, we provide a set of recommendations to inform the design of immersive VLEs. We do this using a collaborative VLE created in line with these recommendations, and which was shown to yield learning gains in two prior studies. Additionally, we present a novel method of automated performance analysis as an example of the way in which immersive VLEs can be applied to yield further benefits for both learners and educators.

Theoretical framework

Framework of Dalgarno and Lee (2010) and Fowler (2015)

Dalgarno and Lee (2010) present a comprehensive framework of potential learning benefits of VLEs, consisting of spatial knowledge representation, experiential learning, engagement, contextual learning and collaborative learning. The framework details how learning benefits may arise from tasks facilitated by features of VLEs. The features they outline are grouped into representational fidelity (i.e. "realistic display of environment", "smooth display of view changes and object motion", "user representation"), and learner interaction (i.e. "embodied actions", "embodied verbal and non-verbal communication", "control of environment attributes and behavior"). Thus, IVR may provide additional benefits to learning over non-IVR by amplifying representational fidelity and learner interaction. Examples of how representational fidelity may be increased using IVR are affordances of increased field of view and display size (McMahan et al., 2012) and improved realism of avatars (digital selves). Additionally, learner interaction quality may for example be enhanced through increased embodiment, reflecting natural motions into the virtual environment.

Fowler (2015) extended Dalgarno and Lee's (2010) framework by including a more explicit pedagogical approach. Fowler first introduced the learning stages of Mayes and Fowler (1999) to explain how they may be represented in learning experiences within VLEs. In the first stage of conceptualization, the learner is familiarized with the topic and gains a basic understanding. In virtual environments, this is, for instance, a situation where a given concept is presented visually and which the learner can freely examine and interact with. The second stage of construction involves active interaction with the topic, leading to enhanced understanding. In virtual environments, this is associated with higher realism, with possibilities for increased hands-on interactivity with the topic. The third and last stage, dialogue, concerns discussion with others to verify and further solidify understanding. In virtual environments, this can for example be facilitated through the representation of oneself and others as avatars.

Fowler introduces a pedagogical "design for learning" approach in which intended learning outcomes are first determined, to arrive at learning requirements for each of the three learning stages. These learning requirements are akin to the learning benefits of Dalgarno and Lee (2010). Next, it is assessed if the learning requirements for each stage are sufficiently supported by potential learning benefits of VLEs and their underlying features. Thus, the framework requires critical evaluation of whether the potential learning benefits and employed features (e.g. avatars, embodied actions) are in line with the needs of the learning requirements (Fowler, 2015).

The frameworks presented by Dalgarno and Lee (2010) and Fowler (2015) are informative on broadly delineated features of VLEs to employ. However, it provides little guidance on the design choices to consider that exist within these features, nor does it facilitate design choices regarding VLE structure. This is especially relevant for IVR settings with opportunities for amplified representational fidelity and learner interaction, such as CAVEs. Three aspects need further discussion because they seem critical for IVR settings: (1) cognitive load, (2) collaborative learning and (3) gamification.

Cognitive load

Compared to conventional non-immersive media, IVR can provide stimuli with higher representational fidelity and more involved learner interaction. Educational content can for example be displayed life-size and in stereoscopic 3D, and maybe interacted with using naturalistic hand and body motions. Given the possibility for rich stimuli and involved learner interactions, careful consideration of appropriate instructional design of immersive VLEs thus becomes all the more important so as not to add unnecessary cognitive load to the user, thereby diminishing the potential for learning.

Cognitive load theory posits that mental effort associated with a specific cognitive task consists of three types: (1) intrinsic cognitive load, resulting from the inherent or intrinsic difficulty of a subject of study; (2) extraneous cognitive load, imposed when the quality of the learning material is not optimal by containing elements irrelevant or extraneous to learning. This is problematic as cognitive resources are limited, leaving less capacity to direct learning (Sweller et al., 2011); (3) germane (essential) cognitive load. Different from intrinsic and extraneous cognitive load, germane cognitive load refers to resources spent on learning and is desirable (Sweller et al., 1998, 2011).

The cognitive theory of multimedia learning (Mayer, 2009, 2014) draws from cognitive load theory and presents ways to address processing related to cognitive load when using multimedia. In the context of IVR, Parong and Mayer (2018) highlight two principles. In the coherence principle, audiovisual elements not contributing to learning are removed to prevent unwanted extraneous processing (Mayer et al., 2001). This may apply to VR environments as well, where a violation of the coherence principle includes the unnecessary addition of visual effects (Parong & Mayer, 2018). In the segmenting principle, educational content is divided or segmented by letting the learner decide when to progress, allowing information to be absorbed before moving on to the next segment and thus serves to manage germane processing (Mayer, 2009, 2014). Additionally, based on generative learning theory, Parong and Mayer (2018) describe the use of generative processing, where information to be learned is organized and integrated into memory to enhance understanding, fostered using various activities including creating summaries, drawing and self-explaining (Fiorella & Mayer, 2015, 2016). An alternative way of managing germane processing, in addition to those highlighted by Parong and Mayer (2018), is pre-training. In pre-training, learners are familiarized with the names and workings of the components of a system and only thereafter proceed to engage in the task of learning the educational content contained within the system. Not having to do these two processes simultaneously may reduce cognitive load when learning (Mayer & Moreno, 2003).

The notion that cognitive load is to be considered in the use of VR for learning is supported by the work of several researchers. Meyer et al. (2019) compared IVR and video conditions with and without pre-training on cell biology in university students. Pre-training as a method of decreasing cognitive load had a positive impact on knowledge and transfer, and only for IVR and is indicative of the benefit of pre-training in immersive settings.

In a first experiment, Parong and Mayer (2018) failed to obtain learning gains for factual information of an IVR simulation over a PowerPoint condition, and concluded that the immersive condition was not purpose-built and violated the coherence and segmenting principles. Consequently, Parong and Mayer (2018) highlighted the need for studies verifying whether the principles can be applied to IVR to increase learning.

In a second experiment, Parong and Mayer (2018) employed generative processing by adding a summarizing task leading to higher learning benefits, demonstrating the potential of generative processing in immersive VLEs. Regarding generative processing, Fiorella and Mayer (2016) suggested that for topics with a high amount of spatial information, generative learning strategies other than summarizing may be more suited. The presented alternative generative learning strategies include mapping (involving the generation of spatial configurations and relationships between received parts of information), drawing, and enacting of educational content using embodied activities.

The relevance of cognitive load in virtual environments is also supported by Lee and Wong (2014), who examined biology learning in non-IVR and suggested the use of VLEs might lead to reduced extraneous- and increased germane cognitive load. This could occur by enabling dynamic interaction with educational content more in line with learner needs and by presenting objects in their natural form, thereby removing the need to mentally convert 2D images into their 3D equivalents. Echoing Lee and Wong (2014), Lin et al. (2019) observed higher germane cognitive load after an IVR condition compared to a non-immersive map condition, yet this did not coincide with a difference in extraneous cognitive load and provides additional support for the relevance of considering cognitive load in immersive settings.

Summarizing, design choices for VLEs may be informed by aspects of cognitive load, with (1) removal of superfluous elements beneficial for reducing extraneous cognitive load; (2) segmentation and pre-training beneficial for managing germane cognitive load; and (3) the use of relevant VR features to foster generative processing. For collaborative settings, cognitive load is specifically important because overload in one individual has the potential to affect performance of multiple interdependent others.

Collaborative learning

Collaborative learning occurs when groups of two or more persons learn something together (Dillenbourg, 1999). Empirical evidence points towards a positive effect of collaboration on learning-, affective- and other outcomes (Casey & Goodyear, 2015; Johnson et al., 2000; Kyndt et al., 2013; Lou et al., 1996; Slavin, 1980). IVR may benefit collaborative learning by enabling higher levels of embodiment, (non-)verbal communication and avatar realism, which are all components of learner interaction, one of two feature groups potentially affording learning benefits in the framework of Dalgarno and Lee (2010). Indeed, empirical studies show higher embodiment leads to higher levels of collaborative learning (Malinverni & Burguès, 2015), while Fowler (2015) suggested avatars to be of importance in the learning stages of Mayes and Fowler (1999).

From the preceding paragraph, the relevance of collaborative learning becomes apparent as an opportunity to further extend potential learning gains in IVR. In addition, designing for collaborative learning is relevant to educational institutions faced with the issue of how to provide immersive learning to large numbers of students when available time and resources are limited. In this situation, collaborative learning is beneficial by allowing groups of students to efficiently share VR resources and engage in immersive learning simultaneously (De Back et al., 2020).

Gamification

Another design consideration for increasing learning in immersive VLEs is the use of gamification, in which game elements are introduced in a non-game learning environment in order to increase engagement (Koivisto & Hamari, 2019). A prominent reason for the use of gamification is its ability to increase motivation, which is predictive of performance (Dichev & Dicheva, 2017). Gamification may be conducive to collaborative learning, as gamification is reported to aid collaboration (Hassan et al., 2021; Knutas et al., 2014; Romero et al., 2012), and the addition of gamification elements such as direct feedback may stimulate self-efficacy, benefiting learning (Parong & Mayer, 2018). Other often employed gamification design principles include clear goal setting, awarding of scores and badges, and levels (Koivisto & Hamari, 2019). Reviews of empirical studies on gamification for education have generally indicated a positive effect on learning outcomes (Clark et al., 2016; Hamari et al., 2014; Wouters et al., 2013; but see Koivisto & Hamari, 2019 for mixed results).

Building on the theoretical framework presented here, the next section provides recommendations for the informed design of immersive VLEs.

Recommendations for designing immersive VLEs

Optimizing cognitive load

Several strategies may be applied for optimizing cognitive load in immersive VLEs. In accordance with the pre-training principle, learners may benefit from receiving information about the names and workings of the educational content in the environment before commencing learning. Immersive environments are known for their ability to induce strong presence, the feeling of existing inside a virtual environment (Heeter, 1992; Steuer, 1992). Both the sense of presence and embodied actions in immersive settings can be novel to learners. We, therefore, recommend that pre-training in immersive settings is combined with a practice period to give learners time to become accustomed to the environment and to familiarize themselves with the embodied interactions to be performed.

Usability in VLEs is reported to be positively related to self-efficacy, which is in turn positively linked to learning (Merchant et al., 2012), yet is deserving of more attention in the design of VLEs (Granić et al., 2020). Usability has been shown to be affected by VR features (Merchant et al., 2012), and can be increased by making interactions simple and intuitive. By making use of pre-existing knowledge, this has the added benefit of requiring less cognitive effort (Greitzer et al., 2007), and can be achieved with a simple control scheme and embodied real-life actions (e.g. push/pull, hold/release, draw/erase).

Throughout the learning experience, it is helpful to provide clear instructions and direct feedback (Koivisto & Hamari, 2019). Especially in learner-centered designs, a rule set, one of the key components in gamified scenarios (Morris et al., 2013) can be implemented into game logic to disallow interactions with the environment which stray from the goal of the task, preventing learner frustration. Distinct discrepancies between real and virtual locomotion have been shown to increase cognitive demand (Marsh et al., 2013; Zanbaka et al., 2005). Consequently, when locomotion is required for the learning task, it is recommended that real and virtual locomotion is matched. Where physical limitations prevent this, alternative options can be considered (e.g. redirected walking, teleportation, joystick-based locomotion, Boletsis, 2017) while taking possible effects on usability into account.

Fostering collaborative learning

The optimization of cognitive load is of even greater importance for collaborative learning, as learners are interdependent in collaborative settings, and are affected if even a single learner experiences high extraneous cognitive load. To ensure equal learning opportunities, it is suggested to make the instructions and educational content equally accessible to all, irrespective of the position of the learners within the environment. Collaborative learning benefits from active and equal participation (Vuopala et al., 2016) and can be fostered using interdependence (Johnson et al., 2014). In situations where one learner is interacting with others observing, discussion and learning can be stimulated by enhancing the saliency of the actions of the interacting learner with audiovisual effects. This is congruent with studies using awareness tools to focus attention, with Povis and Crowley (2015) indicating the use of flashlights enhanced joined attention and was linked to increased learning dialogue, and with Schneider and Pea (2013) showing that gaze awareness using eye-tracking improved joint attention and learning outcomes.

Leveraging platform-specific affordances

Both VR headsets and CAVEs have platform-specific affordances which can be leveraged to benefit learning through afforded tasks. Consistent with Fowler (2015), the intended learning outcomes and learning requirements determine which affordances may potentially enhance learning. For VR headsets, one consideration is to employ embodied interactions to facilitate natural exploration of the educational content, in support of generative processing/germane cognitive load. Voice chat and embodied expressive avatars may be employed to facilitate collaboration (Dalgarno & Lee, 2010). In CAVEs, the level of embodiment is naturally high as learners can see their physical body and that of others while immersed in the virtual environment. Embodied interactions with the environment in modern CAVEs are most commonly enabled using a 3D mouse (Muhanna, 2015).

Mitigating platform-specific limitations

To enhance learning in virtual environments further, one strategy is to mitigate platform-specific limitations. For commodity VR headsets, facial expressions important for social interaction are presently not translated into the virtual environment (Hickson et al., 2019). To improve collaboration, expressive avatars can be used as detailed above (Neji & Ben Ammar, 2007). CAVEs typically provide correct perspective and correct stereoscopic parallax on the virtual environment to one learner at a time (DeFanti et al., 2011). Where possible, the perspective issue can be mitigated by making objects positionally static when object movement is not required for learning. A similar strategy is to reduce parallax by placing virtual objects further away (Gillam, 2007), yet still allow objects of interest to be examined up close when necessary. By providing each learner with position tracked 3D glasses, on-the-fly perspective switching between learners can be enabled in CAVEs, which we will detail in the next section.

Additional strategies

Several additional strategies may be considered to obtain further benefits of VLEs. To increase motivation and learning, one strategy is to use gamification elements. The use of direct feedback mentioned previously may be expanded with other gamification elements if needed. Examples of these elements are the use of levels and scores, while ensuring that this does not take attention away from learning. Additionally, a strategy to increase time- and cost-effectiveness in the design of VLEs is to consider whether the environment needs to be adaptable to fit other topics and/or features. If so, a modular design may be preferred so educational content can be interchanged and new functionality added. A final strategy is to use automated performance analysis to obtain additional benefits of VLEs. The implementation of automated performance analysis in immersive settings is novel and will be discussed in detail later.

The proposed recommendations presented here serve to facilitate the informed design of VLEs. Table 1 presents an overview of both the theoretical framework and our recommendations, which are an operationalization of the framework specific to immersive settings. How to best implement the recommendations to meet specific learning goals may remain non-trivial given the vast

Table 1. Condensed theoretical framework (left) and its operationalization (right) to facilitate design for learning in immersive virtual learning environments.

Condensed overview of theoretical framework	**Recommendations to facilitate design for learning in immersive virtual environments** (De Back et al., 2021)
Main steps in the design for learning approach (Fowler, 2015): Determine: • Intended learning outcomes • Learning requirements (objectives: what, activities: how) for the learning stages of Mayes and Fowler (1999): Conceptualization, construction, dialogue • Teaching and learning approach • Yes/no sufficient match between learning requirements and potential learning benefits of virtual environments in the framework of Dalgarno and Lee (2010): Spatial knowledge representation, experiential learning, engagement, contextual learning, collaborative learning • If match is sufficient: Representational fidelity and learner interaction elements aligned with learning requirements, yielding an optimal benefit-cost trade-off • Learning specification	Broad design choices: Apply frameworks of Dalgarno and Lee (2010) and Fowler (2015) *Apply strategies to optimize cognitive load* • Coherence: Restrict the environment to elements likely to benefit learning • Segmenting (if applicable): Let learner decide the receipt of new content • Pre-training: Implement an introduction and practice phase • Use an intuitive user interface with natural embodied interactions • Generative processing: Foster generative actions using embodiment • Guide learning: Clear instructions, direct feedback and game logic • Match real and virtual locomotion. Otherwise: Provide locomotion options *Consider strategies to foster collaborative learning* • Design the content and instructions to be equally accessible to all learners • Build active participation and interdependence into the instructional design • Make learner interactions clearly visible to foster discussion and learning
Cognitive theory of multimedia learning (Mayer, 2009, 2014) Highlights of possibilities for reducing cognitive load: Segmenting principle: Allow learner to decide when to receive new information Pre-training principle: Create understanding of subparts before commencing learning Coherence principle: Remove information extraneous to learning Generative processing: Employ activities deepening understanding	*Leverage platform specific affordances* • VR: Embodiment, voice, avatars. CAVE: Embodied actions with 3D mouse *Mitigate platform specific limitations* • VR: No facial expressions: Foster collaboration with other means (see above) • CAVE: Leverage single-user interaction to foster generative collaboration. Limit parallax if extraneous to learning, implement perspective switching *Strategies for additional benefits* • Explore the use of gamification elements while retaining focus on learning • Reduce effort, time & cost: Use a modular design when adaptability is desired • Use automated performance analysis to benefit learners and educators

Note. VR = virtual reality.

possibilities and the large amount of design choices to be made. To facilitate this process, the next section demonstrates the implementation of the recommendations using a concrete example of an actual operating immersive VLE that was shown to yield learning gains.

Implementation

Following the frameworks of Dalgarno and Lee (2010) and Fowler (2015) as well as Mayer (2009, 2014), while considering the recommendations outlined above, we developed a collaborative VLE for CAVEs as well as VR headsets. The environment resulted from a two-year iterative design process and has since been operational in a cognitive science course at Tilburg University, The Netherlands. The subject of the collaborative VLE was neuroanatomy. Anatomy is known to be a subject that students struggle with due to the complex 3D nature of the different body parts it involves as well as due to the vast unfamiliar terminology which comes into play (Hackett & Proctor, 2016). Neuroanatomy in particular has been perceived by undergraduate students to be difficult and sometimes even leads to instances of neurophobia, or fear of neurology (Javaid et al., 2018). The choice to use neuroanatomy was made in an effort to alleviate the inherent difficulty of the subject by leveraging the affordances of collaborative VLEs. The environment created included the understanding of brain structures, their interconnections and broader spatial relationships. The information contained in the environment was obtained from a textbook chapter on neuroanatomy in a book by Friedenberg and Silverman (2006). 3D models of brain areas were retrieved from an anatomy database (BodyParts3D/Anatomography, The Database Center for Life Science, CC Attribution-Share Alike 2.1 Japan). The environment incorporated free exploration, knowledge construction and collaboration to foster generative learning.

The intended learning outcomes require spatial understanding and match with the learning benefit in the framework of Dalgarno and Lee (2010) of spatial knowledge representation. Additionally, the representational fidelity of CAVEs allows a realistic spatial presentation of brain areas in full stereoscopic 3D. For anatomy learning specifically, the display of anatomical structures in 3D is further supported by review studies indicating the use of 3D display technologies over conventional 2D methods can yield cognitive benefits (Hackett & Proctor, 2016; Yammine & Violato, 2015). Furthermore, highly embodied learner interaction allows brain areas to be observed from different angles using natural motions, supporting intuitive exploration of their complex structure and thus fostering generative processing. Collaborative learning may be facilitated in CAVEs as learners are represented physically and thus retain expressiveness, enhancing non-verbal communication. The intended learning outcomes may thus be facilitated as there is a match with learning tasks, potentially afforded by the features of CAVEs. Needless to say, even if the tasks performed here in the VLE of the CAVE can also be successfully carried out in non-3D or non-CAVE environments, questions regarding the design and implementation of learning in immersed collaborative VLEs remain valid.

To attain the intended learning outcomes, the broad learning requirements were set to include (spatial) knowledge construction and collaborative learning, and more generally for learning to be engaging and experiential. Accompanying learning activities included free exploration, reproducing educational content through active discussion as well as reviewing information to further consolidate learning. Storyboards and iterative experimentation were used to determine how to best attain the learning requirements. The virtual environment held up to six learners at a time and consisted of four walls corresponding to the physical walls of the CAVE, depicted in Figure 1.

Pre-training was implemented in an introduction and practice phase in which the experimenter explained the workings of the environment, and the learners familiarized themselves with the use of embodied interaction. In short, the task involved learning the size, shape, name and functions of individual brain areas, as well as their location within the brain. These different elements were sorted and presented on separate walls of the environment, with large connecting lines indicating which elements belonged together. The elements were presented on separate walls as to show them at such a size that they would not be easily obstructed from view by one or more learners in a group

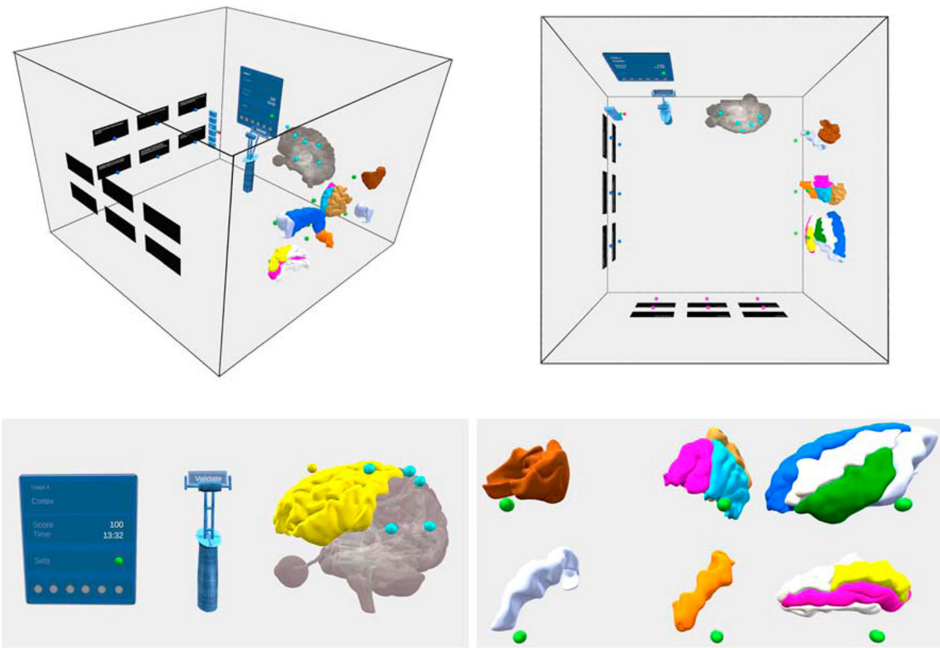

Figure 1. Structure of virtual environment inside the CAVE: side and bird's-eye view (top left, top right), scoreboard, lever to verify answers, model of whole brain (bottom left), individual brain areas (bottom right). Colored dots indicate areas where connections can be drawn between different elements.

and was in support of collaborative use. The core task was to learn the correct connections between the elements. The segmentation principle was implemented by dividing (segmenting) the educational content into different stages, which themselves were each compartmentalized into three phases: content exploration, reconstruction and review. This task differentiation bears surface-level similarities to the conceptualization, construction and dialogue learning stages of Mayes and Fowler (1999) and is consistent with the concept of generative processing. In the exploration phase, all elements and their links indicated with large interconnecting lines were to be memorized through free exploration. In the reconstruction phase, the connecting lines were removed, and learners were to engage in discussion to determine which elements belonged to each other, and to generatively reconnect these by physically redrawing (i.e. reconstructing) their interconnecting lines.

To foster the embodied and generative aspect further, this task of redrawing lines needed to be performed in 3D space and thus required the learners to walk around in the CAVE in order to connect the lines to the correct elements shown on the four CAVE walls. Even though it is possible to consider interfaces where elements are connected in a non-3D space, such interfaces would lack the embodiment of the students and miss the collaborative nature of the learning experience. Lastly, the review phase provided all information as an integrated whole on one wall of the environment, used to reflect on the material and further solidify learning before proceeding to the next stage. Consistent with the segmentation principle, progression to the next phase of a stage was initiated by the learners, ensuring that new information was only presented when the learners were ready for it.

With the aim of reducing extraneous cognitive load and in line with the coherence principle, we refrained from using superfluous audiovisual effects and restricted visual stimuli to the content to be learned and to features likely to facilitate learning. Several measures increased the ease of use. Points of interaction were clearly designated using large dots. Interaction methods were designed to be intuitive and used a 3D mouse and a virtual pair of scissors to respectively draw and cut connecting lines between elements. A pull on a virtual lever verified

Table 2. Examples of the technical implementation of the recommendations into the collaborative virtual learning environment.

Recommendations	Implementation examples
Optimization of cognitive load	
Coherence	Stimuli and interactions limited to those consistent with the intended learning outcomes
Segmenting	Educational content divided into stages and phases. Progression under learner control
Pre-training	Introduction and practice phase before start of learning
Intuitive interface with embodied interactions	Tracked 3D mouse and glasses, virtual scissors and lever for interactions, large dots to indicate points of interaction, perspective switching with single button push
Generative processing	Exploration with natural motions, embodied line drawing and cutting in 3D space, task split into exploration, reconstruction and review
Guide learning	Instructions, direct feedback on answers, game logic to keep learners on track
Match real and virtual locomotion	Natural movement using four wall CAVE system
Foster collaborative learning	
Content and instructions equally accessible to all	Content prominently displayed on CAVE walls, instructions in both text and audio, interactions relevant to learning kept visible until no longer needed
Design for active participation and interdependence	Drawing and walking required to finish stage, turn-based alternation between direct interaction and group discussion
Clearly visualize learner interactions	Interactions indicated using bright colored lines
Leverage platform specific affordances	
Embodied actions with 3D mouse	Tracked 3D mouse used for embodied interaction with the educational content
Mitigate platform specific limitations	
Leverage single-user interaction to foster generative collaboration	Tracked glasses for all learners to overcome single-user perspective limitation of CAVE, reduced parallax by placing virtual objects at distance from learners
Strategies for additional benefits	
Gamification to retain focus on learning	Clear learning goals, stages, direct feedback, scores
Reduce effort, time & cost	Generic task and modular structure to easily reconfigure the virtual learning environment
Automated performance analysis	Enabled using data of tracked 3D mouse and glasses

whether connecting lines were drawn correctly. Perspective switching (discussed more in detail later) was made to be intuitive as well by having learners hold the 3D mouse like a smartphone and push a single button. To support multi-user collaborative settings here the use of a 3D mouse instead of alternative interaction methods such as automated gesture recognition additionally allowed to differentiate between gestures made with the mouse to interact with the environment and regular non-mouse gestures made between users as part of social interaction. Game logic was implemented to disallow actions not conducive to learning, such as unnecessary redrawing of lines. As a result of these design decisions, the use of the game was straightforward, facilitating attention to remain on the learning task. Natural interaction (such as gesture recognition) instead of the use of a 3D mouse is an alternative option for these environments. An overview of the technical implementation of the recommendations into the collaborative VLE is provided in Table 2, while Figure 2 presents visual examples.

To facilitate equal information access for all learners, the educational content was prominently displayed, and instructions were provided both auditorily and in text on all four walls. Providing the same information both auditorily and in text is in violation of the redundancy principle of Mayer (2009), which posits that this may result in unnecessary cognitive load. In our specific case, informal piloting and feedback sessions, however, showed that using both audio and text was helpful in ensuring all learners took notice of the instructions, irrespective of their physical location inside the environment. Coordinated attention on the task was further facilitated by the set stage structure of exploration, reconstruction and review. Active participation was built into the instructional design by requiring learners to walk through the environment and draw lines between elements in order to successfully complete each stage. To further establish equal information access, drawn lines remained visible as a point of reference for active discussion until no longer

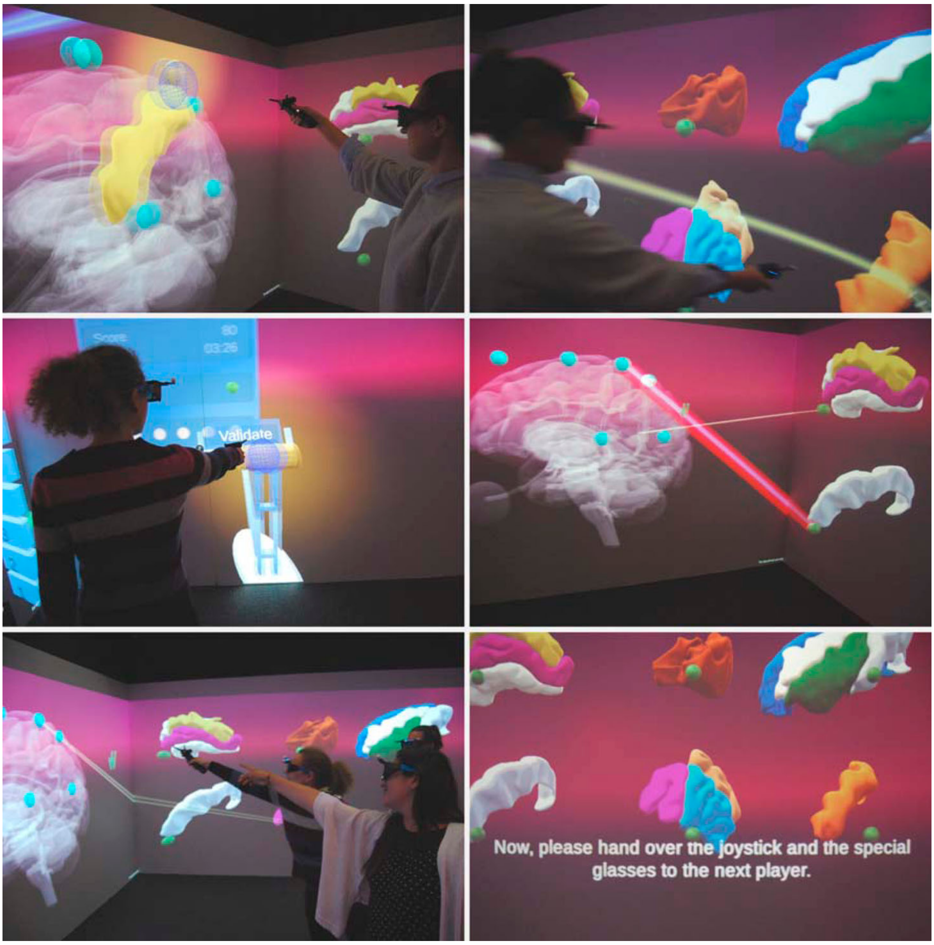

Figure 2. Examples of implemented strategies to foster learning. Intuitive interaction: using 3D mouse to highlight areas of interest (top left), provide answer by drawing connection (top right), pull lever to verify answer (middle left). Feedback on right/wrong answers (middle right), collaborative learning (bottom left), turn taking instruction to foster active participation (bottom right).

necessary. Active participation does not preclude the possibility of social loafing, which may negatively affect collaborative learning (Kyndt et al., 2013; Slavin, 1990). To reduce the likelihood of social loafing, users were prompted to switch turns after a set number of interactions, ensuring all learners directly engaged with the environment.

The design of the virtual environment was iteratively improved to address issues potentially affecting learning. The first was to incorporate gamification elements, but to strike the right balance to maintain a focus on learning. For example, direct feedback clearly benefitted learners by showing which answers were correct. Yet this was not the case for several other gamification elements. In some pilot studies, scores of correct answers were largely ignored by learners. Consistent with the coherence principle, audiovisual elements of the score system were, therefore, removed such that it only served as a reminder of performance, attended to on an individual basis. Similarly, a set time limit per stage was removed, reasoning that this might diminish learning in groups needing more time to master all information. Besides gamification elements, efforts were made to prevent cheating. Part of the learning task was to learn the name and function of brain areas, described in labels placed in the environment. To prevent learners from simply memorizing the spatial location of these labels without memorizing their content, label and brain area location was randomized in the reconstruction phase. Lastly, several measures were taken to increase comfort. For example, addressing the limitation of conventional CAVEs of displaying a perspective correct view on the environment to a single learner at a time, on-the-fly perspective switching was implemented by having all learners wear 3D glasses whose position was tracked in space. The design choices combined resulted in a collaborative VLE in support of the study of neuroanatomy, within which users could intuitively explore complex brain structures and other educational content in 3D space and where collaboration allowed users to learn from each other using natural face-to-face interactions and embodied gestures, all elements which may be hard to accomplish for groups of users when using conventional 2D media.

Empirical support

The implemented collaborative VLE has yielded learning gains, as evidenced in two empirical studies (De Back et al., 2020; De Back et al., 2021). The first study compared learning gains in persons recruited from a participant pool ($n = 40$) who were all unfamiliar with the neuroanatomy subject of the VLE. Two conditions were conducted: (1) collaborative use of the virtual environment in two- to four-person groups, and (2) textbook learning. Learning was modulated by individual differences in spatial ability (the ability to accurately perceive a scene and mentally reconstruct and reconfigure it, Carroll, 1993; Höffler, 2010). The findings of the first study showed that learning gains after collaborative use of the virtual environment exceeded those of conventional textbook learning, especially in participants with low spatial ability. It was additionally shown that the virtual environment with its implementation of the recommendations including pre-training, segmenting, guided learning and incorporated strategies for collaborative learning together enabled an engaging learning experience and supported use by multiple learners at a time.

The results of the first study of learning gains in groups of up to four persons allowed the same virtual environment to be applied in a second study as part of the real-life educational setting of an undergraduate cognitive science course ($n = 158$). For the virtual environment to be used as part of a course with relatively large student numbers, the implementation of collaborative learning was essential. Without it, providing VR sessions for students individually would have been difficult to accomplish given the limited time and resources that were available. This study replicated the learning gains of the first study. Two factors of practical relevance for applying immersive collaborative environments in large academic courses were investigated, namely the potential effect on learning gains of group size (one-person, two- to four-person, five- to six-person groups) and period of application (before, during or after the course). Results indicated learning gains were present for all group sizes, with higher gains for smaller groups. Period of application did not affect learning.

Both studies indicated the virtual environment central to this paper yielded learning gains at all examined levels of spatial ability, group size and period of application.

Automated performance analysis

Central to the design for learning approach of Fowler (2015) is to increase learning benefits while considering the trade-off of increased costs. Here, we examine the use of automated performance analysis to extend benefits of VLEs.

Virtual environments collect data on learner actions to make the environment interactive. This data is conventionally discarded upon use, yet is of potential value to educators and learners and can be retained by adding a logging feature. The potential benefit for learning is especially exciting since little effort is required to enable logging, even if the environment was not designed with this functionality in mind. Chodos et al. (2014) described how learner data in virtual environments may be logged and analyzed for non-immersive 2D desktop environments. This leaves out immersive environments, which gather more body-related information such as gestures, body posture and learner position, potentially indicative of learning. Fardinpour and Reiners (2014) recognized the

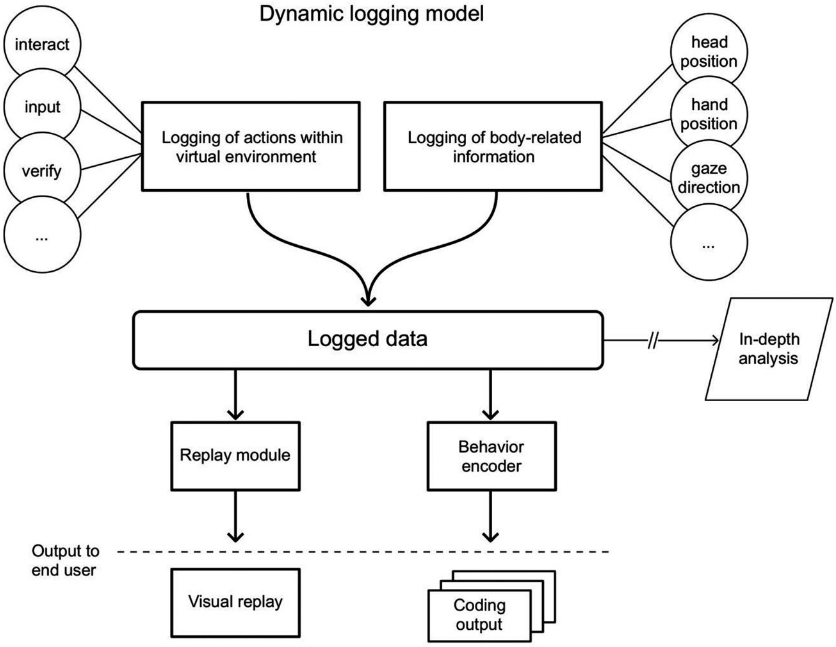

Figure 3. Dynamic logging model. The model shows how incoming data streams on learner actions (indicated with circles in the figure) with the virtual environment and body-related information are logged, making the information available for in-depth analysis, and is automatically processed by a replay module and a behavior encoder, respectively outputting: (1) a visual replay of learner behavior; (2) an analysis overview of individual and collaborative behavior types of choice.

rich behavioral information of immersive technologies and conceptualized a system of automated analysis of goal-oriented user actions, yet did not implement the system.

We implemented a model of automated performance analysis in immersive environments. The model dynamically adapts to the feature set of different immersive technologies and allows analysis of various learning-related behavior types, recognizing that their availability depends both on the immersive technology used and the nature of the topic at hand. This is different from Chodos et al. (2014), who coded for a fixed number of user actions and traces. The model, presented in Figure 3, explains how learner interactions and other behavioral data are logged and automatically analyzed. Next, the model shows how the result of the analysis is presented to the end-user. Finally, it is shown how the log data can be utilized to create a visual replay of learner behavior.

Implementing the model, Figure 4 (top) presents example results of an automated analysis of learner actions. In addition to obvious assessment variables such as the number of correct and incorrect answers and total time on task, additional variables are displayed which may be indicative of learning processes and/or behavioral engagement, reported to be positively related to academic achievement (Lei et al., 2018). Examples are time spent directly engaging with the environment by the user controlling the game at that time (leader time) and time spent paying attention to those leaders by other users in the group (leaders observed). Measures of more general activity and engagement are total meters walked (distance walked) and head rotations expressed as multiples of 360-degrees (full head rotations). Additionally, face-to-face time with others (face time) may for instance function as an indicator of the level of social interaction, relevant to learning effectiveness (Roberts, 2005). Besides these data, Figure 4 additionally depicts learners interacting with the environment (bottom: left), and a demonstration of the replay feature (bottom: right) visualizing the same time segment yet now using the logged data.

Consistent with Chodos et al. (2014) and Fardinpour and Reiners (2014), the implemented model enables automated performance monitoring of groups and individuals for evaluation by educators.

Group total

Correct answers:	62
Incorrect answers:	8
Total time:	43 min

Individual totals

User	Leader time	Correct answers	Incorrect answers	Leaders observed	Face time	Distance to others	Distance walked	Full head rotations
1	1023 sec	21	2	254 sec	4.20 sec	1.56 m	757 m	648
2	964 sec	16	1	339 sec	6.20 sec	1.35 m	683 m	818
3	579 sec	11	3	514 sec	8.30 sec	1.78 m	546 m	589
4	435 sec	14	2	498 sec	7.90 sec	1.26 m	495 m	452

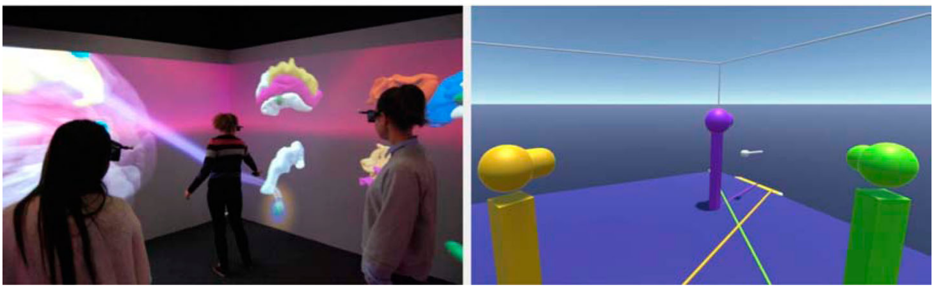

Figure 4. Implementation of the model. Top: Example results of automated (behavioral) analysis, showing group statistics as well as individual performance metrics. The metrics are explained in the section preceding this figure. Bottom: Actual scene inside CAVE (left) and corresponding visual replay, with colored lines indicating gaze direction of the learners (right).

For example, in Figure 4 (top) under group total, we see that the group as a whole performed well in having 62 correct and 8 incorrect answers. Looking at the individual totals, we can additionally see that user 1 had the highest number of correct answers of the group and was the most engaged (highest scores for leader time, distance walked and full head rotations). By contrast, user 3 is shown to have had the fewest correct answers, stood at the largest distance to others and observed others most by having the highest values for the leaders observed and face time metrics. The respective high and low levels of engagement of users 1 and 3 indicated by the aforementioned metrics may be of interest to educators to assess a possible connection with differential learning outcomes. The insights thus gained may be utilized to determine whether stimulating specific activities may be beneficial to learning. By extension, an example of a practical use case of the automated performance monitoring metrics is to verify if there is evidence of social loafing during the immersive learning experience when learning gains are low as determined using pre- and post-tests. These same metrics may be employed to assess learning processes. For example, the engagement level of individual learners expressed using metrics of choice such as distance walked and leader time could be used to explore the relationship with learning gains. Additionally, such metrics as distance walked and head rotations are indicative of the extent of physical and/or virtual motion and may be examined to verify a possible relationship with self-reported instances of motion sickness, potentially affecting learning (Lackner, 2014).

In addition to automated performance monitoring, the implemented model incorporates a replay feature, allowing educators to "relive" learner actions from any vantage- or time point to gain further insight into their learning process. This is different from Chodos et al. (2014), who describe the use of static video recordings which do not allow dynamic perspective changes beneficial to enhance understanding of learner performance.

On a more practical level, the replay feature enables quick and intuitive verification of tracking data quality, allowing early detection and mitigation of potential issues. This is especially useful in collaborative settings where the possibility of tracking errors is higher as there are more data streams to log. The replay feature and the results of the automated performance analysis additionally allow learners to reflect on their individual and group level performance directly after exiting the immersive learning experience. Finally, position and gaze direction data can be utilized to verify if a virtual environment sufficiently guides attention to important hotspots, thus serving as an informative tool for the design of effective educational virtual environments.

Discussion

The current paper operationalized the presented theoretical frameworks and provided a set of practical recommendations to increase potential learning gains. Different from the frameworks of Dalgarno and Lee (2010) and Fowler (2015), our recommendations contribute to existing work by targeting IVR settings specifically. Because of the focus on immersive settings, our recommendations encapsulate a range of opportunities to increase the benefits of immersive learning. A highlighted opportunity is the design of VLEs to support collaborative learning, allowing for readily applying immersive learning in education in a time and resource-efficient manner. To facilitate the application of the recommendations, a collaborative VLE was presented as an example of their implementation in practice. The VLE is currently used in a large undergraduate course and was shown to yield learning gains in two empirical studies.

A largely unexplored advantage of VLEs is that they can be applied to yield further learning benefits given the right choices. One example is to design VLEs to fit multiple topics, extending their range of use. The VLE of the current paper serves to demonstrate this by employing a generic learning task applicable to different topics and by using a modular environment structure, allowing the educational content to be swapped out with ease. A second example of further learning benefits showed the potential of adding an automated performance analysis feature for unlocking novel ways to improve learning while reducing cost. Fardinpour and Reiners (2014) previously detailed the potential for automated analysis of user actions, yet did not implement the proposed system.

The VLE presented in the current paper was minimalistic in accordance with the needs of the topic. Intended learning outcomes of other topics and contexts will require different configurations. This is consistent with Fowler (2015), stressing the need to critically assess whether elements to be potentially included serve the intended learning outcomes. We posit that immersive technologies are no universal panacea for increasing learning and instead yield benefits for specific areas when applied effectively, to which the presented theoretical framework and the recommendations of the current paper are informative. Designing for learning also implies that if there is no match between intended learning outcomes and potential learning benefits of VLEs, conventional methods may be preferred over the use of VR (Fowler, 2015). The virtual environment of the current paper applied four ways of optimizing cognitive load of the cognitive theory of multimedia learning of Mayer (2009, 2014). These possibilities such as pre-training and generative processing were highlighted as they benefit from additional strategies and considerations in immersive settings, which were compiled in our recommendations.

The implemented virtual environment contained identical instructions in both text and audio, in violation of the redundancy principle of Mayer (2009, 2014). This was done as instruction text shown on a CAVE wall could be blocked from view if other learners in the group stood directly in front of it. The addition of audio ensured that all learners received the instructions regardless of where they were standing in the CAVE. This illustrates the importance of weighing the pros and cons of individual design decisions and their potential impact on learning outcomes.

Grounded in a theoretical framework, informed decisions were made regarding cognitive load of the virtual environment. Future studies should investigate the feasibility of taking physiological measures of cognitive load during learning to further facilitate the design of effective VLEs. The potential of measuring physiology during learning in order to gain insight into cognitive load is discussed extensively in Tinga et al. (2019, 2020), who also concluded that including these measures is especially interesting for VLE. Eye-tracking measures such as blink rates can be indicative of cognitive load (Valtchanov & Ellard, 2015) and are potentially interesting as eye-tracking has been integrated into multiple VR headsets and is thus becoming more accessible (Stein et al., 2021). In cases where taking physiological measures is not feasible, the use of cognitive load questionnaires specific to VR such as the simulation task load index can be a viable alternative (Harris et al., 2020). The presented recommendations were formulated in line with the strengths and limitations of contemporary VR headsets and CAVEs. As immersive technologies evolve, the recommendations are to be updated accordingly. Future versions should incorporate considerations specific to head-mounted augmented reality devices which at present are not readily available to educational institutions yet hold the exciting potential to support immersed collaborative learning by retaining non-verbal cues similarly to CAVEs.

Our presented model and implementation of dynamic logging creates new possibilities for immersive settings to support learning. One avenue is to explore novel ways of testing theories of embodiment and constructivism in connection with learning gains using the rich log data afforded by the model. Another exciting possibility is to investigate how automated detection of learner behaviors can be leveraged to determine which actions in immersive settings are especially conducive to learning improvements.

Conclusion

The recommendations presented in the current paper aim to support the informed design of immersive VLEs. This is further facilitated with the presented example of the implementation of the recommendations in a collaborative VLE, shown to yield learning benefits in two studies. Together, the recommendations and the concrete example of their implementation remove barriers for adoption and serve as a catalyst for the effective use of VR in education.

Acknowledgements

The authors wish to thank Anton Sluijtman and Maarten Horden for developing the VLE and Rob Vossen for creating the automated performance analysis and replay functionality. Additionally, we thank Hans van den Dool and Peter van Trier for their help with storyboarding.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was partially funded by the PACCAR foundation and DAF Trucks and on a grant to the last author by the European Union, OP Zuid, the Ministry of Economic Affairs, the Province of Noord-Brabant and the municipalities of Tilburg and Gilze Rijen [PROJ-00076].

Notes on contributors

Tycho T. De Back is a PhD student at Tilburg University, The Netherlands. His work focuses on the use of immersive technologies for learning and education.

Angelica M. Tinga was a PhD student at Tilburg University, The Netherlands. Her interests include behavioral and neurophysiological processes during learning in virtual reality. She currently is Researcher Human Factors in Vehicle Automation at SWOV, The Netherlands.

Max M. Louwerse is a Professor of Cognitive Psychology and Artificial Intelligence at Tilburg University, The Netherlands. His interests include virtual and mixed reality for learning and education and the development and use of embodied conversational agents.

ORCID

Tycho T. De Back http://orcid.org/0000-0002-2667-608X

References

Boletsis, C. (2017). The new era of virtual reality locomotion: A systematic literature review of techniques and a proposed typology. *Multimodal Technologies and Interaction*, *1*(4), 24. https://doi.org/10.3390/mti1040024

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge University Press.

Casey, A., & Goodyear, V. A. (2015). Can cooperative learning achieve the four learning outcomes of physical education? A review of literature. *Quest*, *67*(1), 56–72. https://doi.org/10.1080/00336297.2014.984733

Chodos, D., Stroulia, E., King, S., & Carbonaro, M. (2014). A framework for monitoring instructional environments in a virtual world. *British Journal of Educational Technology*, *45*(1), 24–35. https://doi.org/10.1111/j.1467-8535.2012.01370.x

Clark, D. B., Tanner-Smith, E. E., & Killingsworth, S. S. (2016). Digital games, design, and learning: A systematic review and meta-analysis. *Review of Educational Research*, *86*(1), 79–122. https://doi.org/10.3102/0034654315582065

Cruz-Neira, C., Sandin, D. J., DeFanti, T. A., Kenyon, R. V., & Hart, J. C. (1992). The CAVE: Audio visual experience automatic virtual environment. *Communications of the ACM*, *35*(6), 65–72. https://doi.org/10.1145/129888.129892

Dalgarno, B., & Lee, M. J. W. (2010). What are the learning affordances of 3-D virtual environments? *British Journal of Educational Technology*, *41*(1), 10–32. https://doi.org/10.1111/j.1467-8535.2009.01038.x

De Back, T. T., Tinga, A. M., & Louwerse, M. M. (2021). CAVE-based immersive learning in undergraduate courses: Examining the effect of group size and time of application. *International Journal of Educational Technology in Higher Education*, *18*(1), 56. https://doi.org/10.1186/s41239-021-00288-5

De Back, T. T., Tinga, A. M., Nguyen, P., & Louwerse, M. M. (2020). Benefits of immersive collaborative learning in CAVE-based virtual reality. *International Journal of Educational Technology in Higher Education*, *17*(1), 51. https://doi.org/10.1186/s41239-020-00228-9

DeFanti, T. A., Acevedo, D., Ainsworth, R. A., Brown, M. D., Cutchin, S., Dawe, G., Doerr, K.-U., Johnson, A., Knox, C., Kooima, R., Kuester, F., Leigh, J., Long, L., Otto, P., Petrovic, V., Ponto, K., Prudhomme, A., Rao, R., Renambot, L., … Wickham, G. (2011). The future of the CAVE. *Central European Journal of Engineering*, *1*(1), 16–37. https://doi.org/10.2478/s13531-010-0002-5

Dichev, C., & Dicheva, D. (2017). Gamifying education: What is known, what is believed and what remains uncertain: A critical review. *International Journal of Educational Technology in Higher Education*, *14*(1), 9. https://doi.org/10.1186/s41239-017-0042-5

Dillenbourg, P. (1999). What do you mean by 'collaborative learning'? In P. Dillenbourg (Ed.), *Collaborative-learning: Cognitive and computational approaches* (pp. 1–19). Elsevier.

Fardinpour, A., & Reiners, T. (2014). The taxonomy of goal-oriented actions in virtual training environments. *Procedia Technology*, *13*, 38–46. https://doi.org/10.1016/j.protcy.2014.02.007

Fiorella, L., & Mayer, R. E. (2015). *Learning as a generative activity: Eight learning strategies that promote understanding*. Cambridge University Press.

Fiorella, L., & Mayer, R. E. (2016). Eight ways to promote generative learning. *Educational Psychology Review*, *28*(4), 717–741. https://doi.org/10.1007/s10648-015-9348-9

Fowler, C. (2015). Virtual reality and learning: Where is the pedagogy? *British Journal of Educational Technology*, *46*(2), 412–422. https://doi.org/10.1111/bjet.12135

Friedenberg, J., & Silverman, G. (2006). *Cognitive science: An introduction to the study of the mind*. Sage Publications.

Gillam, B. (2007). Stereopsis and motion parallax. *Perception*, *36*(7), 953–954. https://doi.org/10.1068/p3607ed

Granić, A., Nakić, J., & Marangunić, N. (2020). Scenario-based group usability testing as a mixed methods approach to the evaluation of three-dimensional virtual learning environments. *Journal of Educational Computing Research*, *58*(3). https://doi.org/10.1177/0735633119859918

Greitzer, F. L., Kuchar, O. A., & Huston, K. (2007). Cognitive science implications for enhancing training effectiveness in a serious gaming context. *Journal on Educational Resources in Computing*, *7*(3), 1–16. https://doi.org/10.1145/1281320.1281322

Hackett, M., & Proctor, M. (2016). Three-dimensional display technologies for anatomical education: A literature review. *Journal of Science Education and Technology*, *25*(4), 641–654. https://doi.org/10.1007/s10956-016-9619-3

Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does gamification work? – A literature review of empirical studies on gamification. In *Proceedings of the 47th Hawaii International Conference on System Sciences* (pp. 3025–3034).

Harris, D., Wilson, M., & Vine, S. (2020). Development and validation of a simulation workload measure: The simulation task load index (SIM-TLX). *Virtual Reality*, *24*(4), 557–566. https://doi.org/10.1007/s10055-019-00422-9

Hassan, M. A., Habiba, U., Majeed, F., & Shoaib, M. (2021). Adaptive gamification in e-learning based on students' learning styles. *Interactive Learning Environments*, *29*(4), 545–565. https://doi.org/10.1080/10494820.2019.1588745.

Heeter, C. (1992). Being there: The subjective experience of presence. *Presence: Teleoperators and Virtual Environments*, *1*(2), 262–271. https://doi.org/10.1162/pres.1992.1.2.262

Hickson, S., Dufour, N., Sud, A., Kwatra, V., & Essa, I. (2019). Eyemotion: Classifying facial expressions in VR using eye-tracking cameras. In *2019 IEEE Winter Conference on Applications of Computer Vision (WACV)* (pp. 1626–1635).

Höffler, T. N. (2010). Spatial ability: Its influence on learning with visualizations – A meta-analytic review. *Educational Psychology Review*, *22*(3), 245–269. https://doi.org/10.1007/s10648-010-9126-7

Javaid, M. A., Chakraborty, S., Cryan, J. F., Schellekens, H., & Toulouse, A. (2018). Understanding neurophobia: Reasons behind impaired understanding and learning of neuroanatomy in cross-disciplinary healthcare students. *Anatomical Sciences Education*, *11*(1), 81–93. https://doi.org/10.1002/ase.1711

Johnson, D. W., Johnson, R. T., & Smith, K. A. (2014). Cooperative learning: Improving university instruction by basing practice on validated theory. *Journal on Excellence in University Teaching*, *25*(4), 1–26.

Johnson, D. W., Johnson, R. T., & Stanne, M. B. (2000). *Cooperative learning methods: A meta-analysis*. University of Minnesota Press.

Knutas, A., Ikonen, J., Nikula, U., & Porras, J. (2014). Increasing collaborative communications in a programming course with gamification: A case study. In *Proceedings of the 15th International Conference on Computer Systems and Technologies* (pp. 370–377). Association for Computing Machinery.

Koivisto, J., & Hamari, J. (2019). The rise of motivational information systems: A review of gamification research. *International Journal of Information Management*, *45*, 191–210. https://doi.org/10.1016/j.ijinfomgt.2018.10.013

Kyndt, E., Raes, E., Lismont, B., Timmers, F., Cascallar, E., & Dochy, F. (2013). A meta-analysis of the effects of face-to-face cooperative learning. Do recent studies falsify or verify earlier findings? *Educational Research Review*, *10*, 133–149. https://doi.org/10.1016/j.edurev.2013.02.002

Lackner, J. R. (2014). Motion sickness: More than nausea and vomiting. *Experimental Brain Research*, *232*(8), 2493–2510. https://doi.org/10.1007/s00221-014-4008-8

Lee, E. A.-L., & Wong, K. W. (2014). Learning with desktop virtual reality: Low spatial ability learners are more positively affected. *Computers & Education*, *79*, 49–58. https://doi.org/10.1016/j.compedu.2014.07.010

Lei, H., Cui, Y., & Zhou, W. (2018). Relationships between student engagement and academic achievement: A meta-analysis. *Social Behavior and Personality: An International Journal*, *46*(3), 517–528. https://doi.org/10.2224/sbp.7054

Lin, H. C.-S., Yu, S.-J., Sun, J. C.-Y., & Jong, M. S. Y. (2019). Engaging university students in a library guide through wearable spherical video-based virtual reality: Effects on situational interest and cognitive load. *Interactive Learning Environments*, 1–16. https://doi.org/10.1080/10494820.2019.1624579

Lou, Y., Abrami, P. C., Spence, J. C., Poulsen, C., Chambers, B., & d'Apollonia, S. (1996). Within-class grouping: A meta-analysis. *Review of Educational Research*, *66*(4), 423–458. https://doi.org/10.3102/00346543066004423

Malinverni, L., & Burguès, N. P. (2015). The medium matters: The impact of full-body interaction on the socio-affective aspects of collaboration. In *Proceedings of the 14th International Conference on Interaction Design and Children* (pp. 89–98). ACM.

Marsh, W. E., Kelly, J. W., Dark, V. J., & Oliver, J. H. (2013). Cognitive demands of semi-natural virtual locomotion. *Presence: Teleoperators and Virtual Environments*, *22*(3), 216–234. https://doi.org/10.1162/PRES_a_00152

Mayer, R. E. (2009). *Multimedia learning*. Cambridge University Press.

Mayer, R. E. (2014). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 43–71). Cambridge University Press.

Mayer, R. E., Heiser, J., & Lonn, S. (2001). Cognitive constraints on multimedia learning: When presenting more material results in less understanding. *Journal of Educational Psychology*, *93*(1), 187–198. https://doi.org/10.1037/0022-0663.93.1.187

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist*, *38*(1), 43–52. https://doi.org/10.1207/S15326985EP3801_6

Mayes, J. T., & Fowler, C. J. (1999). Learning technology and usability: A framework for understanding courseware. *Interacting with Computers*, *11*(5), 485–497. https://doi.org/10.1016/S0953-5438(98)00065-4

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

Merchant, Z., Goetz, E. T., Keeney-Kennicutt, W., Kwok, O., Cifuentes, L., & Davis, T. J. (2012). The learner characteristics, features of desktop 3D virtual reality environments, and college chemistry instruction: A structural equation modeling analysis. *Computers & Education*, *59*(2), 551–568. https://doi.org/10.1016/j.compedu.2012.02.004

Meyer, O. A., Omdahl, M. K., & Makransky, G. (2019). Investigating the effect of pre-training when learning through immersive virtual reality and video: A media and methods experiment. *Computers & Education*, *140*, 103603. https://doi.org/10.1016/j.compedu.2019.103603

Morris, B., Croker, S., Zimmerman, C., Gill, D., & Romig, C. (2013). Gaming science: The "gamification" of scientific thinking. *Frontiers in Psychology*, *4*, 607. https://doi.org/10.3389/fpsyg.2013.00607

Muhanna, M. A. (2015). Virtual reality and the CAVE: Taxonomy, interaction challenges and research directions. *Journal of King Saud University – Computer and Information Sciences*, *27*(3), 344–361. https://doi.org/10.1016/j.jksuci.2014.03.023

Neji, M., & Ben Ammar, M. (2007). Agent-based collaborative affective e-learning framework. *Electronic Journal of E-Learning*, *5*(2), 123–134.

Parong, J., & Mayer, R. E. (2018). Learning science in immersive virtual reality. *Journal of Educational Psychology*, *110*(6), 785–797. https://doi.org/10.1037/edu0000241

Povis, K. T., & Crowley, K. (2015). Family learning in object-based museums: The role of joint attention. *Visitor Studies*, *18*(2), 168–182. https://doi.org/10.1080/10645578.2015.1079095

Roberts, T. S. (2005). Computer-supported collaborative learning in higher education: An introduction. In T. S. Roberts (Ed.), *Computer-supported collaborative learning in higher education* (pp. 1–18). IGI Global.

Romero, M., Usart, M., Ott, M., Earp, J., de Freitas, S., & Arnab, S. (2012). Learning through playing for or against each other? Promoting collaborative learning in digital game based learning. In *Proceedings of the European Conference on Information Systems (ECIS 2012)*.

Šašinka, Č, Stachoň, Z., Sedlák, M., Chmelík, J., Herman, L., Kubíček, P., Šašinková, A., Doležal, M., Tejkl, H., Urbánek, T., Svatoňová, H., Ugwitz, P., & Juřík, V. (2018). Collaborative immersive virtual environments for education in geography. *ISPRS International Journal of Geo-Information*, *8*(1), 3. https://doi.org/10.3390/ijgi8010003

Schneider, B., & Pea, R. (2013). Real-time mutual gaze perception enhances collaborative learning and collaboration quality. *International Journal of Computer-Supported Collaborative Learning*, *8*(4), 375–397. https://doi.org/10.1007/s11412-013-9181-4

Slavin, R. E. (1980). Cooperative learning. *Review of Educational Research*, *50*(2), 315–342. https://doi.org/10.3102/00346543050002315

Slavin, R. E. (1990). *Cooperative learning: Theory, research, and practice*. Prentice Hall.

Stein, N., Niehorster, D. C., Watson, T., Steinicke, F., Rifai, K., Wahl, S., & Lappe, M. (2021). A comparison of eye tracking latencies among several commercial head-mounted displays. *i-Perception*, *12*(1). https://doi.org/10.1177/2041669520983338

Steuer, J. (1992). Defining virtual reality: Dimensions determining telepresence. *Journal of Communication*, *42*(4), 73–93. https://doi.org/10.1111/j.1460-2466.1992.tb00812.x

Sweller, J., Ayres, P., & Kalyuga, S. (2011). The redundancy effect. In J. M. Spector & S. P. Lajoie (Eds.), *Cognitive load theory* (pp. 141–154). Springer.

Sweller, J., van Merrienboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, *10*(3), 251–296. https://doi.org/10.1023/A:1022193728205

Tinga, A. M., De Back, T. T., & Louwerse, M. M. (2019). Non-invasive neurophysiological measures of learning: A meta-analysis. *Neuroscience & Biobehavioral Reviews*, *99*, 59–89. https://doi.org/10.1016/j.neubiorev.2019.02.001

Tinga, A. M., De Back, T. T., & Louwerse, M. M. (2020). Non-invasive neurophysiology in learning and training: Mechanisms and a SWOT analysis. *Frontiers in Neuroscience*, *14*, 589. https://doi.org/10.3389/fnins.2020.00589

Valtchanov, D., & Ellard, C. G. (2015). Cognitive and affective responses to natural scenes: Effects of low level visual properties on preference, cognitive load and eye-movements. *Journal of Environmental Psychology*, *43*, 184–195. https://doi.org/10.1016/j.jenvp.2015.07.001

Vuopala, E., Hyvönen, P., & Järvelä, S. (2016). Interaction forms in successful collaborative learning in virtual learning environments. *Active Learning in Higher Education*, *17*(1), 25–38. https://doi.org/10.1177/1469787415616730

Wouters, P., van Nimwegen, C., van Oostendorp, H., & van der Spek, E. D. (2013). A meta-analysis of the cognitive and motivational effects of serious games. *Journal of Educational Psychology*, *105*(2), 249–265. https://doi.org/10.1037/a0031311

Yammine, K., & Violato, C. (2015). A meta-analysis of the educational effectiveness of three-dimensional visualization technologies in teaching anatomy. *Anatomical Sciences Education*, *8*(6), 525–538. https://doi.org/10.1002/ase.1510

Zanbaka, C. A., Lok, B. C., Babu, S. V., Ulinski, A. C., & Hodges, L. F. (2005). Comparison of path visualizations and cognitive measures relative to travel technique in a virtual environment. *IEEE Transactions on Visualization and Computer Graphics*, *11*(6), 694–705. https://doi.org/10.1109/TVCG.2005.92