



Hands-on or hands-off: Deciphering the impact of interactivity on embodied learning in VR



Sara Khorasani ^{a,*}, Brandon Victor Syiem ^b, Sadia Nawaz ^c, Jarrod Knibbe ^a, Eduardo Velloso ^a

^a University of Melbourne, Australia

^b Queensland University of Technology, Australia

^c Monash University, Australia

ARTICLE INFO

Keywords:

Virtual reality
Embodied learning
VR learning
VR education
VR training
Interaction techniques
Interaction fidelity
Movement in VR
Learning analytics

ABSTRACT

Studies suggest that Sense of Embodiment (SoE) enabled by VR promotes embodied and active learning. However, it is unclear what features of VR learning environments tap into the concept of embodied learning. For example, interaction techniques, movement and purely observational scenarios in VR can all play a role in facilitating embodied learning. To understand how these mechanisms impact learning, we conducted 2 studies with a total of 64 participants who had no prior experience in the training task. Participants were taught how to use a table saw in 4 conditions and were tested on their task performance in a fully interactive VR assessment. The conditions were analyzed in pairs; 2 conditions with different interaction techniques, 2 conditions with differing ability to move and a cross-study analysis comparing conditions with purely observational learning to interactive learning. We used a mixed methods approach; Analysis of Variance (ANOVA), pairwise comparison of the learning outcomes in each condition as well as thematic analysis of the interview results. We found that some types of “hands-on” interactions can have a detrimental impact on learning and that observational learning can be as impactful as a fully interactive experience. Based on participant interviews, we explored how these mechanisms of the learning environment can impact participants’ learning ability.

1. Introduction

VR is lauded for its ability to create immersive, hands-on learning environments, facilitating a sense of embodied learning Petersen, Petkakis, and Makransky (2022). Specifically, the Sense of Embodiment (SoE) - the phenomenon where one feels a virtual body as their own - is often credited as a significant enhancer of learning in VR (e.g., Makransky and Petersen (2021); Petersen et al. (2022); Johnson-Glenberg (2018)). However, the underlying mechanisms as to why, and the insights into exactly how these environments should be designed to maximise cognitive learning, remain contentious.

For example, some believe that, as a result of the embodied nature of VR, simply being immersed in a virtual environment and *observing* content facilitates active learning (Mayer, Makransky, and Parong (2022); Cheng, Yang, and Andersen (2017)). However, others believe *physical interaction with the environment* is necessary for active learning (Johnson-Glenberg (2018); Checa and Bustillo (2020)). Some researchers go further still, suggesting that *interactions need to closely mirror their real-world counterparts* in VR for embodied learning to occur

(e.g., if the real-world tasks require a lever to be pulled, then that pulling interaction needs to also occur in VR) (Johnson-Glenberg (2018); Skulmowski and Rey (2018)). Counter to these arguments, however, research has also suggested VR introduces additional cognitive burden, which can detract participants from learning (Mayer et al. (2022)).

Some may suggest the relationship between interactivity and learning in VR is simply on a continuum, with more interactivity leading to enhanced learning? We argue that more interactivity does not automatically translate to improved learning outcomes, especially in a procedural learning context. The complexity of VR interactions, their congruence with the learning tasks, and the cognitive demands they place on learners must also be considered. As such, this study explores these nuances, attempting to better understand the mechanisms that influence embodied learning in VR.

How, then, should we design interactions to maximise learning in VR? If physical interaction, for example, truly is more effective for learning, how much better is it than purely observational scenarios? Do these interaction styles facilitate drastically different behaviours that might further impact learning outcomes?

* Corresponding author.

E-mail address: s.khorasani@unimelb.edu.au (S. Khorasani).

In an attempt to provide clarity to this debate, we explore the effect of different interaction mechanisms on learning a procedural task in VR – how to safely operate a table saw in VR. The training was followed by a fully interactive test in VR which examined participants' understanding of the procedures within the VR environment. Our work does not explore the transfer of learning from VR to real life, thus no tests were done outside of VR. We are purely focused on improving learning outcomes within a VR environment.

We conducted 2 studies comparing 2 conditions per study, Study 1 compared Conditions 1 & 2, while Study 2 compared Conditions 3 & 4. (C1) *Interactions mirroring the real world*: A fully interactive environment where the student carries out the specified tasks on their own. (C2) *Observing, with indirect interactions with objects*: students use clicking interactions to trigger pre-recorded animations that demonstrate the task. (C3) *Observing, but freely moving*: The student can freely navigate the environment and watch animations of the tasks. (C4) *Observing from a fixed location*: The student watches animations from a seated position. Following a survey and a distraction task, participants in all conditions were tested on their skills in an identical VR environment with a slightly modified task.

The 2 studies enabled us to ask three fundamental questions about learning in VR. (1) What is the impact of movement (room-scale vs seated)? (2) What is the impact of the interaction technique used for manipulating the learning content? (3) What is the impact of actively observing versus actively manipulating?

We did not find a significant difference in learning outcomes as a result of movement, in other words, between a room-scale (C3) and seated (C4) VR learning setting. However, we found that interaction techniques used for manipulating learning content can have a significant detrimental impact on learning. We found that non-mirroring interactions (C2) led to worse learning outcomes compared to C1. Finally, by comparing conditions across studies, we found no significant difference between actively observing (C3) versus actively manipulating the content (C1). The interview data and NASA-TLX results reveal that (a) the embodied nature of VR may result in "active" learning even when the user is not interacting physically but simply observing, (b) the "wrong" interaction technique may cause extraneous processing, and finally, (c) low interaction fidelity can increase the mental workload on learners.

To the best of our knowledge, our work is the first to examine the impacts of the aforementioned interaction features on a procedural learning task. This has been called for by the community (see, e.g., [Makransky and Petersen \(2021\)](#)) as many existing interaction design studies use declarative tasks, which are less suited to examining embodied learning. We discuss this further in the related work section. Additionally, our work provides foundational knowledge to inform the design of VR learning systems going forward. For example, as VR gains traction in industry, companies may spend hundreds of thousands of dollars on fully interactive experiences when a more cost effective approach (observing a recorded animation) may produce similar results.

2. Related work

2.1. The theory - Embodiment, agency and embodied learning

For decades, studies on Virtual Reality in the education domain have focused on its positive impacts on learning ([Oberdörfer, Heidrich, and Latoschik \(2019\)](#); [Rogers, El-Mounaryi, Wasfy, and Satterwhite \(2018\)](#); [Murphy and Higgins \(2019\)](#); [Oagaz, Schoun, and Choi \(2021\)](#); [Rettinger et al. \(2021\)](#); [Petersen, Mottelson, and Makransky \(2021\)](#)). Benefits of VR learning environments include the ability to turn abstract and intangible concepts into concrete interactive spaces ([Strangman and Hall \(2003\)](#)) and the ability to develop models that allow students to visit historical places, experience different cultures and even immerse themselves in fictional spaces ([Hedberg and Dalgarno \(2002\)](#)). However, the aspects or features of VR learning environments that lead to

better learning are not yet clear ([Keifert et al. \(2017\)](#)).

Sense of Embodiment (SoE), sense of agency, and embodied learning are factors that have been shown to impact learning in the classroom and in virtual settings ([Keifert et al. \(2017\)](#); [Abrahamson et al. \(2020\)](#); [Bergström, Knibbe, Pohl, and Hornbæk \(2022\)](#)). *Sense of Embodiment* is defined by "three subcomponents—the sense of self-location, agency, and body ownership" ([Kilteni, Grotens, and Slater \(2012\)](#)). Lindgren et al. among others believe that "embodiment is a powerful underpinning of cognition" ([Lindgren and Johnson-Glenberg \(2013\)](#)). However in practice, it is hard to say, how we can leverage these concepts to enhance learning outcomes in VR. For example, a VR environment that has no interactive elements, or the ability to move (room-scale), is often still referred to as an embodied learning environment. But how does it compare to a fully interactive room-scale experience in terms of learning outcomes?

Differing from the above concepts, *Embodying Learning* is defined as the involvement of the whole body in the learning process (mind, body, physical action, cognition, emotions) ([Robbie \(2020, pp. 225–249\)](#); [Keifert et al. \(2017\)](#)). While related (both derived from embodied cognition theory), embodied learning and embodiment are different concepts. In practice, embodied learning is when a user uses their body in the learning process, for example using their arms or fingers to trace changes in population of a certain animal when numbers of their predators rise or fall ([Andrade \(2017\)](#)). These factors are sometimes but not always present in VR, and differ in the degree to which they are present (eg. fully interactive versus seated with no interaction). But they are often referred to (along with level of immersion) as key factors in creating an effective exploratory, experiential and embodied learning environment in VR ([Keifert et al. \(2017\)](#)). So does that mean a hands-on approach always increases embodied learning, and does that always mean better learning outcomes?

Other factors that support learning in VR include the concept of *active learning*. Active learning is defined as engaging with the learning content to construct knowledge or understanding whether through discussion or through physical interaction ([Waldrop \(2015\)](#)). Some believe that simply being immersed in a virtual environment and observing learning content facilitates active learning given the embodied nature of VR ([Mayer et al. \(2022\)](#); [Liao, Sung, Wang, Lin, and Cherng \(2019\)](#); [Ponder et al. \(2003\)](#)). Others consider physical interaction with the environment as necessary for active learning ([Johnson-Glenberg \(2018\)](#)). So as the interaction fidelity increases in VR, do we make the learning more active and thus more impactful? Or does the embodied nature of VR create an active learning environment without the need for complex interactions in the environment?

Johnson et al. argue that for embodied learning to take place, it needs the interactions or movements to be congruent with the learning content ([Johnson-Glenberg \(2018\)](#); [Skulmowski and Rey \(2018\)](#)). For example, if students are learning about orbits they should move their hands or bodies in circular motions. However, this can be complicated to achieve and is highly dependent on the way the students perceive the movements thus can result in actually confusing students ([Lindgren and Johnson-Glenberg \(2013\)](#)).

Additionally, even *without* interactive elements, researchers find that VR introduces extraneous processing (i.e., cognitive processing that does not support the instructional objective) ([Mayer et al. \(2022\)](#)) and work towards minimizing extraneous processing in VR learning settings ([Carpentier and Lourdeaux \(2014, pp. 245–260\)](#); [Mackenzie and Harris \(2014\)](#); [Matthias and Beckhaus \(2012\)](#)). So will more active learning elements and interaction solve the problem of extraneous processing by facilitating embodied learning, or will it exacerbate the problem?

While VR is typically considered an environment where embodied learning happens ([Lindgren and Johnson-Glenberg \(2013\)](#)), that may or may not be true depending on the design of the experience. Some experiences are stationary, where the user sits, simply observes their surroundings, and has little to no control over what happens next ([Mayer et al. \(2022\)](#)). Alternatively some environments allow interacting with

objects using raycasting (laser pointing) (Seo et al. (2017)). This is a common interaction technique, but its effectiveness for supporting embodied learning has not been tested. Comparatively, an environment can allow the user to walk around freely, interact with objects (pick them up, manipulate them), and determine the outcome of the experience based on the choices they make and their behaviour. However, beyond testing their usability, studies have overlooked interaction techniques and their impact on learning outcomes (Moher, Johnson, Ohlsson, and Gillingham (1999); Seo et al. (2017); Yoo, Kim, and Lee (2020)). Usability testing, in the form of evaluating user engagement and learner confidence may not be sufficient measures of system design effectiveness due to the illusion of understanding (Seo et al. (2017)). The illusion of the understanding states that, in some cases, the student may feel confident during the learning stage, but in practice, they may find that they have not mastered the content to the level they had *felt* they had. So learner confidence does not always positively correlate with skill mastery.

Additionally, the interaction design literature tends to focus on using VR for conceptual and declarative training (such as historical facts or mathematical operations), where suitable interaction techniques, level of interactivity and movement may have different impacts on learning outcomes (Mayer et al. (2022)). The principles applied to declarative training cannot be generalized to a procedural environment (Slavova and Mu (2018); Baceviciute, Mottelson, Terkildsen, and Makransky (2020)). A procedural task includes the learning of physical movements and processes. It introduces a setting where the application of embodied learning principles may be particularly important. For example, the learning literature tells us that for embodied learning to take place the learning task must involve movements that mirror the correct actions required to perform the task (Johnson-Glenberg (2018)).

2.2. The mechanisms - Interaction technique, movement & observational learning

Though the embodied learning theory is well established, the challenge of how to apply it in VR training environments still remains. Due to limited evidence-based best practices for procedural VR training, the design and development of VR applications can at times be erratic (Biermann, Ajisafe, and Yoon (2022)). The lack of clarity on the *mechanisms* that may facilitate the embodied learning process makes the design of embodied learning systems even harder. For example, while the importance of interaction for enhanced embodied cognition and design is understood, the practical application of the theories to procedural training in VR is lacking (Jin, Liu, Yarosh, Han, and Qian (2022); He et al. (2020); Mayer et al. (2022)).

2.2.1. Interaction techniques

It is surprising how little is known about the impact of different interaction techniques on learning outcomes. Particularly in the case of procedural training where object manipulation techniques are considered fundamental to learning (Reeves (2012)). However, the human-computer interaction research community has predominantly focused on how interaction techniques in VR impact motivation and engagement instead of learning outcomes. This is due to the belief that one of the main advantages of VR is its ability to enhance engagement and an understanding that an increase in engagement indirectly impacts learning (Oberdörfer et al. (2019); Wirth, Grndl, Sembdner, Kuhrt, and Eskofier (2018)). But while interaction techniques may increase engagement they may have a significant impact on extraneous processing which is one of the main challenges in VR learning environments (Mayer et al. (2022)).

2.2.2. Movement

Studies have found that node-based movement (tele-porting from one location to another) in VR is not helpful in remembering factual information but powerful in enhancing spatial memory (where objects

are and how to get to them) (Ferguson, van den Broek, van Oostendorp, de Redelijkhed, and Giezeman (2020)). In the embodied learning literature different types of movement have been associated with enhanced embodied cognition (Abrahamson et al. (2020)). But in the context of VR not much else is known around how moving and walking in VR impacts learning outcomes.

2.2.3. Observational learning

Studies have conflicting results about how the level of interactivity impacts learning. For example, comparing a participant that only observes a demonstration in VR versus one that interacts with the scene. Some believe interactivity enhances user experience and immersion (Seo et al. (2017)), others show that purely observational environments have comparable results to an interactive one (Rowe et al. (2017); Watson and Livingstone (2018); Snyder et al. (2011); Koßmann, Straatmann, Mueller, and Hamborg (2023)). Once again, the research evidence is inconclusive with studies showing that simplifying the interaction had mixed impact on user performance, preference, and learning outcomes (Bertrand (2016)).

Adding a more nuanced perspective, recent research demonstrates that VR learning experiences' distinguishing features – high levels of immersion and interactivity – contribute more than mere entertainment value. Specifically, both immersion and interactivity augment the physical presence experienced by learners. Interactivity becomes more crucial for experiencing agency and embodied learning under conditions of low immersion. High interactivity is also seen to mitigate extraneous cognitive load from the environment (Petersen et al. (2022)). However, self-reported embodied learning negatively predicted participants' declarative memory of the learning topic, possibly due to incongruity between bodily actions and learning content (Petersen et al. (2022)).

This debate leaves us with many unanswered questions. If interactivity is better for reducing cognitive load and enhancing embodied learning, are all types of interactivity created equal? Do different interaction techniques influence immersion, memory, and embodied learning differently? How does the ability to move impact observational learning?

We believe that our study provides design considerations that apply embodied design principles to VR learning environments for optimizing the aforementioned mechanisms. We also suggest future work to further understand the intricacies introduced by these mechanisms in a learning context.

3. System design and implementation

In order to test how interaction techniques, interaction fidelity, and movement play out in procedural training, we designed and built a VR application to teach participants how to safely operate a *SawStop - Professional Cabinet Model* table saw. The exact saw is shown in Fig. 1 (Saw). The saw and all of its associated buttons and levers were fully interactive in our simulation.

The task was chosen specifically due to its ecological validity; this task is a safety training procedure mandatory for people intending to use maker spaces at many institutions. Our research primarily focused on the cognitive aspect of learning - specifically, the acquisition and application of procedural knowledge. This specific task was chosen because it's a practical skill that requires both knowledge and precise physical interactions, thus making it a fitting task to evaluate the impact of VR and the potential benefits of hands-on versus hands-off approaches to learning.

We identified 18 necessary tasks required to operate the machine based on our university's online training module for this particular saw and implemented them as part of the VR training protocol in our study. The steps outlined align with industry-standard health and safety protocols, which are widely accepted for operating similar types of machinery across various contexts, not merely within our institution. We broke down major tasks into multiple smaller tasks to increase the



Fig. 1. 3D Model of a Table Saw showing the red lever used to adjust the fence, the wheels used to adjust the blade height & angle and the on and off switch. These were all parts of the machine that had to be used to safely setup and cut the piece of wood (SawStop). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

granularity of the procedure presented in [Table 1](#). The major tasks in detail:

1. Removing all tripping hazards - Participants must ensure there are no objects on the ground that could result in tripping, and must remove these objects by placing them in a safe location.
2. Putting safety glasses on - Participants must wear their safety glasses before turning the saw on.
3. Installing a riving knife to prevent kickback - Participants must ensure that they install the riving knife behind the blade of the saw, this is to prevent material from kicking back and causing injuries during the cutting procedure.
4. Unlocking and locking the fence - Participants must ensure that they unlock the fence before attempting to adjust it, and locking it following adjustment to prevent the wood being cut from moving.
5. Moving the fence in position against the wood being cut - Participants must ensure that the fence is set to an appropriate position for cutting the wood at the required length, and that the wood is secure.
6. Adjusting the height of the blade to the safe cutting height - Participants must ensure that the height of the blade is roughly 1 cm above the height of the wood being cut to ensure a safe and even cut.
7. Adjusting the angle of the blade - Participants must ensure they set the angle of the blade to the correct cutting angle required.
8. Using a Mitre Gauge and push stick to push the material safely into the blade - Participants must ensure a mitre gauge is used in their left hand, and a push stick in their right hand to push the wood through the blade evenly on both sides.
9. Not picking up cut parts before the machine is off - Participants must ensure they do not attempt to remove any cut materials until the saw is turned off.
10. Pushing the material fully past the blade to cut evenly - Participants must ensure they push the wood all the way through to the other side of the blade to ensure an even cut.
11. Not putting the push-stick into the blade - Participants must ensure they do not accidentally or intentionally cut the push stick or any materials in their hand.
12. Turning the machine on and off at appropriate times - Participants must ensure they turn the machine on once they have completed all the steps for setting up, and turned off after completing a cut.

3.1. The VR application

All conditions in our study were deployed in the same virtual environment. The minimalist design of our VR environment, including the simple background, was indeed intentional. Drawing from cognitive load theory per the work of ([Sweller, van Merriënboer, & Paas, 1998](#)), we aimed to prevent any extraneous cognitive processing that could potentially distract from the primary learning task at hand.

The text prompts that the participants received with instructions on how to operate the machine were identical in content and order of appearance among all conditions. Text prompts are shown in [Fig. 2](#). For example, one text prompt read “always put your safety glasses on before cutting”. Finally, the object that the text prompts referred to was outlined in red in all conditions because we did not assume that participants knew the names of the objects and tools as shown in [Fig. 2](#) (c). For example, when the text prompt said “install the riving knife behind the blade”, the riving knife would become outlined.

We conducted two studies comparing 2 conditions in each study. Study one compares C1 & C2 while Study 2 compares C3 & C4. Conditions C2, C3 & C4 were presented with pre-recorded animations of the appropriate steps that must be taken to operate the machine. These animations were similar to receiving a demonstration by another user within VR.

The 4 training conditions were as follows.

1. **C1 - Interactions mirroring the real world:** Participants received text prompts telling them how to operate the machine step-by-step which they had to learn. The interactions were designed to be realistic, so all objects could be grabbed, picked up, and moved. Levers and wheels could be adjusted and rotated with the use of the controllers. The participants used the A key on the controller to proceed to the next instruction board. This was identical in all conditions. For example, when the text prompt reading “Install the riving knife behind the blade to prevent kickback” appears, the riving knife object is simultaneously outlined in red. Participants must then use the controllers to pick up the object, and place it behind the blade of the saw, before moving to the next text prompt.
2. **C2 - Observing, with indirect interactions with objects:** Participants received the same text prompts, but were only able to click on the object in question to activate a pre-recorded animation that demonstrated the correct procedure. They were only able to click on objects when the object in question was outlined in red per the instructions. When they pointed at the outlined object with their laser pointer the object would turn red to signify that the pointer was hovering on it, and then it would change colour again when clicked on. Once clicked on, each object moved per a pre-recorded demonstration animation showing the correct steps to be followed. For example, when the text prompt reading “Install the riving knife behind the blade to prevent kickback” appears, the riving knife object is simultaneously outlined in red. Participants must then use their laser pointer to click on the riving knife, this will result in the riving knife animation to play, showing the object moving to the correct position behind the blade.
3. **C3 - Observing, but freely moving:** Participants received the same text prompts as the previous conditions but pre-recorded animations of the appropriate steps automatically started playing with each new text prompt. No clicking was necessary to activate the animations. Participants were able to walk around and watch the animations from any angle or as closely as they liked. Participants could also go back to the previous animation or click to go next. To go back and forth between the text prompts the buttons A and B on the controller were used. This was the same for all conditions, and not to be confused with the pointer clicking that was exclusive to condition C2, used to activate animations. For example, when the text prompt reading “Install the riving knife behind the blade to prevent kickback” appears, the riving knife object is simultaneously outlined in

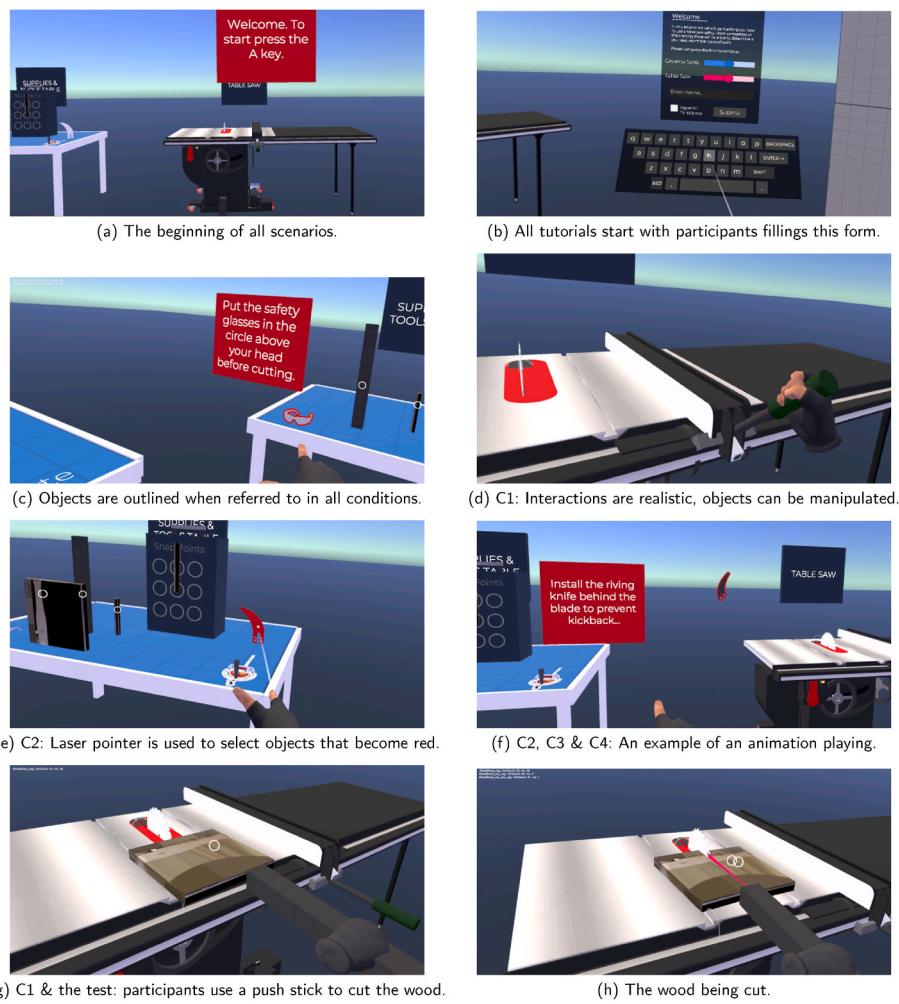


Fig. 2. Images from the VR environment for all conditions.

red. Participants do not need to do anything, the animation showing the riving knife moving to the correct position behind the blade is automatically played. Participants may walk around to see the animation from different angles.

4. C4 - Observing from a fixed location: This condition is identical to the walking animation condition (C3), except participants were seated. They could move their head or stand to get a better look, but were not able to walk. For example, when the text prompt reading “Install the riving knife behind the blade to prevent kickback” appears, the riving knife object is simultaneously outlined in red. Participants do not need to do anything, the animation showing the riving knife moving to the correct position behind the blade is automatically played. Participants can observe from a seated or standing position without moving.

3.2. The physical space & apparatus

The physical space was a large 5×10 m room and we mapped the VR space to the exact dimensions of the room. This mapping allowed participants in the movement conditions to freely move around the space without the need to teleport in VR to get to the desired location. To prevent distraction and encourage free movement, we did not implement a teleport feature. Participants used a Meta Quest 2 with a 5 m long link cable. The research team insure participants can freely and safely walk around the space why tethered to the PC.

In the experimental setup, the term “operating” the machine is pertaining to the interaction within the VR environment. Participants were

equipped with Meta Quest 2 controllers, serving as their primary interaction devices as shown in Fig. 3. The trigger button was used for “clicking” with the laser pointer, and the grip button was used for grabbing and manipulating objects within the virtual space, where such interactions were allowed. All levers, wheels, and other components of the machine within the simulation were all rendered representations.

The images provided in Fig. 2(d) show the 3D rendered hand interacting with a lever, which mirrors the participant’s real-world action of squeezing the grip button on their controller. This action is what allows the participants to “grab” the lever within the simulation. All participants were thoroughly practiced with these controls in tutorial scenarios before they moved to the actual training and testing scenarios.

We made a conscious decision to use VR controllers instead of gesture-based controls for several reasons. First, the maturity of the gesture recognition technology still presents challenges that might influence the learning outcomes in ways that are external to the factors we intended to study. Second, hand tracking does not allow for haptic feedback, which is crucial in tasks involving manipulation of virtual objects and tools. Therefore, our decision to use VR controllers was based on their capacity to provide both reliable interaction and haptic feedback.

4. Method

4.1. Participants

We recruited 64 participants (40 women, 23 men, 1 other) between



Fig. 3. A photo of a participant wearing the Oculus Quest 2 headset and controllers in the physical space interacting with the virtual environment. The virtual view and hand interactions in VR are simultaneously shown in the bottom right snapshot.

the ages of 18–55 through online university channels and notice boards. We determined the sample size by following the guidance of Caine et al., balancing the sample size against estimates of effect sizes' (Caine (2016)). Participants were given a brief description of the study and were informed that they would receive a \$15 gift voucher for participating in the experiment and an additional \$5 gift voucher for satisfactory performance. The intention was to ensure participants were engaged in the training and attempted to perform well in the test, but all participants received the full \$20.

We undertook several steps to control for potential pre-intervention influences. (1) Participants were carefully selected to be completely novice to the field of machining. Only those who had no prior experience with machining tools were included in the study, which was verified not only through self-reporting in surveys but also via an additional interview process on the day of the experiment. (2) Prior to the intervention, all participants received an identical briefing to ensure they all started from the same baseline knowledge. (3) The assignment of participants to different intervention groups was done randomly to ensure that any unknown or uncontrolled pre-intervention factors were equally distributed among groups. We believe these measures have controlled for potential pre-intervention factors influencing the results of our study.

4.2. Procedure

The study took between 40 min and 1 h, with an average of 50 min per participant. It consisted of training in one of the four conditions, followed by a test. The step by step details of the study procedure are as follows and represented in Fig. 4.

Upon arrival, participants were first required to read and sign the informed consent form. They were then randomly assigned to one of the 4 conditions (between-participants).

The study started with a brief overview of the procedure. We familiarized participants with the 4 buttons they would use on the controllers. These included buttons A and B to cycle through the instructions, the trigger button to click, and the grip button to grip/hold objects. To ensure that participants had mastered the controls, they went through a 2–3 min tutorial to learn them. The tutorial started with participants filling a virtual form that asked them about their name and asked for their consent again. We only used this task for participants to gain practice with the controls and we did not collect any data from this stage. The first author was in the room with participants at all times to ensure they were safe and the wire attached to the headset was not a tripping hazard.

Upon completion of the practice with controls, participants completed a tutorial that corresponded to their assigned condition. Participants were only given instructions for the controls they would need in the training phase. For example, only participants in condition C2 learnt how to use the laser pointer to click on objects as this was not necessary for any other condition. We ensured that participants were fully comfortable with the controls before proceeding. The tutorial was also an opportunity for participants to become comfortable with walking around the space. This was new to many and required some training to get them comfortable especially with the wire attached to the headset.

Following the tutorials, participants began the training scenario assigned to them. All participants were allowed to ask for help if they didn't understand a step, and a scripted instruction was read out to them. Most participants simply required the text prompt read out loud, while others required rephrasing the text prompts. We recorded the number of times participants asked a question. If participants skipped a step and did not complete a task, or skipped an animation, they were told to go back and complete the step. If a step was done incorrectly (in the case of C1 condition), they were asked to repeat the step and were offered the pre-made scripts if asked. This modification was introduced as we observed a problem where some participants misunderstood which direction to push the wood and repeated the incorrectly learnt procedure in the test during our pilot. Given that the conditions with animations (C2, C3 & C4) could see the correct direction through the animations, we wanted to ensure all conditions had a correct understanding of the procedure and did not go into the test repeating incorrectly learnt steps. The training took between 10 and 15 min, depending on whether participants chose to go back to previous steps or dwell on certain steps.

After completing the training, participants were asked to fill in a survey that included NASA-TLX to measure mental workload and a Likert scale survey to capture their perceptions about the experience and confidence in the material taught (see appendix) (Peck and Gonzalez-Franco (2021); Eubanks, Moore, Fishwick, and McMahan (2021); Kim and Shute (2015); Hart and Staveland (1988, pp.

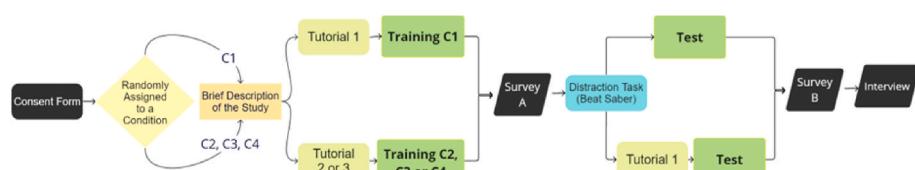


Fig. 4. The study procedure for all conditions showing the 10 (9 in the case of C1) steps participants go through during the study.

139–183). This took approximately 5 min.

Participants were then given a brief introduction to how to play the game Beat Saber and asked to play for 2.5 min. The intention here was to clear their working memory before the test.

Participants were then asked if they needed a break, before beginning a tutorial on how to interact with objects relevant to the test (pick them up, hold them etc.). This tutorial was only given if they were in conditions C2, C3, or C4 where they had not yet manipulated the objects. If they were the C1 condition this step was skipped, and participants were asked if they were ready for the test. This tutorial took 2–3 min.

Participants were given the VR test which was the baseline all participants were compared against. This study did not explore the concept of transfer of learning from VR to real life, it was focused purely on optimizing learning outcomes within a VR environment. During the test if participants faced any difficulty with the controls, they were offered support, but no help or clues were given as to how to complete the task. The tests consisted of all the same steps as the training plus an additional step that was not included in the training. This additional step was to adjust the angle of the blade using the angle wheel to cut the wood at an angle. Participants were asked (using text prompts) to repeat all of the exact same steps in the same order with the addition of this extra step. This task took 5–10 min.

Participants were asked to complete a second survey which included a NASA-TLX and a questionnaire about their perceptions of the test and their confidence in their performance. The survey also included 5 multiple choice questions on the correct table saw operating procedures as well as a brief 4 question spatial reasoning test. This step took 5–10 min.

Participants were then interviewed in a semi structured manner about the experience. They were asked about the training and how it helped or did not help them. They were asked about how they typically learn procedural tasks. Some interviews were brief (1 min) while others were longer (10 min+) as participants expressed everything that was helpful or distracting to them. We focused on understanding their challenges and factors that helped them.

All participants were emailed 2 vouchers to thank them for their support regardless of their performance in the test.

4.3. Measures

We recorded participants' performance during the test. We operationalised performance by giving 1 point per correct implementation of the 18 identified tasks. We recorded their number of interactions and the order in which they conducted those tasks. For example, locking the fence after the wood was cut was recorded as an interaction but considered incorrect. In this case, while the participant performed the required tasks, they mixed the order of activities which could be reflective of their lack of understanding of the correct procedure.

4.4. Analysis

We used a mixed methods approach. We used one-way analyses of variance (ANOVA) to test for the effects of the conditions on learning outcomes. We used Bartlett's test to identify any violation of homogeneity of variances. We carried out a Tukey's HSD for pair-wise comparison among all groups as our data met all assumptions. To evaluate effect size we used eta squared which is required to be greater than 0.14 for a large effect size (Yatani (2016, pp. 87–110)). NASA TLX was also used to understand the cognitive load after the training and after the test.

We used a loose interpretation of the reflexive thematic analysis approach to analyse the interviews. We transcribed and coded the relevant quotes from the interviews that we believed helped shed light on the research questions and the themes we identified in the process.

5. Results

We had a 2 study design, where we compared the learning performance of participants in Conditions C1 to C2 and C3 to C4, and an additional cross-study analysis of C1 to C3. We used one-way ANOVA to compare the test scores of all 4 conditions. Table 2, 3 and 4 and Fig. 5 show the summary of the results side by side. It is important to reiterate the baseline used to examine participants' learning of the procedures taught was a fully interactive VR test. This study did not attempt to answer any questions relating to the transfer of learning from VR to real life.

Bartlett's test did not show a violation of homogeneity of variances ($\chi^2(2) = 6.79$, $p = 0.08$). With the one-way ANOVA, we found a significant effect ($F(3,60) = 7.45$, $p < 0.001$, partial $\eta^2 = 0.27$) with CI = [0.07, 0.41]. An eta squared (η^2) of greater than 0.14 shows a large effect size (Yatani (2016, pp. 87–110)).

A Tukey's pairwise comparison revealed significant differences between C1 and C2 ($p < 0.001$), and between C3/C4 and C2 ($p < 0.01$). Table 4 illustrates this comparison. We see that participants in the C2 condition performed much worse than the C1 condition, while C3, C4, and C1 perform similarly. This is also represented in the violin plot shown in Fig. 5.

Table 5 shows the NASA-TLX means and standard deviations (within parenthesis) of each subscale for all conditions after the training and after the test. The C1 condition has an overall higher NASA TLX score for the training and lower for the test, whereas in all other conditions (i.e., C2, C3 and C4) the trend is in the opposite direction.

6. Discussion

For each research question we discuss what the quantitative results tell us and draw quotes from the interview to attempt to explain these findings.

6.1. Study 1: C3 & C4 – What is the impact of movement (room-scale vs seated) on learning in VR?

6.1.1. What does the analysis of variance tell us?

We can see that C3 and C4 conditions perform similarly (means of participant scores 14.8 vs 14.8 and medians 15 vs 16, respectively). However, while their averages are similar, this data was not equally spread (C3 versus C4 standard deviation: 1.91 versus 2.65). This could highlight that the seated condition (C4) works well for many, but it does not lead to the same consistent performance of the walking condition (C3) for a significant group of the population.

6.1.2. How do the interviews explain the results?

Participants that were able to walk in the VR environment mentioned that this ability allowed them to remember the order or sequence of tasks better. Participants mentioned that their movement during the test

Table 1

The 18 tasks participants were scored on, and the appropriate time to do each step.

Before turning the saw ON	While saw is ON	Checks
Hazard Removal	Saw On	Do Not Cut Hands
Glasses On	Push Both	Do Not Cut Stick
Riving Knife	Push Through	Do Not Grab Before Off
Wood Placement	Saw Off	
Fence Unlock		
Fence Position		
Fence Lock		
Blade Height		
Blade Angle		
Push Stick		
Mitre Gauge		

Table 2

Test score mean, median and standard deviation for each condition. As well as proportions of each population that performed the angle adjustment of the wheel correctly.

	C1	C2	C3	C4
Mean test score (out of 18)	15.9	11.8	14.8	14.8
Median test score (out of 18)	16.5	12.0	15.0	16.0
Standard deviation of test scores	1.95	3.40	1.91	2.65
Angle wheel correct proportion	56%	31%	38%	56%
Count	16	16	16	16

Table 3

ANOVA Results showing that there was a difference between the conditions compared with a p-value smaller than 0.005.

Source	SS	df	MS	F	Prob > F
Groups	155.30	3	51.77	7.45	0.0003
Error	416.81	60	6.95		
Total	572.11	63			

Table 4

A Tukey's pairwise comparison revealed the significant differences between C1 and C2 ($p < 0.001$), and between C3/C4 and C2 ($p < 0.01$).

Condition A	Condition B	Lower Limit	A-B	Upper Limit	P-value
C1	C2	1.74	4.19	6.65	0.0002
C1	C3	-1.34	1.13	3.59	0.6247
C1	C4	-1.34	1.13	3.59	0.6247
C2	C3	-5.53	-3.06	-0.60	0.0090
C2	C4	-5.53	-3.06	-0.60	0.0090
C3	C4	-2.46	0.00	2.46	1.0000

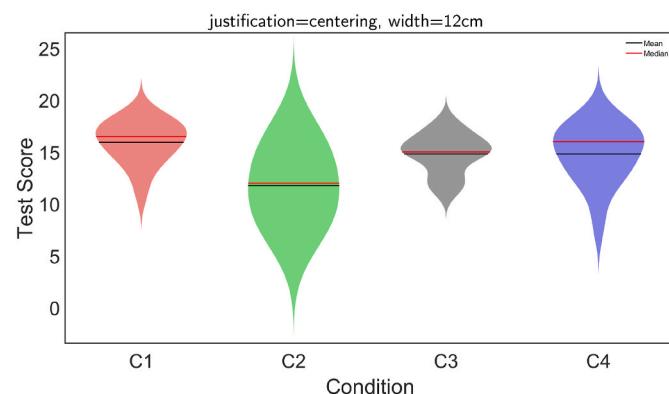


Fig. 5. Violin plot of the distribution of the test scores of each condition. It can be seen that while C1 performed the best, their median score is similar to the that of C3 & C4. C2, on the other hand seemed to perform poorer than all other conditions. Data spread seems higher for C2 and C4 and relatively lower for C1 and C3. (Holger Hoffmann (2022)).

could be mapped to associated tasks. They mentioned that walking in the test then triggered their memory to perform the associated task (eg. at the start walking to the *left* reminded them to pick up the safety glasses located on the far *left*). This may be a form of embodied learning, where the action of walking to one side or towards an object is congruent to what needs to be done in the test thus helping the participant remember those actions. Alternatively, it could be a form of spatial memory. Virtual reality has been shown to significantly improve spatial memory specifically when the participant is moving in VR (Ferguson et al. (2020)). C3 Participant: “I remembered the sequence of tasks using spatial relationships, for example when I move to Table 1 need to remove the hazard, or left table tasks first. It’s like arrows or a flow chart in my mind.” C3 Participant: “Movement stimulated my memory to remember things and the order.” C2 Participant: “Walking helped me compartmentalize tasks, when I move here this is what I need to do.”

However, given that those that were seated performed similarly, perhaps being able to look around also facilitated spatial memory. C4 Participant: “More engaging than a YouTube video. Being able to look around helped me remember where to go.”

Some mentioned that being able to be seated allowed them to just focus on the task and nothing else: C4 Participant: “Sitting down was better so I can focus on what is going on.” In contrast, some C3 participants noted that walking around at times resulted in them missing the action behind them. C3 Participant: “Sometimes I was looking one way and then the animation had started behind me.”

To conclude, our participants’ accounts suggested that the ability to physically move within the VR environment aided their memory recall and task sequencing. This finding aligns with theories of embodied cognition, which argue that our bodily experiences profoundly shape our cognitive processes. In particular, our results suggest that embodied learning may facilitate cognitive mapping in VR, allowing learners to associate specific physical movements or locations with particular learning tasks. This points to the potential for utilizing movement-based cues as a cognitive scaffold in VR learning environments, enhancing learners’ ability to remember and sequence tasks correctly.

However, the findings also highlight that even without physical movement, VR learning can still provide a rich and engaging learning environment. Some participants felt that being seated allowed them to concentrate more on the tasks, suggesting that the immersive nature of VR, even when seated, can foster focused attention. This suggests that VR’s immersive characteristics can facilitate sustained attention and cognitive engagement, crucial for successful learning outcomes.

6.2. Study 2: C1 & C2 – What is the impact of different interaction techniques on learning in VR?

6.2.1. What does the analysis of variance tell us?

The analysis showed that C2 led to significantly worse performance worse (mean of test scores for C2 = 11.8/18, versus 15.9/18 for C1) compared to the C1 condition (and the non-interactive conditions). This may be a surprising finding for those that believe more interaction in VR always leads to better engagement and learning outcomes. However the embodied learning literature might help shed light on this. This body of work highlights the importance of congruent movements with the task

Table 5

NASA TLX Mean and Standard Deviation (within parenthesis) for all conditions after the training and after the test.

Condition	Mental Demand	Physical Demand	Temporal Demand	Performance	Effort	Frustration	Overall
C1 Training Mean	39.7 (22.4)	27.2 (22.8)	31.3 (24.1)	45.9 (23.8)	55.0 (22.7)	44.1 (28.9)	40.5 (9.3)
C1 Test Mean	41.6 (17.6)	21.3 (18.8)	24.7 (20.5)	50.3 (21.2)	46.6 (24.3)	35.0 (24.2)	36.6 (10.7)
C2 Training Mean	34.4 (15.2)	22.8 (19.7)	19.4 (15.9)	60.3 (19.8)	30.0 (16.9)	20.3 (20.6)	31.2 (14.1)
C2 Test Mean	46.9 (20.3)	32.5 (24.0)	22.8 (20.1)	39.1 (22.7)	37.5 (19.3)	30.3 (28.0)	34.8 (7.5)
C3 Training Mean	41.9 (22.1)	21.9 (28.0)	20.6 (22.3)	61.3 (22.3)	41.3 (30.6)	26.6 (30.8)	35.6 (14.3)
C3 Test Mean	54.7 (28.6)	34.7 (28.7)	26.6 (28.2)	48.4 (29.8)	60.3 (24.7)	44.7 (32.7)	44.9 (11.4)
C4 Training Mean	51.6 (18.4)	16.6 (22.2)	14.7 (14.2)	63.4 (26.7)	31.9 (22.8)	17.2 (22.6)	32.6 (18.8)
C4 Test Mean	50.9 (27.6)	41.3 (26.1)	30.6 (23.0)	58.8 (25.3)	58.1 (18.0)	30.9 (29.5)	45.1 (11.7)

being learnt (Johnson-Glenberg (2018)). The pointer interaction was not congruent with the actual actions required for the task thus not an appropriate interaction technique for learning, potentially introducing extraneous processing of the pointing task. In contrast, the C1 condition was more likely to have removed extraneous processing, as participants were solely focused on the task.

6.2.2. How do the interviews explain the results?

Some participants mentioned that the prompts told them where to click rather than explain the task that they were required to learn. Thus they failed to capture the reasoning for why they needed to perform this step and in what sequence. This is despite the fact that these prompts were the same as in the conditions where participants just passively observed the animations and this effect was not present. C2 Participant: “*Didn’t pay attention, just kept clicking. Clicking the buttons felt like a task to complete. Then next task, complete. There was no flow of the sequence of events to follow.*” C2 Participant: “*I paid too much attention to how to click and finish the task.*”

Multiple participants mentioned they did not pay attention to the learning tasks. This lack of attention appeared several times in the interview responses. This suggests that the pointer clicking resulted in extraneous processing of the clicking action, distracting from the real task. C2 Participant: “*I didn’t pay attention during the training. I couldn’t imagine the whole picture, what exactly I need to do.*”

This participant similarly could not remember tasks that were repeated as often as four times in the training (the process of pushing the wood). C2 Participant: “*I thought it would be easier than it was. I should have paid attention more. I remembered some tasks, but had no idea what to do when. Couldn’t remember that I had to push the wood to the blade, I thought I had to bring the blade towards the wood.*”

Some participants mentioned that clicking was an additional learning task that took cognitive effort. C2 Participant: “*The clicking adds another learning element for us to learn. Also, by clicking we are making sure this thing is picked up, but we overlook what it is meant to do, what goes before and after this. When in the test I was grabbing things I didn’t know where it actually goes, because previously I just had to click and it would automatically happen.*”

Some participants mentioned that clicking through the pre-recorded animations did not feel like they are interacting with objects. Animations refer to the pre-recorded steps that were shown to participants within the VR environment to demonstrate the correct steps. Participants in C2, C3 and C4 received animations. C2 Participant: “*It was confusing, it was just animations, I didn’t interact with the objects.*”

Some participants attempted to click during the test when they were told and given a tutorial on gripping objects instead. This may indicate that they had embodied the clicking interaction during the training. C2 Participant: “*The fact that the training’s laser pointing is different from how you interact in the test, that is hard. I tried to laser point to turn the machine one [during the test].*”

To conclude, in our comparison of different interaction techniques, we found that the use of a pointer (raycast) led to significantly lower performance than direct manipulation. Given the widespread use of pointer-based interactions in many VR systems, it highlights the importance of research on interaction design choices. Our study illustrates the potential pitfalls of using interaction techniques, such as the pointer, that may seem intuitive but do not align (are not congruent) with the physical actions associated with the learning task. Furthermore, the excessive cognitive effort associated with the in-congruent pointer interactions likely led to extraneous cognitive load, diverting cognitive resources away from the learning task itself. Hence, our study provides additional evidence for the importance of minimizing extraneous cognitive load in VR learning environments, especially when designing interaction techniques.

6.3. Cross-study comparison: C1 & C3 - What is the impact of active observation versus active manipulation on learning in VR?

6.3.1. What does the analysis of variance tell us?

We were surprised to find how similar the performance of those in the C3 condition (observation) was compared to the C1 condition (manipulation). So more interaction is not always justified when designing a VR learning experience. However, we believe there are a number of important things to note. While those in the C1 condition did not perform significantly better than those that had no interaction with the learning content, there were a larger group of participants that did not perform as well in the C3 condition. To be more specific, the C1 condition participants consistently performed well, while those in the C3 condition had a wider distribution. Perhaps, observation in VR in itself does result in embodied learning, even without interactions, given the sense of presence it creates (compared to simply watching a video, for example).

6.3.2. How do the interviews explain the results?

Participants implied that the training helped them embody the knowledge. C1 Participant: “*Once I do it once it is like a video in my mind. The instructions are clear about using your left hand to do what, I associated things with my hands and with my body, so I remembered how many steps to walk to the left side and right side. I had a lot of interaction with the spatial environment, so [in the test I remembered] when I am in this space this is what I need to do, when I am there I need to do that.*”

But those in the non-interactive conditions also talked about the sense of presence helping them connect to the physical actions. C4 Participant: “*This VR training [with animations] eliminates the experience part, but connects with the physical part, it felt like I was there.*”

While some participants felt the inability to interact was a huge disadvantage, others felt the sense of presence was helpful. C3 Participant: “*Because I couldn’t touch the things or interact with them, it just went through. It went through my head, it didn’t stick, because I couldn’t feel what I should be doing, physically trying to do these activities was different from just reading what to do.*”

C3 Participant: “*It wasn’t the order that I remembered the most, I remembered where the pieces went. It feels like you are in the space with them, and there is the actual object in front of you, you have the push stick and gauge here and this is how they move together.*”

While the conditions C3 & C4 performed similarly to the C1 condition, they felt they had to be fully focused to learn. The NASA TLX results help us to understand that those in the C3 and C4 conditions felt they had higher mental demand during the training, while the participants in the C1 condition did not have a high mental demand comparatively. Implying the need to really focus during the training to ensure they are learning. C4 Participant: “*Because it was not practical [interactive] I really had to use my brain. It helped that I went back and re-watched the animations as I remembered the safety glasses.*”

Participants in the C4 and C3 conditions mentioned that they were distracted by objects in the scene (extraneous processing), but C1 condition did not appear to have this problem. Perhaps the congruence of the interactions and the task being learnt helped them focus on the task only. C4 Participant: “*Having objects I didn’t need made me wonder if there will be animations associated with them.*”

Some participants mentioned that the animations created an illusion of learning that was not accurate or was mismatched with the expectations of the test. This is backed by studies that found watching videos often creates an illusion of understanding but the skills is not actually mastered (Seo et al. (2017)). C3 Participant: “*During the training I felt I had a good grasp of things, but forgot things during the test. I remembered what the instructions said but as I was doing it it felt more difficult.*”

One participant used embodied learning principles during the animated condition, they moved their body/hands tracing the steps in the animations. C3 Participant: “*When the animations played, I imagined the tasks, like putting the glasses on. It is easier to understand the process if I*

try to imitate the steps. I think it is pretty helpful, if you need to perform an action you simulate the steps without doing it so you can remember what sort of actions you need to perform.”

“It is different from watching a video as the view is fixed [in a video], in VR it makes it tempting to imitate the steps, also you can look around and wonder what different objects are used for. Walking around was helpful to look at other objects and wonder what they are used for, like the stop button next to the start button.”

To conclude, our findings highlight that active observation in VR can lead to similar learning outcomes as active manipulation, calling into question traditional assumptions about the superiority of interactive learning. These findings have significant theoretical implications for the understanding of embodied learning in VR. Particularly, they challenge the conventional notion that physical interactions are always necessary for embodied learning. Indeed, it seems that mere observation in an immersive VR environment can trigger a similar sense of embodiment, leading to effective learning.

6.4. Do the NASA TLX results confirm the findings?

The C1 condition has an overall higher NASA TLX score (Table 5) for the training and lower for the test, whereas in all other conditions (i.e., C2, C3 and C4) the trend is the opposite. It seems to suggest that while the fully interactive training (condition C1) may have a higher cognitive demand, it gave the participants an “easier” test experience. In contrast, though participants in the other conditions reported the training to be cognitively less demanding, they found the test to be cognitively more demanding. This is reiterated in some of the interview responses, where the participants reported that the training content was easy and that they felt confident in their knowledge. However, in the test, they found themselves not as skilled as they had previously thought. This is in accordance with the Challenge Point Framework (Guadagnoli, Morin, and Dubrowski (2012)), which suggests that some level of challenge during practice may cause a certain degree of failure in performance, and that failure or ‘struggle’ (Nawaz, Kennedy, Bailey, Mead, and Horodyskyj (2018)) during practice may eventually result in success during the test conditions. It is, therefore, important to note that the level or degree of difficulty of the tasks can have implications on both ‘practice’ and ‘learning’. As the difficulty of tasks increases, the ‘learning’ can also increase but with a corresponding decrease in practice performance. Once the learners are optimally challenged, learning occurs efficiently. Beyond the optimal challenge point, both practice performance and the learning performance can suffer. Therefore, while designing the VR learning environments, an effort should be made to ensure that the tasks offer some level of challenge or difficulty to the participants but not so difficult that it impedes their learning (Nawaz et al. (2021); Nawaz, Srivastava, et al. (2020)).

From Table 5 the mental demand numbers are also interesting, as participants stated in their interviews that they had to really “use their brain” to remember everything in the training and similarly in the test. For C1, however, the trend seems different - the mental effort scores under test condition are the lowest for this group. It may be indicative of embodied learning where the participants needed less mental demand to learn and recall.

6.5. Discussion of survey findings

The results from Survey A (given after training) and B (given after test) are presented in the Appendix. They further validate the findings. In particular, Fig. 6(a) and (b) show that participants in the C1 condition felt less competent in their performance after the training, and felt more confident after the test (illusion of understanding). This suggests that a hands-on approach that is congruent with the training while mentally demanding produces more accurate perceptions of competence. Additionally, the higher sense of embodiment in C3 and C4 compared to C1 may be interesting to note (Fig. 6(c)). As this sense of embodiment may

have supported their performance. Finally, it is important to note the results represented by Fig. 6(d) showing that the controls did not appear to be particularly challenging for any group as this could lead to artificially poor performance in one group.

6.6. Design considerations

From our findings we propose the following design considerations that should be taken into account when designing VR experiences for procedural training:

1. It is better to build a VR application with no interactions than one with interactions that are not congruent with the actions required to learn the task.
2. If creating a fully interactive environment with congruent interaction techniques is not possible, a seated or room-scale learning environment that is purely observational (with pre-recorded animations) can produce comparable learning outcomes.
3. A room-scale training environment is not necessarily superior to a seated one. Although users believe the ability to walk does trigger their spatial memory.

6.7. Limitations

It is important to acknowledge that the use of Quest 2 controllers may not fully replicate real-life “hands-on” interaction. The interaction fidelity may not be as high as other possible “hands-on” methods (e.g., using hand tracking alone, or controllers that more closely mimic the shape and feel of objects) and therefore the conclusions drawn are within the context of this specific interaction method and may only partially represent a high fidelity truly hands-on experience. However, due to limitations in current hand tracking technology (difficulties in interacting with objects and lacking any haptic feedback), and given that the Quest 2 controllers are the most common use case scenario we chose to use them to represent “hands-on” interaction.

One of the limitations of this study was the small sample size that prevented us from understanding differences between the various conditions. Additionally, the 18 tasks identified were perhaps too easy not allowing for us to see a real difference between conditions C1 and C3/C4. A more complex task (where for example, a device suddenly breaks down - something for which the participants were not prepared for) may highlight bigger differences in performance given that we already see a wider distribution in conditions C3 and C4 (D'Mello and Graesser (2014); Graesser and Olde (2003)). Additionally, due to a lack of certain conditions (eg. sitting and pointing, sitting and interactive) we were not able to do a comparison across all these factors in a 3×6 study design format.

Additionally, while the use of audio instructions was mentioned as future work, we believe the lack of it was a limitation of this study. Some researchers have found that text prompts do not work well in VR learning environments while others argue the opposite (Johnson-Glenberg (2018); Baceviciute et al. (2020)). Several participants needed the researcher to simply read out the text prompts word for word before they understood what to do, and this need was reiterated in the interviews. The addition of the option to listen to the audio instructions could have allowed for the evaluation of a more robust learning system.

6.8. Future work

A question for future studies is whether giving users more autonomy over how they learn by allowing participants to freely explore without constant prompts and no time limits would have an impact on the learning experience. In the future, we can further explore the provision of tools such as whiteboards and other note-taking tools that may support participant learning and allow us to better understand how different interaction techniques, interaction fidelity, and movement in

general play a role in training. While our environment included a whiteboard, participants mentioned that they either did not notice it, or felt they were not allowed to use it. But many wanted to be able to take notes.

Attentional tunneling is a problem in VR ([Syiem, Kelly, Goncalves, Velloso, and Dingler \(2021\)](#)) and so is extraneous processing. By giving users more agency we can see the impact of these play out and explore their relevance and if they impact embodied learning. One way to explore this is by removing or adding clutter (items user won't need). We can use gaze tracking to explore how much they impact or distract them. Several participants mentioned unused objects in the VR environment. Adding objects that were not part of the training to the test may also be a way to make the task harder and look more closely at the impact of interaction techniques, interaction fidelity and movement on learning. Or by perhaps putting all the tools in a tool box such that it is not within site anymore.

Text prompts have been shown not to be effective in some VR studies ([Johnson-Glenberg \(2018\)](#)), while others say text is ideal for remembering facts in VR ([Baceviciute et al. \(2020\)](#)). If the study design could be modified to include audio prompts (which many participants asked for), we could further test whether the training content is being optimally processed with audio or text prompts. The findings of these study may not generalize to all environments, given the diversity in VR learning experiences such as procedural learning versus declarative learning tasks.

The use of multimodal learning analytics to better understand and observe behaviours and actions of the participants in each condition would have helped us to quantitatively validate the interview results. By looking at gaze tracking data for example we could observe how often participants in each condition were distracted, or not looking at the relevant object. We could measure the distance covered by walking and correlate that to learning gains and so on.

In the future, an alternate way to assess participants' learning could be through a breakdown scenario within the VR environment which would require the participants to take charge of their learning and apply their learnt concepts in unknown situations ([D'Mello and Graesser \(2014\)](#); [Graesser and Olde \(2003\)](#)). This would allow us to explore how participants respond emotionally e.g., get confused, bored or frustrated ([Nawaz, Kennedy, Bailey, and Mead \(2020\)](#)). This is further emphasised

by prior studies ([Lee and Wong \(2008, pp. 231–241\)](#)) suggesting to consider psychological factors as they can influence the effectiveness of learning in VR based learning environments.

7. Conclusion

In conclusion, we have found that hands-on is not always better for learning. Not all types of interaction techniques support embodied learning, highlighting the importance of congruent movement in the design of interaction techniques. We found that higher interaction fidelity does not necessarily lead to better learning performance but the sense of embodiment provided by a purely observational experience in VR may create an embodied learning environment. Finally, we found that although participants felt that walking in VR stimulated their memory, there were no real difference in learning outcomes. Just being immersed in a seated position produces similar results to a walking experience, perhaps by reducing distractions. We believe that the fully interactive condition will produce more consistent results across a bigger population. This is due to the fully interactive condition leaving no room for error and perhaps minimizing distraction and extraneous processing. Whereas the conditions C3 and C4 require us to rely on participants paying attention for the training to be effective evident by NASA TLX mental load scores. However, producing a fully interactive condition like C1 requires a bigger investment of time and money, so institutions can choose a purely observational experience and achieve similar learning outcomes.

Statements on open data and ethics

The participants were protected by hiding their personal information in this study. They were voluntary and they knew that they could withdraw from the experiment at any time. The data can be provided upon requests by sending e-mails to the corresponding author.

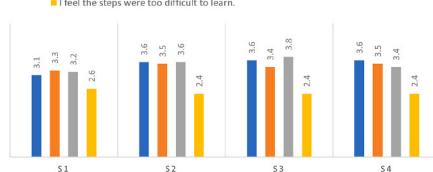
Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix

SURVEY A - COMPETENCE (TRAINING)

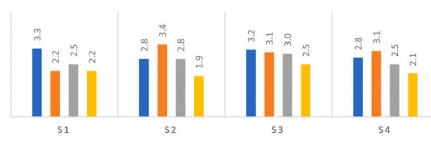
- I feel competent at the tasks taught in this VR training experience.
- I feel capable and effective when using this VR method of training.
- My ability to learn this material is well matched with the VR training's challenges.
- I feel the steps were too difficult to learn.



(a)

SURVEY B - QUESTIONS AFTER TEST

- I feel confident that I performed well on the VR test.
- The test was much more difficult than the training.
- The test left me feeling anxious.
- The controls for the test were hard to learn.

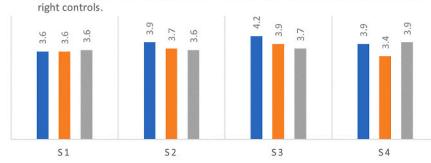


(b)

SURVEY A - INTUITIVE CONTROLS (TRAINING)

- Learning the controls of the VR experience was easy.
- The controls of the experience were intuitive.

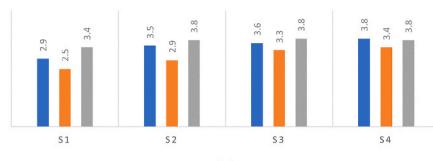
- When I wanted to do something in the VR experience, it was easy to remember the right controls.



(d)

SURVEY A - SENSE OF EMBODIMENT (TRAINING)

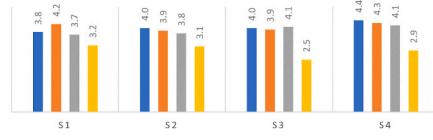
- It felt like I was in control of the movements during the task
- I felt like I could control the virtual hands as if they were my own hands
- I felt I could interact with the VR experience around me



(c)

SURVEY A - ENJOYMENT (TRAINING)

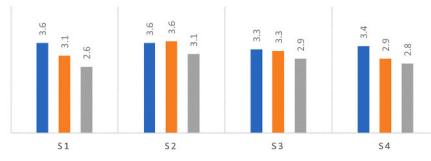
- I enjoyed the VR training experience.
- I found the VR training experience fun.
- I would recommend this type of VR training to my friends.
- I felt tense or stressed during the learning experience.



(e)

SURVEY A - AGENCY (TRAINING)

- The VR training experience provided me with interesting options and choices.
- The VR training experience let me explore freely.
- I experienced a lot of freedom in this VR training experience.



(f)

Fig. 6. Survey A results showing participants sense of competence, embodiment, agency during the training. Additionally how intuitive the controls of the experience felt, as well as how enjoyable the training was. Survey B included 4 post-test questions about similar concepts.

References

- Abrahamson, D., Nathan, M. J., Williams-Pierce, C., Walkington, C., Ottmar, E. R., Soto, H., et al. (2020). The future of embodied design for mathematics teaching and learning. *Frontiers in Education*, 5. <https://doi.org/10.3389/feduc.2020.00147>
- Andrade, A. (2017). Understanding student learning trajectories using multimodal learning analytics within an embodied-interaction learning environment. In *Proceedings of the seventh international learning analytics & knowledge conference* (pp. 70–79). New York, NY, USA: ACM. <https://doi.org/10.1145/3027385.3027429>.
- Baceviciute, S., Mottelson, A., Terkildsen, T., & Makransky, G. (2020). Investigating representation of text and audio in education VR using learning outcomes and EEG. In *Proceedings of the 2020 CHI conference on human factors in computing systems*. New York, NY, USA: ACM. <https://doi.org/10.1145/3313831.3376872>.
- Bergström, J., Knibbe, J., Pohl, H., & Hornbæk, K. (2022). Sense of agency and user experience: Is there a link? *ACM Transactions on Computer-Human Interaction*, 29, 1–22. <https://doi.org/10.1145/3490493>
- Bertrand, J. W. (2016). *Examining the effects of interaction fidelity on task performance and learning in virtual reality*. Ph.D. thesis, Clemson University.
- Biermann, O. C., Ajisafe, D., & Yoon, D. (2022). Interaction design for VR applications: Understanding needs for university curricula. In *CHI conference on human factors in computing systems extended abstracts* (pp. 1–7). New York, NY, USA: ACM. <https://doi.org/10.1145/3491101.3519859>.
- Caine, K. (2016). Local standards for sample size at CHI. In *Proceedings of the 2016 CHI conference on human factors in computing systems* (pp. 981–992). New York, NY, USA: ACM. <https://doi.org/10.1145/2858036.2858498>.
- Carpentier, K., & Lourdeaux, D. (2014). Generation of learning situations according to the learner's profile within a virtual environment. https://doi.org/10.1007/978-3-662-44440-5_15
- Checa, D., & Bustillo, A. (2020). A review of immersive virtual reality serious games to enhance learning and training. *Multimedia Tools and Applications*, 79, 5501–5527. <https://doi.org/10.1007/s11042-019-08348-9>
- Cheng, A., Yang, L., & Andersen, E. (2017). Teaching Language and culture with a virtual reality game. In *Proceedings of the 2017 CHI conference on human factors in computing systems*. New York, NY, USA: ACM. <https://doi.org/10.1145/3025453.3025857>.
- D'Mello, S., & Graesser, A. (2014). Confusion and its dynamics during device comprehension with breakdown scenarios. *Acta Psychologica*, 151, 106–116. <https://doi.org/10.1016/j.actpsy.2014.06.005>
- Eubanks, J. C., Moore, A. G., Fishwick, P. A., & McMahan, R. P. (2021). A preliminary embodiment short questionnaire. *Frontiers in Virtual Reality*, 2. <https://doi.org/10.3389/frvr.2021.647896>
- Ferguson, C., van den Broek, E. L., van Oostendorp, H., de Redelikheid, S., & Giezeman, G. J. (2020). Virtual reality aids game navigation: Evidence from the hypertext lossiness measure. *Cyberpsychology, Behavior, and Social Networking*, 23, 635–641. <https://doi.org/10.1089/cyber.2019.0435>
- Graesser, A. C., & Olde, B. A. (2003). How does one know whether a person understands a device? The quality of the questions the person asks when the device breaks down. *Journal of Educational Psychology*, 95, 524–536. <https://doi.org/10.1037/0022-0663.95.3.524>
- Guadagnoli, M., Morin, M. P., & Dubrowski, A. (2012). The application of the challenge point framework in medical education. *Medical Education*, 46, 447–453. <https://doi.org/10.1111/j.1365-2923.2011.04210.x>
- Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (task load index): Results of empirical and theoretical research. [https://doi.org/10.1016/S0166-4115\(08\)62386-9](https://doi.org/10.1016/S0166-4115(08)62386-9)
- Hedberg, J., & Dalgarno, B. (2002). The contribution of 3D environments to conceptual understanding. In O. J. Mc Kerrow (Ed.), *Winds of change in the sea of learning: Proceedings of the 19th annual conference of the australasian society for computers in learning in tertiary education* (pp. 149–158).
- He, L., Wang, R., Shi, X., Liang, Q., Fang, K., & Li, J. (2020). VR educational game design and research based on multi-modal interaction from the perspective of embodied cognition. In *2020 4th annual international conference on data science and business analytics (ICDSBA)* (pp. 329–331). IEEE. <https://doi.org/10.1109/ICDSBA51020.2020.00091>.

- Holger Hoffmann. (2022). Violin plot. <https://www.mathworks.com/matlabcentral/fileexchange/45134-violin-plot>.
- Jin, Q., Liu, Y., Yarosh, S., Han, B., & Qian, F. (2022). How will VR enter university classrooms? Multi-Stakeholders investigation of VR in higher education. In *CHI conference on human factors in computing systems* (pp. 1–17). New York, NY, USA: ACM. <https://doi.org/10.1145/3491102.3517542>.
- Johnson-Glenberg, M. C. (2018). Immersive VR and education: Embodied design principles that include gesture and hand controls. *Frontiers in Robotics and AI*, 5. <https://doi.org/10.3389/frobt.2018.00081>
- Keifert, D., Lee, C., Dahn, M., Illum, R., DeLiema, D., Enyedy, N., et al. (2017). Agency, embodiment, & affect during play in a mixed-reality learning environment. In *Proceedings of the 2017 conference on interaction design and children* (pp. 268–277). New York, NY, USA: ACM. <https://doi.org/10.1145/3078072.3079731>.
- Kiltinen, K., Groten, R., & Slater, M. (2012). The sense of embodiment in virtual reality. *Presence: Teleoperators and Virtual Environments*, 21, 373–387. https://doi.org/10.1162/PRES_a_00124
- Kim, Y. J., & Shute, V. J. (2015). The interplay of game elements with psychometric qualities, learning, and enjoyment in game-based assessment. *Computers & Education*, 87, 340–356. <https://doi.org/10.1016/j.compedu.2015.07.009>
- Koßmann, C., Straatmann, T., Mueller, K., & Hamborg, K. C. (2023). Effects of enactment in virtual reality: A comparative experiment on memory for action. *Virtual Reality*, 27, 1025–1038. <https://doi.org/10.1007/s10055-022-00701-y>
- Lee, E. A. L., & Wong, K. W. (2008). A review of using virtual reality for learning. https://doi.org/10.1007/978-3-540-69744-2_18
- Liao, M. Y., Sung, C. Y., Wang, H. C., Lin, W. C., & Cherng, F. Y. (2019). Embodying historical learners' messages as learning companions in a VR classroom. In *Extended abstracts of the 2019 CHI conference on human factors in computing systems*. New York, NY, USA: ACM. <https://doi.org/10.1145/3290607.3312861>.
- Lindgren, R., & Johnson-Glenberg, M. (2013). Embodied by embodiment. *Educational Researcher*, 42, 445–452. <https://doi.org/10.3102/0013189X13511661>
- Mackenzie, A. K., & Harris, J. M. (2014). Characterizing visual attention during driving and non-driving hazard perception tasks in a simulated environment. In *Proceedings of the symposium on eye tracking research and applications*. New York, NY, USA: ACM. <https://doi.org/10.1145/2578153.2578171>.
- Makransky, G., & Petersen, G. B. (2021). The cognitive affective model of immersive learning (CAMIL): A theoretical research-based model of learning in immersive virtual reality. *Educational Psychology Review*, 33, 937–958. <https://doi.org/10.1007/s10648-020-09586-2>
- Matthias, H., & Beckhaus, S. (2012). Adaptive generation of emotional impact using enhanced virtual environments. *Presence: Teleoperators and Virtual Environments*, 21, 96–116. https://doi.org/10.1162/PRES_a_00092
- Mayer, R. E., Makransky, G., & Parong, J. (2022). The promise and pitfalls of learning in immersive virtual reality. *International Journal of Human-Computer Interaction*, 1–10. <https://doi.org/10.1080/10447318.2022.2108563>
- Moher, T., Johnson, A., Ohlsson, S., & Gillingham, M. (1999). Bridging strategies for VR-based learning. In *Proceedings of the SIGCHI conference on Human factors in computing systems the CHI is the limit - CHI '99* (pp. 536–543). New York, New York, USA: ACM Press. <https://doi.org/10.1145/302979.303153>.
- Murphy, D., & Higgins, C. (2019). Secondary inputs for measuring user engagement in immersive VR education environments, Article 01586. [ArXiv abs/1910.04080](https://arxiv.org/abs/1910.04080).
- Nawaz, S., Kennedy, G., Bailey, J., & Mead, C. (2020a). Moments of confusion in simulation-based learning environments. *Journal of Learning Analytics*, 7, 118–137. <https://doi.org/10.18608/jla.2020.7.9>
- Nawaz, S., Kennedy, G., Bailey, J., Mead, C., & Horodyskyj, L. (2018). Struggle town? developing profiles of student confusion in simulation-based learning environments. *OpenOceans: Learning Without Borders*, 224.
- Nawaz, S., Srivastava, N., Yu, J. H., Baker, R. S., Kennedy, G., & Bailey, J. (2020b). Analysis of task difficulty sequences in a simulation-based POE environment. In *International conference on artificial intelligence in education* (pp. 423–436). Springer.
- Nawaz, S., Srivastava, N., Yu, J. H., Khan, A. A., Kennedy, G., Bailey, J., et al. (2021). How difficult is the task for you? Modelling and analysis of students' task difficulty sequences in a simulation-based POE environment. *International Journal of Artificial Intelligence in Education*. <https://doi.org/10.1007/s40593-021-00242-6>
- Oagaz, H., Schoun, B., & Choi, M. H. (2021). Performance improvement and skill transfer in table tennis through training in virtual reality. *IEEE Transactions on Visualization and Computer Graphics*. <https://doi.org/10.1109/TVCG.2021.3086403>
- Oberdörfer, S., Heidrich, D., & Latoschik, M. E. (2019). Usability of gamified knowledge learning in VR and desktop-3D. In *Proceedings of the 2019 CHI conference on human factors in computing systems* (pp. 1–13). New York, NY, USA: ACM. <https://doi.org/10.1145/3290605.3300405>.
- Peck, T. C., & Gonzalez-Franco, M. (2021). Avatar embodiment. A standardized questionnaire. *Frontiers in Virtual Reality*, 1. <https://doi.org/10.3389/frvir.2020.575943>
- Petersen, G. B., Mottelson, A., & Makransky, G. (2021). Pedagogical agents in educational VR: An in the wild study. In *Proceedings of the 2021 CHI conference on human factors in computing systems* (pp. 1–12). New York, NY, USA: ACM. <https://doi.org/10.1145/3411764.3445760>.
- Petersen, G. B., Petkakis, G., & Makransky, G. (2022). A study of how immersion and interactivity drive VR learning. *Computers & Education*, 179, Article 104429. <https://doi.org/10.1016/j.compedu.2021.104429>
- Ponder, M., Herbelin, B., Molet, T., Schertenlieb, S., Ulicny, B., Papagiannakis, G., et al. (2003). Immersive VR decision training. In *Proceedings of the workshop on Virtual environments 2003 - EGVE '03*. New York, New York, USA: ACM Press. <https://doi.org/10.1145/769953.769965>.
- Reeves, T. C. (2012). Interactive learning techniques. In *Encyclopedia of the sciences of learning* (pp. 1610–1611). Boston, MA: Springer US. https://doi.org/10.1007/978-1-4419-1428-6_331.
- Rettinger, M., Müller, N., Holzmann-Littig, C., Wijnen-Meijer, M., Rigoll, G., & Schmaderer, C. (2021). VR-Based equipment training for health professionals. In *Extended abstracts of the 2021 CHI conference on human factors in computing systems*. New York, NY, USA: ACM. <https://doi.org/10.1145/3411763.3451766>.
- Robbie, S. (2020). Literacies of the body. <https://doi.org/10.4018/978-1-7998-2588-3.ch010>
- Rogers, C., El-Mounayri, H., Wasfy, T., & Satterwhite, J. (2018). Assessment of STEM e-learning in an immersive virtual reality (VR) environment. *The ASEE Computers in Education (CoED) Journal*.
- Rowe, E., Asbell-Clarke, J., Baker, R. S., Eagle, M., Hicks, A. G., Barnes, T. M., et al. (2017). Assessing implicit science learning in digital games. *Computers in Human Behavior*, 76. <https://doi.org/10.1016/j.chb.2017.03.043>
- SawStop.
- Seo, J. H., Smith, B. M., Cook, M. E., Malone, E. R., Pine, M., Leal, S., et al. (2017). Anatomy builder VR. In *Proceedings of the 2017 CHI conference extended abstracts on human factors in computing systems*. New York, NY, USA: ACM. <https://doi.org/10.1145/3027063.3053148>.
- Skulmowski, A., & Rey, G. D. (2018). Embodied learning: Introducing a taxonomy based on bodily engagement and task integration. *Cognitive Research: Principles and Implications*, 3, 6. <https://doi.org/10.1186/s41235-018-0092-9>
- Slavova, Y., & Mu, M. (2018). A comparative study of the learning outcomes and experience of VR in education. In *2018 IEEE conference on virtual reality and 3D user interfaces (VR)* (pp. 685–686). IEEE. <https://doi.org/10.1109/VR.2018.8446486>.
- Snyder, C. W., Vandromme, M. J., Tyra, S. L., Porterfield, J. R., Clements, R. H., & Hawn, M. T. (2011). Effects of virtual reality simulator training method and observational learning on surgical performance. *World Journal of Surgery*, 35, 245–252. <https://doi.org/10.1007/s00268-010-0861-1>
- Strangman, N., & Hall, T. (2003). *Virtual reality/simulations*. Wakefield, MA: National Center on Accessing the General Curriculum.
- Sweller, J., van Merriënboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive Architecture and Instructional Design. *Educational Psychology Review*, 10, 251–296. <https://doi.org/10.1023/A:1022193728205>
- Syiem, B. V., Kelly, R. M., Goncalves, J., Velloso, E., & Dingler, T. (2021). Impact of task on attentional tunneling in handheld augmented reality. In *Proceedings of the 2021 CHI conference on human factors in computing systems* (pp. 1–14). New York, NY, USA: ACM. <https://doi.org/10.1145/3411764.3445580>.
- Waldrop, M. M. (2015). Why we are teaching science wrong, and how to make it right. *Nature*, 523, 272–274. <https://doi.org/10.1038/523272a>
- Watson, P., & Livingstone, D. (2018). Using mixed reality displays for observational learning of motor skills: A design research approach enhancing memory recall and usability. *Research in Learning Technology*, 26. <https://doi.org/10.25304/rlt.v26.2129>
- Wirth, M., Gradi, S., Sembdner, J., Kuhrt, S., & Eskofier, B. M. (2018). Evaluation of interaction techniques for a virtual reality reading room in diagnostic radiology. In *Proceedings of the 31st annual ACM symposium on user interface software and technology* (pp. 867–876). New York, NY, USA: ACM. <https://doi.org/10.1145/3242587.3242636>.
- Yatani, K. (2016). Effect sizes and power analysis in HCI. https://doi.org/10.1007/978-3-319-26633-6_5
- Yoo, S., Kim, S., & Lee, Y. (2020). Learning by doing: Evaluation of an educational VR application for the care of schizophrenic patients. In *Extended abstracts of the 2020 CHI conference on human factors in computing systems*. New York, NY, USA: ACM. <https://doi.org/10.1145/3334480.3382851>.