



Increasing the immersivity of 360° videos facilitates learning and memory: implications for theory and practice

Tom Peney¹ · Paul A. Skarratt¹

Accepted: 21 June 2024
© The Author(s) 2024

Abstract

Recent years have seen an increase in the use of immersive virtual reality (IVR) technology in education and training. Studies examining the efficacy of IVR-based interventions have shown improved performance compared to traditional training programmes; however, little is known about whether such improvements can be detected at the level of individual cognitive abilities. The present study sought to examine the effect of IVR on memory using an incidental learning paradigm. Undergraduate volunteers viewed a three-minute 360° video clip under immersive and non-immersive conditions—respectively, using a Head Mounted Display (HMD) or a 2D flat screen monitor—followed by a surprise recall task. Although both devices supported active exploration of the scene in 360°, recall was significantly improved for learners in the immersive condition. These findings suggest that IVR has a facilitative effect on cognition, and that learners may naturally engage with IVR-delivered content without any special instruction or preparation.

Keywords Immersive technologies · Virtual reality · Media in education · Informal learning · Improving classroom teaching

Introduction

Immersive virtual reality (IVR) refers to a computer-simulated environment that is rendered in three-dimensional space, allowing users to feel like they exist in an authentic world instantly and without constraint (Chittaro et al., 2018; Rodriguez et al., 2018). Although various devices can be used to create immersive environments, by far the most common are head-mounted displays (HMDs; Jensen & Konradsen, 2018). Due to recent advances in the complexity and affordability of HMDs, it is now possible for the general public to experience events and environments that would otherwise be inaccessible. Indeed, non-specialists can now effectively ‘climb’ Mount Everest, explore the streets of an historically accurate Pompeii, tour the International Space Station, and even take *Fantastic Voyage*-style field trips inside the human body. Although the content of these experiences can be accurately reproduced in two-dimensional videos, the immersive nature of

✉ Paul A. Skarratt
p.skarratt@hull.ac.uk

¹ School of Psychology, University of Hull, Cottingham Road, Hull HU6 7RX, UK

VR stimulates an experience that is phenomenologically different to watching a 2D video or a 3D machinima (Makransky & Lilleholt, 2018). Thus the idea that IVR can improve learning and teaching practises seems almost self-evidently true. With that in mind, the aim of the present study was threefold: first, to establish whether the effects of IVR are detectable at the level of individual cognitive abilities; second, whether the hypothesised effects manifest as improved learning outcomes; and third, whether these effects can be obtained under conditions that would make IVR useful to educators balancing differentiated learning practises with general classroom management (Roiha, 2014). Therefore, the present experiment was designed to provide the simplest possible test of IVR's efficacy as an educational tool by comparing the recall of learners viewing a 360° video clip via a HMD or a 2-D flatscreen monitor.

Comparing IVR with traditional learning methodologies

A growing number of studies have examined the use of IVR in a wide range of educational settings, assessing outcomes in terms of cognitive re-education following traumatic brain injury (De Luca et al., 2019), skills development (Howells et al., 2007; Xie et al., 2021), and health and safety training (Clevenger et al., 2015). Evidence suggests that IVR-delivered educational content can engage learners on emotional, motivational and empathetic levels (Calvert & Abadia, 2020; Makransky & Lilleholt, 2018). A typical examination of how IVR compares to other modes of learning was provided by Howells et al. (2007). They compared the performance outcomes for junior orthopaedic trainees following either a simulation-based training protocol rendered in IVR or a traditional in-theatre training method. Those in the simulation group performed virtual knee arthroscopies, while their in-theatre counterparts followed the traditional protocol of theory, observation and assistance during live surgery. Following six training sessions taken over seven days, the trainees were assessed in a live setting and rated on several competencies by experienced surgeons. Results were twofold. First, trainees in the simulator group were rated significantly higher on standardised measures of orthopaedic competency. Second, the simulator group outscored their counterparts even on kinematic measures of performance, suggesting that virtual interaction can improve real-world physical dexterity.

Although Howells et al. (2007) demonstrate the advantages of IVR over an extended training period, other studies have assessed learning outcomes following a single IVR exposure. In this regard, results are mixed. A recent study by Parong and Mayer (2021) had undergraduate students watch a video depicting a series of battles between Australian and Japanese forces in World War II. The video contained historically accurate battle simulations and narrations, as well as interviews with veterans who took part in the conflict. Importantly, participants viewed the video either on a standard 2D desktop monitor or while fully immersed using a HMD. Participants were asked to rate their subjective experience on variables such as presence, enjoyment and motivation. Learning outcomes were objectively measured using questions related to the displayed content, while the neural and physiological correlates of cognition were obtained through electroencephalography (EEG) and skin conductance responses (SCR), respectively. Results showed that although virtual immersion improved affective processing, with higher levels of enjoyment and lower levels of boredom, the benefits were not reflected in measures of cognitive processing or performance outcomes.

In contrast, Christopoulos et al. (2022) examined the learning outcomes of 10th and 11th graders taking part in an *Escape Room*-style game designed to facilitate learning

about cell biology. The game required students to solve a series of puzzles relating to the internal structures of a cell, thus enabling them to understand the role of enzymes and catalysts in normal cell function. Students performed the activity while fully immersed in the escape room via a HMD, or accessed the machinima via a PC or portable device. Importantly, the interactive and exploratory nature of the activity was preserved in both viewing conditions, as the gameplay mechanics allowed players to respond using computer peripherals such as touchpads and touchscreens.

In accordance with Parong and Mayer's study, Christopoulos et al. also recorded subjective and objective measures of learning, indexing knowledge acquisition and retention at different points pre- and post-intervention. However, unlike the earlier study, a positive effect of IVR was obtained on learning outcomes: Relative to students viewing the machinima, IVR learners showed significantly greater knowledge retention following a lag of one week, although this had dropped to the same level as the machinima group by the four-week mark. On the affective measures, the IVR group reported greater levels of engagement than did their machinima counterparts, experiencing greater enjoyment, motivation, and learning satisfaction. Despite these clear benefits, however, the IVR group did not rate VR as offering improved cognitive benefits or more effective learning. This is an interesting finding that raises questions about the role of IVR in the classroom; a point to which we shall return below and in our Discussion.

Empirical challenges when testing immersion effects

The above studies have been described in detail to highlight the challenges faced by researchers investigating immersion effects on performance. Indeed, one of the foremost difficulties happens to arise from IVR's principal strength—that is, its ability to create environments in which users feel *immersed*. The issue is that immersion is not a single variable that can be switched on or off according to the demands of an empirical test. Rather, it is the combination of multiple variables that determine how information sampled from the external environment in analogue form by the human sensory system can be replicated in digital form. Taking examples from only the visual modality, for instance, the human visual system uses a variety of cues to construct three-dimensional models of an environment. These include important binocular cues such as ocular vergence angle (Collewijn & Erkelens, 1990; Mon-Williams et al., 2000; Richard & Miller, 1969; Ritter, 1977; Viguier et al., 2001) and retinal disparity (Bishop, 1989; Mayhew & Longuet-Higgins, 1982), but also monocular cues such as retinal blurring (Mather, 1997; O'Shea et al., 1997), ocular accommodation (Mon-Williams & Tresilian, 1999, 2000), and the surface texture, contrast and shading of observed objects (Gonzalez & Perez, 1998; Johnston, 1991; Johnston et al., 1993; O'Shea et al., 1994). Thus, the problem is that immersion is actually a composite of all of these variables, whereas a rigorous scientific test demands the manipulation of only one.

When comparing learning under varying conditions of immersion, therefore, the challenge is in finding a trade-off that will allow researchers to retain as many of the features as possible that make an IVR experience so compelling, while selecting a control that is matched closely enough to make the comparison meaningful. This trade-off is evident in all of the studies cited above. For instance, the careful stimulus matching in Parong and Mayer's (2021) viewing conditions is laudable, as learners received the same visual content and audio narration in both IVR and 2D video. However, as the authors themselves point out, some objects were interactable in IVR but not in the 2D video. Thus the critical

differences between the two treatments extended to more than just the level of immersivity—the active versus passive engagement with the content may also have impacted on the learners’ experience and outcomes (Chrastil & Warren, 2012; James et al., 2002; Minhas et al., 2012).

Christopoulos et al. (2022) also made a commendable attempt to match immersive and non-immersive escape rooms—and show a positive effect of immersion on knowledge retention. However, their chosen task leaves one unable to establish precisely which aspect (or aspects) of cognition benefit from immersion. Indeed, one of the reasons why escape rooms are increasingly popular among educators is because they exercise a number of basic cognitive and executive functions, such as spatial, divided and temporal attention, encoding, retention and retrieval of information, reasoning, numerosity, reading, pattern recognition, logic and reasoning (e.g., Makri et al., 2021; Veldkamp et al., 2020). As such, a sensitive examination of IVR’s effects on cognition requires isolation of a single ability.

A final challenge concerns how positive research findings of IVR on cognition can be successfully translated into applied use within the classroom. The research cited above is important in identifying the conditions under which IVR can be effective, but if the benefits of IVR require extended periods of training (Howells et al., 2007), improve only the affective experience of learners (Parong & Mayer, 2021), or yield short-lived performance gains (Christopoulos et al., 2022), then the day-to-day practicalities of classroom management may not allow educators to commit the resources required to make IVR-based activities a viable part of school curricula.

With those issues in mind, the purpose of the present experiment was to examine the effect of IVR on a single cognitive ability that has direct bearing on educational attainment, and in a way that has maximum utility in a classroom setting. This is in contrast with the studies cited above, whose experimental designs and selected paradigms make it impossible to isolate individual cognitive abilities. To that end, forty undergraduate learners viewed a short (three-minute) video clip depicting a fictional scene taking place in an art gallery. Participants viewed the clip via an (‘immersive’) HMD or on a (‘non-immersive’) touchscreen monitor. The clip appeared in 2D in both cases but, as it was filmed in 360°, participants were able to control where they looked by moving their head in the immersive condition or by swiping the laptop touchscreen in the non-immersive condition. Following the film’s conclusion, participants were given a surprise recall test comprising 20 questions about the clip. There is a long history of using surprise recall as a means to measure incidental learning processes (e.g., McLaughlin, 1965). Thus the strengths of the present method were threefold: first, that the visual stimuli were identical in both viewing conditions; second, both devices allowed learners the same level of interaction with the content; and finally, the surprise recall test measured basic memory processes in the absence of instruction provided by the experimenter, and preparatory and motivational differences in the learners. As such, if immersivity has a positive effect on learning, then recall rates ought to be higher for IVR learners than for their 2-D counterparts.

Method

Participants

A total of 40 Psychology undergraduate participants were recruited from the University of Hull, UK. All were volunteers who responded to a research advertisement posted in

the researchers' home department, with the first eligible 20 male and 20 female respondents selected to take part. The only eligibility criterion was for participants to have (self-reported) normal or lens-corrected binocular vision. To ensure an equal number of males and females in each group, participants were pseudorandomly assigned to the two immersivity conditions. Participants in the 'immersive' condition had a mean age of 20.7 (SD=0.87), and 20.9 (SD=1.16) in the 'non-immersive' condition. This difference was not significant, $t(38)=0.77, p=.45$.

Apparatus and stimuli

The immersive condition was carried out using an Oculus Rift CV1 HMD. The HMD contains two OLED panels, resulting in a resolution of 1080×1200 pixels per eye, at a refresh rate of 90Hz, and with a field of view of 110° . The device was powered by Pentium PC running Windows 10. The stimulus video was presented on the Rift using Skybox (Skybox, 2019, version 0.2.0.1), a free piece of software capable of rendering and outputting 360° video to VR devices.

The non-immersive condition was carried out on a Microsoft Surface laptop also operating on Windows 10. The 13.5" touchscreen display had a native resolution of 2256×1504 and a refresh rate of 60Hz. The stimulus was presented on the laptop using the native Windows application 'Films & TV' (version 10.19031.1141.0), which allowed for playback and interaction with the stimulus via the touchscreen. The film was displayed at a quality of 1080s on both devices.

The stimulus used in both viewing conditions was a three-minute film entitled *Do Not Touch* (Leigh & Wickert, 2018; See Fig. 1). The video was shot in 360°, placing observers in the centre of the scene and allowing them to pan the video in any direction. The film depicts a fictional scenario set in an art gallery. The protagonist, after finding out he can 'enter' into the paintings, triggers an alarm and is subsequently pursued by two guards

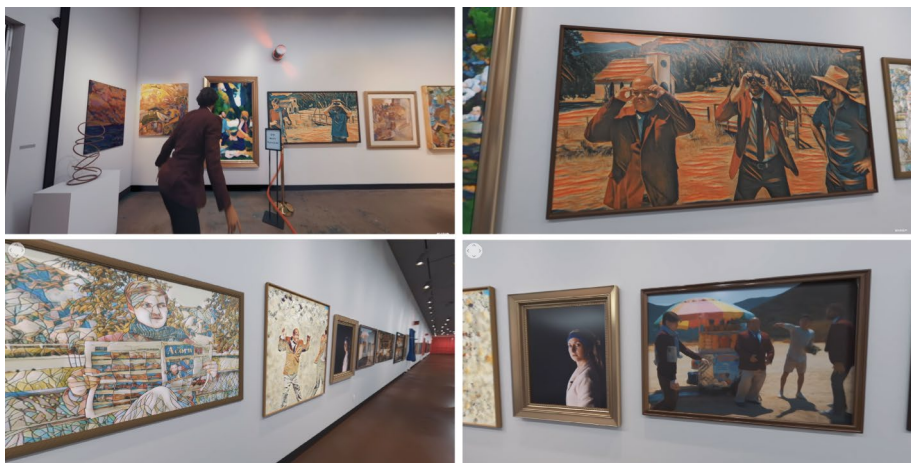


Fig. 1 Four stills from the learning content used in the present study—a 360° three-minute short film titled 'Do Not Touch', which depicts two security guards chasing a protagonist through the various painting within an art gallery. See Leigh and Wickert (2018; <https://www.youtube.com/watch?v=ecajlIKwCg>) for the full clip (best viewed in context using a HMD device or laptop)

through the various artworks as he attempts to flee, with the visual style and editing of the characters mirroring the painting in which they move around. The viewer does not fully enter the painting themselves, but rather observes them as if they are walking along the corridor of the art gallery.

A surprise cued-recall test was also administered, comprising 20 questions based on details from the video stimulus, such as vocalizations, physical actions and observable character features. Fifteen of these questions related to events at specific timepoints within the video, with the intervals between these spaced relatively equally. An example of one such target question is “What is the name of the newspaper the protagonist pretends to read?”. The remaining five items were foils, concerning details and events that did not occur in the film, e.g., “What colour is the horse in the first ‘wild-west’ painting?” (with no horse being present). In these cases, responses indicating the negative were coded as correct. Questions were presented in a unique random order for each participant on the same device on which they viewed the film, with answers inputted via a keyboard. The full list of questions can be found in the “[Appendix](#)”.

Design and procedure

The experiment employed a single-factor, between-subjects design comparing the effect of immersion on cued recall (see Fig. 2).

Participants were tested individually while seated in a dimly-lit laboratory. Following informed consent, participants were given slightly different instructions according to the viewing condition to which they were assigned. Participants in the immersive condition were instructed to put on the headset and perform an acuity check to ensure the HMD lenses were positioned correctly. This consisted of two green horizontal lines with edges

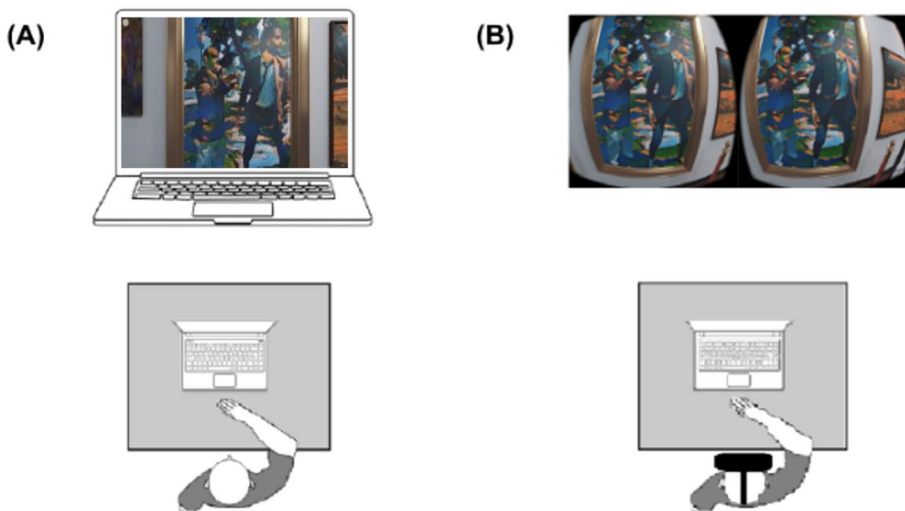


Fig. 2 The non-immersive (panel **A**) and immersive (panel **B**) viewing conditions in the experiment. Participants viewed and interacted with a three-minute film depicting events taking place in an art gallery. Following the film’s conclusion, participants completed a surprise recall test designed to measure incidental learning. Images are not shown to scale

that only appeared sharp when the HMD was in the optimal medial and lateral positions. When completed, the camera position was then centred, and the title scene of the stimulus film was presented. Participants were asked to read this aloud as to verify the acuity check was valid and, if not, the check was repeated. In the ‘non-immersive’ condition participants did not have to undergo any such calibration tasks. Instead, they were presented with the title screen of the stimulus and shown how to use the touchscreen to interact with and pan the video in different directions. Once comfortable with this, the stimulus was reset, such that the title was once again in the centre of the display.

After set-up was completed, participants in both conditions were informed that they would be shown a short 360° video clip and, while free to look around (either physically or by using the touchscreen), they should keep the scene focused on the protagonist. The video was then played in its entirety, with the ability to pause, fast-forward or rewind the video disabled for participants. Once completed, participants were informed that they would be presented with 20 questions based upon what they had just seen. It was emphasized that they should attempt to answer each question as accurately as possible, but that if a particular answer could not be remembered, or if they had no memory of the event taking place, to respond with ‘don’t know’ or ‘N/A’. Participants were then presented with the recall test. Upon completion of this task, participants were given a debriefing and thanked for their time.

Results

Participant responses were coded as either correct (1) or incorrect (0), yielding a maximum correct score of 20. Some responses required a subjective judgement to ascertain their accuracy. For example, the question “What colour shoes was the protagonist’s (real) friend wearing?” was potentially answerable as any synonym of “brown”, such as “tan”, or “beige”. Therefore, all participant responses were scored independently by two different scorers who were blinded to the immersivity conditions of the participants. One author of this study served as one scorer, while the other, a doctoral student in biomedical science, had no other involvement in the research. The unambiguous design of the questions and answers led to very high inter-scorer reliability, $r(38) = 0.997$, $p = .00001$.

The purpose of this study was to examine the influence of IVR on basic memory processes. Cued-recall scores were submitted to an independent samples t-test which showed that recall was significantly higher in the immersive condition ($M = 9.30$, $SD = 1.75$) than in the non-immersive condition ($M = 7.55$, $SD = 2.21$), $t(38) = 2.77$, $p = .009$. As such, greater immersivity improves incidental learning and recall ability.

Discussion

The present study sought to examine the effect of IVR on recall using an incidental learning paradigm. In the context of learning and education, previous research has demonstrated greater efficacy for IVR interventions over traditional training methods, at least in the case of extended training programmes (e.g., De Luca et al., 2019; Howells et al., 2007). In the case of single exposures, however, evidence suggests that while IVR improves affective feelings of engagement and enjoyment, it offers no improvement (Parong & Mayer, 2021) or relatively short-lived improvement (Christopoulos et al.,

2022) on learning outcomes. The present study examined the effect of immersion on incidental learning and recall memory. Participants viewed a short film via either an immersive HMD or a non-immersive 2D touchscreen, and were able to interact with the film using head movements or touch gestures, respectively. Following this, they were given a surprise cued recall task covering events in the film, and results showed that learners in the immersive condition recalled significantly more content than did their non-immersive counterparts. These results suggest a useful role for IVR as a way of delivering educational content, and further that learners naturally engage in VR-delivered content without any special instruction or preparation.

Our findings raise interesting questions about the mechanisms through which IVR can facilitate human cognition, and candidates are likely to include both arousal and attention. The ability of IVR to elicit physiological arousal cannot be overstated. Even though HMDs capable of displaying VR content have been in use for several decades, it is only within the last few years that their cost-per-unit has made them accessible to casual users (Hodgson et al., 2015). Thus for most people, IVR remains a highly novel experience and one that is eminently capable of increasing arousal relative to experiencing the same content on other devices (see Tian et al., 2021). The relationship between arousal and performance is famously depicted as an inverted U-shaped curve (Teigen, 1994; Yerkes & Dodson, 1908): Performance is maximal with moderate levels of arousal, but quickly drops off when arousal is too high or too low. Easterbrook (1959) further advanced the arousal-performance model by positioning attention as an intermediary factor. In other words, optimal attentional focus is maintained during moderate levels of arousal, but becomes narrowed when arousal passes thresholds at either end of the inverted curve. Thus, by controlling the gatekeeper to incoming information, arousal can modulate the amount of information that is encoded, stored and subsequently retrieved (see Storbeck & Clore, 2008, and Checa et al., 2021, for reviews on the contrasting effects of arousal on performance).

Because the present study focused on the effect of IVR on only cognitive performance, we elected not to measure levels of subjective affect. However, participants in both Parong and Mayer (2021) and Christopoulos et al.'s (2022) studies did show IVR-induced improvements in the psychological correlates of increased arousal—namely, higher enjoyment, motivation and learning satisfaction, and lower levels of frustration and boredom. Hence we may be seeing a picture in which IVR reliably improves the subjective and affective experience of learners, while stimulating only subtle and possibly limited effects on objective learning outcomes. Thus, returning to a point raised in our General Introduction, which we pose as an open question: What expectations do we have of IVR technology in the classroom? Would it be considered useful if it improves feelings of affect, but has only small, transient effects on knowledge retention?

While the use of IVR in education may ultimately rest on matters of school funding and finance, we suggest the following steps in order to submit the question to full scientific scrutiny. First, a proper empirical test requires a consensus of terms and concepts to be agreed. As Gonçalves and colleagues point out, we currently lack an accepted taxonomy of the various terms and definitions used in VR research, which can hinder clear communication (Gonçalves et al., 2022). Immersion, presence and fidelity are related concepts that are fundamental to creating a believable VR experience, and are widely used in the literature (e.g., Agrawal et al., 2020; Howard, 2019; Slater, 2018). However, many authors identify other concepts such as *coherence* (Skarbez et al., 2017), *authenticity* (Malliet, 2006), *credibility* (Boucaud et al., 2019), and *realism* (Ferwerda, 2003). While the definitions of the latter concepts seem to provide some overlap with the generally adopted terms of immersion, presence and fidelity, we need some consensus on whether they are merely synonyms

to be used interchangeably, or independent constructs that are best measured subjectively or objectively.

Second, having agreed these concepts, education researchers need to pursue two separate lines of enquiry: (i) whether IVR improves learning in practice; and, if so, (ii) *how* does IVR improve learning? These different questions require comparison treatments to be set at different levels of control. In the former, a macro-level of control is sufficient to compare the efficacy of IVR interventions with traditional methods of learning (De Luca et al., 2019; Howells et al., 2007; Xie et al., 2021). Such studies can examine IVR operating at its maximal capacity; that is, exploiting its capacity to create genuinely novel and even fantastical learning environments without any constraint whatsoever. In this context, the question of IVR efficacy is a simple empirical question that asks whether learners undertaking an IVR-based programme are more successful than are those following a traditional programme. Thus, matching IVR and control treatments requires only that the training content of two interventions are valid to the abilities under investigation. The second line of enquiry—understanding *how* IVR works at the *micro* level—must examine the effect on individual cognitive processes, but doing so demands that researchers set out precisely which elements of IVR are hypothesised to drive an effect. If it is *immersion*, the objectives of this line of research have to be to identify which subvariables constitute a sense of immersion, reducing them to a manageable number via a process of factor analysis, and manipulating immersivity *within a HMD* to maintain a very high level of experimental control. With that in mind, the present findings offer an early foray into that line of enquiry, and seek to offer suggestions for how we might match modes of delivery, manipulate immersivity, and isolate individual cognitive abilities of interest. We note that other researchers are making similar attempts and showing promising results (e.g., Sauter et al., 2020; Schubring et al., 2020).

This study has several limitations which should be noted. First, immersion was manipulated coarsely, and on a binary level. As discussed above, immersivity is supported by multiple constituent elements that can be difficult to separate, and the design of the present study does not allow us to identify which aspect(s) of immersion gave rise to the observed differences in recall performance. In accordance with our proposed ‘roadmap’ for how we think immersion effects can be empirically manipulated and tested, we are currently undertaking work in our laboratory in which we are attempting to manipulate immersivity at a more granular level.

Second, the study assessed only the immediate effects of immersivity on passive learning. As such, it was not possible to determine if the observed results occurred solely due to increased arousal resulting from the novelty of the experience, which we would argue are likely to be short-term in nature. Future research will explore the longevity of the facilitatory effects of IVR on cognition, resulting from both a single-shot exposure as in the present experiment, or as a longer-term training programme involving repeated exposures.

Finally, there are limitations relating to how closely the two presentation modes can be compared. The overall resolution and effective screen size of the laptop was different to that of the HMD, and of course the source of content interaction was different (touch vs. head movement). While this may elicit caution when interpreting the results from a strictly empirical position, such presentation mediums are indicative of what would be normally found in the classroom. Therefore, if such research findings are to translate well into use in the classroom, they ought to be achievable using equipment that is currently in use in school settings.

In conclusion, the contribution of the present study is threefold. First, it seeks to identify the challenges faced by researchers investigating the use of IVR in education

and pedagogy. Second, it offers a ‘roadmap’ for how researchers approach investigations into the effects of IVR on learning. Finally, it provides new evidence IVR can facilitate learning and memory in just a single exposure, and without special instruction from the instructor or preparation from the learner.

Appendix: Cued recall questions

Question	Target (T)/Foil (F)	Correct response
1. What did the sign at the start of the exhibition say?	T	Do Not Touch
2. What colour shoes was the protagonists (real) friend wearing?	T	Brown
3. What did the guard say into his walkie-talkie?	T	We’ve got a breach
4. How many people are in the first painting (before anyone has entered)	T	Two
5. In the second painting, what was borrowed from the man wearing a sunhat?	T	Binoculars
6. What is the name of the newspaper the protagonist pretends to read?	T	The Acorn
7. In one painting, a woman slaps the protagonist. What colour eyes did she have?	T	Blue
8. What was the bystander next to the food cart holding in his left hand?	T	Pineapple
9. Which cowboy wins the shootout—the one on the left or right?	T	Right
10. In the second ‘wild-west’ style painting, how many cowboys are in the scene?	T	Two
11. Next to the painting which falls from the wall, there was a set of curtains. What colour were they?	T	Blue
12. A collage of squares featuring a flexing man was displayed. How many squares were there in total?	T	Six
13. How many pieces of furniture were on the floating jetty?	T	Three
14. What does the protagonist use to swim away on?	T	Surfboard
15. What is the final action performed by the protagonist?	T	Wink
16. What colour was the feather attached to the protagonist’s hat in the final painting?	F	No feather/Blank
17. What was used to break the glass of the first painting?	F	Nothing/Blank
18. What colour is the horse in the first ‘wild-west’ painting?	F	No horse/Blank
19. What happens to the young boy in the American Gothic (farmer) style painting?	F	Nothing/There is none/Blank
20. How many dogs are in the beach-side painting?	F	None/Blank/No Dogs

Acknowledgements The authors would like to thank Zoe Booth for acting as an independent scorer.

Data availability The data used to support the findings of this study are available from the corresponding author upon request.

Declarations

Conflict of interest The authors declare that they have no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Agrawal, S., Simon, A., Bech, S., Bærentsen, K., & Forchhammer, S. (2020). Defining immersion: Literature review and implications for research on audiovisual experiences. *Journal of the Audio Engineering Society*, 68, 404–417. <https://doi.org/10.17743/jaes.2020.0039>
- Bishop, P. O. (1989). Vertical disparity, egocentric distance and stereoscopic depth constancy: A new interpretation. *Proceedings of the Royal Society of London b: Biological Sciences*, 237, 445–469. <https://doi.org/10.1098/rspb.1989.0059>
- Boucaud, F., Tafari, Q., Pelachaud, C., & Thouvenin, I. (2019). Social touch in human-agent interactions in an immersive virtual environment. In 3rd international conference on human computer interaction theory and applications (pp. 129–136). <https://doi.org/10.5220/0007397001290136>
- Calvert, J., & Abadia, R. (2020). Impact of immersing university and high school students in educational linear narratives using virtual reality technology. *Computers & Education*, 159, 104005. <https://doi.org/10.1016/j.compedu.2020.104005>
- Checa, D., Miguel-Alonso, I., & Bustillo, A. (2021). Immersive virtual-reality computer-assembly serious game to enhance autonomous learning. *Virtual Reality*. <https://doi.org/10.1007/s10055-021-00607-1>
- Chittaro, L., Corbett, C., McLean, M., & Zangrando, N. (2018). Safety knowledge transfer through mobile virtual reality: A study of aviation life preserver donning. *Safety Science*, 102, 159–168. <https://doi.org/10.1016/j.ssci.2017.10.012>
- Chrastil, E. R., & Warren, W. H. (2012). Active and passive contributions to spatial learning. *Psychonomic Bulletin & Review*, 19, 1–23. <https://doi.org/10.3758/s13423-011-0182-x>
- Christopoulos, A., Mystakidis, S., Cachafeiro, E., & Laakso, M.-J. (2022). Escaping the cell: Virtual reality escape rooms in biology education. *Behaviour & Information Technology*. <https://doi.org/10.1080/0144929X.2022.2079560>
- Clevenger, C., Lopez del Puerto, C., & Glick, S. (2015). Interactive BIM-enabled safety training piloted in construction education. *Advances in Engineering Education*, 1–14.
- Collewin, H., & Erkelens, C. J. (1990). Binocular eye movements and the perception of depth. *Reviews of Oculomotor Research*, 4, 213–261.
- De Luca, R., Maggio, M. G., Maresca, G., Latella, D., Cannavò, A., Sciarone, F., et al. (2019). Improving cognitive function in traumatic brain injury: A clinical trial on the potential use of the semi-immersive virtual reality. *Behavioural Neurology*. <https://doi.org/10.1155/2019/9268179>
- Easterbrook, J. A. (1959). The effect of emotion on cue utilization and the organization of behavior. *Psychological Review*, 66, 183–201. <https://doi.org/10.1037/h0047707>
- Ferwerda, J. A. (2003). Three varieties of realism in computer graphics. In *Proceedings human vision and electronic imaging VIII* (Vol. 5007). <https://doi.org/10.1117/12.473899>
- Gonçalves, G., Coelho, H., Monteiro, P., Melo, M., & Bessa, M. (2022). Systematic review of comparative studies of the impact of realism in immersive virtual experiences. *ACM Computing Surveys*. <https://doi.org/10.1145/3533377>
- Gonzalez, F., & Perez, R. (1998). Neural mechanisms underlying stereoscopic vision. *Progress in Neurobiology*, 55, 191–224. [https://doi.org/10.1016/S0301-0082\(98\)00012-4](https://doi.org/10.1016/S0301-0082(98)00012-4)
- Hodgson, E., Bachmann, E. R., Vincent, D., Zmuda, M., Waller, D., & Calusdian, J. (2015). WeaVR: A self-contained and wearable immersive virtual environment simulation system. *Behavior Research Methods*, 47, 296–307. <https://doi.org/10.3758/s13428-014-0463-1>

- Howard, M. C. (2019). Virtual reality interventions for personal development: A meta-analysis of hardware and software. *Human-Computer Interaction, 34*, 205–239. <https://doi.org/10.1080/07370024.2018.1469408>
- Howells, N. R., Gill, H. S., Carr, A. J., Price, A. J., & Rees, J. L. (2007). Transferring simulated arthroscopic skills to the operating theatre: A randomised blinded study. *Journal of Bone and Joint Surgery, 90B*, 494–499. <https://doi.org/10.1302/0301-620X.90B4.20414>
- James, K. H., Humphrey, G. K., Vilis, T., Corrie, B., Baddour, R., & Goodale, M. A. (2002). “Active” and “passive” learning of three-dimensional object structure within an immersive virtual reality environment. *Behavior Research Methods, Instruments & Computers, 34*, 383–390. <https://doi.org/10.3758/BF03195466>
- Jensen, L., & Konradsen, F. (2018). A review of the use of virtual reality head-mounted displays in education and training. *Education and Information Technologies, 23*(4), 1515–1529. <https://doi.org/10.1007/s10639-017-9676-0>
- Johnston, E. B. (1991). Systematic distortions of shape from stereopsis. *Vision Research, 31*, 1351–1360. [https://doi.org/10.1016/0042-6989\(91\)90056-b](https://doi.org/10.1016/0042-6989(91)90056-b)
- Johnston, E. B., Cumming, B. G., & Parker, A. J. (1993). Integration of depth modules: Stereopsis and texture. *Vision Research, 33*, 813–826. [https://doi.org/10.1016/0042-6989\(93\)90200-G](https://doi.org/10.1016/0042-6989(93)90200-G)
- Leigh, E., & Wickert, S. [SoKrispyMedia]. (2018, October 15). Do Not Touch (360 VR) [Video file]. Retrieved from <https://www.youtube.com/watch?v=ecajlIKwICg>
- Makransky, G., & Lilleholt, L. (2018). A structural equation modelling investigation of the emotional value of immersive virtual reality in education. *Educational Technology Research and Development, 66*, 1141–1164. <https://doi.org/10.1007/s11423-018-9581-2>
- Makri, A., Vlachopoulos, D., & Martina, R. A. (2021). Digital escape rooms as innovative pedagogical tools in education: A systematic literature review. *Sustainability, 13*, 4587. <https://doi.org/10.3390/su13084587>
- Malliet, S. (2006). An exploration of adolescents’ perceptions of videogame realism. *Learning, Media & Technology, 31*, 377–394. <https://doi.org/10.1080/17439880601021983>
- Mather, G. (1997). The use of image blur as a depth cue. *Perception, 26*, 1147–1158. <https://doi.org/10.1068/p261147>
- Mayhew, J. E., & Longuet-Higgins, H. C. (1982). A computational model of binocular depth perception. *Nature, 297*, 376–378. <https://doi.org/10.1038/297376a0>
- McLaughlin, B. (1965). “Intentional” and “incidental” learning in human subjects: The role of instructions to learn and motivation. *Psychological Bulletin, 63*, 359–376. <https://doi.org/10.1037/h0021759>
- Minhas, P. S., Ghosh, A., & Swanzy, L. (2012). The effects of passive and active learning on student preference and performance in an undergraduate basic science course. *Anatomical Sciences Education, 5*, 200–207. <https://doi.org/10.1002/ase.1274>
- Mon-Williams, M., & Tresilian, J. R. (1999). Some recent studies on the extraretinal contribution to distance perception. *Perception, 28*, 167–181. <https://doi.org/10.1068/p2737>
- Mon-Williams, M., & Tresilian, J. R. (2000). Ordinal depth information from accommodation? *Ergonomics, 43*, 391–404. <https://doi.org/10.1080/001401300184486>
- Mon-Williams, M., Tresilian, J. R., & Roberts, A. (2000). Vergence provides veridical depth perception from horizontal retinal image disparities. *Experimental Brain Research, 133*, 407–413. <https://doi.org/10.1007/s002210000410>
- O’Shea, R. P., Blackburn, S. G., & Ono, H. (1994). Contrast as a depth cue. *Vision Research, 34*, 1595–1604. [https://doi.org/10.1016/0042-6989\(94\)90116-3](https://doi.org/10.1016/0042-6989(94)90116-3)
- O’Shea, R. P., Govan, D. G., & Sekuler, R. (1997). Blur and contrast as pictorial depth cues. *Perception, 26*, 599–612. <https://doi.org/10.1068/p260599>
- Parong, J., & Mayer, R. E. (2021). Learning about history in immersive virtual reality: Does immersion facilitate learning? *Educational Technology Research & Development, 69*, 1433–1451. <https://doi.org/10.1007/s11423-021-09999-y>
- Richard, W., & Miller, J. F. (1969). Convergence as a cue to depth. *Perception & Psychophysics, 5*, 317–320. <https://doi.org/10.3758/BF03209573>
- Ritter, M. (1977). Effect of disparity and viewing distance on perceived depth. *Perception & Psychophysics, 22*, 400–407. <https://doi.org/10.3758/BF03199707>
- Rodriguez, C., Hudson, R., & Niblock, C. (2018). Collaborative learning in architectural education: Benefits of combining conventional studio, virtual design studio and live projects. *British Journal of Educational Technology, 49*, 337–353. <https://doi.org/10.1111/bjet.12535>

- Roiha, A. S. (2014). Teachers' views on differentiation in content and language integrated learning (CLIL): Perceptions, practices and challenges. *Language and Education*, 28, 1–18. <https://doi.org/10.1080/09500782.2012.748061>
- Sauter, M., Stefani, M., & Mack, W. (2020). Towards interactive search: Investigating visual search in a novel real-world paradigm. *Brain Sciences*, 10, 927. <https://doi.org/10.3390/brainsci10120927>
- Schubring, D., Kraus, M., Stolz, C., Weiler, N., Keim, D. A., & Schupp, H. (2020). Virtual reality potentiates emotion and task effects of alpha/beta brain oscillations. *Brain Sciences*, 10, 537. <https://doi.org/10.3390/brainsci10080537>
- Skarbez, R., Brooks, F. P., Jr., & Whitton, M. C. (2017). A survey of presence and related concepts. *ACM Computing Surveys*, 50, 1–39. <https://doi.org/10.1145/3134301>
- Slater, M. (2018). Immersion and the illusion of presence in virtual reality. *British Journal of Psychology*, 109, 431–433. <https://doi.org/10.1111/bjop.12305>
- Storbeck, J., & Clore, G. L. (2008). Affective arousal as information: How affective arousal influences judgments, learning, and memory. *Social & Personality Psychology Compass*, 2, 1824–1843. <https://doi.org/10.1111/j.1751-9004.2008.00138.x>
- Teigen, K. H. (1994). Yerkes-Dodson: A law for all seasons. *Theory & Psychology*, 4, 525–547. <https://doi.org/10.1177/0959354394044004>
- Tian, F., Hua, M., Zhang, W., Li, Y., & Yang, X. (2021). Emotional arousal in 2D versus 3D virtual reality environments. *PLoS ONE*, 16, e0256211. <https://doi.org/10.1371/journal.pone.0256211>
- Veldkamp, A., van de Grint, L., Knippels, M.-C.P.J., & van Joolingen, W. R. (2020). Escape education: A systematic review on escape rooms in education. *Educational Research Review*, 31, 100364. <https://doi.org/10.1016/j.edurev.2020.100364>
- Viguiet, A., Clement, G., & Trotter, Y. (2001). Distance perception within near visual space. *Perception*, 30, 115–124. <https://doi.org/10.1068/p3119>
- Xie, B., Liu, H., Alghofaili, R., Zhang, Y., Jiang, Y., Lobo, F. D., Li, C., Li, W., Huang, H., Akdere, M., Mousas, C., & Yu, L.-F. (2021). A review on virtual reality skill training applications. *Frontiers in Virtual Reality*, 2, 645153. <https://doi.org/10.3389/frvir.2021.645153>
- Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology & Psychology*, 18, 459–482. <https://doi.org/10.1002/cne.920180503>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Tom Peney is an experimental psychologist in the School of Psychology, University of Hull. His specialisms include VR high-level vision.

Paul A. Skarratt is an experimental psychologist in the school of Psychology, University of Hull. His specialisms include high-level visual and motoric processing.