Behaviour & Information Technology

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tbit20

Escaping the cell: virtual reality escape rooms in biology education

Athanasios Christopoulos, Stylianos Mystakidis, Enrique Cachafeiro & Mikko-Jussi Laakso

To cite this article: Athanasios Christopoulos, Stylianos Mystakidis, Enrique Cachafeiro & Mikko-Jussi Laakso (2022): Escaping the cell: virtual reality escape rooms in biology education, Behaviour & Information Technology, DOI: 10.1080/0144929X.2022.2079560

To link to this article: https://doi.org/10.1080/0144929X.2022.2079560

© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

Published online: 27 May 2022.

Submit your article to this journal

Article views: 1136

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=tbit20

BEHAVIOUR & INFORMATION TECHNOLOGY
https://doi.org/10.1080/0144929X.2022.2079560

Taylor & Francis
Taylor & Francis Group

RESEARCH ARTICLE

OPEN ACCESS Check for updates

Escaping the cell: virtual reality escape rooms in biology education

Athanasios Christopoulos [a], Stylianos Mystakidis [b,c], Enrique Cachafeiro[d] and Mikko-Jussi Laakso [e]

[a]Faculty of Science, Centre for Learning Analytics, University of Turku, Turku, Finland; [b]School of Natural Sciences, University of Patras, Patras, Greece; [c]School of Humanities, Hellenic Open University, Patras, Greece; [d]OESO, Duke Health, Durham, NC, USA; [e]Faculty of Technology, Department of Computing, University of Turku, Turku, Finland

ABSTRACT

Individuals' conceptual knowledge of contemporary biological issues influences ethical attitudes and life decisions around health, as evidenced by the recent worldwide debate on vaccinations. The reasons for which the public opinion is so critically divided can be partially explained from the didactic approaches used in biology education. In view of the rapid evolution of Virtual Reality (VR) as well as the increasing interest that youths have shown toward Escape Rooms (ER), an unexplored research and development avenue is identified. In this experimental study, we present the design and development steps as well as the evaluation of a VRER dedicated to biology education. For comparison purposes, an educational instructional video utilising the storytelling technique was also prepared. Both approaches were examined in the upper secondary education context with participants ($N = 50$) emerging from the Applied Sciences academic path. The key-findings suggest that the active learning approach leads to increased knowledge acquisition, in the short term, as opposed to the passive learning approach but no significant differences were identified in the long term (knowledge retention) across the methods. Based on the instructional decisions made and the key-findings emerged we provide implications and guidelines for the design, development, and integration of VRER in education.

ARTICLE HISTORY
Received 23 December 2021
Accepted 14 May 2022

KEYWORDS
Biology education; educational digital storytelling; escape room; machinima; virtual reality

1. Introduction

Researchers agree that procedural or higher-order cognitive skills are best rehearsed and attained through hands-on experimentation (Jones et al. 2019; Semilarski and Laius 2021). However, teacher-oriented instruction is still dominant in biology education, despite the serious limitations that such an approach inhibits on knowledge acquisition and construction even when basic scientific concepts are explored (Xu, Ye, and Wang 2021). This can be primarily attributed to the fact that schools are often lacking the tools or even the primary resources to support complex or abstract science-related experiments (Reeves and Crippen 2021). A proposed solution to counter this inadequacy of the educational systems comes from the review work that Gawlik-Kobylińska, Walkowiak, and Maciejewski (2020) conducted in a related field (Chemistry education) where educators are highly recommended to consider the adoption of alternative, digital-based, didactic approaches – such as instructional videos or serious games – as a means of supporting interdisciplinary learning and promoting conceptual understanding.

Educational or instructional multimedia videos of diverse forms (interactive, animated, storytelling, 360°) and styles (presentations/lectures, tutorials, demonstrations, scenarios, and simulations) are considered to be one of the most widely adopted and effective learning methods. Studies conducted especially in the STEM (Science, Technology, Engineering, Mathematics) education fields report a variety of applications that academic content videos have in the modern classroom (e.g. complimentary to the course delivery, asynchronous learning). The review that Otchie et al. (2020) performed has concluded that videos with educational content promote authentic learning and bring multiple benefits on learners' psychology (motivation, engagement, enjoyment), academic performance (knowledge construction, retention), and skill acquisition (psychomotor, cognitive, affective). Nevertheless, several issues are also reported with the most prominent being the technological challenges that content creators face (material collection, content design, quality assurance) and the increased cognitive load that learners usually report which impacts negatively the perceived learning experience (Brame 2016).

CONTACT Athanasios Christopoulos atchri@utu.fi

© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

The other recommendation concerns the introduction of the so-called 'Serious Games' to the educational practices. One of the key objectives of Serious Games is to provide reliable and, where possible, cost-effective instructional alternatives through which learners can experience the knowledge in a more vivid, interactive, and realistic way (Ravyse et al. 2017). Not surprisingly, the newly emerged digital Serious Games are heavily influenced by the wider evolution of emerging technologies, such as the ones emerging from the Mixed Reality spectrum. Researchers who have examined systematically the application of immersive technologies in STEM education have concluded that such applications provide fertile ground for the design and development of situated, problem-based, educational activities (Pellas, Dengel, and Christopoulos 2020; Mystakidis, Christopoulos, and Pellas 2021). To this end, the meta-analysis that Girard, Ecalle, and Magnan (2013) performed a few years ago had already highlighted the gradual growth of educational games dedicated to the STEM education fields. However, a more recent review has revealed that there is still lack of research and development efforts related to the subject of biology (Adita, Nugraheni, and Srisawasdi 2020).

Reviews of empirical studies that have explored computer-supported instruction in STEM education underline the added-value that gamification brings in the training process (Pellas, Mystakidis, and Kazanidis 2021). Escape Rooms (ER) constitute a relatively new recreational concept which has also inspired educators, from different disciplines and settings, to explore (Veldkamp et al. 2020; Makri, Vlachopoulos, and Martina 2021). Nevertheless, the findings of the review study of Lathwesen and Belova (2021), concerning the integration of ER in STEM education, identified the lack of practical applications and empirical evaluations in the fields of physics and biology.

Wu and Chen (2020) performed a systematic review on an alternative instructional approach – the Educational Digital Storytelling (EDS) – which has been receiving increasing interest within the last decade. According to the authors, EDS has been applied both as a stand-alone and supplementary pedagogy with overly positive results. Regardless of the intended use, storytelling is considered to be an effective approach to communicate science-related matters to the public (e.g. health campaigns) as it 'bridges the gap' between facts and understanding (Csikar and Stefaniak 2018). Nevertheless, despite the reported benefits that EDS entails, there is a scarcity of empirical works exploring its potential in STEM education subjects and even less so in the secondary education setting.

For the design and development of the story narration, content creators can utilise a wide range of multimedia tools including text, images, animations, and music (Robin 2016). To this end, a method that can potentially combine all the aforementioned techniques is called 'machinima'. Sagri, Sofos, and Mouzaki (2019) define machinima as 'a contemporary storytelling and filmmaking practice involving the capture of existing animated characters […] in order to create a film'. The wider acceptance that 3D (three-dimensional) environments have seen in the last decades has brought the opportunity to turn the virtual space into a 'canvas' wherein educational stories can be designed, unrolled, and recorded (Checa-Romero and Pascual Gómez 2018). Nevertheless, there is a scarcity of machinima videos in science education, despite its added-value to practice-based learning (Harwood 2013).

1.1 Significance of the study

Based on the future work recommendations elicited from the pertinent literature, the following unresolved issues related to biology education and digital instruction in the biology courses were detected. First, biology education is still relying heavily on the passive, teacher-oriented, learning model with restricted opportunities for practical exercises (Xu, Ye, and Wang 2021). Second, limited attempts can be identified where VR-supported Serious Games for biology education are integrated, especially in the secondary school context (Adita, Nugraheni, and Srisawasdi 2020). Third, there is an acute need for easily adaptable ER for STEM education and more empirical evidence on their effects on learners (Lathwesen and Belova 2021). Finally, studies that explore the impact of digital instruction usually assess the effect of an intervention immediately after its completion, without considering the time required for students to consolidate the newly acquired knowledge or skills (Girard, Ecalle, and Magnan 2013). In consideration of the latter, delayed assessments can be utilised as a means to control for the short-term memory or the novelty effect while also offering researchers the opportunity to gather robust evidences regarding the sustained impact of the training approach (All, Castellar, and Van Looy 2016). Therefore, conducting studies that evaluate participants' knowledge, also in the long-term, is highly recommended (Räsänen et al. 2009).

1.2 Contribution of the study

The aforementioned recommendations set the research objectives of this experimental study which are as follows. First, we sought out to investigate the integration of a novel instructional method (Virtual Reality Escape Room – VRER) in the context of biology education, as it seems to be the most under-supported and under-evaluated field when alternative instructional approaches are considered (Adita, Nugraheni, and Srisawasdi 2020; Xu, Ye, and Wang 2021). The integration was done in the physical classroom context, in contrast to the norm that relevant empirical studies usually follow —i.e. in fully controlled, Mixed Reality, laboratory environments where optimal conditions are assumed (c.f., Pirker et al. 2018; Makransky, Wismer, and Mayer 2019). Second, to strengthen the impact of the intervention, we chose one of the most interesting yet challenging subjects; biochemistry. According to Bretz and Linenberger (2012), Higher Education students have multiple misconceptions around this particular topic. As the authors report, students in the field have oftentimes inability to define even the basic characteristics of enzymes or to determine the impact that the alternation of environmental factors has on them. Although no relevant research could be identified in secondary school settings, it can be argued that the lack of conceptual understanding or – even worse – the development of misconceptions, has emerged well before tertiary education. Third, we explored and compared the educational potential of this alternative instructional method from diverse perspectives (active *versus* passive instruction – playing the VRER *versus* observing an actor playing the VRER via a machinima video) and points of view (cognitive benefits, motivation, engagement, enjoyment, satisfaction). Finally, we examined the effects and impact of these instructional methods in multiple intervals (i.e. short- and long-term) in an effort to identify their impact on both knowledge construction and retention.

In line with these objectives, the following research questions (*RQ*) were formed:

> *RQ1.* Is there any difference between the academic performance of students who followed the active learning instruction and their counterparts who followed the passive learning method on near and far transfer assessments?

> *RQ2.* Is there any difference between the perceived value of the learning experience of students who followed the active learning instruction and their counterparts who followed the passive learning method?

2. Background

2.1 Educational virtual reality escape rooms

Nicholson (2018) defines escape or breakout rooms as 'live-action, team-based, games where players discover clues, solve puzzles, and execute tasks in one or more physical spaces to accomplish a specific goal, usually within a limited time frame'. Although escape rooms emerged as an alternative leisure approach, the fact that players are required to use a wide range of cognitive (problem-solving, situational awareness, creativity, memory retention) and motor skills (dexterity, spatial awareness) to complete the challenge, attracted educators' interest and motivated their adoption in the school context. Relevant review studies report the successful integration of such educational activities in both formal (Makri, Vlachopoulos, and Martina 2021) and informal (Veldkamp et al. 2020) educational settings.

Returning to their origins, computer-generated ER have been implemented recently for pedagogical purposes both locally and online across all the educational levels (Makri, Vlachopoulos, and Martina 2021). Educational digital escape room games are considered to be a special genre of Serious Games as they provide an elaborated example of puzzle-based, playful learning, aligned to the following dimensions: theme, activities, narrative, and components (Mystakidis 2021). Escape rooms are usually positioned within a specific genre or theme, a common semiotic domain that dictates the game's components, aesthetics, and props. A story around specific characters creates the context of the mission, the role of the players in it and the array of problem-based activities and actions that propel active learning. Educational digital escape room games in VR environments are considered to be appropriate for both distance (online) and face-to-face instruction. This is mainly attributed to the flexibility that instructional designers have to align the contextual characteristics of such VR-based experiences to the cognitive needs and motivational incentives of learners of all age groups.

Key-findings of indicative empirical studies in Science education (Ang, Ng, and Liew 2020; Monnot et al. 2020) have concluded that digital ER have the potential to help students deepen their understanding and comprehend complex scientific concepts and procedures while at the same time recording positive effects on behavioural, affective, and social skills (engagement, interest, motivation, satisfaction, collaboration). Monnot et al. (2020) developed a website to emulate an ER experience with riddles on Chemical Engineering placed in the French revolution era. Results revealed high levels of satisfaction as students were able to create tangible connections between theory and its practical applications. Moreover, they proceeded to place students as designers of physical ER puzzles. Ang, Ng, and Liew (2020) created a physical collaborative ER on Chemistry and used Google Forms as vehicle to create a digital version for online teaching. Interestingly, the study emphasised the need of immersive environments for digital ER. As a matter of fact, although digital ER have been implemented successfully in various STEM fields, there is a dearth of evidence from empirical studies in VR-based ER in Science education. This scarcity can be attributed to the novelty of VR and the perceived complexity of designing an ER as well as on the design and programming skills that are required for the development of immersive environments (Pellas, Mystakidis, and Kazanidis 2021). Elford, Lancaster, and Jones (2021) used an immersive VR application to allow students to visualise and manipulate chemical molecules in an ER experience. This affordance led to in-depth discussions around the structural properties and increased engagement even from learners with lower interest in the subject.

2.2 Educational digital storytelling

Thomas and Schneider (2018) have mapped the uses of EDS in the following distinct ways: (a) as a complimentary material to the course books, (b) as a complimentary instructional activity, and (c) as a learning activity. The added-value of this approach is identified on the possibility to simulate various situations, that are difficult to perform in the real-world context, as well as on its low production cost, as the acquisition of physical resources is not required (Gawlik-Kobylińska, Walkowiak, and Maciejewski 2020). The integration of EDS in the classroom setting can support a wide variety of educational tasks including: (a) introduction of new topics, (b) review, discussion, and reflection on already explored subjects, and (c) provision of feedback, when assessment activities are performed. Besides, machinima videos can reach a wider audience when registered as open educational resources (Butler 2018).

Despite the potential of this approach to attract learners' attention (active participation, engagement, reflection) and the high expectations that teachers have in terms of facilitating students' field knowledge construction and retention (Karakoyun and Yapici 2016), the recent systematic review of Wu and Chen (2020) revealed that empirical works related to EDS have been mostly conducted in humanities and social sciences, with only a few examples identified in the field of science (computing and technology subjects, in particular). A possible explanation to this may be the fact that teachers lack the skills and the competencies required to create such sophisticated experiences (Tiba et al. 2015). Another noteworthy drawback that hinders educators' willingness in integrating this method concerns the time required to design and develop the storyline including its actual integration in the classroom context. In view of these obstacles, Anastasiadis et al. (2018) developed an authoring tool that blends the attributes of immersive technologies (Augmented / Virtual Reality) with storytelling elements and learning analytics. The platform has been developed in accordance with the needs that STEM educators have, as it provides fertile ground for the design and development of scientific inquiry learning activities, with the preliminary evaluation findings reporting overly positive results in students' collaborative problem-solving competences.

As far as the integration of EDS in biology instruction is concerned, only a handful of relevant works (Efthimiou 2016; Karakoyun and Yapici 2016; Cross 2017) could be identified of which only one (Mwelwa and Soko 2020) reports findings from the secondary school setting. In consideration of the latter, the adoption of this instructional approach in the classroom context had significant impact on both students' achievements and attitude towards biology. The aforementioned positive effects are attributed to the inherently vivid nature that EDS videos have as well as to their potential in transforming abstract scientific concepts and information into understandable and engaging visual narratives. Besides, as follow-up discussions with the participating students revealed, the adoption of this alternative method bridged the teacher-student relationship gap; an essential prerequisite for meaningful learning.

3. Materials and methods

3.1 Instructional design

Educational digital escape room games can be designed using a transdisciplinary approach (Arnab and Clarke 2017) wherein pedagogy and gameplay are unified (Lameras et al. 2017). One method to design such experiences is the 'multi-layer template' where all the essential dimensions are described and detailed (Clarke et al. 2017). Another recommendation comes from Zhang, Lu, and Park (2018) who explored the aspect of interactivity from the user experience perspective. According to the authors, the interplay design should be done carefully to avoid technical pitfalls as well as in a way that players can build trust and suspend disbeliefs. Christopoulos and Sprangers (2021) urge educational developers to consider carefully the intensity of gamification, especially when it comes to educational activities beyond the primary school setting. Under the term 'gamification' we define the application of game-design principles in non-gaming contexts (Arnab et al. 2015). Gamification in education enables educators and instructional designers to transform the learning activities – be it in the physical classroom or online – or even an entire course into a game (Romero-Rodriguez, Ramirez-Montoya, and González 2019). In view of these, we introduced gamification elements only in the gameplay, as recommended by Pellas, Mystakidis, and Christopoulos (2021), whereas, the visual representations are replicas of the equipment used in the laboratories.

The implementation of the above was done in collaboration with a biology educationalist who facilitated the procedures. The produced experience is an open educational resource for biology with a theme and a story related to the subject of enzymes. The general objective of the VRER is to assist students in building a mental model around the ways enzymes operate under different conditions and circumstances. The key educational objectives emerged from the curriculum-specified intended learning outcomes (Table 1) which are: (a) to help learners establish the necessity of enzymes and the role that catalysts play during the biochemical reactions and (b) to enable them comprehend how enzymes function. Moreover, learners were expected to develop a 'cause-and-effect' model related to biochemistry substances (enzymes) and the factors that impact their reaction speed.

For the technical development of the VRER the following decisions and respective actions were made. The multimedia elements and the gameplay mechanics were designed and developed using the Amazon Sumerian platform. The main reasons for choosing this development kit were the wide variety of content creation and animation tools that are natively offered to the developers as well as the cross-platform (Windows, iOS, Android) compatibility support it offers for the distribution of the produced product. The aforementioned elements make the creation of such experiences easily attainable, also from non-technically oriented individuals, such as the secondary school teachers. The second requirement taken into consideration was the time-pressure that the classroom timetable naturally applies. To overcome this obstacle, we had to ensure that the compiled experience would be immediately available to the students. For this reason, the Amazon Web Service 'Amazon CloudFront' was utilised. As a cloud service it enables developers to store the VR assets (3D scenes, animations, scripted interaction, special effects) online so that the VR environment can later be loaded directly into any WebVR-compatible browser (Google Chrome, Firefox, Microsoft Edge). The next requirement concerned the cost efficiency. In contrast to the added financial cost that is incurred for the acquisition of specialised equipment – such as headsets and handheld devices – when immersive-VR experiences are conducted, the present experience was communicated under the 'Bring Your Own Device' model. In simple terms, participating students could engage with the learning activities using either their personal computers or portable devices. For this reason, we ensured that all the interplay mechanics can be performed using either a computer mouse/touchpad (laptops) or the touchscreen (tablets). The final decision concerned the player setting (single player *versus* multiplayer). Although we acknowledge that ER are designed with the collaborative element in mind, adopting such an approach in the present study would have introduced uncontrolled bias (e.g. lack of equal engagement or participation across the team members). For this reason, the experience was delivered deliberately as a single player activity.

Concerning the design elements, the VRER features multiple gameplay levels (Figure 1) which include brief but linear challenges (Table 2) that need to be solved prior to proceeding to the next level. In the first level, players are introduced to the game objectives and gameplay mechanics. Accordingly, they collect keys that unlock basic theoretical information about enzymes. Information is presented in both readable (written text) and auditory narration form in line with the theory of Multimedia Learning (Mayer 2002). In the final stage, players are presented with a series of

Table 1. Connection of the intended learning objectives and outcomes with the instructional design decisions.

Bloom's Taxonomy	Intended Learning Objectives	Instructional Design	Intended Learning Outcome
Know	Recognise the structural elements of enzymes.	Experimental puzzle	Relate the fact that enzymes are proteins.
Comprehend	Identify which functions of enzymes are affected by other factors.	Experimental puzzle	Associate the impact that the alteration of pH and temperature have on the functions of the enzymes.
Apply	Formulate different enzyme structures.	Simulation/Metaphoric puzzle	Distinguish the different enzyme structures by matching the different shapes.
Analyse	Interpret the observable changes when the activation energy alters.	Simulation/Metaphoric puzzle	Explore how catalysts affect the biochemical decomposition reaction time.

Figure 1. Upper frames: interactive puzzles. Lower frame: content presentation of the VRER.

Table 2. Breakdown of the instructional design features of the VRER.

Phase	Cognitive Outcome	Approach / Rationale	Methods / Tools
Introduction	Removing technological barriers, understanding gameplay.	Technology Acceptance	Informative panels
Exposition	Presenting main facts and principles around the function on enzymes.	Multimedia Theory	Artefact collection Text and audio narration
Application	Four levels, realistic challenges of increasing complexity related to aspects of enzymes operation.	Problem-Based Learning	Tactile manipulation of virtual instruments and objects Puzzle solving

puzzles requiring them to apply the obtained information. Puzzles have been implemented in two ways; namely experimental and metaphoric. Experimental puzzles provide a realistic manipulation of equipment and digital objects. On the other hand, simulated metaphoric puzzles, provide visual equivalents as central vehicles for meaning-making, sophisticated reflection, and critical thinking (Mystakidis et al. 2021). Upon completion of each puzzle, players uncover a clue or an object that helps them advance to the next challenge in an adjacent virtual space.

Following completion of the developmental process, a machinima instructional video was rendered in line with the recommendations made by Lambert (2010) for digital storytelling design (integration of emotional content, inclusion of noticeable questions, incorporation of background music or narration audio, pacing of event changes).

3.2 Educational setting

In line with the principles of the convenience sampling approach (Cohen, Manion, and Morrison 2018), prospect participants were sought from a local upper-secondary school which, coincidentally, emphasises in subjects related to STEM education. While considering the experimental nature of the study as well as the relevance of the subject under investigation (enzymes) against the curricula structure for each grade, it was deemed wiser to invite students only from the entry/middle grades. Students from the final grade were not considered also out of an abundance of respect to their preparation efforts for the university-entry exams.

Students who expressed their wish to partake in the study were following the Applied Sciences academic path. In view of this, we can assume that the target group had an inherent interest (*intrinsic motivation*) toward the so-called 'hard sciences'. Although participation to the study was voluntary, the educator in charge agreed to elevate the assignments (homework) of this particular topic, as a means of compensation to students' time for participating in the study (applicable only to 10th grade students). Considering this, a positive boost on participants' *extrinsic motivation*, may as well be assumed.

3.3 Participants

The experiment took place at the school environment, during school hours. The intervention was pre-authorised from the school head whereas, the guardians of those students ($N = 57$) who expressed interest in participating were informed in writing about the nature of the study as well as the experimental procedures that were to be followed. Prospect participants countersigned an informed consent document prior to engaging in any research-related activities. However, not all the participants completed the required tasks and were, thus, excluded from the study. The final sample consists of fifty (50) students stemming from both the 10th ($N = 24$) and the 11th ($N = 26$) school grades.

3.4 Research design

The primary aims of both instructional approaches (EDS, VR) were similar: (a) to enhance learning of a specific biology subject, (b) to motivate students learn a new topic, and (c) to facilitate students' critical thinking skills. On a secondary level, we were able to: (a) examine the integration of alternative instructional approaches in the physical classroom setting, and (b) explore the impact of diverse digital instruction methods on knowledge acquisition and retention as well as on the perceived value of the learning experience. The research-related activities were conducted over the course of five classroom sessions across the time-span of eight calendar weeks. The research design is illustrated in Figure 2 and detailed in Table 3.

For the conduct of the intervention, the two non-equivalent groups design approach with repeated measurements, as recommended by Cohen, Manion, and Morrison (2018), was followed. The initial plan was to divide students in groups at classroom level. However, in an effort to eliminate the sample distribution bias, the non-randomised sample method, as described by Shadish, Cook, and Campbell (2002), was deemed more appropriate. To achieve this goal, participants filled in a personal background survey (Appendix 1) during the first stage of the study. After analysing the responses, they were assigned to one of the experimental groups (Group 1: EDS video; Group 2: VR experience) based on the answers provided regarding their: (a) gender, (b) experience with digital learning resources (educational videos / VR applications), and (c) English capacity (self-reported). Based on the anecdotal conversations, none of the participants had prior experience with VRER; be it for leisure or educational purposes.

To demonstrate our appreciation over students' willingness to support this initiative, the principal researcher delivered a short lecture related to the integration of emerging technologies (interactive books, 360° videos, Virtual/Augmented Reality) in the school setting. During the same session, the classroom teacher primed students with some key-information about the subject under investigation.

For the conduct of the learning activity, students from the VR group were gathered together and the researcher performed a short demonstration related to the gameplay of the ER. At the same time, the EDS group engaged with the machinima learning resource. Contrary to the fundamental element of ER (i.e. time pressure), no time restrictions were applied in this intervention; in fact, students from both groups were encouraged to interact with the digital learning resources for as long as deemed necessary. Therefore, the impact effect between the employed instructional methods derived only from their nature (active *versus* passive learning). Following completion of the learning activity, a debriefing session was held between the teacher/principal researcher and the students discussing the subject under investigation and the potential of each instructional approach.

The final stage included the knowledge acquisition and retention assessments as well as the evaluation of the learning experience. To eliminate the impact of the 'short-term memory effect' (All, Castellar, and Van Looy 2016), the posttest was performed one week later. At the same session, students were requested to fill in a psychometric survey related to their learning experience (Makransky and Lilleholt 2018). The delayed posttests were conducted four weeks later as recommended by Brown, Irving, and Keegan (2008).

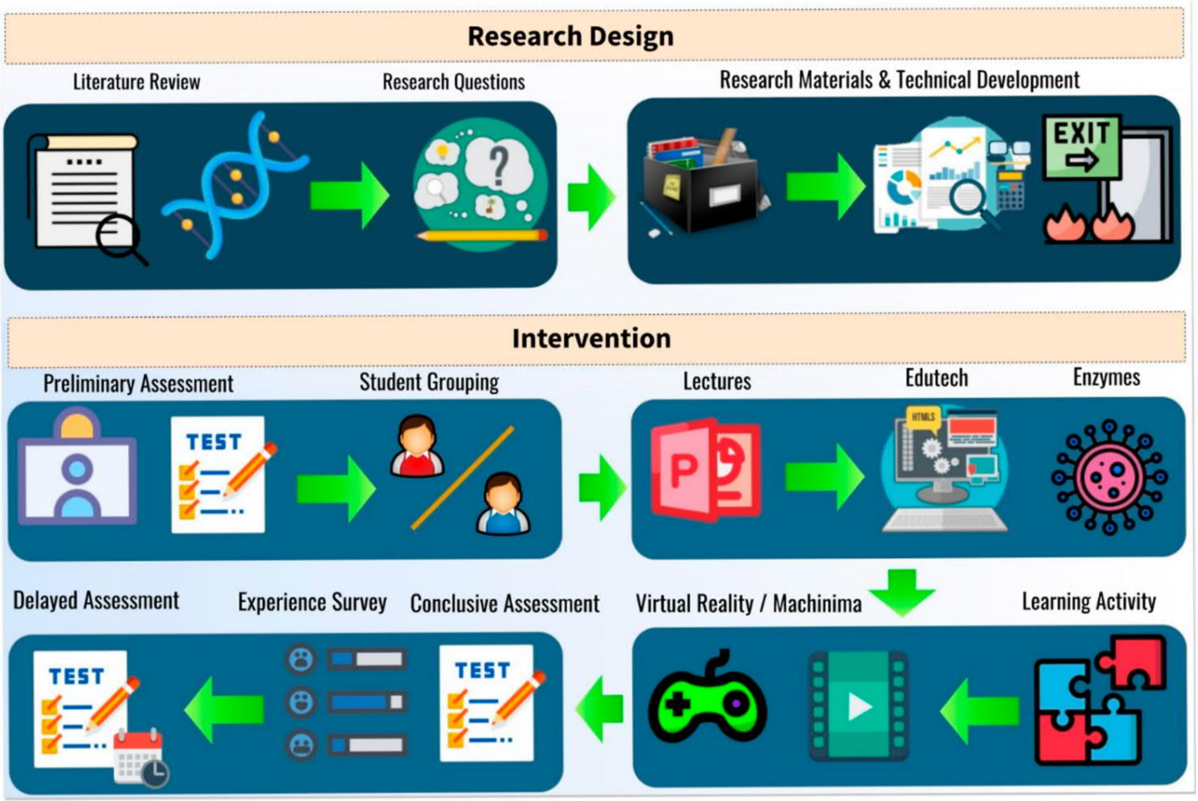

Figure 2. Overview of the research design.

Table 3. Overview of the research procedures.

Week	Procedures	Operations
–	*Research design*	Literature review on studies evaluating instructional practices in biology education with particular focus on the secondary school level. Identification of the open research gaps and generation of the research questions.
		Design of the research procedures (i.e. identification of appropriate data collection methods and materials). Design and development of the Virtual Reality biology escape room / machinima instructional video.
1	*Introductory session*	The principal researcher was introduced to the participants by the classroom teacher. Following a brief self-presentation, students were informed about the purpose of the visit and the nature of the experimental study. Accordingly, they were requested to: (a) counter sign a consent form, (b) complete the pre-intervention survey (demographics and background information), and (c) undertake the knowledge quiz (independently). This activity lasted approximately 15 min per visit paid to each cohort.
	Grouping	Students were assigned in one of the two groups based on the responses provided regarding their gender, experience with digital learning resources (educational videos / Virtual Reality applications), and English capacity (self-reported).
2	*Lecture*	A brief session related to different innovative technologies (interactive books, 360° videos, Virtual/Augmented Reality) and their potential in education was given to all students as a compensation for their willingness to participate in this study. This activity lasted approximately 20 min.
	Presentation	Students were given a brief presentation about the subject under investigation (enzymes) from the class teacher. This activity lasted approximately 15 min.
3	*Learning activity*	The researcher performed a brief demonstration to the Virtual Reality cohort related to the gameplay mechanics. At the same time, the EDS group was allowed to engage with the machinima resource. No time limit was applied; students were able to play the game / replay the video for as long as they wished.
		The classroom teacher and the researcher supervising the process were at students' disposal during this process. These activities lasted approximately 25 min, for the VR group, and 15 min, for the EDS group, respectively.
	Reflection activity	Students from both groups were encouraged to reflect on the learning activity. An informal discussion was held between the students and the classroom teacher/researcher regarding the pros and the cons of each instructional method. This activity lasted approximately 15 min.
4	*Posttest*	Students were requested to complete the same test used for the pretest. The order of the questions/items has been altered to prevent memorisation of answers. This activity lasted approximately 10 min.
	Questionnaire	Students were requested to complete the learning experience psychometric survey. This activity lasted approximately 15 min.
8	*Delayed posttest*	Students were requested to complete the same test used for the pretest/posttest. The order of the questions/items has been altered to prevent memorisation of answers. This activity lasted approximately 15 min.

Table 4. Overview of the constructs of the adopted psychometric instrument (Makransky and Lilleholt 2018).

Constructs	Explanation	Item Example	Items	Cronbach's *a*
(C1) Enjoyment	'Perceived enjoyment is the degree to which a student finds a virtual environment pleasant, fun, and enjoyable' 'VR fosters a higher level of immersion than standard media, which in turn could facilitate learning through positive emotions such as enjoyment'.	'I find watching educational videos / interacting with Virtual Reality simulations enjoyable'	3	.90
(C2) Motivation	'There is limited empirical evidence of the affective value of VR, and even less research that investigates the psychological process by which immersion impacts students' interest and motivation in the learning process' 'Understanding the underlying mechanisms that impact learners' perceptions and motivations can guide the optimal development of VR learning simulations'.	'This educational video / Virtual Reality Escape Room was boring'.	7	.91
(C3) Cognitive benefits	'Cognitive benefits are described as improved understanding and application as well as a more positive perception of the learned material'. 'Cognitive factors, such as the appraisal of cognitive benefits and the amount of control or autonomy, also play an important role in the learning process'	'Educational videos / Virtual Reality applications help me to better analyse the problems'.	4	.83
(C4) Learning effectiveness	'Perceived learning represents the degree to which the student perceives the simulation as educational'. 'Research into the effectiveness of using VR in education has been inconclusive'.	'I learned a lot of factual information in the topics'.	8	.83
(C5) Satisfaction	'Intrinsic motivation is defined as performing an action for the inherent satisfaction of the performance itself'. 'Satisfaction refers to the degree to which the student finds the simulation satisfactory'.	'A wide variety of learning materials was provided in this educational video / Virtual Reality escape room'.	7	.87

3.5 Instruments

To investigate the impact of the integrated instructional approaches on students' academic performance and gauge the differences on the learning experience, we adopted a quantitative approach. To measure participants' academic performance, a small set of multiple-choice questions (Appendix 2) was prepared by the biology teacher who supported this initiative. The quiz items (10) emerged after carefully scrutinising the multimedia content and further consulting the schoolbook. This allowed us to ensure that the whole spectrum of knowledge that students could acquire during the training process would be assessed both adequately and sufficiently. To facilitate the comparison process (repeated measurements) the same evaluation content was used across all the stages (pretest, posttest, delayed posttest), but in different order, as a precaution to the 'same-set' response effect (Sanchez 1992).

As for the learning experience, we adopted and adjusted the psychometric instrument developed by Makransky and Lilleholt (2018). According to the authors, the instrument has been specifically designed for and validated in studies that employ VR applications for subjects akin to the STEM education fields. In total, 29 items, distributed across five constructs (C1-C5), were relevant to the context of the present study (Appendix 3). To mitigate the impact of the language barrier, the data-collection items were translated to participants' native language by language professionals (back-translation method). Participants provided their responses, on a five-point Likert scale (5: 'Strongly Agree' to 1: 'Strongly Disagree'), anonymously (Table 4).

To facilitate the conduct of the study, a custom course was created in the Learning Management System that the school generally utilises. Participating students undertook the tests without, however, receiving any feedback on their performance (be it overall or in specific items) at the end of each assessment. Besides, in consideration of the experimental nature of the study, the performance results of these custom

Table 5. Descriptive statistics for the demographic and background data.

Statistic		Machinima		Virtual Reality	
		n	Percent	*N*	Percent
Gender	Males	12	48.0%	11	44.0%
	Females	13	52.0%	14	56.0%
Age	16 years old	13	52.0%	11	44.0%
	17 years old	12	48.0%	14	56.0%
English Level	Intermediate	6	24.0%	6	24.0%
	Advanced	13	52.0%	16	64.0%
	Proficient	6	24.0%	3	12.0%
Use of Digital Learning Resources	Slightly familiar	8	32.0%	9	36.0%
	Somewhat familiar	14	56.0%	16	64.0%
	Moderately familiar	3	12.0%	0	0.0%
Interest in Biology	Slightly interested	5	20.0%	5	20.0%
	Somewhat interested	14	56.0%	12	48.0%
	Moderately interested	6	24.0%	8	32.0%

Table 6. Descriptive statistics for the academic performance.

Statistic		Machinima			Virtual Reality		
		Pretest	Posttest	Delayed Posttest	Pretest	Posttest	Delayed Posttest
M		.36	.49	.44	.34	.57	.46
Mdn		.40	.50	.40	.30	.50	.40
Std. dev.		.11	.12	.11	.11	.13	.12
Skewness		−.48	.53	.70	.28	.78	2.24
Kurtosis		−.17	−.73	−.12	.81	.47	7.43
Cronbach's *a*		.79	.75	.72	.85	.74	.78
Shapiro-Wilk	W	.90	.87	.86	.91	.90	.73
	p	.021	.004	.003	.048	.021	< .01

Note: To ease the presentation of the results, the scores have been scaled (min = 0, max = 1).

assessments were not accounted toward the participants' final course grade.

3.6 Data analysis

The *R* programming language was used to analyse the primary data and a significance level of $p < .05$ was adopted. The datasets were preliminary explored via descriptive statistics. Accordingly, Cronbach's *alpha* was calculated to examine the reliability of the quiz items across each cohort as well as that of the educational experience instrument subscales. To determine which tests to use to analyse academic performance the Shapiro–Wilk test of normality was performed wherein skewness and kurtosis were examined (Orcan 2020). Given that the data were not normally distributed as well as the small sample size, nonparametric tests were adopted.

Specifically, a Mann–Whitney *U* test was used to determine differences in academic performance by gender, whereas Spearman's Coefficient of Rank Correlation (*rho*) was used to identify any relationships between the remaining demographic characteristics and academic performance at each administration of the quiz (Oti, Olusola, and Esemokumo 2021). The latter is a non-parametric, rank-based, measure that is a good option when ties (i.e. equal scores) are presented in the data (Puth, Neuhäuser, and Ruxton 2015). A Wilcoxon Signed Rank test was used to determine differences from pretest to posttest, posttest to delayed posttest, and pretest to delayed posttest for each group. In case of statistical significance, Cohen's *d* was calculated to measure effect size (Sawilowsky 2009).

Finally, a Mann–Whitney *U* test was used to determine differences in academic performance between groups (EDS; VR) at pretest, posttest, and delayed posttest. Concerning the impact of each instructional approach on the educational experience, Kruskal–Wallis *H* tests were used to determine differences by group, since the data were ordinal and the study employed a small sample. Differences were determined by each subscale of the instrument.

4. Results

4.1 Demographics

The distributions were fairly similar for both groups. There were approximately equal numbers of boys and girls. Likewise, an almost equal variation was achieved in both groups regarding participants' age. Most students had an intermediate English level and were somewhat interested in digital learning tools. Most students were somewhat interested in biology, despite the assumption that interest would be high due to voluntary participation and the greater inclination students in this particular school have as far as the STEM education fields are concerned. In relative terms, there were more students in the EDS group with proficient levels of English, as well as students who were moderately familiar with the use of digital learning tools (which were not present at all in the VR group). However, these are quite small groups and in absolute terms the differences are only 3 students. Table 5 shows the demographic and background data for each group.

4.2 Academic performance

The quiz utilised shows adequate to good internal consistency at all time points and for both groups (Table 6). However, the scores are not normally distributed, as evidenced by the significant effects on the Shapiro–Wilk tests. The distributions seem fairly symmetrical, except for the delayed post-test in the VR condition, where there is clearly positive skewness. Although most kurtosis values seem close to zero, the distribution in the machinima condition is somewhat more platykurtic (flatter with thinner tails), yet the distribution in the VR condition is more leptokurtic (has a higher peak and taller tails), especially at the delayed post-test.

Mann–Whitney *U* tests were performed to test for gender differences in performance. Boys and girls showed equal test performance at the pretest ($U = 58$, $z = 1.132$, $p = .257$), posttest ($U = 61$, $z = .965$, $p = .334$), and delayed post-test ($U = 61.5$, $z = .958$, $p = .338$) in

Table 7. Correlations between academic performance, demographics and background information.

Statistic	Machinima			Virtual Reality		
	Pretest	Posttest	Delayed Posttest	Pretest	Posttest	Delayed Posttest
Age	.22	.42*	.34	.02	−.25	−.35
Grade	.02	−.17	−.07	−.06	.19	.14
English	−.19	−.28	−.27	.12	−.01	−.02
Familiarity	−.08	−.12	−.07	−.14	−.07	−.08
Interest	.18	.14	.03	.20	.46*	.30

* $p < .05$.

the machinima condition. Similarly, in the VR condition, boys and girls showed equal performance at the pre-test ($U = 72$, $z = .29$, $p = .771$), post-test ($U = 51.5$, $z = 1.445$, $p = .148$), and delayed post-test ($U = 58$, $z = 1.136$, $p = .255$). This indicates that boys and girls did not differ on prior performance and that the effects of the two types of instructional methods had a similar effect for boys and girls.

Table 7 shows the Spearman *rho* correlations between the background variables and performance by group for each of the three measurement occasions. At the immediate post-intervention measurement, there was a positive effect of age on performance within the machinima condition ($\rho = .42$, $p = .035$). This indicates that students within this condition generally showed more short-term learning effect the older they were. On the other hand, in the VR condition, there was a positive effect of interest on the immediate post-intervention measurement ($\rho = .46$, $p = .020$). This makes it clear that the short-term learning effect was greater the more interested students were in digital learning tools.

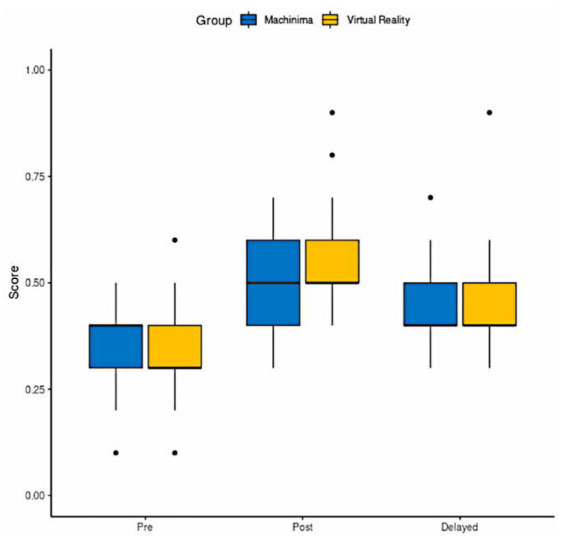

Figure 3. Comparison of participants' pretest, posttest, and delayed posttest performance.

As would be expected, there is an increase in the averages at the post-measurements compared to the pre-measurements. The Wilcoxon Signed Rank test was used to determine if there were significant differences in performance at the three measurement points. Since a condition effect was expected, this was done separately for both conditions (Figure 3).

Within the machinima condition, students' performance improved at the posttest compared to the pretest ($z = -3.84$, $p < .001$, $d = 1.13$), but at the delayed posttest the performances became somewhat worse compared to the direct posttest ($z = 2.80$, $p < .001$, $d = .37$). The scores at the delayed posttest nevertheless still showed an improvement compared to the pretest (z = −3.00, $p < .001$, $d = .81$). In the VR condition, a similar pattern was found. Again, the performance improved at the posttest compared to the pretest ($z = -4.106$, $p < .001$, $d = 1.95$), but then deteriorated at the delayed posttest compared to the direct posttest ($z = 3.82$, $p < .05$, $d = .90$). Compared to the pretest, there was still a strong improvement in scores on the delayed posttest ($z = -3.37$, $p < .05$, $d = 1.05$).

In order to compare the performance between the two conditions, Mann–Whitney *U* tests were carried out. There were no significant differences between the two conditions at the pretest ($U = 269$, $z = .882$, $p = .377$). However, a significant difference was identified on students' performance at the posttest ($U = 203$, $z = -2.19$, $p = .028$) with a medium effect size and in favour of the VR group ($d = .31$). At the delayed posttest, the students showed, once more, similar performance ($U = 296$, $z = -.345$, $p = .730$). Thus, even though there was a stronger learning effect for the students in the VR condition, immediately after the learning phase, the positive, longer-term learning effect was similar in both conditions.

4.3. Learning experience

All subscales had satisfactory to good reliability (Table 8). Kruskal–Wallis *H* tests were carried out to examine if there were any significant differences between the two conditions. Significant differences were identified in three out of the five constructs. In greater detail,

Table 8. Descriptive statistics for the learning experience survey.

Construct	Group	Min	Max	*M*	Med	*Std. Dev.*	Cronbach's *α*
(C1) Enjoyment*	*Machinima*	2	5	3.39	3	.59	.83
	Virtual Reality	2	5	3.61	4	.71	.77
(C2) Motivation*	*Machinima*	1	5	3.45	3	.98	.82
	Virtual Reality	2	5	3.67	4	.76	.78
(C3) Cognitive benefits	*Machinima*	2	5	3.50	3	.78	.80
	Virtual Reality	2	5	3.65	4	.77	.84
(C4) Learning effectiveness	*Machinima*	1	5	3.49	3	.97	.85
	Virtual Reality	2	5	3.62	4	.73	.80
(C5) Satisfaction*	*Machinima*	1	5	3.21	3	.97	.82
	Virtual Reality	1	5	3.38	4	.99	.82

Note: *Significant difference between the conditions at $p < .05$ as indicated by the Kruskal–Wallis *H* tests.

students in the VR condition ($Mdn = 4$) reported a greater degree of enjoyment while playing the VRER as opposed to their counterparts ($Mdn = 3$) who simply observed the machinima video ($H = 4.06$, $\chi^2(1) = 3.302$, $p = .043$). An even stronger difference was identified on participant's motivation for learning ($H = 5.509$, $\chi^2(1) = 4.913$, $p = .018$), with the students in the VR condition ($Mdn = 4$) providing, once again, more positive feedback as opposed to the machinima cohort ($Mdn = 3$). Finally, a difference was also identified in participants' learning satisfaction ($H = 4.112$, $\chi^2(1) = 3.677$, $p = .042$). Students in the VR condition were again more positive ($Mdn = 4$) compared to students in the machinima condition ($Mdn = 3$). No further differences were found. In other words, the perceived cognitive benefits and learning effectiveness were similar for students in the machinima and the VR condition ($H = 1.869$, $\chi^2(1) = 1.608$, $p = .171$ and $H = 2.510$, $\chi^2(1) = 2.218$, $p = .113$). These results indicate that students have similar experiences when it comes to the learning effect in both conditions, which is consistent with their actual longer-term performance, but that had more positive affective experience in the VR condition, which is reflected in enjoyment, motivation, and satisfaction.

5. Discussion

5.1 Empirical contribution

The main argument put forward in this study challenged the inadequacy of the traditional, teacher-centered, instructional approach to facilitate deep and meaningful learning (Xu, Ye, and Wang 2021). Likewise, the scarcity identified in the international literature regarding the limited Technology-Enhanced Learning efforts to support and facilitate the didactic of biology – especially in the secondary school level (Adita, Nugraheni, and Srisawasdi 2020; Gawlik-Kobylińska, Walkowiak, and Maciejewski 2020) – motivated further the conduct of such intervention.

The novelty of the proposed approach goes beyond the instructional design elements that typical Science simulators usually present wherein educators and learners perform predefined tasks and observe the outcomes of their actions (c.f., Pirker et al. 2018). In view of these, we proposed, designed, and developed an alternative instructional experience – mediated via VR and framed within an ER – and further enlisted and outlined the steps followed to achieve this outcome. This is effectively the first contribution of the present work in the field of instructional science. Another sign of innovation concerns the conduct of the VR-supported activities to the school context (Reeves and Crippen 2021) as opposed to extending the classroom setting to a Mixed-Reality laboratory; a practice which is common in such studies (Pellas, Mystakidis, and Kazanidis 2021).

While considering the fundamentally diverse didactic methods that have been generally adopted in Science-related subjects (active *versus* passive learning), we introduced an instructional approach which combines and integrates elements that learners are conceptually and practically familiar with (i.e. interaction with replicas of laboratory equipment, learning from an instructional video). Accordingly, we sought to identify any substantial differences between the academic performance of participants who followed the active learning instruction, as opposed to their counterparts who followed the passive learning method, on near and far transfer assessments. To contextualise the findings related to academic performance, we also explored the most notable differences between the perceived educational value of each experience across the two groups.

The academic performance of both groups improved after the intervention, recording a strong improvement in the direct and delayed posttest, a phenomenon reported also in the literature (Lathwesen and Belova 2021; Makri, Vlachopoulos, and Martina 2021). However, the VR group significantly outperformed the EDS group in the immediate posttest, administered one week after the intervention, but this effect deteriorated in the delayed posttest at the four-week mark.

As for the reason of these outcomes, it can be assumed that the relatively short duration of the intervention as well as the limited number of exercises that learners performed withheld the opportunity for spaced practice and repetition to enhance retention. An alternative explanation comes from Gligora Marković, Plantak Vukovac, and Kliček (2015) who have correlated the quality of the multimedia instructional materials with knowledge acquisition. Transferring their findings to the present study, it can be argued that the VR experience was of higher quality, when compared to the video learning resource. Under this consideration, the knowledge-level difference gap observed in the short-term assessment becomes more justifiable. In either case, an agreement across both studies is identified with regard to the gradual and proportional decrease in the long-term knowledge retention, regardless of the multimedia instructional methods utilised or their quality. Finally, the learning gains were influenced by students' interest in the used technological medium, a conclusion which is also in agreement with the findings that Lathwesen and Belova (2021) draw, but not their gender, an outcome which opposes the findings that Makransky, Wismer, and Mayer (2019) report.

In terms of learning experience evaluation, students rated the VR experience in a more positive way across all five psychometric dimensions. Significant differences were noted in favour of the VR group in the affective dimensions of enjoyment, satisfaction, and learning motivation. According to Chittaro and Buttussi (2015), interactive and non-interactive instructional methods – which are equally vivid in terms of visual and auditory elements – have similar impact on learners' outcomes and emotions. The authors of the present study believe that both multimedia resources were equally evocative, especially after considering that the machinima video contained the actual VR experience. Therefore, a possible explanation to the distinction identified in the affective dimensions can be given after accounting the role that the instructional approaches (i.e. active learning *versus* passive learning) played in the perceived experience (Pereira-Santos, Prudêncio, and de Carvalho 2019). In other words, the positive attitude that participants maintained over the VR supported instructional approach may as well be due to the fact that they were actively engaged with the experience (Huang et al. 2021). These findings can also be grounded to well-established theories – such as the Cognitive Theory of Multimedia Learning (Mayer 2014) or the interest theory (Harackiewicz, Smith, and Priniski 2016) – which distinguish the impact that diverse interactive media have on motivation, engagement, and performance. Elford, Lancaster, and Jones (2021) attribute the added value that VR brings in education to its ludic nature. According to the authors, this affective amplification was in fact stronger in students who reported lower levels of interest in the studied subjects. Indeed, the impact that motivation enhancement had on participants' academic performance improvement is also in agreement with the findings that Makri, Vlachopoulos, and Martina (2021) report. Nevertheless, the perceived cognitive benefits and learning effectiveness were similar in both conditions, as confirmed by the delayed academic achievement measures.

5.2 Implications

Notwithstanding the technology selection, the role of educators is to orchestrate situations and experiences for students to construct their own meaning through active engagement with content rather than transmit packaged information or promote rote learning. Hence, attracting teachers' interest and curiosity to experiment with educational technology is the prime goal. For this purpose, access to techno-pedagogical competences' training and technical support are equally important, particularly for the majority of teachers who possess no programming skills, as Christopoulos and Sprangers (2021) also recommend.

The incorporation of immersive technologies, enriched with gamification elements, in education can ineradicably increase students' curiosity and engagement by triggering their intrinsic motivation and interest in the studied subject. However, researchers and instructional designers who consider ER as a substitute supplementary learning resource need to place special attention on the alignment of gameful dimensions (e.g. narrative, theme, player actions, problems/puzzles, environment aesthetics) with the subject matter and the intended learning outcomes to optimise learners' enjoyment and cognitive achievement. Indicatively, a successful alignment is exhibited in Monnot et al. (2020) where chemistry laboratory procedures are integrated seamlessly into the escape room's narrative and puzzles around a historical figure's scientific contribution.

Another implication is the dawn of the democratisation of ER creation; the introduction of a new avenue for exploration where high visual fidelity of the 3D environment is no longer the prime driver. Instead, lower fidelity environments, with high task interactivity and intuitiveness can be equally – if not more – effective for learning. Thus, cognitive and affective gains can be attained with a lower effort and in a faster pace through WebXR (cloud-based Virtual/Augmented Reality platforms). For this direction, the exploration of metaphoric visual elements and simulated game mechanics is recommended. Metaphors encourage better visualisation and deeper comprehension in education (Mystakidis et al. 2021). In the present VRER this was implemented through representational analogies.

6. Limitations and future directions

As with every research work, the limits of the present study provide recommendations for future research and development. The first barrier concerns the experimental and exploratory nature of the study which prevented us from recruiting a large number of participants. This impacts the validity of the results outside the context in which the intervention took place (i.e. school environment, country). However, the described VRER is registered as open educational resource and thus, researchers and educators are highly recommended to explore its application and potential in multiple settings, with larger samples, and in different scientific contexts.

Moreover, research works related to the integration of alternative instructional solutions usually report the impact that the so-called 'novelty effect' may have had on the observed outcomes. During the execution phase, every possible action was taken to mitigate this effect. However, a variable that has not been accounted for concerns the long-term use of technological (learning) resources, as reiterated since the emergence of the COVID-19 pandemic. In view of this, future studies can also explore the impact that the recent, continuous, and systematic use of (educational) technology has on learners, prior to evaluating new educational technology platforms or tools.

The study employed a single data collection method (quantitative) wherein both objective (performance) and subjective (psychometrics) data were collected. Although this is the most common approach to evaluate such interventions (Checa and Bustillo 2020), future works can consider the adoption of mixed-methods research designs or the integration of Learning Analytics to facilitate the collection of objective data (such as digital traces) in a more consistent and systematic way. The complimentary integration of such methods can support further the development of clearer understanding, especially with regard to the instructional design and gameplay elements.

Finally, although not a limitation, another frontier for future research is the development of ER in standalone and interconnected social VR platforms (Metaverse) where multiple geographically dispersed learners will be able to communicate and operate simultaneously in the shared virtual environment.

Disclosure statement

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

ORCID

Athanasios Christopoulos http://orcid.org/0000-0002-1809-5525
Stylianos Mystakidis http://orcid.org/0000-0002-9162-8340
Mikko-Jussi Laakso http://orcid.org/0000-0001-9163-2676

References

Adita, A., A. R. E. Nugraheni, and N. Srisawasdi. 2020. "Trends of Instructional Research Using Biology Game: A Systematic Review of the Evidence During 2010-2019." In *Proceeding of the 28th International conference on Computers in Education*, 418–424.

All, A., E. P. N. Castellar, and J. Van Looy. 2016. "Assessing the Effectiveness of Digital Game-Based Learning: Best Practices." *Computers & Education* 92-93: 90–103. doi:10.1016/j.compedu.2015.10.007

Anastasiadis, M., G. Tsatsanifos, F. Terzakis, N. V. Karadimas, S. Sotiriou, A. Lazoudis, J. Koslowsky, et al. 2018. "A Storytelling Platform for Deeper Learning in STEM Combined with art-Related Activities." *International Journal of Education and Learning Systems* 3: 84–92.

Ang, J. W. J., Y. N. A. Ng, and R. S. Liew. 2020. "Physical and Digital Educational Escape Room for Teaching Chemical Bonding." *Journal of Chemical Education* 97 (9): 2849–2856. doi:10.1021/acs.jchemed.0c00612

Arnab, S., and S. Clarke. 2017. "Towards a Trans-Disciplinary Methodology for a Game-Based Intervention Development Process." *British Journal of Educational Technology* 48 (2): 279–312. doi:10.1111/bjet.12377

Arnab, S., T. Lim, M. B. Carvalho, F. Bellotti, S. De Freitas, S. Louchart, N. Suttie, R Berta, and A. De Gloria. 2015. "Mapping Learning and Game Mechanics for Serious Games Analysis." *British Journal of Educational Technology* 46 (2): 391–411. doi:10.1111/bjet.12113

Brame, C. J. 2016. "Effective Educational Videos: Principles and Guidelines for Maximizing Student Learning from Video Content." *CBE—Life Sciences Education* 15 (4): es6.

Bretz, S. L., and K. J. Linenberger. 2012. "Development of the Enzyme–Substrate Interactions Concept Inventory." *Biochemistry and Molecular Biology Education* 40 (4): 229–233. doi:10.1002/bmb.20622

Brown, G., E. Irving, and P. Keegan. 2008. *An Introduction to Educational Assessment, Measurement and Evaluation*. Auckland, NZ: Pearson Education.

Butler, D. 2018. *Utilising Second Life Machinima-Facilitated Narratives to Support Cognitive and Imaginative Engagement Across an Undergraduate Curriculum. Authentic Virtual World Education* (pp. 153–173). Singapore: Springer.

Checa-Romero, M., and I. Pascual Gómez. 2018. "Minecraft and Machinima in Action: Development of Creativity in the Classroom." *Technology, Pedagogy and Education* 27 (5): 625–637. doi:10.1080/1475939X.2018.1537933

Checa, D., and A. Bustillo. 2020. "A Review of Immersive Virtual Reality Serious Games to Enhance Learning and

Training." *Multimedia Tools and Applications* 79 (9): 5501–5527. doi:10.1007/s11042-019-08348-9
Chittaro, L., and F. Buttussi. 2015. "Assessing Knowledge Retention of an Immersive Serious Game vs. A Traditional Education Method in Aviation Safety." *IEEE Transactions on Visualization and Computer Graphics* 21 (4): 529–538. doi:10.1109/TVCG.2015.2391853
Christopoulos, A., and P. Sprangers. 2021. "Integration of Educational Technology During the Covid-19 Pandemic: An Analysis of Teacher and Student Receptions." *Cogent Education* 8 (1): 1964690. doi:10.1080/2331186X.2021.1964690
Clarke, S., D. J. Peel, S. Arnab, L. Morini, H. Keegan, and O. Wood. 2017. "escapED: A Framework for Creating Educational Escape Rooms and Interactive Games For Higher/Further Education." *International Journal of Serious Games* 4 (3): 73–86. doi:10.17083/ijsg.v4i3.180
Cohen, L., L. Manion, and K. Morrison. 2018. *Research Methods in Education*. [8th ed.]. New York: Routledge.
Cross, C. J. 2017. "An Analysis of Storytelling on Student Content Acquisition." *The American Biology Teacher* 79 (8): 628–634. doi:10.1525/abt.2017.79.8.628
Csikar, E., and J. E. Stefaniak. 2018. "The Utility of Storytelling Strategies in the Biology Classroom." *Contemporary Educational Technology* 9 (1): 42–60.
Efthimiou, O. 2016. "Breaking Disciplinary Boundaries: Heroism Studies and Contemporary Research Practices." *Limina: A Journal of Historical and Cultural Studies* 22 (1): 19–33.
Elford, D., S. J. Lancaster, and G. A. Jones. 2021. "Stereoisomers, not Stereo Enigmas: A Stereochemistry Escape Activity Incorporating Augmented and Immersive Virtual Reality." *Journal of Chemical Education* 98 (5): 1691–1704. doi:10.1021/acs.jchemed.0c01283
Gawlik-Kobylińska, M., W. Walkowiak, and P. Maciejewski. 2020. "Improvement of a Sustainable World Through the Application of Innovative Didactic Tools in Green Chemistry Teaching: A Review." *Journal of Chemical Education* 97 (4): 916–924. doi:10.1021/acs.jchemed.9b01038
Girard, C., J. Ecalle, and A. Magnan. 2013. "Serious Games as New Educational Tools: How Effective are They? A Meta-Analysis of Recent Studies." *Journal of Computer Assisted Learning* 29 (3): 207–219. doi:10.1111/j.1365-2729.2012.00489.x
Gligora Marković, M., D. Plantak Vukovac, and B. Kliček. 2015. "A Study on Knowledge Gain and Retention When Using Multimedia Learning Materials of Different Quality." *Journal of Information and Organizational Sciences* 39 (1): 33–48.
Harackiewicz, J. M., J. L. Smith, and S. J. Priniski. 2016. "Interest Matters: The Importance of Promoting Interest in Education." *Policy Insights from the Behavioral and Brain Sciences* 3 (2): 220–227. doi:10.1177/2372732216655542
Harwood, T. 2013. "Machinima as a Learning Tool." *Digital Creativity* 24 (3): 168–181. doi:10.1080/14626268.2013.813375
Huang, W., R. D. Roscoe, M. C. Johnson-Glenberg, and S. D. Craig. 2021. "Motivation, Engagement, and Performance Across Multiple Virtual Reality Sessions and Levels of Immersion." *Journal of Computer Assisted Learning* 37 (3): 745–758. doi:10.1111/jcal.12520
Jones, S. M., P. Katyal, X. Xie, M. P. Nicolas, E. M. Leung, D. M. Noland, and J. K. Montclare. 2019. "A 'KAHOOT!' Approach: The Effectiveness of Game-Based Learning for an Advanced Placement Biology Class." *Simulation & Gaming* 50 (6): 832–847. doi:10.1177/1046878119882048
Karakoyun, F., and IÜ Yapici. 2016. "Use of Digital Storytelling in Biology Teaching." *Universal Journal of Educational Research* 4 (4): 895–903. doi:10.13189/ujer.2016.040427
Lambert, J. 2010. *Digital Storytelling Cookbook and Traveling Companion*. Berkeley, CA: Digital Diner Press.
Lameras, P., S. Arnab, I. Dunwell, C. Stewart, S. Clarke, and P. Petridis. 2017. "Essential Features of Serious Games Design in Higher Education: Linking Learning Attributes to Game Mechanics." *British Journal of Educational Technology* 48 (4): 972–994. doi:10.1111/bjet.12467
Lathwesen, C., and N. Belova. 2021. "Escape Rooms in STEM Teaching and Learning—Prospective Field or Declining Trend?" *Education Sciences* 11 (6): 308. doi:10.3390/educsci11060308
Makransky, G., and L. Lilleholt. 2018. "A Structural Equation Modeling Investigation of the Emotional Value of Immersive Virtual Reality in Education." *Educational Technology Research and Development* 66 (5): 1141–1164. doi:10.1007/s11423-018-9581-2
Makransky, G., P. Wismer, and R. E. Mayer. 2019. "A Gender Matching Effect in Learning with Pedagogical Agents in an Immersive Virtual Reality Science Simulation." *Journal of Computer Assisted Learning* 35 (3): 349–358. doi:10.1111/jcal.12335
Makri, A., D. Vlachopoulos, and R. A. Martina. 2021. "Digital Escape Rooms as Innovative Pedagogical Tools in Education: A Systematic Literature Review." *Sustainability* 13 (8): 4587. doi:10.3390/su13084587
Mayer, R. E. 2002. *Multimedia Learning. Psychology of Learning and Motivation* (Vol. 41, pp. 85–139. Santa Barbara: Academic Press.
Mayer, R. E. 2014. "*Cognitive Theory of Multimedia Learning.*" In *The Cambridge Handbook of Multimedia Learning*, 2nd ed., edited by R. E. Mayer, 43–71. Cambridge: Cambridge University Press.
Monnot, M., S. Laborie, G. Hébrard, and N. Dietrich. 2020. "New Approaches to Adapt Escape Game Activities to Large Audience in Chemical Engineering: Numeric Supports and Students' Participation." *Education for Chemical Engineers* 32: 50–58. doi:10.1016/j.ece.2020.05.007
Mwelwa, S., and J. Soko. 2020. "Story Telling in Secondary Biology Classrooms for Enhanced Understanding: A Case of Kitwe District." *International Journal of Research and Scientific Innovation* 7 (2): 153–161.
Mystakidis, S. 2021. "Combat Tanking in Education: The TANC Model for Playful Distance Learning in Social Virtual Reality." *International Journal of Gaming and Computer-Mediated Simulations* 13 (4): 1–20. doi:10.4018/IJGCMS.291539
Mystakidis, S., A. Christopoulos, and N. Pellas. 2021. "A Systematic Mapping Review of Augmented Reality Applications to Support STEM Learning in Higher

Education." *Education and Information Technologies* 27: 1–45.

Mystakidis, S., G. Filippousis, D. Tolis, and E. Tseregkouni. 2021. "Playful Metaphors for Narrative-Driven E-Learning." *Applied Sciences* 11 (24): 11682. doi:10.3390/app112411682

Nicholson, S. 2018. "Creating Engaging Escape Rooms for the Classroom." *Childhood Education* 94 (1): 44–49. doi:10.1080/00094056.2018.1420363

Orcan, F. 2020. "Parametric or Non-Parametric: Skewness to Test Normality for Mean Comparison." *International Journal of Assessment Tools in Education* 7 (2): 255–265.

Otchie, W. O., M. Pedaste, E. Bardone, and I. A. Chounta. 2020. "Can YouTube Videos Facilitate Teaching and Learning of STEM Subjects in High Schools?" *Bulletin of the Technical Committee on Learning Technology* 20 (1): 3–8.

Oti, E. U., M. O. Olusola, and P. A. Esemokumo. 2021. "Statistical Analysis of the Median Test and the Mann-Whitney U Test." *International Journal of Advanced Academic Research* 7 (9): 44–51.

Pellas, N., A. Dengel, and A. Christopoulos. 2020. "A Scoping Review of Immersive Virtual Reality in STEM Education." *IEEE Transactions on Learning Technologies* 13 (4): 748–761. doi:10.1109/TLT.2020.3019405

Pellas, N., S. Mystakidis, and A. Christopoulos. 2021. "A Systematic Literature Review on the User Experience Design for Game-Based Interventions via 3D Virtual Worlds in K-12 Education." *Multimodal Technologies and Interaction* 5 (6): 28. doi:10.3390/mti5060028.

Pellas, N., S. Mystakidis, and I. Kazanidis. 2021. "Immersive Virtual Reality in K-12 and Higher Education: A Systematic Review of the Last Decade Scientific Literature." *Virtual Reality* 25 (3): 835–861. doi:10.1007/s10055-020-00489-9

Pereira-Santos, D., R. B. C. Prudêncio, and A. C. de Carvalho. 2019. "Empirical Investigation of Active Learning Strategies." *Neurocomputing* 326-327: 15–27. doi:10.1016/j.neucom.2017.05.105

Pirker, J., I. Lesjak, M. Parger, and C. Gütl. 2018. *An Educational Physics Laboratory in Mobile Versus Room Scale Virtual Reality-a Comparative Study. Online Engineering & Internet of Things* (pp. 1029–1043). Cham: Springer.

Puth, M. T., M. Neuhäuser, and G. D. Ruxton. 2015. "Effective use of Spearman's and Kendall's Correlation Coefficients for Association Between Two Measured Traits." *Animal Behaviour* 102: 77–84. doi:10.1016/j.anbehav.2015.01.010

Räsänen, P., J. Salminen, A. J. Wilson, P. Aunio, and S. Dehaene. 2009. "Computer-assisted Intervention for Children with Low Numeracy Skills." *Cognitive Development* 24 (4): 450–472. doi:10.1016/j.cogdev.2009.09.003

Ravyse, W. S., A. Seugnet Blignaut, V. Leendertz, and A. Woolner. 2017. "Success Factors for Serious Games to Enhance Learning: A Systematic Review." *Virtual Reality* 21 (1): 31–58. doi:10.1007/s10055-016-0298-4

Reeves, S. M., and K. J. Crippen. 2021. "Virtual Laboratories in Undergraduate Science and Engineering Courses: A Systematic Review, 2009–2019." *Journal of Science Education and Technology* 30 (1): 16–30. doi:10.1007/s10956-020-09866-0

Robin, B. R. 2016. "The Power of Digital Storytelling to Support Teaching and Learning." *Digital Education Review* 30 (30): 17–29.

Romero-Rodriguez, L. M., M. S. Ramirez-Montoya, and J. R. V. González. 2019. "Gamification in MOOCs: Engagement Application Test in Energy Sustainability Courses." *IEEE Access* 7: 32093–32101. doi:10.1109/ACCESS.2019.2903230

Sagri, M., F. Sofos, and D. Mouzaki. 2019. "Digital Storytelling, Comics and new Technologies in Education: Review, Research and Perspectives." *International Education Journal: Comparative Perspectives* 17 (4): 97–112.

Sanchez, M. E. 1992. "Effects of Questionnaire Design on the Quality of Survey Data." *Public Opinion Quarterly* 56 (2): 206–217. doi:10.1086/269311

Sawilowsky, S. S. 2009. "New Effect Size Rules of Thumb." *Journal of Modern Applied Statistical Methods* 8 (2): 26. doi:10.22237/jmasm/1257035100

Semilarski, H., and A. Laius. 2021. "Exploring Biological Literacy: A Systematic Literature Review of Biological Literacy." *European Journal of Educational Research* 10 (3): 1181–1197. doi:10.12973/eu-jer.10.3.1181

Shadish, W., T. Cook, and D. Campbell. 2002. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston, MA: Houghton Mifflin Company.

Thomas, M., and C. Schneider. 2018. "Language Teaching in 3D Virtual Worlds with Machinima: Reflecting on an Online Machinima Teacher Training Course." *International Journal of Computer-Assisted Language Learning and Teaching* 8 (2): 20–38. doi:10.4018/IJCALLT.2018040102

Tiba, C., J. Condy, A. Chigona, and N. Tunjera. 2015. "Digital Storytelling as a Tool for Teaching: Perceptions of Pre-Service Teachers." *TD: The Journal for Transdisciplinary Research in Southern Africa* 11 (1): 82–97.

Veldkamp, A., L. van de Grint, M. C. P. Knippels, and W. R. van Joolingen. 2020. "Escape Education: A Systematic Review on Escape Rooms in Education." *Educational Research Review* 31: 100364. doi:10.1016/j.edurev.2020.100364

Wu, J., and D. T. V. Chen. 2020. "A Systematic Review of Educational Digital Storytelling." *Computers & Education* 147: 103786. doi:10.1016/j.compedu.2019.103786

Xu, W., T. Ye, and X. Wang. 2021. "The Effectiveness of the Problem-Based Learning in Medical Cell Biology Education: A Systematic Meta-Analysis." *Medicine* 100 (39): 1–24.

Zhang, X., J. Q. Lu, and K. Park. 2018. "Interaction Design for Room Escape Virtual Reality Games." *TECHART: Journal of Arts and Imaging Science* 5 (4): 24–29. doi:10.15323/techart.2018.11.5.4.24

Appendices

Appendix 1. Demographics and background information.

1. Please identify your gender.	(a) Male(b) Female(c) Other(d) Prefer to not answer
2. Please specify your age group.	(a) 16 years old(b) 17 years old(c) Other, please specify
3. Please specify your school grade.	(a) Grade 10(b) Grade 11
4. Please determine your English level.	(a) Beginner(b) Intermediate(c) Advanced(d) Proficient
5. Please determine your familiarity with Video Learning Resources / Virtual Reality.	(a) Not at all familiar(b) Slightly familiar(c) Somewhat familiar(d) Moderately familiar(e) Extremely familiar
6. Please determine your interest toward Biology.	(a) Not at all interested(b) Slightly interested(c) Moderately interested(d) Very interested(e) Extremely interested

Appendix 2. Knowledge assessment quiz questions.

Q1. Which of the following best describes enzymes in the human body?	A. CarbohydrateB. ***Catalyst***C. LipidsD. Nucleic acid	Q6. The sucrase enzyme is added to water and sucrose. What is happening?	A. ***Sucrose decomposes e.g. glucose***B. Sucrose decomposes e.g. glucose and fructoseC. Sucrose decomposes into sucroseD. Nothing
Q2. Why is the activity of enzymes often described by the word pair lock & key?	A. Because the enzyme acts as a key in the cellB. Because the inhibitor acts as a key to the enzymeC. ***Because enzymes are specialised***D. Because the enzyme closes the cell membrane	Q7. How do enzymes accelerate chemical reactions?	A. The substances they produce degrade substances more quicklyB. ***They calculate the amount of energy needed for the reaction***C. They raise the temperature to a very high levelD. They produce a very acidic environment
Q3. Indicators are substances whose colour changes can be used to determine which substances are involved. What do you find if you add biuret solution?	A. StarchB. Nucleic acidsC. ***Proteins***D. Lipids	Q8. How can the biochemical role of enzymes best be explained?	A. They control metabolismB. They slow down the breakdown of substances in cellsC. ***They speed up metabolism***D. They carry substances in and out of the cell
Q4. To which category do enzymes belong?	A. CarbohydratesB. LipidsC. Nucleic acids***D. Proteins***	Q9. What affect does extremely high pH have on enzymes?	A. Destroys themB. Kills themC. ***Denatures them***D. Enhances them
Q5. Which enzyme deficiency causes lactose intolerance?	A. LactaseB. LipaseC. ***Lactose***D. Liposis	Q10. Which of these statements about enzymes is true?	A. They work best at low pHB. The work best at high pHC. Temperature does not affect their activity***D. They can be reused***

Appendix 3. Learning experience psychometric survey (adjusted).

Group	Machinima	Virtual Reality
C1. Perceived Enjoyment		
S1	I find watching educational videos enjoyable.	I find using Virtual Reality educational applications enjoyable.
S2	Watching educational videos is pleasant.	Using Virtual Reality educational applications is pleasant.
S3	Watching educational videos is fun.	I have fun using Virtual Reality educational applications.
C2. Motivation		
S4	I enjoy learning from educational videos very much.	I enjoy interacting with Virtual Reality educational applications very much.
S5	Educational videos are fun to watch.	Virtual Reality educational activities are fun to do.
S6	This educational video was boring.	This Virtual Reality escape room was boring.
S7	This educational video did not hold my attention at all.	This Virtual Reality escape room did not hold my attention at all.
S8	I would describe this educational video as useful source of information.	I would describe this Virtual Reality escape room as useful source of information.
S9	I thought that the educational video was quite enjoyable.	I thought that the Virtual Reality escape room was quite enjoyable.
S10	While I was watching the educational video, I was thinking about how much I enjoyed it.	While I was playing the Virtual Reality escape room, I was thinking about how much I enjoyed it.
C3. Cognitive Benefits		
S11	Educational videos make the comprehension easier.	Virtual Reality educational applications make the comprehension easier.
S12	Educational videos make the memorisation easier.	Virtual Reality educational applications make the memorisation easier.
S13	Educational videos help me to better apply what was learned.	Virtual Reality educational applications help me to better apply what was learned.
S14	Educational videos help me to better analyze the problems.	Virtual Reality educational applications help me to better analyse the problems.
C4. Perceived Learning		
S15	I was more interested to learn the topics.	I was more interested to learn the topics.
S16	I learned a lot of factual information in the topics.	I learned a lot of factual information in the topics.
S17	I gained a good understanding of the basic concepts of the enzymes/ catalysts.	I gained a good understanding of the basic concepts of the enzymes/ catalysts.
S18	I learned to identify the main and important issues of the topics.	I learned to identify the main and important issues of the topics.
S19	I was interested and stimulated to learn more.	I was interested and stimulated to learn more.
S20	I was able to summarise and concluded what I learned.	I was able to summarise and concluded what I learned.

(Continued)

Appendix 3. Continued.

Group	Machinima	Virtual Reality
S21	The learning process was meaningful.	The learning activities were meaningful.
S22	What I learned, I can apply in real context.	What I learned, I can apply in real context.
C5. Satisfaction		
S23	I was satisfied with this type of learning experience.	I was satisfied with this type of learning experience.
S24	A wide variety of learning materials was provided in this educational video.	A wide variety of learning materials was provided in this type of Virtual Reality escape room.
*(R)*S25	I don't think this educational video would benefit my learning achievement.	I don't think this type of Virtual Reality escape room would benefit my learning achievement.
S26	I was satisfied with the immediate information gained via this educational video.	I was satisfied with the immediate information gained in this type of Virtual Reality escape room.
S27	I was satisfied with this teaching method.	I was satisfied with this teaching method.
S28	I was satisfied with this type of educational activity.	I was satisfied with this type of educational activity.
S29	I was satisfied with the overall learning effectiveness.	I was satisfied with the overall learning effectiveness.