RSE2R

Évaluation de la charge cognitive et de la performance dans la réalité virtuelle immersive et non immersive : Une étude croisée dans le domaine de l'éducation à la santé

Examining Cognitive Load and Performance in Immersive vs. Non-Immersive Virtual Reality: A Cross-Over Study in Health Education

Léa LONJOU

PT

ILFOMER, Institut Limousin de Formation aux Métiers de la Réadaptation, Limoges, France

Anaïs C. AUGRAS

PT

ILFOMER, Institut Limousin de Formation aux Métiers de la Réadaptation, Limoges, France

Nathan GROSBOILLOT

PhDS

HAVAE UR 20217, Limoges University, 123 Avenue Albert Thomas, Limoges, France

Anaïck PERROCHON

Prof

anaick.perrochon@unilim.fr

ILFOMER, Institut Limousin de Formation aux Métiers de la Réadaptation, Limoges, France

HAVAE UR 20217, Limoges University, 123 Avenue Albert Thomas, Limoges, France

Abstract: Background: Virtual Reality (VR) is utilized in health simulations as a method for presenting clinical cases. VR experiences offer numerous advantages such as interactivity and a high level of immersion, which enhance performance compared to conventional teaching methods. The extent of immersive VR's impact on cognitive load remains insufficiently investigated. This experimental cross-over study aimed (a) to assess students' cognitive load, (b) to evaluate the usability, intrinsic motivation, and cybersickness of the system, and (c) to compare students' performance in resolving two clinical cases between immersive VR and non-immersive conditions.

Method: Twenty students were included in this study. We developed two physiotherapy clinical cases (musculoskeletal and respiratory) as 360° videos. The clinical cases were randomized between exposure conditions: immersive VR using a head-mounted display (HMD) and non-immersive VR using a laptop. Performance was evaluated through multiple-choice questions, cognitive load was measured using functional near-infrared spectroscopy, and usability, intrinsic motivation, and cybersickness were assessed using the System Usability Scale, Intrinsic Motivation Inventory, and Simulator Sickness Questionnaire, respectively.

Results: There was no significant difference between the scores obtained with the HMD and the laptop ($p = 0.245$). Results indicated a higher activation of the prefrontal cortex with the laptop condition ($p = 0.007$). Usability was significantly better ($p = 0.005$), and the number of reported side effects was lower for the computerized condition, whereas intrinsic motivation was similar.

Conclusion: Immersive VR led to a lower cognitive load compared to non-immersive VR. Despite similar

performance between the two exposure conditions, usability was superior, and side effects were fewer for the computerized condition.

Keywords: Virtual reality simulation, Cognitive load, Immersive VR, Functional neuroimagery

Résumé : Contexte : La réalité virtuelle (RV) est utilisée dans la simulation médicale pour présenter des cas cliniques. Les expériences en RV offrent de nombreux avantages tels que l'interactivité et un haut niveau d'immersion, mais l'impact de la RV immersive sur la charge cognitive reste insuffisamment exploré. Cette étude expérimentale visait (a) à mesurer la charge cognitive des étudiants, (b) à évaluer la facilité d'utilisation, la motivation intrinsèque et le cybersickness du système, et (c) à comparer les performances des étudiants dans la résolution de deux cas cliniques entre les conditions de RV immersive et non immersive.

Méthode : Vingt étudiants ont été inclus dans cette étude. Nous avons développé deux cas cliniques de kinésithérapie (musculo-squelettique et respiratoire) sous forme de vidéos 360 degrés. Les cas cliniques ont été répartis de manière aléatoire entre RV immersive à l'aide d'un visocasque et RV non immersive à l'aide d'un ordinateur portable. Les performances ont été évaluées à l'aide d'un QCM, la charge cognitive a été mesurée à l'aide de la spectroscopie proche infrarouge fonctionnelle, et la facilité d'utilisation, la motivation intrinsèque et le cybermalaise ont été évalués respectivement à l'aide de la System Usability Scale, de l'Intrinsic Motivation Inventory, et du Simulator Sickness Questionnaire.

Résultats : Aucune différence significative n'a été observée entre les scores obtenus avec le visocasque et l'ordinateur portable (p = 0,245). Les résultats ont indiqué une activation plus élevée du cortex préfrontal avec la condition de l'ordinateur portable (p = 0,007). La facilité d'utilisation était significativement meilleure (p = 0,005), et les effets secondaires rapportés était plus faible pour la condition informatique, tandis que la motivation intrinsèque était similaire.

Conclusion : La RV immersive a entraîné une charge cognitive inférieure par rapport à la RV non immersive. Malgré des performances similaires entre les deux conditions d'exposition, la facilité d'utilisation était supérieure et les effets secondaires étaient moins nombreux pour la condition sur ordinateur.

Mots clés : Simulation virtuelle, Charge cognitive, Réalité virtuelle, Neuroimagerie fonctionnelle

Auteur correspondant
Anaïck Perrochon
UR HAVAE 20217
123 Av Albert Thomas
87060 LIMOGES
FRANCE
Email: anaick.perrochon@unilim.fr

Introduction

Case-Based Learning (CBL) serves as a pivotal method in healthcare education, using genuine clinical cases within a controlled learning environment to prepare students for real-world clinical practice (1). It allows students to apply theoretical knowledge to tangible clinical scenarios. CBL exists in various forms, from text-based to simulation with acting or virtual patients. In this regard, virtual reality (VR) is an effective tool for medical and paramedical education, enabling the simulation of specific situations in virtual environments (2). VR facilitates learning from real-life scenarios, allowing practice, error learning, and skill repetition without endangering patients (3). In this sense, the design of VR scenarios can focus on specific skills such as communication, decision-making, critical thinking, or clinical reasoning skills (4).

Environments within VR can derive from real-life sources (videos and photos) or artificial content typically generated by computer (5). VR environment encompasses different levels of immersion, from non-immersive experiences introduced on a 2D screen, such as a laptop, to immersive experiences, integrated with a Head-Mounted Display (HMD). The degree of immersion varies based on the device used, defining the subjective sensation of "being in the environment" (6). Interactivity with the environment might also fluctuates; for instance, in a 360° video, students have the ability to explore but not interact within the environment (7). Medical fields like surgery employ interactive content, allowing students to actively influence the ongoing task and simulate real-life decision-making (2).

These different aspects of VR significantly impact learners' motivation and the cognitive load (CL) associated with the learning process (8,9). In essence, learning requires the assimilation of new information that the working memory (WM) processes to effectively store it in long-term memory (10). However, WM has limited capacities for temporary storage and manipulation of information (10,11), making it susceptible to the CL of the ongoing task. CL is categorized into intrinsic, extraneous, and germane loads, according to the CL theory (12,13). Intrinsic load involves task complexity modulated by user knowledge, while extraneous load relates to task presentation and procedure execution. Germane load encompasses WM for managing the learning process. Therefore, it is crucial to identify and assess these factors by considering the devices and levels of immersion in a learning task, to prevent WM from overloading and optimize CBL in VR scenarios.

CL assessment methods include indirect, subjective, efficiency measures, or secondary task assessments (12). Frederiksen and colleagues evaluated CL in laparoscopic surgery simulation using immersive and non-immersive VR via secondary task reaction time, noting higher CL and lower performance in immersive VR (14). Another study from Chao and colleagues compared CL of medical students during history-taking and physical examination, finding higher intrinsic load in 360° video but superior procedural skill evaluation results (15).

Additionally, CL can be assessed physiologically by monitoring variations in brain activity through neuroimaging techniques such as functional Magnetic Resonance Imaging (fMRI), Electroencephalography (EEG), or functional Near-Infrared Spectroscopy (fNIRS). Specifically, the prefrontal cortex (PFC) plays a crucial role in WM and task response (16). Studies using fNIRS reported increased PFC activation with heightened WM task difficulty (17,18). Aksoy and colleagues evaluated PFC activation during a healthcare training course with VR, finding decreased fNIRS signals with increased familiarity and practice, potentially indicating decreased CL (19). However, no studies comparing CL across different VR devices using fNIRS have been reported.

Therefore, this experimental cross-over study aimed to (a) assess students' CL using fNIRS and (b) evaluate performance during two clinical cases in immersive versus non-immersive VR conditions. We hypothesized that PFC activation would be higher in immersive VR irrespective of clinical case; and that performance would be superior in immersive VR. Secondary objectives were to assess VR experience through intrinsic motivation, usability and cybersickness. We hypothesized that immersive VR would show higher motivation and usability, but more cybersickness than non-immersive environment.

Method

Population

This original research took place between November 25 and December 7, 2022, at the ILFOMER (Institut Limousin de Formation aux Métiers de la Réadaptation) in Limoges, France. The study was presented as a mandatory class for 5ᵗʰ year physiotherapist (PT) students from ILFOMER, with voluntary participation in the study. The inclusion criteria were being a 5ᵗʰ PT student, having a normal or corrected to normal vision, having a normal or corrected to normal hearing and agreeing to participate in this study. The exclusions criteria were having a neurological or cerebral pathology, having known effects on VR, having a contraindication to VR, putting the VR device down during the exercise, leaving the clinical case interface, and not finishing the clinical case. Before the start of the study, informed consent was obtained from all participants. They were informed of their right to refuse participation and stop the simulation at any time if they had to.

All participants begin by VR and continued with laptop. Then, they were assigned to one of the two groups, one group did the musculoskeletal clinical case with the HMD first and then the respiratory, and second group does the inverse.

Intervention

The intervention timeline is visually represented in Figure 1. Participants underwent a 15-minute briefing, which included instructions on how to navigate the online 360° platform, an overview of the clinical cases' objectives, guidance on what to anticipate, and instructions on responding to multiple-choice questions (MCQs). Additionally, we provided a comprehensive presentation on the fNIRS device, covering its installation, utility, and characteristics

Subsequently, we proceeded to install both the fNIRS device and the Head-Mounted Display (HMD), as depicted in Figure 2. During the first clinical case, participants were positioned at the center of a fictitious circle with a diameter of 2 meters, ensuring an obstacle-free environment. Upon completion of the HMD case, the device was removed, and participants were directed to complete an online questionnaire on a laptop, focusing on their experience with the simulation device

For the second clinical case, participants stood in front of a stand with the laptop set at eye level. After concluding the case, we removed the fNIRS device and instructed participants to respond to the online questionnaire for the second time.

Prior to each condition (HMD and laptop), a 25-second baseline recording was conducted, during which participants were instructed to stand quietly without any movement.
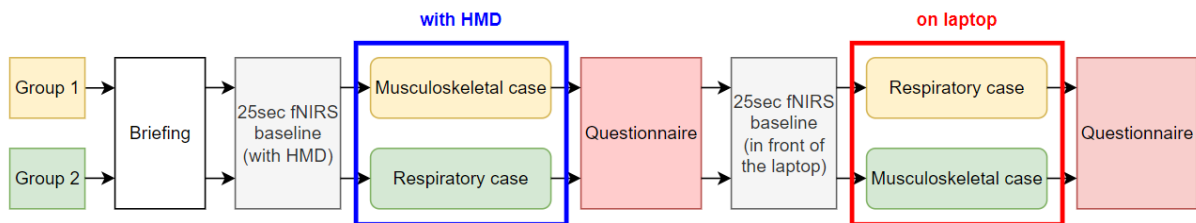
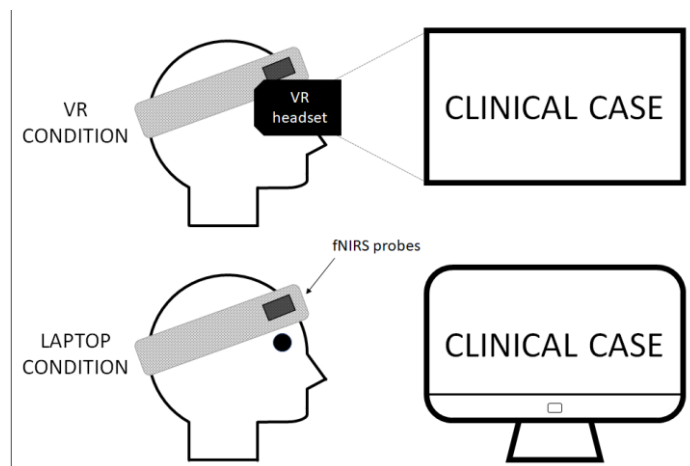Figure 1: The intervention timeline

Figure 2: Immersive VR and non-immersive VR device installation

Clinical case production

We developed two clinical cases in two physiotherapy fields: respiratory (initial assessment of a patient with chronic obstructive pulmonary disease (COPD)) and musculoskeletal (initial consultation with a patient with a knee sprain). Both cases adhered to the French recommendations for physiotherapy assessments (Ministry of Social Affairs, Health, and Women's Rights 2015) and the International Classification of Functioning, Disability and Health (World Health Organization 2001). The scenarios were designed based on specialized literature (20,21). For the respiratory case, we enlisted a patient from Limoges University Hospital to participate in the scenes, while for the musculoskeletal case, we engaged an actor. All actors provided signed authorization for image rights. Both scenarios underwent review by registered French physiotherapists.

The clinical cases included 360° videos and photos captured with an Insta360 Pro (Insta360, Shenzhen, China), 2D videos and photos recorded with an iPhone 12 mini (Apple, Cupertino, USA), and audio recordings using a Tascam DR-40X microphone and an EW 100 G4-ME2-B wireless set (Sennheiser electronic GmbH & Co. KG, Wedemark, Germany). Insta360Stitcher version 3.0.0 (Insta360, Shenzhen, China) was used for the stitching of 360° media. The footage underwent editing in Adobe Premiere Pro (Adobe Inc., San Jose, USA) and was transformed into an interactive virtual environment using Uptale (Uptale, Paris, France). Additional documents, such as prescriptions and further tests, were displayed.

In the clinical cases, participants assumed the role of physiotherapy trainees observing the interactions between the physiotherapist and the patient. Each clinical case comprised four segments: i) patient's file, ii) anamnesis, iii) clinical and functional tests, and iv) multiple-choice questions (MCQ) (Figure 3). These MCQ focused on the patient's characteristics and the participant's reasoning regarding the International Classification of Functioning (ICF) and the treatment plan. Each case had a duration ranging from 20 to 30 minutes.

The VR modules were presented on two supports: a HMD Meta Quest 2 (Menlo Park, California, USA) with its two controllers and an HP laptop (Palo Alto, California, USA) with a wired mouse.

Figure 3: Musculoskeletal clinical case. A: clinical test and B: MCQs

Assessment

Performance and cognitive load

To assess students' performance, we developed a scale to mark the MCQ's answers, which resulted in scores out of 20. Each question was out of 1 point, with 0.25 points awarded for each item correctly checked.

To assess the CL, we used the PortaLite fNIRS system (Artinis Medical Systems, Elst, the Netherlands) to measure changes in HbO2 and HHb concentrations. The PortaLite uses near infrared light, transmitted at 2 wavelengths, 760 and 850 nm. Data were sampled with a frequency of 10 Hz. One probe was positioned on the left side of the forehead the other one on the right side, above the VR HMD, to measure PFC activity. The probes were held with a headband and shielded from ambient light by a black patch covering the forehead. Oxysoft version 3.0.97.1 (Artinis Medical Systems, Elst, the Netherlands) was used for data collection. Concentration changes of HbO2 and HHb in the PFC were calculated from the changes in detected light intensity using the modified Lambert-Beer law, assuming constant scattering. The PortaLite probes are both composed of 3 transmitters and 1 receiver, with transmitter-receiver distances of 30, 35, and 40 mm. The concentrations of HbO2 and HHb were exported and processed in HOMER3 (MATLAB and Statistics Toolbox Release 2012b, The MathWorks, Inc, Natick, MA). Because HbO2 has been reported to be a more accurate indicator of cortical activity, we only used statistical analysis of HbO2 (22). Regarding the signal processing, the same procedure as the one applied in Maidan and colleagues (2016) was used (23). A bandpass filter with frequencies of 0.01 to 0.14 Hz was used to reduce physiological noise such as heartbeat and drift of the signal. A wavelet filter was used to remove motion artefacts, followed by correlation-based signal improvement (CBSI). HbO2

concentration signals of the three channels of each probe were averaged, resulting in an HbO2 signal for the left PFC and one for the right. The two baselines (HMD and laptop) consisted of 25 seconds of standing, with participants instructed to stand quietly, with no movement. The last 10 seconds of the baseline, just before the task, were averaged and referred to as the baseline's concentration. Time markers were put at the beginning of each part of the clinical case (patient's file, anamnesis, clinical and functional tests and MCQ). Since task duration differed between each participant and for each task in both conditions, average HbO2 concentrations were made from the minimal common time of activity for each part, regardless of condition or case. The minimum common time for each part was: 23s for the patient's file, 86s for the anamnesis, 199s for the tests and 154s for the MCQs. In addition, the first 6s of activity of each task were excluded to take into account the hemodynamics response delay. Each baseline concentration was subtracted from the average concentration during the task performance, resulting in the relative activity during each specific task.

VR experience

To assess participants' intrinsic motivation for each clinical case, we translated in French the subscale Interest/Pleasure from the Intrinsic Motivation Inventory by Ryan (24) validated in English by (25). We used the 6 first items of this subscale. They are marked on a 7-point Likert scale from "1 strongly disagree" to "7 strongly agree".

System's usability was assessed with the validated French System Usability Scale (SUS) (26), consisting of 10 items marked on a 5-point Likert scale from 1 "strongly disagree" to 5 "strongly agree". We also translated the Single Item Adjective Rating of Usability scale from Bangor et al (27), consisting of 7 adjectives from "the worst imaginable" to "the best imaginable".

We used the Simulator Sickness Questionnaire (SSQ) items to assess the symptoms felt after each clinical case (28,29). For each item, participants answered "yes" or "no" depending on whether they felt the symptom or not.

Statistical analysis

The data were analyzed using, Jasp version 0.17.1.0. We checked the normality distribution by conducting Shapiro-Wilk tests and checked variance homogeneity with Levene's tests. To analyze performance, we used an two-way ANOVA with the condition (HMD and laptop) and clinical cases (musculoskeletal and respiratory) as fixed factors. A Shapiro-Wilk test confirmed the non-normality of our fNIRS data. A Kruskal-Wallis test was then conducted to compare the differences in HbO2 levels between conditions (HMD and laptop) and clinical cases (musculoskeletal and respiratory). To analyze IMI, SUS and SSQ scores we used paired sample t-tests between the HMD and laptop data. The significance level was set up at a p value <0.05.

Results

A total of 20 participants were recruited, of which 10 were males (50%) and 10 were females (50%). The median age was 24 years (IQR 23-24).

Performance and cognitive load

There is no significant difference between the scores obtained with the HMD and the laptop (HMD: 15.6 points; laptop: 16.1 points. p= 0.245).

There was a significant effect of condition (p = 0.007) with a higher PFC activation in the laptop condition compared to the HMD condition (laptop condition mean HbO2 concentration: -22.64 ±1.10 μmol.L^{-1}; HMD condition mean HbO2 concentration: -49.12±7.13 μmol.L^{-1}). There was no effect of conditions on HbO2 levels according to different parts (see Table 1).

Table 1: Mean HbO2 concentration (μmol.L^{-1}) by task and by condition and p values

	HMD	Laptop	p value
Patient's file	-23.10±1.15	-14.40±1.42	0.488
Anamnesis	-69.00±1.03	-33.70±1.10	0.113
Clinical and functional test	-63.10±9.94	-36.30±1.24	0.279
Multiple choice questions	-41.30±7.61	-62.30±1.18	0.107
Total	-49.12±7.13	-22.64±1.10	0.007

VR experience

The IMI, SUS and SSQ outcomes data are presented in Table 2. For the IMI, there was no significant difference between the scores with 6.30±0.68 for the HMD condition and 6.05 ±1.03 for the laptop one (t-test p= 0.258).

For the SUS, the laptop score (91.71±8.08) was a significantly higher (p=0.05) than the HMD score (81.28±9.70). For the laptop condition it was "Ok": 2 participants, "Good": 10 participants and "Excellent": 8 participants. For the HMD condition, the adjectives chosen on the Single item Adjective Rating of Usability were "Ok": 3 participants, "Good": 11 participants and "Excellent": 6 participants.

Concerning the SSQ, all the symptoms were felt more during the HMD condition than the laptop condition, except difficulty concentrating (HMD: 11; laptop: 11), stomach awareness (HMD: 1; laptop: 1) and burping (HMD: 1; laptop: 1) that were equally felt in both conditions. With the HMD condition,

more than half of the participants felt general discomfort (13), fatigue (12), eyestrain (19), difficulty focusing (11), difficulty concentrating (11), fullness of head (15) and blurred vision (17). With the HMD condition, all participants felt at least 2 symptoms. With the laptop one, the most felt symptoms were fatigue (9), eyestrain (12) and difficulty concentrating (11), and 4 participants felt no symptoms.

Table 2: IMI, SUS and SSQ results

	HMD	LAPTOP	p-value
IMI Mean score /7 (SD) n = 20	6.30±0.68	6.05±1.02	0.258
SUS Mean score /100 (SD) n = 20	81.28±9.70	91.71±8.1	<0.05
SSQ nb (%) n = 20			
General discomfort	13 (65)	5 (25)	/
Fatigue	12 (60)	9 (45)	/
Headache	9 (45)	6 (30)	/
Eyestrain	19 (95)	12 (60)	/
Difficulty focusing	11 (55)	5 (25)	/
Increased salivation	2 (10)	1 (5)	/
Sweating	7 (35)	3 (15)	/
Nausea	4 (20)	0 (0)	/
Difficulty concentrating	11 (55)	11 (55)	/
Fullness of head	15 (75)	6 (30)	/
Blurred vision	17 (85)	3 (1)	/
Dizzy (eyes open)	6 (30)	1 (5)	/
Dizzy (eyes closed)	1 (5)	0 (0)	/
Vertigo	4 (20)	1 (5)	/
Stomach awareness	1 (5)	1 (5)	/
Burping	1 (5)	1 (5)	/

Legend: IMI: Intrinsic Motivation Inventory; SUS: System Usability Scale; SSQ: Simulator Sickness Questionnaire; nb: number

Discussion

Our results revealed higher PFC activity in non-immersive VR on a 2D screen compared to immersive VR through an HMD. Participants exhibited similar performances both setups.

Interestingly, our results deviate from the current literature, which support a higher CL in immersive VR conditions. A potential explanation for this difference originates from the content displayed.

Frederiksen and colleagues used two distinct VR conditions with substantial differences: a simulated surgery on a 2D screen (non-immersive VR) for the control group and the same procedure on in a fully-presented operating room environment in a HMD (immersive VR) for the intervention group (14). The additional environment available only in the immersive VR condition increased the quantity of information processed by the participants, potentially leading to heightened CL and worse performances. Likewise, Chao et al. used HMD in both conditions but with different fields of view (15). Specifically, they presented an operating room with a 360° view in one condition and a 2D video limited to a 120° view from another perspective in the other condition. The increased quantity of information in the 360° condition might explain the related higher PFC activity. In contrast, our study presented identical 360° video with the same quantity of information in both the immersive and the non-immersive conditions. The consistency of the intrinsic load could explain the absence of difference in performance for our participants between the two conditions. Regarding PFC activity, the use of HMDs facilitated natural and effortless spatial exploration of the environment through head movements, while spatial navigation on a laptop requires using a computer mouse. This difference in access to information might have contributed to reducing the extraneous load in immersive VR, thus decreasing the overall CL and PFC activity.

Concerning VR experience, no significant difference of intrinsic motivation was observed between the conditions. While the level of immersion varied, 360° videos did not allow for direct interaction with the environment in both setups (7). Students could not ask the patients additional questions, gather more information, conduct additional clinical tests, or perform them differently. This lack of interaction may explain the similar intrinsic motivation, irrespective of the device used. Regarding the usability, both HMD and laptop SUS scores demonstrated acceptable usability based on the scale proposed by Bangor and colleagues (27). Novelty of an HMD device for participants could explain the lower score in contrast to the familiarity of a laptop. Regarding the SSQ, the number of participants presenting each symptom was higher with the HMD than with the laptop. Cybersickness can result in discomfort, fatigue, headache, nausea, disorientation, difficulty concentrating and blurred vision (6). While this difference did not affect performance scores, futures studies should assess the impact of cybersickness on usability and motivation, as well as the impact of exposition on cybersickness, with multiple training sessions.

This study presents some limitations. Participants were all recruited from the same PT institute and had followed the same training. It would be interesting to assess CL with a similar multicentric study, increasing the number of participants and the background from several PT institutes. Regarding brain activity, while the current study used filtering techniques to minimize motion artefacts and physiological noise, future studies should use fNIRS devices with short-separation channels to reduce artefacts more efficiently.

Conclusion

In our study we found that CL was higher during the clinical case in non-immersive VR compared to the immersive one. A possible explanation could be that with a HMD information is more easily accessible than on a computer.

The device used to introduce the clinical case had no impact on students' performance. This may suggest that in terms of performance, the HMD may be as relevant a choice of medium as the computer.

References

1. Thistlethwaite JE, Davies D, Ekeocha S, Kidd JM, MacDougall C, Matthews P, et al. The effectiveness of case-based learning in health professional education. A BEME systematic review: BEME Guide No. 23. Med Teach. 2012;34(6):e421-444.

2. Barteit S, Lanfermann L, Bärnighausen T, Neuhann F, Beiersmann C. Augmented, Mixed, and Virtual Reality-Based Head-Mounted Devices for Medical Education: Systematic Review. JMIR Serious Games. 2021 Jul 8;9(3):e29080.

3. Pottle J. Virtual reality and the transformation of medical education. Future Healthc J. 2019 Oct;6(3):181–5.

4. Bonnin C, Pejoan D, Ranvial E, Marchat M, Andrieux N, Fourcade L, et al. Immersive virtual patient simulation compared with traditional education for clinical reasoning: a pilot randomised controlled study. J Vis Commun Med. 2023;46(2):66–74.

5. Saxena N, Kyaw BM, Vseteckova J, Dev P, Paul P, Lim KTK, et al. Virtual reality environments for health professional education. Cochrane Database Syst Rev. 2018 Oct 2;2018(10):CD012090.

6. Rebenitsch L, Owen C. Review on cybersickness in applications and visual displays. Virtual Real. 2016 Jun;20(2):101–25.

7. Roche L, Kittel A, Cunningham I, Rolland C. 360° Video Integration in Teacher Education: A SWOT Analysis. Front Educ. 2021 Nov 1;6:761176.

8. Blair C, Walsh C, Best P. Immersive 360° videos in health and social care education: a scoping review. BMC Med Educ. 2021 Nov 24;21(1):590.

9. Kyaw BM, Saxena N, Posadzki P, Vseteckova J, Nikolaou CK, George PP, et al. Virtual Reality for Health Professions Education: Systematic Review and Meta-Analysis by the Digital Health Education Collaboration. J Med Internet Res. 2019 Jan 22;21(1):e12959.

10. Baddeley AD, Hitch G. Working Memory. In: Bower GH, editor. Psychology of Learning and Motivation. Academic Press; 1974. p. 47–89.

11. León-Domínguez U, Martín-Rodríguez JF, León-Carrión J. Executive n-back tasks for the neuropsychological assessment of working memory. Behav Brain Res. 2015 Oct 1;292:167–73.

12. Sweller J, Ayres P, Kalyuga S. Intrinsic and Extraneous Cognitive Load. In: Sweller J, Ayres P, Kalyuga S, editors. Cognitive Load Theory. New York, NY: Springer; 2011. p. 57–69. (Explorations in the Learning Sciences, Instructional Systems and Performance Technologies).

13. Van Merriënboer JJG, Sweller J. Cognitive load theory in health professional education: design principles and strategies. Med Educ. 2010;44(1):85–93.

14. Frederiksen JG, Sørensen SMD, Konge L, Svendsen MBS, Nobel-Jørgensen M, Bjerrum F, et al. Cognitive load and performance in immersive virtual reality versus conventional virtual reality

simulation training of laparoscopic surgery: a randomized trial. Surg Endosc. 2020 Mar 1;34(3):1244–52.

15. Chao YP, Kang CJ, Chuang HH, Hsieh MJ, Chang YC, Kuo TBJ, et al. Comparison of the effect of 360° versus two-dimensional virtual reality video on history taking and physical examination skills learning among undergraduate medical students: a randomized controlled trial. Virtual Real [Internet]. 2022 Aug 16 [cited 2023 Feb 1]; Available from: https://link.springer.com/10.1007/s10055-022-00664-0

16. Miller EK, Cohen JD. An integrative theory of prefrontal cortex function. Annu Rev Neurosci. 2001;24:167–202.

17. Fishburn FA, Norr ME, Medvedev AV, Vaidya CJ. Sensitivity of fNIRS to cognitive state and load. Front Hum Neurosci [Internet]. 2014 [cited 2023 Feb 7];8. Available from: http://journal.frontiersin.org/article/10.3389/fnhum.2014.00076/abstract

18. Molteni E, Contini D, Caffini M, Baselli G, Spinelli L, Cubeddu R, et al. Load-dependent brain activation assessed by time-domain functional near-infrared spectroscopy during a working memory task with graded levels of difficulty. J Biomed Opt. 2012 May;17(5):056005.

19. Aksoy E, Izzetoglu K, Baysoy E, Agrali A, Kitapcioglu D, Onaral B. Performance Monitoring via Functional Near Infrared Spectroscopy for Virtual Reality Based Basic Life Support Training. Front Neurosci. 2019;13.

20. Antonello M, Delplanque D. Comprendre la kinésithérapie respiratoire. Elsevier Masson. 2009.

21. Dufour M, Tixa S, Acedo del Valle S. Démarche clinique et diagnostic en kinésithérapie. Elsevier Masson. 2018.

22. Strangman N, Hall T, Meyer A, National Center on Accessing the General Curriculum. Virtual Reality/Computer Simulations and the Implications for UDL mplementation. 2003.

23. Maidan I, Nieuwhof F, Bernad-Elazari H, Reelick MF, Bloem BR, Giladi N, et al. The Role of the Frontal Lobe in Complex Walking Among Patients With Parkinson's Disease and Healthy Older Adults: An fNIRS Study. Neurorehabil Neural Repair. 2016 Nov;30(10):963–71.

24. Ryan RM. Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory. J Pers Soc Psychol. 1982 Sep;43(3):450–61.

25. McAuley E, Duncan T, Tammen VV. Psychometric Properties of the Intrinsic Motivation Inventory in a Competitive Sport Setting: A Confirmatory Factor Analysis. Res Q Exerc Sport. 1989 Mar;60(1):48–58.

26. Gao M, Kortum P, Oswald FL. Multi-Language Toolkit for the System Usability Scale. Int J Human–Computer Interact. 2020 Dec 13;36(20):1883–901.

27. Bangor A, Kortum PT, Miller JT. An Empirical Evaluation of the System Usability Scale. Int J Hum-Comput Interact. 2008 Jul 29;24(6):574–94.

28. Bouchard S, Renaud P. Revising the factor structure of the simulator sickness questionnaire. Annual Review of CyberTherapy and Telemedicine. 2007 Jan 1;5:128–37.

29. Kennedy RS, Lane NE, Berbaum KS, Lilienthal MG. Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness. Int J Aviat Psychol. 1993 Jul;3(3):203–20.