



**HAL**  
open science

# Using virtual reality for enhancing neuroanatomy learning by optimizing cognitive load and intrinsic motivation

Matisse Poupard, Florian Larrue, Martin Bertrand, Dominique Liguoro,  
André Tricot, Hélène Sauzéon

## ► To cite this version:

Matisse Poupard, Florian Larrue, Martin Bertrand, Dominique Liguoro, André Tricot, et al.. Using virtual reality for enhancing neuroanatomy learning by optimizing cognitive load and intrinsic motivation. 2024. hal-04814452

**HAL Id: hal-04814452**

**<https://hal.science/hal-04814452v1>**

Preprint submitted on 2 Dec 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Using virtual reality for enhancing neuroanatomy learning by optimizing cognitive load and intrinsic motivation

M. Poupard<sup>\*1,2,3</sup>, F. Larrue<sup>2</sup>, M. Bertrand<sup>4</sup>, D. Liguoro<sup>5</sup>, A. Tricot <sup>†6</sup>, and H. Sauzéon<sup>1,3,7</sup>

<sup>1</sup>Flowers Team, Inria centre at the University of Bordeaux, Talence, France

<sup>2</sup>Centre Technologique Nouvelle-Aquitaine du numérique (CATIE), Talence, France

<sup>3</sup>University of Bordeaux, BPH Bordeaux Population Health, Bordeaux, France

<sup>4</sup>Surgery Department, Nîmes University Hospital, University of Montpellier, Nîmes, France

<sup>5</sup>Department of Neurosurgery, University Hospital of Bordeaux, Bordeaux, France

<sup>6</sup>Université Paul Valéry Montpellier 3, Epsilon EA 4556, Montpellier, France

## Keywords

Virtual reality  
Learning  
Neuroanatomy  
Intrinsic motivation  
Cognitive load

## Abstract

This research investigates the effectiveness of virtual reality (VR) for enhancing neuroanatomy learning among medical students, focusing on optimizing cognitive load, intrinsic motivation, and user experience. Conducted with 77 second-year medical students, the study compares traditional video-based instruction with three VR interaction conditions: active, guided, and passive. VR significantly improved anatomical learning performance compared to traditional methods, particularly in the passive and active conditions. Specifically, VR enhanced intrinsic motivation, optimized cognitive load by reducing extraneous load, and improved germane load. Additionally, the guided VR condition yielded the poorest learning performance, though it did not significantly differ from the other two VR conditions, suggesting that interactivity is not the primary factor driving VR's effectiveness. The findings highlight the complex relationship between cognitive load, intrinsic motivation, and learning performance, supporting a balanced approach to using VR for educational purposes. The role of domain expertise and interaction modes within VR are discussed as critical factors in VR's instructional efficacy.

## Highlights

- VR leads to better anatomical learning than traditional teaching.
- Intrinsic motivation, cognitive load and user experience were optimized by VR.
- Connections between intrinsic motivation and cognitive load can help understand VR effectiveness.
- Interactivity is not the VR feature that explains its effectiveness for anatomy learning.

---

\*Corresponding author: [matisse.poupard@inria.fr](mailto:matisse.poupard@inria.fr)

<sup>†</sup>Hélène Sauzéon and André Tricot are last authors due to an equal supervision

## Artificial intelligence- generated content

During the writing process of this work, the authors used ChatGPT 4.0 (<https://chat.openai.com>) to improve readability and language. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

## Funding information

The authors thank the Association Nationale de la Recherche et de la Technologie (CIFRE 2021/1881) for their financial support of this project.

## Conflict of interest

All authors declare no conflict of interest.

## Data availability statement

Data supporting the findings of this study are publicly available at this link.

## Ethic statement

Studies involving human participants were reviewed and approved by the Operational Committee for the Evaluation of Legal and Ethical Risks (OCELER). Participants provided their written informed consent to participate in this study.

## Acknowledgements

We would like to thank the Bordeaux and Montpellier Faculties of Medicine for their help in setting up this experiment. The authors also thank the CATIE for the human and financial support that enabled us to carry out this work. We also acknowledge the contributions of Alice Mérel and Gabrielle Bou Habib for their dedicated involvement and assistance in conducting the experiments.

## CRedit authorship contribution statement

**M. Poupard:** Conceptualization, Methodology, Software, Investigation, Formal analysis, Writing – original draft, Visualization. **F. Larrue:** Conceptualization, Methodology, Supervision. **A. Tricot and H. Sauzéon:** Conceptualization, Methodology, Supervision, Writing – review & editing, Validation. **D. Liguoro and M. Bertand :** Conceptualization & Resources

Anatomy is the cornerstone of medical education, providing a crucial understanding of the human body's structure that is essential for future clinicians. However, mastering anatomical knowledge, especially in the complex area of neuroanatomy, poses significant challenges for medical students. Neuroanatomy requires the integration of extensive details, a deep understanding of spatial relationships, and robust 3D visualization skills (Hall et al., 2018). Despite its critical importance, recent years have seen a notable reduction in the hours dedicated to teaching anatomy in medical education.

To address these challenges, educators are increasingly incorporating digital technologies to complement traditional teaching methods. Virtual reality (VR) offers an immersive and interactive learning environment that allows students to visualize and manipulate anatomical structures in three-dimensional space. These innovative approaches not only enhance students' understanding of neuroanatomy but also promote active engagement and participation in the learning process (Newman et al., 2022).

A core aspect of VR frequently cited in the literature is its interactive nature, which allows students to actively manipulate and explore anatomical structures in a 3D space Sinha et al. (2023). This feature not only makes learning more engaging and motivating, but also impacts cognitive load. However, recent studies suggest that this presumed positive effect on intrinsic motivation does not consistently translate into improved learning performance (Poupard et al., 2024). In fact, VR can cognitively overload learners, resulting in reduced performance compared to traditional teaching methods (Newman et al., 2022). Thus, understanding the impact of interactivity on anatomical learning, intrinsic motivation, and cognitive load remains a significant challenge.

Furthermore, most studies on VR for learning adopt a comparative media design (Buchner, 2023; Glaser and Moore, 2023). Media comparison involves comparing a new technology with a more traditional and established medium or technology (e.g., video). However, comparative studies are situational and highly context-dependent (Honebein and Reigeluth, 2021; Buchner and Kerres, 2023; Clark and Feldon, 2014). Furthermore, such methods do not identify the precise variables that lead to improved teaching in a particular method. Addressing these limitations, recent reviews have called for more studies aimed at improving VR effectiveness for learning by identifying key features of the technology and the contexts where VR can enhance learning (Honebein and Reigeluth, 2021; Clark and Feldon, 2014; Buchner, 2023; Glaser and Moore, 2023).

Hence, this study focuses on interactivity and exploration levels in the context of anatomical lessons, with the aim of pinpointing the determinants of VR effectiveness for learning. Cognitive load theory and the intrinsic motivation learning model, two fundamental learning theories in psychology, were used to explain VR's effects on learning. The main research questions addressed in this study are :

RQ1 - How does virtual reality influence anatomy learning?

RQ2 - What role does interaction play in the effectiveness of virtual reality for learning?

RQ3 - Can the effects of virtual reality on learning be better understood through the lens of cognitive load theory and intrinsic motivation?

To address these research questions, a virtual reality application for neuroanatomy was designed and evaluated in second-year medical students. The effectiveness of the VR application was evaluated in terms of anatomical knowledge, cognitive load, intrinsic motivation and user experience. Additionally, to evaluate the role of interactivity in VR's effect on learning, three different interaction modes of the VR application were used: passive, guided, and active.

## 1 Background

Several recent literature reviews demonstrate promising results for the use of virtual reality in anatomy and medical education. Barteit et al. (2021) highlighted that 3D visualization of anatomical structures significantly improves learning outcomes. Duarte et al. (2020) showed that visual aids improve memory retention compared to traditional methods, while Taylor et al. (2022) reported an increased effectiveness of virtual reality due to

interactivity and repetitive practice. Furthermore, Wang et al. (2024) found that 3D visualization stimulates personal exploration and improves learning performance, although some methodological issues prevent a definitive conclusion on VR's effectiveness. However, a recent meta-analysis by García-Robles et al. (2024) showed a positive and significant effect of VR on anatomy learning compared to traditional methods. This study emphasized that VR is an effective and flexible tool for anatomy education, facilitating practical remote learning. Based on this literature, we can formulate the following hypothesis:

(H1) *The use of virtual reality enhances neuroanatomical learning effectiveness*

## 1.1 VR for learning : interactivity and exploration

Virtual reality, particularly immersive VR, is often characterized by its high level of immersion. Technically, immersion is defined as a dynamic between sensory richness (or vividness) and interactivity provided by a system (Moreau et al., 2018). Interactivity refers to the user's ability to interact with and manipulate the digital environment. The more the user can act on the system and environment, and the more natural this interaction feels (i.e., similar to real interaction in a natural environment), the higher the system's degree of interactivity. According to Fuchs (2006), VR should immerse the user in a virtual environment composed of virtual objects that interact in real-time with each other and with one or more users through sensory-motor channels. From this perspective, interactivity is a major distinguishing feature of VR compared to more traditional digital devices.

Interactivity is emphasized because it can enhance agency (i.e., the feeling of control) and presence (i.e., feeling to be immersed), engaging the user in the activity (Won et al., 2023), which could be beneficial for learning. García-Robles et al. (2024) suggest that VR helps students take an active role in their own learning, and Sinha et al. (2023) identified interactivity as one of the six determinants of anatomy learning that can be manipulated to improve VR as an anatomy teaching tool. Moreover, some studies demonstrated that interactive VR can lead to better learning gains and recall than less interactive VR setups (Mahmoud et al., 2020; James et al., 2002). However, other studies have found no significant difference in students' performance when using interactive versus passive VR (Chua et al., 2019; Zhang et al., 2019), suggesting that the level of interactivity may not be a determining factor in the effectiveness of VR for learning (Harris and Sun, 2022).

These inconsistencies can be explained by the nature of interactions in VR, which involve both motor and cognitive activities. Indeed, Meade et al. (2019) distinguished between physical and cognitive interaction in VR. Physical interaction involves the motor dimension of interaction, such as gestures to manipulate the system (e.g., grabbing a virtual object or turning the head to control direction). Cognitive interaction, on the other hand, is related to the mental manipulation of information and decision-making regarding actions or focus. Hence, guided interaction allows physical interaction, but without the freedom to choose how to act, while active VR allows both motor and cognitive interaction.

However, questions remain about whether learning outcomes are better when students explore content independently or receive guided instruction. Some studies suggest that guiding learners in VR results in better learning gains and factual retention than allowing them to explore the environment independently (Mulders, 2023; Ferguson et al., 2020). Conversely, active navigation is often reported to be beneficial for memory and spatial learning (Chrastil and Warren, 2012; Carr and Koch, 1919). Bakdash et al. (2008) found that decision-making (i.e., the cognitive component of interaction) was the most important factor for learning in virtual environments rather than control (i.e., the motor component of interaction). Recent studies support this, showing that active navigation in virtual environments leads to higher spatial memory performance compared to passive or guided tours, although the strength of this effect varies across studies (Meade et al., 2019; Schomaker and Wittmann, 2021; Sivashankar et al., 2024).

(H2) *Cognitive interaction allows to a better learning than motor interaction only*

The contradictory findings on the impact of interactivity on learning are mirrored in studies that examine intrinsic motivation and cognitive load. While cognitive interaction tends to enhance curiosity and engagement, research on cognitive load reveals that it can introduce excessive cognitive strain, especially for students with limited prior knowledge.

## 1.2 Curiosity-driven learning

According to curiosity-driven models, learners experience intrinsic rewards when acquiring new knowledge, fostering a positive feedback loop between intrinsic motivation and knowledge acquisition (Oudeyer et al., 2016; Murayama et al., 2019). This cycle of curiosity-driven learning behaviors enhances student engagement, driven by intrinsic motivation.

Intrinsic motivation is defined as "the inherent tendency to seek out novelty and challenges, to extend and exercise one's capacities, to explore, and to learn" (Deci and Ryan, 2012). It represents a natural penchant characterized by exploratory behaviors and spontaneous interest, often manifested as curiosity. Intrinsic motivation is generally associated with engaging in an activity for the personal pleasure derived from its completion, contrasting with extrinsic motivation, which is driven by external factors like pressure or rewards. Research consistently demonstrates that students who are intrinsically motivated tend to learn more, achieve better academic results, exhibit improved retention rates in both short-term and long-term memory, and show greater persistence when facing challenges (Oudeyer et al., 2016).

Due to its immersive, interactive, and novel nature, the virtual reality learning experience is often described as engaging and motivating. A recent scoping review on the use of VR for learning anatomy by Sinha et al. (2023) shows that over 80% of the included studies measured student perception, including interest, motivation, and enjoyment. No studies reported negative opinions about VR for learning anatomy from students. In contrast, the literature shows higher levels of curiosity, motivation, and satisfaction reported by students than with traditional teaching methods (Aridan et al., 2024; García-Robles et al., 2024; Mayer et al., 2023; Makransky et al., 2020). The positive perception of VR can be explained by the self-guided nature of immersive VR technology and the first-person perspective that supports a learner-centered approach to learning (Sinha et al., 2023).

However, inconsistent results regarding the assessment of intrinsic motivation in a virtual reality context have recently been highlighted (Poupard et al., 2024). In particular, the positive effect of VR on motivation is often not correlated with better learning performance, challenging the hypothesis of self-determination theory (SDT) of a positive correlation between intrinsic motivation and learning (Deci and Ryan, 2012). Moreover, the novelty effect hypothesis cannot be ruled out in these studies. Recent research suggests that the initial discovery of a new tool may lead to excitement and engagement, which is directed more toward exploring the tool and its entertainment value rather than focusing on the learning task itself (Miguel-Alonso et al., 2024). These challenges underscore the need for better evaluations of VR's effect on intrinsic motivation (Sinha et al., 2023; Poupard et al., 2024).

SDT posits that providing controls and choices to individuals enhances their autonomy and intrinsic motivation (Deci and Ryan, 2012). Indeed, agency and self-regulation are crucial factors in fostering curiosity. Feeling autonomous and capable of controlling one's learning enhances task engagement (Deci and Ryan, 2012; Peters et al., 2018). Consequently, the interactivity features of VR can stimulate active and self-directed learning (Mahmoud et al., 2020). Freedom in decision-making could lead to higher levels of engagement and curiosity (Großmann and Wilde, 2020), while controlling the environment reduces spontaneous and autonomous exploration and motivation (Ferguson et al., 2020; Oudeyer et al., 2016; Bonawitz et al., 2011). Recent models such as the Cognitive Affective Model of Immersive Learning (CAMIL; Makransky and Petersen, 2021) and the Cognitive-Affective Theory of Learning with Media for Virtual Reality (CATLM-VR; Huang et al., 2022) posit that VR features, especially interactivity, lead to feelings of presence and agency, promoting intrinsic motivation and curiosity (Petersen et al., 2022). In a related manner, Zhang et al. (2019) used VR with low, medium, and high levels of interactivity and found no significant difference in learning gains, but increasing interactivity levels significantly promoted students' engagement, attention, and focus on the learning material.

Based on this review of VR’s effect on motivation to learn, we hypothesize that:

(H3a) Virtual reality enhances enjoyment and can foster motivation to learn.

(H3b) Interactive versions of VR, especially active VR, should improve intrinsic motivation.

(H3c) Intrinsic motivation is positively correlated with learning performance.

### 1.3 Cognitive load in virtual reality learning

Cognitive Load Theory (CLT) posits that learning requires cognitive resources in working memory (Sweller et al., 2019). According to this theory, optimal learning conditions are achieved when the complexity and presentation of the task do not exceed the learner’s available resources. CLT distinguishes between three types of cognitive load: intrinsic, extraneous, and germane (Kalyuga, 2011).

Intrinsic cognitive load (IL) represents the cognitive resources necessary for learning and depends on the complexity of the task (number and interactivity of elements to learn) and the learner’s prior knowledge. Extraneous cognitive load (EL) involves processing irrelevant information during learning, such as decorative elements in learning materials. Unlike intrinsic cognitive load, extraneous cognitive load can be reduced by improving instructional design. Both types of load are additive and represent the total cognitive load. Germane cognitive load (GL), related to knowledge acquisition, is currently understood as the mental effort dedicated to processing and understanding the intrinsic aspects of the task. Rather than adding to the total cognitive load, the germane cognitive load redistributes cognitive resources from extraneous activities to intrinsic processing, thus improving learning efficiency (Sweller et al., 2019; Kalyuga, 2011).

Numerous studies highlight the importance of considering cognitive load in VR learning. The literature suggests that VR can impose irrelevant cognitive loads and potentially overload students (Makransky and Petersen, 2021; Andersen and Makransky, 2021), especially those with low subject background knowledge (Poupard et al., 2024). Specifically, the realism and richness of virtual environments create irrelevant information that can distract users from essential content (Makransky and Petersen, 2021; Parong and Mayer, 2018). Another well-documented source of EL is the complexity of interaction features (Makransky and Petersen, 2021; Makransky et al., 2019). Due to the novelty of these technologies, users are not familiar with the interactions, and learning how to use VR can impose additional EL. Based on this finding, Andersen and Makransky (2021) distinguished three different dimensions of EL when using VR: related to instruction and explanation (instruction), related to interacting in the virtual learning environment (interaction), and related to learning in the 3D virtual environment (environment). Similar to motivation variables, CAMIL (Makransky and Petersen, 2021) and CATLM-VR (Huang et al., 2022) suggest an effect of interactivity on cognitive load, particularly on irrelevant load.

(H4a) Virtual reality, especially for interactive versions, induces extraneous cognitive load that can reduce learning.

Beyond the motor interaction that appears to impose an irrelevant cognitive load, the cognitive dimension of interactivity also impacts the cognitive demand. CLT Sweller et al. (2019) and the Cognitive Theory of Multimedia Learning (CATLM; Mayer and Estrella, 2014) both suggest that giving the autonomy to make choices could increase the extraneous cognitive load, especially for novice learners. Conversely, providing guidance with signaling helps focus the learner’s attention on the essential elements of the activity, reducing EL, and thereby promotes generative learning (Parong and Mayer, 2021; Han et al., 2023). This so-called signaling effect, in the VR context, has shown promising results in reducing irrelevant cognitive load (Ali et al., 2022; Wen et al., 2024; Schneider et al., 2018) and fostering germane cognitive load (Lehikko et al., 2024; Albus et al., 2021).

Furthermore, some studies suggest that the positive effect of guidance on cognitive load depends on the learner’s prior knowledge. Specifically, while signaling allows students with low prior knowledge to reduce cognitive

load by focusing on essential information, learners with high prior knowledge benefit more from autonomous environments (Han et al., 2023). This different effect depending on expertise is called the expertise reversal effect (Kalyuga, 2007).

(H4b) Guided version of VR induce less extraneous cognitive load than active one, especially for learners with lower prior knowledge

## 1.4 Integrating Cognitive Load and Intrinsic Motivation in VR Learning

Recent literature suggests considering the interactions between intrinsic motivation and cognitive load when evaluating the effectiveness of virtual reality (VR) for learning (Makransky and Petersen, 2021; Huang et al., 2022). This perspective emerges from the inconsistent effects observed between theoretical expectations and actual learning outcomes in VR environments. Specifically, recent studies challenge the assumed negative impact of extraneous cognitive load in VR learning contexts, indicating that the relationship may be more complex than previously thought (Skulmowski and Xu, 2022). Similarly, many studies report positive effects of intrinsic motivation on VR learning without corresponding improvements in learning performance (Poupard et al., 2024).

In response to these inconsistencies, the integration of cognitive and motivational dimensions into analyses appears to be a promising approach (Mayer and Estrella, 2014). From one perspective, CLT and recent research assume motivation as a prerequisite to engage in a cognitively demanding task (Paas et al., 2005). Despite extensive research suggesting that humans generally seek to avoid cognitive effort, recent studies indicate that curiosity and intrinsic interest motivate students to engage in challenging tasks, thereby fostering deeper learning and understanding (Feldon et al., 2019; Serki and Bolkan, 2024; Spitzer et al., 2024). For example, some studies have shown that when given a choice, participants tend to select intermediate difficulty tasks over easier ones, seeking learning progress and intrinsic rewards (Sayali et al., 2023; Ten et al., 2021).

However, these studies also assume that participants avoid tasks that are excessively demanding. This assumption aligns with the expectancy–value–cost theory (Flake et al., 2015). According to this theory, human decision-making is based on a cost-benefit analysis that considers the cognitive demands of a task (cost), self-efficacy (expectancy), and the value of the task. Learners are more likely to engage in tasks where perceived benefits, such as learning progress and intrinsic or extrinsic rewards, outweigh the cognitive costs. If a task is perceived as too difficult or cognitively demanding, it can surpass the learner’s intrinsic motivation, leading to amotivation and disengagement. Feldon et al. (2019) encapsulates this idea by describing the cognitive demand of a task as a form of motivational cost.

This complex relationship could correspond to a negative quadratic relationship (an inverted U-shape) between intrinsic motivation and cognitive load (Sayali et al., 2023; Ten et al., 2021). Moderate levels of cognitive load paired with high intrinsic motivation can lead to optimal learning conditions, while too much or too little cognitive load can result in task abandonment (Feldon et al., 2023).

In our context, highly motivated participants should report a low perceived cognitive effort (i.e., intrinsic and extraneous cognitive load). Furthermore, if VR, especially in its active form, fosters intrinsic motivation to learn, we can expect a lower perceived cognitive load in these conditions. In contrast, if the control task increases the students’ experiences of extraneous cognitive load, their motivation to learn may diminish due to the perceived cognitive costs associated with the task.

(H5a) High motivated participants report low extraneous cognitive load.

(H5b) Intrinsic motivation foster germane cognitive load.



## 2 Material and methods

### 2.1 Sample

To estimate the necessary sample size, an a priori power analysis was conducted using G\*Power 3.1. A recent meta-analysis on studies investigating virtual reality in anatomy learning reported a medium effect size (SMD = 0.57) (García-Robles et al., 2024). Based on this effect size, with an alpha level of 0.05, the projected total sample size required for detecting a within-between interaction in a repeated measures ANOVA was approximately  $N = 76$  for a power of 0.95, and  $N = 48$  for a power of 0.80.

The study was conducted at the French University of Bordeaux and Nîmes, involving second-year medical students. A total of 77 students (55 females and 21 males) participated voluntarily after being recruited via email. One participant experienced technical issues that prevented task completion, while post-test data from two others were lost. In addition, a participant withdrew due to cybersickness. These individuals were excluded from the analysis, resulting in a final sample size of 73 participants. As the participants had completed one year of medical school, they had a basic understanding of anatomy. Furthermore, all participants had previously attended a lesson on the specific topic addressed in this experiment as part of their regular curriculum. Therefore, the participants’ prior knowledge level on the subject can be considered intermediate.

During post-tests, students were questioned about their previous exposure to immersive technologies, with 59% reporting no prior usage, 38% acknowledging occasional usage, and 2% declaring frequent use. Fisher’s exact tests indicated that there were no significant differences between the groups based on gender ( $p = .633$ ), VR experience ( $p = .310$ ) and Vandenberg and Kuse (1978) mental rotation score ( $F(3,68) = 1.03$ ,  $p = 0.385$ ) (see Table 1).

Table 1: Sample demographic characteristics table

	Control	Passive VR	Guided VR	Active VR	Comparison (p)
Gender					
Female	10 (66.7%)	11 (61.1%)	16 (80.0%)	15 (75.0%)	0.633
Male	5 (33.3%)	7 (38.9%)	4 (20.0%)	5 (25.0%)	
Prior VR experience					
No	9 (60.0%)	6 (33.3%)	10 (50.0%)	12 (60.0%)	0.351
Yes	6 (40.0%)	12 (66.7%)	10 (50.0%)	8 (40.0%)	
Mental rotation					
	18.50 (7.5)	21.94 (8.4)	17.40 (8.5)	19.10 (8.0)	0.383

### 2.2 Material

All participants were provided with individual computers on assigned desks. The questionnaires, administered both before and after the neuroanatomy lesson, were completed directly on the computer using a local version of the app.peac2h.io tool. The course focused on the anatomy of the cerebral arterial system, aiming to increase learners’ declarative knowledge (factual and conceptual) of the roles and relationships of cerebral arteries. Four different conditions were designed for this experiment: a control condition, which involved following a video course on anatomy representing the typical teaching method, and three virtual reality conditions.

#### 2.2.1 Video Lesson

Teaching anatomy typically includes lectures in which a teacher explains anatomical structures, often using accompanying drawings. In the control condition of this study, the students watched a 22-minute video lecture on the arterial vascularization of the brain, given by one of the author (DL) and recorded specifically for the experiment (see Fig. 1a). Previous literature suggests that encouraging students to reproduce the professor’s drawings is beneficial for learning in traditional teaching methods (Borrelli et al., 2018; Peart, 2022). The

students were instructed to listen attentively to the course and reproduce the anatomical drawings, simulating a traditional classroom scenario. They were not allowed to pause the video during this process.

### 2.2.2 VR Lesson

The virtual reality (VR) simulation, developed using Unity3D Game Engine, used Oculus Quest Pro and 3 with hand-tracking functionalities (see Fig. 1b). The content of the VR lessons, developed by the research team, was based on the control condition video and multiple focus groups with two anatomy teachers. The virtual environment consisted of a classroom with a desk in the middle. A 3D model of the brain and its arterial system, along with a virtual tablet, were placed on the desk (see Fig. 1c). The purpose of the application was to explore the vascular brain system by interacting with the 3D brain model under three conditions.

In the active VR condition, the students interacted with the 3D brain directly using the Oculus hand-tracking functionality. Users could touch and cut any arteries to collect information about their name, origin, function, and vascular territory. Depending on the user's actions, the VR environment provided verbal explanations accompanied by animations, as well as a written summary on a virtual tablet at the end of the explanation. Each artery (and its collateral) could only be touched and cut once, and its color indicated whether it had already been touched and/or cut. Users could also grab the 3D brain model to move or resize it. In the guided VR condition, the students interacted with the 3D brain, but the system imposed the choice of actions to be performed. A blue hand coach indicated the actions to be performed on each artery in a predetermined order, mirroring the control video. In the passive condition, students observed the animations sequentially without interaction. A demo of the VR application is available at the following address: .

All information from the control video was transferred to the VR lesson, maintaining the same order and modalities: oral information in the video was presented orally in VR, and animations were created to align with the visual information from the teacher's drawings and gesture indications. Additionally, similar to the video where the information on the blackboard was consistently visible, in VR, the information acquired on an artery could be consulted at any time on the virtual tablet by hovering over the artery with the finger.

Furthermore, students had to follow a tutorial to familiarize themselves with the environment before starting the task. For the active and guided conditions, the tutorial also provided practice in interacting with the 3D model. Finally, students were instructed to touch and cut all the system's arteries.

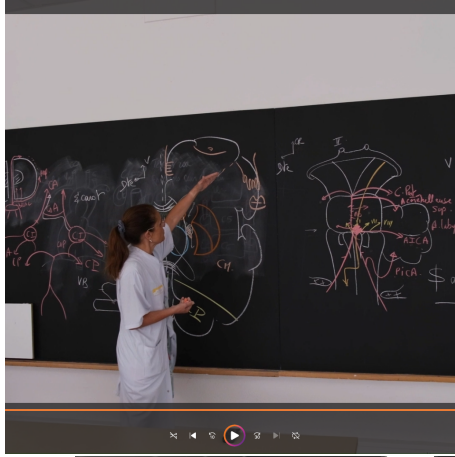
## 2.3 Measures

To assess learning, a pre- and post-knowledge questionnaire was administered, comprising verbal and visuospatial items. According to literature suggesting multidimensional knowledge assessment (Hamilton et al., 2021; Boulet et al., 2011), the questionnaire developed by the research team was designed to assess various aspects of learning. Verbal questions (MCQ and open-ended questions) consisted of 10 items of factual recall, where the answer was explicitly provided in the course, and 6 items requiring inference, which required participants to deduce answers based on the course content. Additionally, the visuospatial component involved a captioning task for six illustrations, assessing participants' ability to identify and label anatomical structures. To mitigate the potential pretesting effect, 19 control questions on general anatomy were included in both the pre- and post-tests to ensure consistency.

Cognitive load during virtual reality (VR) learning was evaluated using the Multidimensional Cognitive Load Scale for Virtual Environments (MCLSVE) (Andersen and Makransky, 2021). This adapted scale incorporates three subscales for extraneous cognitive load (environment, interaction, and instruction), alongside measures of intrinsic and germane cognitive load.

Intrinsic motivation was assessed using the Situational Motivation Scale (SIMS) (Guay et al., 2000), which measures the motivational orientation of the individuals and the situational motivation towards the task. The scale includes a global self-determination indicator calculated from four sub-scales (intrinsic motivation, identified regulation, external regulation, amotivation).

(a) Screenshot of the control video



(b) Participants during the experimentation



(c) Screenshot of the VR environment

Figure 1: Illustrations of experimental setup

Furthermore, the impact of the learning media condition on self-determination was assessed with the Technology-based Experience of Need Satisfaction–Interface (TENS)(Peters et al., 2018). The TENS probes how technology interfaces meet essential user needs, including autonomy and competence according to the self-determination theory (Ryan and Deci, 2020). The competence and autonomy measures are closely related to the intrinsic motivation measures provided by the SIMS.

Learner engagement was assessed with the French version of User Engagement Scale (UES)(Fontaine et al., 2019). The UES assesses engagement through 4 dimensions such as (1) focused attention (feeling absorbed in the interaction and losing track of time), (2) intrinsic reward (a sense of satiety about learning and success), (3) perceived usability, and (4) aesthetic appeal of technology. The first two dimensions are therefore respectively linked to the cognitive load (attention capture) and intrinsic motivation (reward dimension) generated by the content and activities proposed by the technological device, while the last two dimensions are more related to the practicality and aesthetics of the device.

Lastly, to address potential cybersickness, the CyberSickness in Virtual Reality Questionnaire (CSQ-VR) (Kourtis et al., 2023) was utilized. This scale allows for the evaluation of VR-induced symptoms and effects, providing insight into participants' well-being during the VR learning experience.

It should be noted that the cognitive load, intrinsic motivation, cybersickness, and TENS questionnaires were translated into French by two of the authors and validated through discussion with all authors.

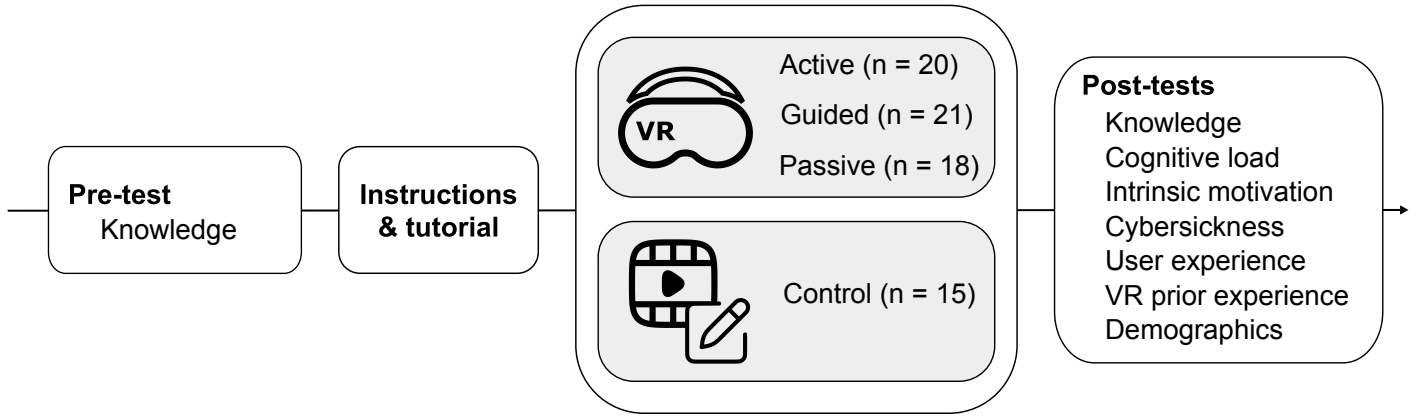


Figure 2: Experimental procedure

## 2.4 Procedure

Figure 2 illustrates the experimental process, including pre-test, experimental conditions and post-test. The experimental design consisted of a four-modality independent variable (video control, passive VR, guided VR and active VR) and learning (pre and post), curiosity, cognitive load and user experience were the dependent variables.

Experimental sessions, lasting approximately 1 hour and 30 minutes, were conducted in a quasi-ecological context, involving 1 to 8 students based on availability in a designated classroom. Upon arrival, each student was assigned an individual desk with a computer. Before commencing the experimental tasks, they read and signed a comprehensive consent form detailing the study’s nature and purpose.

Then, students individually had to answer a set of items designed to provide level of prior knowledge on the topic and general anatomy knowledge. The participants were randomly assigned to one of the four experimental conditions presented in section 2.2. Those in the virtual reality groups received VR headsets, while the control group received headphones, a sheet of paper, and pens. To ensure a standardized understanding and use of VR, students in VR groups were helped with installing and adjusting their headsets. Subsequently, all participants received a tutorial explaining the task objectives. After completing the neuroanatomy lesson, the students received the post-questionnaire and post-test, which they completed at their own pace.

Research assistants were available throughout the procedure to address any inquiries or concerns.

## 2.5 Data analysis

All analyzes were conducted in R version 4.4.1. A p-value less than 0.05 was considered statistically significant, and missing values were excluded from the analyses. Similarly, extreme values (above  $Q3 + 3 \times IQR$  or below  $Q1 - 3 \times IQR$ ) were excluded. Parametric methods were always used when assumptions of normality and homogeneity of variance were met. For data that did not meet these assumptions, non-parametric or robust methods were employed.

In order to investigate the impact of different interventions on participants’ learning, a linear mixed-effects model was used. The model included fixed effects for time (pre-test vs. post-test), groups (Control, Passive VR, Guided VR, Active VR), and their interaction using the lmer function from the lme4 and lmerTest packages from R. Model was fit with restricted maximum likelihood (REML) and p-values were calculated using the Satterthwaites approximation.

Simple ANOVAs were used to compare the groups on cognitive load, intrinsic motivation, and user experience. Effects size were calculated using partial eta-squared ( $\eta^2$ ). According to Cohen (2013),  $\eta^2$  can be interpreted as small ( $\eta^2 = 0.01$ ), medium ( $\eta^2 = 0.06$ ) and large ( $\eta^2 = 0.14$ ).

Post hoc tests were performed using pairwise t tests with Bonferroni correction for multiple comparisons. Cohen's d was employed as the effect size measure. Cohen's d can be interpreted as small ( $d = 0.20$ ), medium ( $d = 0.50$ ), and large ( $d = 0.80$ ).

To explore the relationships between different variables, the normalized learning gain was calculated based on the formula proposed by Hake (Hake, 1998):  $(\text{pretest} - \text{posttest}) / (\text{max} - \text{pretest})$ . Pearson correlation coefficient or Kendall's rank correlation test was utilized, alongside linear regression, to assess correlations between different variable. As multiple correlations were performed, a Bonferonni correction was applied ( $\alpha = 0.004$ ) (Curtin and Schulz, 1998).

All scripts used for the analysis are publicly available at this link.

### 3 Results

An overview of the results can be seen in Table 2.

Since the pre- and post-tests were similar, we included control questions to control the pretesting effect in the pre- and post-tests. Mixed ANOVA did not show a significant main effect of group ( $F(3,69) = 2.6, p = 0.058$ ) or time ( $F(1,69) = 2.3, p = 0.13$ ), nor interaction effect ( $F(3,69) = 1.243, p = 0.301$ ). We therefore conclude that the variations in neuroanatomy learning are not due to the pretesting effect and can be attributed to intervention effect.

Table 2: Summary of results for each indicators

	Conditions				p	$\eta p^2$	Comparison
	Control	Passive VR	Guided VR	Active VR			
Outcomes	M (SD)	M (SD)	M (SD)	M (SD)			
Learning							
Pre-test	0.14 (0.05)	0.13 (0.08)	0.17 (0.06)	0.17 (0.06)	.336	-	
Post-test	0.24 (0.08)	0.33 (0.11)	0.30 (0.08)	0.35 (0.07)	.006	.17	Ctrl < Passive, Active
Cognitive load							
Intrinsic Load	0.63 (0.17)	0.46 (0.22)	0.49 (0.17)	0.47 (0.20)	.045	.11	
Extraneous Load	0.32 (0.24)	0.09 (0.08)	0.08 (0.08)	0.06 (0.06)	.003	.39	Ctrl > Passive, Guided, Active
Germane Load	0.50 (0.23)	0.86 (0.14)	0.81 (0.12)	0.85 (0.13)	< .001	.46	Ctrl < Passive, Guided, Active
Intrinsic motivation							
Self-Determination Index	37.07 (9.21)	54.28 (8.29)	48.70 (10.84)	52.00 (14.38)	< .001	.226	Ctrl < Passive, Active
Intrinsic Motivation	0.62 (0.14)	0.85 (0.08)	0.82 (0.09)	0.87 (0.07)	< .001	.494	Ctrl < Passive, Guided, Active
Identified Regulation	0.76 (0.12)	0.85 (0.08)	0.85 (0.09)	0.88 (0.09)	.004	.179	Ctrl < Passive, Guided, Active
External Regulation	0.33 (0.15)	0.27 (0.15)	0.39 (0.14)	0.28 (0.13)	.038	.079	
Amotivation	0.20 (0.06)	0.17 (0.04)	0.18 (0.06)	0.18 (0.06)	.738	-	
User Engagement							
Focused Attention	3.27 (1.03)	4.28 (0.50)	4.28 (0.61)	4.30 (0.60)	< .001	.273	Ctrl < Passive, Guided, Active
Perceived Usability	3.73 (0.92)	4.48 (0.35)	3.65 (0.55)	4.10 (0.73)	< .001	.233	Ctrl, Guided < Passive
Aesthetic Appeal	3.24 (0.70)	4.61 (0.40)	4.23 (0.48)	4.18 (0.67)	< .001	.420	Ctrl < Passive, Guided, Active
Intrinsic Reward	3.60 (0.86)	4.79 (0.30)	4.65 (0.41)	4.57 (0.47)	< .001	.420	Ctrl < Passive, Guided, Active
Technology acceptance							
Competence	19.93 (2.58)	21.11 (3.16)	18.60 (3.30)	21.05 (3.00)	.040	.077	Guided $\leq$ Passive, Active
Autonomy	19.80 (3.05)	20.56 (2.57)	19.85 (2.83)	21.65 (2.18)	.114	-	
Cybersickness							
Overall	0.16 (0.03)	0.20 (0.03)	0.21 (0.08)	0.22 (0.08)	0.005	.015	Ctrl < Passive, Guided, Active
Nausea	2.00 (0.00)	2.39 (0.61)	2.47 (0.90)	2.40 (0.60)	0.165	-	
Vestibular	2.00 (0.00)	2.72 (0.96)	2.55 (1.10)	3.05 (1.32)	0.006	.014	Ctrl < Passive, Active
Oculomotor	2.23 (0.44)	3.39 (1.14)	3.70 (1.49)	4.00 (1.95)	< 0.001	0.161	Ctrl < Passive, Guided, Active

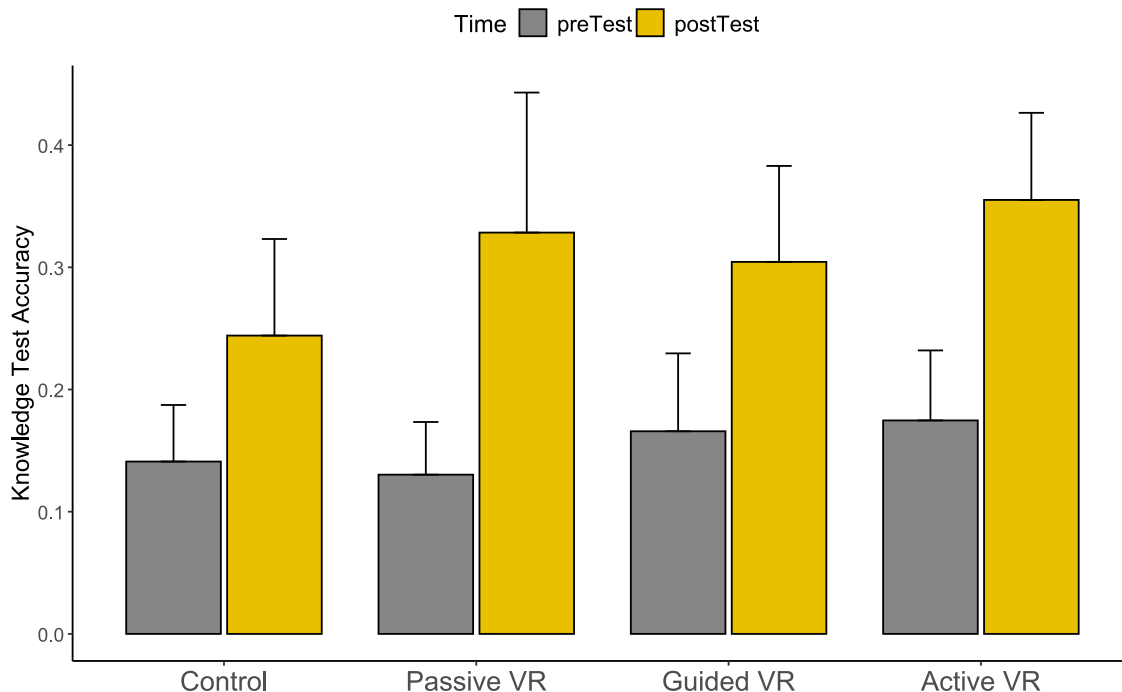


Figure 3: Effect of virtual reality on pre- and post-test overall learning score

### 3.1 How does VR impact learning ?

Table 3 show the results of linear-mixed model used to examine the effects of group (Control, Passive VR, Guided VR, Active VR) and time (Pre-test, Post-test) on learning scores. Marginal (0.56) and conditional (0.76)  $R^2$  suggested that the model provides a good fit to the data.

The results revealed a significant improvement in anatomical knowledge from the pre-test to the post-test in all groups ( $F(1, 60.6) = 273.3, p < 0.001, \eta_p^2 = 0.922$ ). However, the control condition did not significantly differ from Passive VR ( $\beta = -0.0035, t(108.47) = -0.13, 95\% \text{ CI} = [-0.054, 0.048], p = 0.895$ ), Guided VR ( $\beta = 0.0248, t(106.77) = 0.97, 95\% \text{ CI} = [-0.024, 0.074], p = 0.336$ ), and Active VR ( $\beta = 0.0336, t(106.77) = 1.31, 95\% \text{ CI} = [-0.016, 0.083], p = 0.192$ ) on knowledge pretest. Furthermore, ANOVA in the pre-test score did not show significant differences in the pre-test scores between the VR groups ( $F(3, 68) = 2.63, p = 0.057$ ), indicating that all groups were equivalent in knowledge of neuroanatomy before the intervention (see Figure 3).

Of particular interest were the interaction effects between time and each intervention group. The interaction analysis for Passive VR ( $\beta = 0.088, t(60.28) = 3.20, 95\% \text{ CI} = [0.035, 0.141], p = 0.002$ ) and Active VR ( $\beta = 0.074, t(60.78) = 2.74, 95\% \text{ CI} = [0.022, 0.126], p = 0.008$ ) were statistically significant, suggesting that the restitution score of these groups improved significantly more than the control group. Moreover, the interactions for Guided VR ( $\beta = 0.036, t(59.63) = 1.34, 95\% \text{ CI} = [-0.016, 0.087], p = 0.187$ ) did not reach statistical significance. These results support the effectiveness of the learning intervention in all conditions, with VR allowing the greatest improvement in learning outcomes, particularly for Passive and Active modalities, confirming hypothesis 1. Moreover, although the active condition appears to give better learning results than the guided condition, the results did not show significant learning differences between the guidance levels, invalidating hypothesis 2.

Regarding the learning sub-dimensions, analysis did not reveal any significant interaction effect between the groups and the time on verbal ( $F(3, 67) = 2.084, p = 0.111$ ) and visuospatial dimensions ( $F(3, 67) = 2.2, p = 0.093$ ) suggesting an equivalent progression between groups on both dimensions.

Table 3: Knowledge evolution between condition - Linear mixed model results

Fixed Effects					
	Estimate	SE	95% CI	t	p
Intercept	0.141	0.020	0.10 - 0.18	7.17	< <b>.001</b>
TimepostTest	0.103	0.020	0.06 - 0.14	5.05	< <b>.001</b>
GroupePassive VR	-0.0035	0.026	-0.05 - 0.05	-0.13	0.895
GroupeGuided VR	0.025	0.026	-0.02 - 0.07	0.97	0.336
GroupeActive VR	0.034	0.026	-0.02 - 0.08	1.31	0.192
PostTest x Passive VR	0.088	0.027	0.04 - 0.14	3.20	<b>0.002</b>
PostTest x Guided VR	0.036	0.027	-0.02 - 0.09	1.34	0.187
PostTest x Active VR	0.074	0.027	0.02 - 0.13	2.74	<b>0.008</b>
Random Effects					
	Variance		SD		
Participants (intercept)	0.002501		0.050		
Residuals	0.002912		0.054		
Model fit					
R <sup>2</sup>	Marginal		Conditional		
	0.557		0.762		
ANOVA					
	F	p	$\eta_p^2$		
Time	273.3	< <b>0.001</b>	0.955		
Group	3.42	<b>0.023</b>	0.446		
Interaction	4.25	<b>0.009</b>	-		

*Note.* Model equation: Score - Time \* Group + (1 — Participant). P-values were calculated using Satterthwaite’s approximations.

### 3.2 How does VR influence intrinsic motivation to learn ?

ANOVA tests indicated that the groups differed significantly with respect to intrinsic motivation ( $F(3, 67) = 21.8, p < .001, \eta_p^2 = 0.494$ ), and identified regulation ( $F(3, 68) = 4.95, p = 0.004, \eta_p^2 = 0.179$ ) (see Fig. 4). Consistent with H3a, the results showed that the control condition reported significantly lower levels of intrinsic motivation and identified regulation than all VR groups (see Table 4). Moreover, significant difference between groups was found on the self-determination index ( $\chi^2 = 18.4, p < .001, \eta_p^2 = 0.226$ ). Post-hoc Dunn tests with Bonferroni correction indicated that the control condition showed lower self-determination than passive ( $p = 0.0006, d = 0.73$ ) and active ( $p = 0.0013, d = 0.54$ ) VR conditions.

ANOVA also showed significant differences between groups on external regulation ( $\chi^2 = 8.43, p = 0.038, \eta_p^2 = 0.0788$ ), and even though pairwise comparisons showed no significant differences between groups, the control and guided VR groups revealed higher levels than passive and active VR conditions (see Table 4). It should be noted that comparing the three VR modalities among themselves shows a significantly higher level of external regulation for the guided condition compared to the other two modalities ( $F(2, 59) = 8.08, p = 0.018, \eta_p^2 = 0.11$ ). Furthermore, no differences were found between the groups for amotivation ( $\chi^2 = 1.26, p = 0.738$ ).



Table 4: Situational motivation scale significant pairwise comparison between groups

Dependent variable	Control vs.	t stat	p value	effect size
Self-determination	Passive	3.85	< .001	r = 0.73
	Guided	2.55	0.06	-
	Active	3.84	< 0.001	r = 0.55
Intrinsic motivation	Passive	4.59	< 0.001	d = 2.04
	Guided	3.58	< 0.001	d = 1.70
	Active	4.95	< 0.001	d = 2.27
Identified regulation	Passive	0.09	0.03	d = 0.91
	Guided	0.09	0.03	d = 0.84
	Active	0.12	0.001	d = 1.18

Furthermore, the comparison of interaction levels within the VR conditions did not reveal significant differences in self-determination ( $p = 0.237$ ) or intrinsic motivation ( $p = 0.156$ ). Consequently, hypothesis 3b, which proposed an effect of interactivity on motivation, was not supported. Additionally, while the regression analysis indicated a positive but non-significant correlation (after applying Bonferroni adjustment) between intrinsic motivation and learning ( $r(69) = 0.27$ , 95% CI [0.04, 0.47],  $p = 0.022$ ), we cannot fully support H3c or motivational theories, and the findings do not allow us to conclude that the measured intrinsic motivation was entirely directed towards learning.

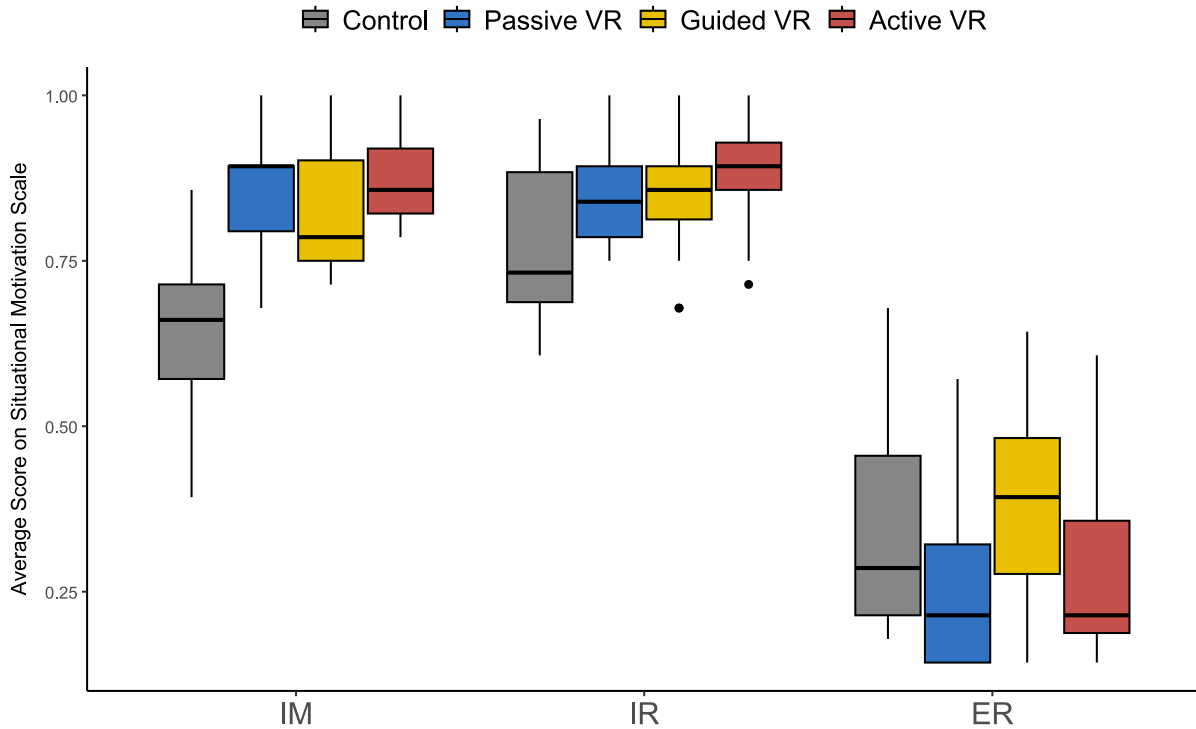


Figure 4: Situational motivation scale results

### 3.3 How does VR influence cognitive load ?

Results on overall perceived cognitive load showed no differences between conditions ( $F(3, 69) = 0.25$ ,  $p = 0.86$ ) but the sub-dimension analysis suggested variations in load distribution across the groups.

In contrast to H4a, the results showed that all VR conditions induced significantly lower levels of extraneous

cognitive load compared to the control condition ( $F(3, 34.6) = 5.55, p = 0.003, \eta p2 = 0.39$ ). Consequently, participants reported significantly higher levels of germane load under VR conditions compared to the control group ( $F(3, 35.6) = 10.25, p < .001, \eta p2 = 0.46$ )(see Table 5). These results revealed that VR optimized learners’ cognitive load by minimizing irrelevant cognitive load and maximizing relevant load. It should be noted that differences in extraneous cognitive load were mirrored in extraneous load relative to instruction ( $p = 0.009$ ) and interaction ( $p = 0.007$ ) but not in EL relative to environment ( $p = 0.005$ ).

Table 5: Multidimensional cognitive load scale for virtual environments pairwise comparison between groups

Dependent variable	Control vs. VR	T	p	d
Intrinsic Load	Passive	-0.17	0.051	-
	Guided	-0.14	0.132	-
	Active	-0.16	0.079	-
Extraneous Load	Passive	-0.23	0.012	1.29
	Guided	-0.24	0.010	1.32
	Active	-0.27	0.004	1.48
Germane Load	Passive	0.36	< .001	1.91
	Guided	0.31	< .001	1.71
	Active	0.35	< .001	1.88

Additionally, a significant difference was found ( $F(3, 69) = 2.83, p = 0.045, \eta p2 = 0.11$ ) between the groups for intrinsic cognitive load, but the adjusted Bonferroni pairwise comparisons did not reveal significant differences between the groups. This result suggested a minor difference between the perceived difficulty of the task reported under the four conditions (see Fig. 5). Furthermore, no differences between VR conditions were observed for the three types of cognitive load, refuting H4b, which indicated that guided interaction would be less cognitively demanding.

To ensure compliance with CLT assumptions, the correlation between measured learning performance and cognitive load was tested. Bonferonni adjusted correlation tests showed that learning gain was not significantly correlated with intrinsic ( $r(71) = -0.288, 95\% \text{ CI } [-0.49, -0.06], p = 0.013$ ), extraneous ( $p = 0.380$ ) and germane cognitive load ( $p = 0.161$ ).

### 3.4 What are the relationships between learning, cognitive load and intrinsic motivation?

Regarding the third research question, the analysis of the relationship between cognitive load and intrinsic motivation to learn revealed different patterns (see Fig.6). While intrinsic motivation was negatively correlated with intrinsic cognitive load ( $r(69) = -0.38, 95\% \text{ CI } [-0.56, -0.15], p = 0.001$ ) and extraneous cognitive load ( $\tau = -0.43, 95\% \text{ CI } [-0.56, -0.3], p < .001$ ), results showed a moderate, positive correlation with germane cognitive load ( $\tau = 0.50, 95\% \text{ CI } [0.38, 0.62], p < .001$ ). Additionally, multiple regression analysis ( $RSE = 0.09, R^2 = 0.52, F(3,67) = 25.86, p < .001$ ) revealed that only germane processing significantly predict intrinsic motivation ( $\beta = 0.033, t = 4.63, p < .001$ ), while IL ( $p = 0.195$ ) and EL ( $p = 0.075$ ) were not significant predictors. These findings support the hypothesis 5a and 5b that motivated participants experience less extraneous and more germane cognitive load.

### 3.5 User experience

Concerning user engagement with task and technology, the results revealed that VR conditions significantly enhance user engagement compared to the traditional learning environment. A Kruskal-Wallis test indicated significant differences in user engagement (UES) between the groups ( $\chi^2 = 24.8, p < 0.0001$ ), with all VR

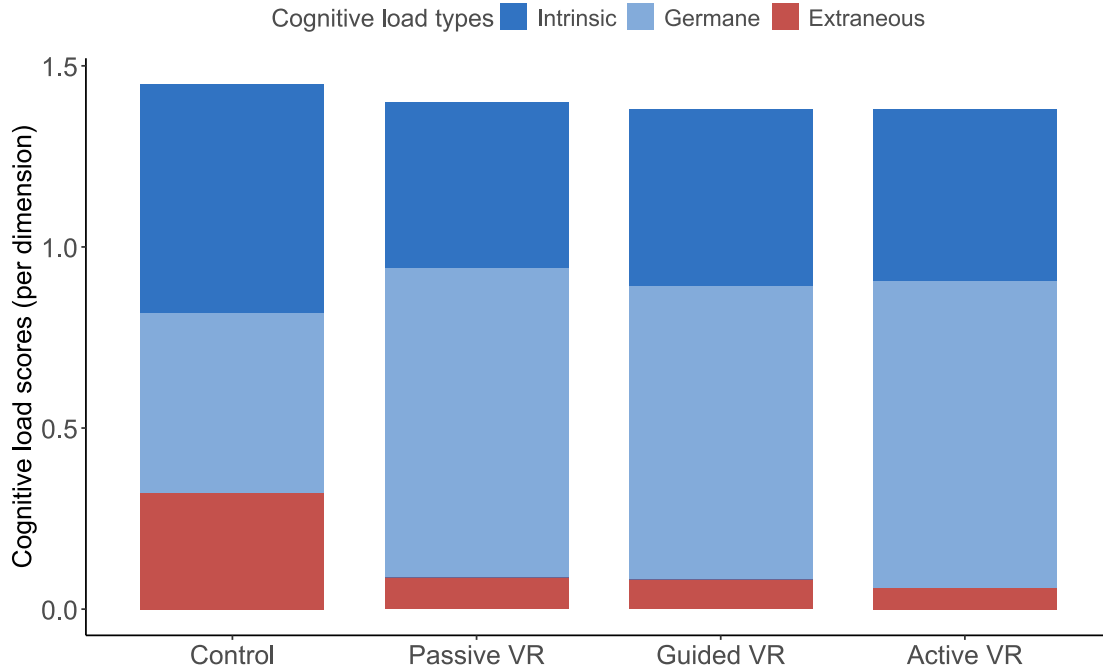


Figure 5: **Cognitive load scores by group.** Maximum score for each cognitive load type is 1. The figure illustrates the overall cognitive load levels across the groups, highlighting the differences in the distribution of cognitive load types.

conditions outperforming the control group. ANOVA results showed significant differences in focused attention ( $F(3, 69) = 8.65, p < 0.001, \eta p^2 = 0.273$ ) and aesthetic appeal ( $F(3, 68) = 16.4, p < 0.001, \eta p^2 = 0.42$ ), with post-hoc tests confirming the superiority of VR conditions over the Control group. Welch ANOVA demonstrated significant differences in reward ( $F(3, 34.7) = 8.62, p < 0.001, \eta p^2 = 0.420$ ), with VR conditions again outperforming the control group. A Kruskal-Wallis test ( $\chi^2 = 19.1, p < 0.001, \eta p^2 = 0.233$ ) and Dunn tests with Bonferroni correction indicated significantly higher perceived usability in Passive VR compared to the Control ( $p = 0.014$ ) and Guided VR ( $p < 0.001$ ) groups.

The findings revealed significant relationships between user engagement, intrinsic motivation, and cognitive load. Intrinsic motivation was positively correlated with focused attention ( $r(69) = 0.55, 95\% \text{ CI } [0.36, 0.69], p < 0.001$ ), intrinsic reward ( $\tau = 0.52, 95\% \text{ CI } [0.39, 0.64], p < 0.001$ ), and aesthetic appeal ( $r(68) = 0.6, 95\% \text{ CI } [0.42, 0.73], p < 0.001$ ), highlighting the strong connection between technology involvement and motivation to learn.

Regarding extraneous cognitive load, usability, aesthetic appeal, and reward all demonstrated negative correlations with extraneous load (usability:  $\tau = -0.31, 95\% \text{ CI } [-0.44, -0.18], p < 0.001$ ; aesthetic appeal:  $\tau = -0.33, 95\% \text{ CI } [-0.49, -0.17], p < 0.001$ ; reward:  $\tau = -0.36, 95\% \text{ CI } [-0.49, -0.22], p < 0.001$ ), indicating that these factors help to minimize unnecessary cognitive strain. Specifically, the negative impact of usability on irrelevant cognitive load was most evident in aspects related to interaction ( $\tau = -0.38, p < 0.001$ ) and environment ( $\tau = -0.32, p < 0.001$ ), rather than instruction ( $p = 0.14$ ), suggesting that VR can present usability challenges primarily in interactive elements.

Finally, all dimensions of user engagement except usability were positively correlated with germane processing, underscoring the role of enhanced engagement in promoting deeper, more generative learning processes (see Fig. 7).

The measurement of technology acceptance related to self-determination feelings revealed significant differences in competence among the groups ( $\chi^2 = 8.47, p = 0.037$ ). Post-hoc Dunn tests with Bonferroni correction

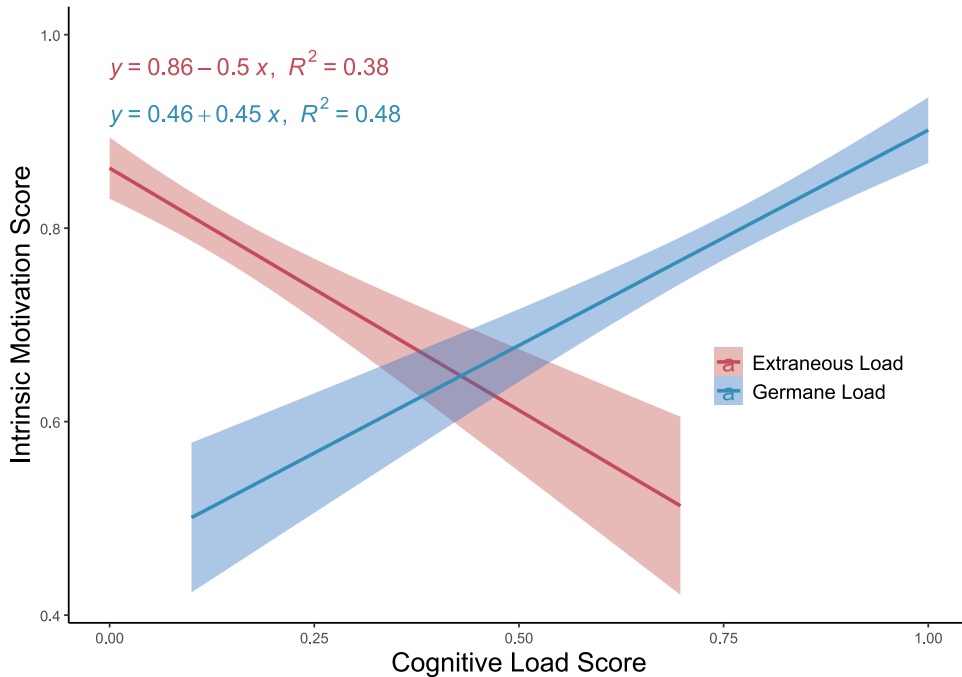


Figure 6: Correlation between intrinsic motivation and cognitive load sub-dimensions

indicated that the Guided VR group had lower competence scores than the Passive VR ( $p = 0.015$ ) and Active VR ( $p = 0.016$ ) conditions, although these differences were not statistically significant after adjustment. In contrast, autonomy did not show significant differences among the groups ( $\chi^2 = 5.76$ ,  $p = 0.124$ ). This suggested that while certain types of VR-based learning environments may enhance users' sense of competence, overall acceptance in terms of autonomy remained consistent across different learning methods.

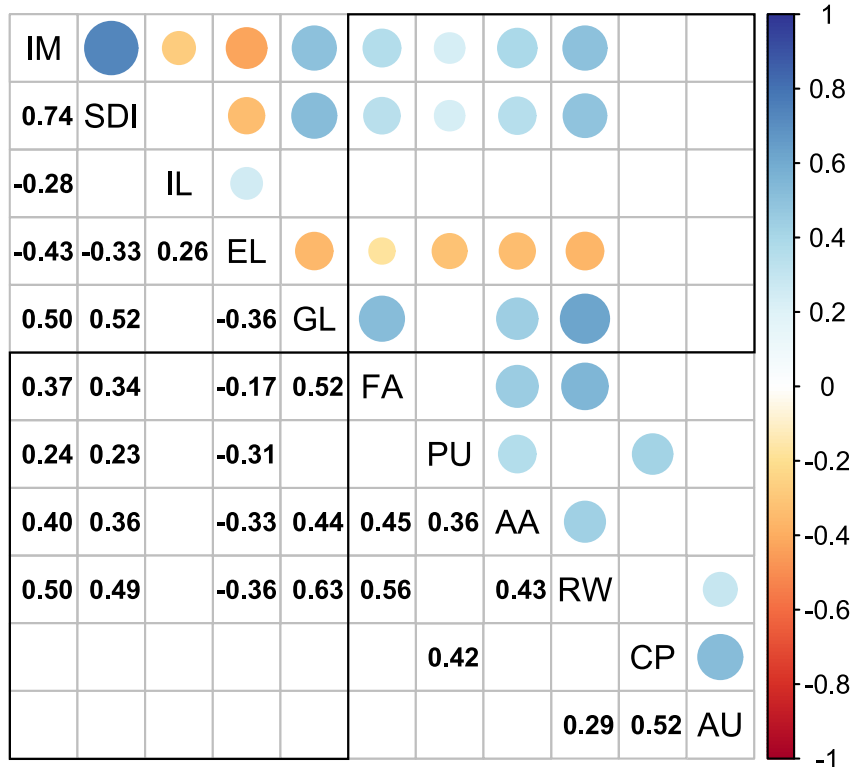
Cybersickness was assessed immediately after the task using the CyberSickness in Virtual Reality Questionnaire (CSQ-VR) (Kourtesis et al., 2023). ANOVA showed significant differences between the groups ( $\chi^2 = 12.09$ ,  $p = 0.005$ ,  $\eta^2 = 0.15$ ). Pairwise comparisons indicated higher levels of cybersickness in the passive ( $p = 0.035$ ), guided ( $p = 0.025$ ) and active ( $p = 0.004$ ) conditions compared to the control group. Moreover, these differences in cybersickness were primarily due to higher levels of disorientation ( $\chi^2 = 12.5$ ,  $p = 0.006$ ,  $\eta^2 = 0.14$ ) and oculomotor discomfort ( $F(3,35.6) = 12.5$ ,  $p < 0.001$ ) in the VR groups, rather than nausea. It should be noted that reported levels of cybersickness symptoms were quite low and that no significant differences were observed among the three VR groups themselves.

## 4 Discussion

### 4.1 Empirical contribution

In the present experiment, the effect of virtual reality on neuroanatomy learning was explored. A stronger improvement in knowledge, associated with more intrinsic motivation, was expected with VR compared to traditional teaching. However, we anticipated a negative impact on cognitive load due to the novelty, the amount of information, and the difficulty of use, all of which could impose extraneous cognitive load. Beyond the general effect of VR, this study investigated the role of interactivity levels in these technological solutions to learning. Three degrees of interaction (i.e. passive, guided and active) were proposed and better learning progress and motivation was expected in high-interactivity modalities.

First, we show that VR improves both the factual and conceptual learning of neuroanatomy compared to traditional teaching methods. These findings align with García-Robles et al. (2024), who observed a moderate to strong positive impact of VR on anatomical learning gains in a recent meta-analysis. The positive results



IM = intrinsic motivation, SDI = self-determination index, IL = intrinsic load, EL = extraneous load, GL = germane load, FA = focused attention, PU = perceived usability, AA = aesthetic appeal, RW = reward, CP = competence, AU = autonomy

Figure 7: Intrinsic motivation, cognitive load and user experience correlation matrix

in our study are further supported by higher intrinsic motivation and a more optimal cognitive load under the intervention conditions. Specifically, participants in the VR conditions reported significantly higher levels of intrinsic motivation and self-determination compared to those in the traditional learning condition. This increased motivation can be attributed to the immersive nature of VR, which has been consistently emphasized in the literature (Makransky and Petersen, 2021; Huang et al., 2022). The first-person perspective and self-guided nature of VR create a strong sense of presence and embodiment (i.e., the sensation of being part of the virtual environment) for the user (Sinha et al., 2023), which in turn enhances self-determination and intrinsic motivation.

Furthermore, recent studies have highlighted methodological concerns regarding the effects of VR on intrinsic motivation (Poupard et al., 2024). Although the results indicate a positive trend between intrinsic motivation and learning, the Bonferroni correction reveals that this relationship is not strong enough to reach statistical significance at the adjusted threshold. Consequently, we cannot conclude that the intrinsic motivation measured was mainly directed towards learning, as it may also have been influenced by participants' curiosity and entertainment with discovering the system and its novelty (Poupard et al., 2024).

Consistent with previous research (Chen et al., 2024), the higher levels of engagement and competence observed in VR in this study suggest that immersive experiences in novel environments can stimulate intrinsic interest and curiosity. Beyond embodiment, agency is also recognized as a crucial factor for the effectiveness of VR learning environments (Makransky and Petersen, 2021). The ability to actively participate in one's learning process and interact within the environment is essential for fostering self-determination (Deci and Ryan, 2012). Supporting this, García-Robles et al. (2024), in a systematic review of anatomy learning in XR, found that manipulating 3D anatomical structures leads to better learning outcomes and promotes active, motivated, and self-determined learning. Although our study did not find significant differences in autonomy across conditions, autonomy scores were generally high, particularly in the Active VR condition, and the self-determination index

showed a significant improvement under VR conditions. The absence of significant differences in autonomy may be due to the variety of learning conditions tested, which might have limited the ability to detect an effect specifically in the Active VR condition. Therefore, it appears more accurate to suggest that the observed improvement in self-determination in VR environments is more closely associated with an increased sense of competence rather than a substantial increase in autonomy or self-regulation.

Moreover, VR was associated with intrinsic reward and aesthetic appeal, indicating that this technology promotes enjoyment, which is consistent with the findings of previous studies (Makransky and Lilleholt, 2018; Huang et al., 2022). Therefore, in line with the CAMIL framework (Makransky and Petersen, 2021), we can conclude that VR fosters positive emotions, thus enhancing intrinsic motivation to learn, which is linked to better self-determination and learning performance. In addition, strong correlations between intrinsic motivation, germane cognitive load, and other factors linked to cognitive involvement and flow (i.e., focused attention and reward) confirm a well-consistent of measures used as requested in the recent literature, to better encompass the effect of virtual reality on intrinsic motivation to learn (Poupard et al., 2024).

The results showed that VR had a positive effect on the perceived cognitive load of the participants compared to the traditional condition. VR students reported less extraneous load than the control condition, resulting in a higher level of germane load, which is directed towards learning. Moreover, results showed that participants in VR conditions reported more focused attention, associated with less extraneous cognitive load and more germane processing. This finding invalidate the assumption that VR induce irrelevant cognitive load due to the realism and richness of virtual environments that can distract users from learning activity (Parong and Mayer, 2021; Poupard et al., 2024). One possible explanation is that the traditional method imposes significant cognitive demands (Hall et al., 2018), as it requires students to divide their attention between active listening, observing anatomical diagrams, and completing drawing tasks. According to CLT, this division of attention is a source of non-relevant cognitive load (i.e., split-effect) (Ayres and Sweller, 2014). In this context, VR could help students focus on essential elements and provide the division of attention. Moreover, EL associated with VR is often linked to high difficulty of use (Han et al., 2021) and novelty of the system. Even if in the present study, most of the participants discovered virtual reality, the result on perceived usability indicated that the technology was easy to use. Hence, a well-designed virtual reality application could minimize EL and allow students to fully benefit for VR advantages. Furthermore, in anatomy, a discipline where understanding the spatial representation of elements is essential, 3D and stereoscopic display of structures in VR can reduce cognitive load (Bogomolova et al., 2021), particularly for participants with lower visuospatial abilities. Poupard et al. (2024) showed that EL induced by VR especially impact low prior knowledge learners. Although knowledge pre-test was rather low, included participants of the study were in second year in medical university, and they already had assisted to several anatomical lessons, including one on the subject approached in this experiment. Hence, we can consider the prior knowledge level of participant as intermediate, explaining the particularly low level of EL reported in VR conditions. Finally, given the quasi-ecological context of the study, immersion and isolation induced by VR can minimize extraneous cognitive load from distraction sources (e.g., others, experimenters explaining, etc.) compared to the control condition.

This study also hypothesized that the interactivity features of VR, especially cognitive ones (in contrast to motor-only interactivity), could improve learning and intrinsic motivation to learn through more agency. However, the results showed that the interactive dimension of VR is not the determining characteristic of VR-related learning effectiveness. According to the present study, the role of interaction in VR learning remains unclear. Only the passive and active VR versions actually produced better learning results than the control condition, and there were no significant differences in terms of learning, motivation, and cognitive load based on the level of interaction. This result does not confirm the literature suggesting that interactivity is an essential and determining component of VR effectiveness for learning anatomy (Sinha et al., 2023; García-Robles et al., 2024) or any other type of learning (Makransky and Petersen, 2021; Mulders et al., 2020). However, several previous studies argued that interactivity had no effect on learning outcomes (Chua et al., 2019; Zhang et al., 2019). In a study on language learning, Harris and Sun (2022) did not show a significant effect of interaction on learning, suggesting that interaction had a distracting and irrelevant effect on the task. For an interaction to be beneficial, it must be consistent with the task and learning-relevant (Mulders et al., 2020). In our case, it is

conceivable that the manipulation of the 3D brain and the interactions did not provide enough information for the learner to have a significant effect. Moreover, the literature showed that the benefits of interaction and motor activity could depend on spatial abilities (Gittinger and Wiesche, 2023). A one-sample t-test revealed that participants ( $M = 19.2$ ,  $SD = 8.2$ ) had significantly higher mental rotation scores compared to the normative data for the same age group ( $M = 17.1$ ,  $SD = 6.9$ ) as reported by Albaret and Aubert (1996) ( $t(70) = 2.16$ ,  $p = 0.034$ , 95% CI [17.29, 21.16]). Given that participants in this study have high spatial abilities, they may benefit less from the interactive dimension of VR. Moreover, studies with similar results have mentioned the difficulty of highlighting the effect of interaction on learning with a pre- and post-test design (Chua et al., 2019; Zhang et al., 2019).

Although no significant differences were demonstrated among the three VR conditions, the active condition appears to yield slightly better results than the guided condition. This tenuous difference is supported by significantly higher levels of external regulation in the guided condition compared to the active (and passive) conditions. Similarly, participants in the passive and active groups reported a higher feeling of gaining competence than the guided condition. These findings suggest that a motor-driven, non-cognitive, and constrained activity shifts intrinsic motivation to more extrinsic motivations, resulting in lower self-determination (Ferguson et al., 2020; Oudeyer et al., 2016). In addition, guidance does not seem to reduce extraneous cognitive load as suggested by the literature (Parong and Mayer, 2021; Han et al., 2023). According to Cognitive Load Theory (CLT), guidance typically reduces extraneous cognitive load compared to free exploration, particularly for learners with limited prior knowledge (Han et al., 2023). However, as learners' expertise increases, the benefits of guidance diminish, and in some cases, the imposition of guidance can even increase cognitive load, a phenomenon known as the expertise reversal effect (Kalyuga, 2007). For more advanced learners, self-directed exploration can be equally or more beneficial than guided learning (Jin et al., 2024). In this study, the expected reduction in cognitive load through guidance was not observed, suggesting that the expertise reversal effect might be at play. The medical students in the sample, likely possessing sufficient prior knowledge, may have been able to benefit more from the freedom of exploration in the active condition, allowing them to engage with the VR content more effectively without needing direct guidance. It appears that the benefits of guidance were neutralized, implying that the learning task was appropriate for the students' prior expertise.

Moreover, this study indicates that guided condition showed lower usability than passive VR, suggesting that VR with interaction, especially motor-only, was less easy to use. This usability was also negatively correlated with an irrelevant cognitive load due to interactions and the VR environment. This finding agrees with previous research suggesting that interaction with virtual environment elements could impose extraneous cognitive load (Makransky and Petersen, 2021; Andersen and Makransky, 2021). Similarly, considering the prior knowledge of the subject of the participants in this study, unnecessary and excessive guidance could lead to more user difficulties and cognitive demands (Xue et al., 2024).

The third aim of this study was to explore the relationship between the cognitive and motivational dimension of learning with virtual reality. We demonstrated that intrinsic motivated students reported more perceived learning (i.e., competence), less intrinsic and extraneous cognitive load, and more germane cognitive load. This result supported previous literature showing that cognitive demanding tasks are perceived easier in curiosity-driven learning situations (Sayali et al., 2023; Spitzer et al., 2024). Similarly, low reported EL was associated with high self-efficacy feeling (i.e., intrinsic reward). Feldon et al. (2023) suggest that cognitive demand could directly impair self-efficacy and motivation by acting as a motivational cost. According to EVCT, the cognitive load imposed by the task and the technology represents a cost that is counterbalanced by the benefits of VR on intrinsic motivation (Serki and Bolkan, 2024; Skulmowski and Xu, 2022). Because the task was perceived as less demanding in VR, the cost-benefit balance was positive, leading to more engaged and motivated learning. This finding supports recent work trying to better understand the complex dynamic of these two essential dimensions of learning (Feldon et al., 2019; Makransky and Petersen, 2021; Huang et al., 2022; Skulmowski and Xu, 2022).

However, the idea of a cost-benefit balance inherently implies the presence of metacognitive mediation, which may be the critical point of connection between intrinsic motivation and cognitive load (Hennecke and Kulkarni, 2024). Metacognition, often described as "thinking about thinking" Flavell (1979) is essential in self-directed

learning and it could play a central role in determining how much cognitive effort learners are willing to invest, based on their self-perceived ability to manage that effort and regulate their learning (Seufert, 2018). Learners with stronger metacognitive skills should be better equipped to regulate their cognitive load, which may lead to a more favorable cost-benefit ratio, enabling them to maintain motivation while managing the cognitive demands of the task. Conversely, learners with weaker metacognitive skills may struggle to balance these factors, potentially leading to cognitive overload or diminished motivation. Future studies should consider individual differences in metacognitive strategies, as widely documented in the literature (Karlen et al., 2014). Such differences may play a key role in how learners navigate the cognitive and motivational challenges of VR-based learning.

This study also reveals a strong relationship between germane cognitive load and intrinsic motivation, as reported by participants. Similar patterns of correlations between these variables and others (e.g., extraneous cognitive load, focused attention, aesthetic appeal, and reward) suggest that intrinsic motivation and germane cognitive load may be closely interconnected in how they influence or are influenced by other factors. Both are critical for effective learning processes: intrinsic motivation drives engagement and the desire to learn (Deci and Ryan, 2012), while germane cognitive load represents the mental resources devoted to processing and understanding information (Sweller et al., 2019), which is essential for meaningful learning.

The strong correlation between these two variables indicates that when learners are intrinsically motivated, they are more likely to invest cognitive resources in understanding the material, leading to increased germane load. This aligns with the recent reconceptualization of germane load as "generative processing" (Mayer and Estrella, 2014; Kalyuga, 2011). In this sense, the germane cognitive load is not a separate type of cognitive load, but the mental effort devoted to productive, meaningful learning, an effort that is largely driven by the motivation of the learner. When learners are intrinsically motivated, they engage more deeply with the material, allocating cognitive resources to internalize and make sense of it. This generative processing is precisely the type of cognitive activity that germane load encompasses, highlighting a close connection between motivation and effective learning processes. From this perspective, the germane cognitive load can be seen as the cognitive dimension of engagement (i.e., learning strategies that learners use to regulate their learning), while intrinsic motivation could correspond to the emotional dimension (i.e., interest and enthusiasm about learning) of Reeve and Tseng's framework (Reeve and Tseng, 2011; Wong and Liem, 2022). In VR learning environments, the immersive and interactive nature of the experience can stimulate curiosity and motivation, encouraging learners to process information more deeply.

## 4.2 Practical implications

Some practical contributions emerge from the present study. First, the study demonstrates that VR significantly improves factual and conceptual learning of neuroanatomy compared to traditional teaching methods. This is supported by increased intrinsic motivation and optimal cognitive load management in VR conditions. Educators and institutions can use VR to make learning more effective and engaging, potentially transforming the way complex subjects like neuroanatomy are taught.

To maximize the educational benefits of VR, it is crucial to design VR activities that are consistent with the learning task and easy to use. The study indicates that perceived usability is higher in passive VR conditions than in guided ones, with usability linked to extraneous cognitive load caused by interactions and the virtual environment. Therefore, VR applications should avoid unnecessary motor interactions that can increase difficulty and cognitive load. A well-designed VR application minimizes extraneous cognitive load and enhances the focus and motivation of students.

Despite its educational benefits, active VR can induce vestibular and oculomotor sickness. This suggests a preference for passive VR versions that achieve similar learning outcomes without causing physical discomfort. When implementing VR in educational settings, it is essential to monitor and address these physical discomforts to maintain an effective and comfortable learning environment for students.

Furthermore, the study highlights the importance of considering students' prior knowledge when designing



VR learning experiences, particularly regarding the level of guidance provided. Providing the right amount of guidance can help students, especially those with low subject background knowledge, to take full advantage of VR learning tools. However, further studies with larger participant samples are needed to refine these considerations.

The findings also showed some link between intrinsic motivation and cognitive load, underscoring the importance of considering both intrinsic motivation and cognitive load in the development and evaluation of VR learning tools. In light of several recent studies (Feldon et al., 2023; Sayah et al., 2023; Skulmowski and Xu, 2022), more research, with a larger sample size, is needed to better characterize this relationship. As a preliminary step in a comprehensive study, it would be pertinent to explore the quadratic relationship between cognitive demand and motivation, taking into account the learner's prior knowledge. By better understanding and balancing these factors, developers can create VR applications that are both engaging and cognitively effective, thereby enhancing the overall learning experience.

Finally, one of the critical considerations for applying VR in educational contexts is deciding whether to use it as a complement to or a replacement for traditional learning resources. García-Robles et al. (2024) suggests that VR can be effectively integrated with existing methods, enhancing the overall learning experience without completely discarding conventional approaches. This hybrid model can provide a balanced approach that combines the best of both worlds to optimize learning outcomes.

### 4.3 Limitations and future research

In this experiment, the interactions were confined to moving and resizing the 3D brain and touching and cutting arteries. Although these interactions were relevant to the activity, they may not have been substantial enough to demonstrate a significant positive effect of VR interactions on learning. An explanation highlighted in the study is the high level of prior knowledge among the participants, which could have minimized the effect of interaction. Future studies should explore more complex and varied interactions by varying participants' prior knowledge to better understand their potential impact on learning outcomes.

Another limitation is the potential novelty effect associated with the use of VR technology. This novelty effect, tied to the excitement of using new technology, differs from the novelty effect related to the discovery of new knowledge, which is more directly connected to intrinsic motivation. While the initial excitement of using VR may temporarily boost engagement, the motivation linked to acquiring new information reflects deeper, more intrinsic learning processes. Huang et al. (2021) indicates that the initial enjoyment and engagement with VR persist over time. Although their research showed sustained motivation and engagement throughout the sessions, more evidence is needed to confirm the long-term persistence of these effects. In addition, although a follow-up assessment was planned four months after the intervention, the participation rate was almost zero, making it impossible to evaluate long-term retention of knowledge. Future research should include longitudinal studies to assess whether the novelty effect diminishes and how it impacts learning motivation over extended periods.

Although motivation and engagement indicators were used, the study did not measure the sense of presence in VR. Presence is a critical factor in VR experiences, potentially influencing both engagement and learning. Future research should include presence measurement to provide a more comprehensive understanding of how VR experience affects the learning process.

Conducting the study in a quasi-ecological context introduced variability in classroom perturbations, which may have affected the results. This variability can include differences in classroom environments, distractions, and other uncontrolled factors. Future studies could aim to minimize these variances or better control the study environment to reduce possible confounding factors. Laboratory vs. real-life comparisons could be used to establish the criticality gradient of factors of interest for VR-based learning activities.

Finally, there was variability in the way participants used VR and control conditions, which could have influenced the results. Analyzing specific activities and interactions within the VR environment is essential

to understand how different usage patterns affect learning. Future research should include detailed activity analysis to identify the most effective VR interactions and ensure consistency across conditions.

## 5 Conclusion

This study evaluated the effectiveness of virtual reality (VR) in neuroanatomy education, focusing on learning outcomes, cognitive load, and intrinsic motivation among second-year medical students. The findings indicate that VR significantly enhances anatomical learning compared to traditional methods, particularly by increasing intrinsic motivation and optimizing cognitive load. VR reduced extraneous cognitive load and increased germane cognitive load, while participants in VR conditions reported significantly higher intrinsic motivation due to the immersive and engaging nature of VR.

However, despite expectations, the research found no significant difference in learning outcomes, cognitive load, or motivation between the different levels of interactivity (passive, guided, and active VR conditions). Active learning by manipulating and exploring 3D models of the brain was no more effective for the learning experience than passive observation, indicating that the positive effects of VR are not due to the interaction level.

A key takeaway is the demonstrated relationship between cognitive load and intrinsic motivation according to reliable measurements. This relationship underscores the importance of considering both cognitive and motivational factors in the design and evaluation of educational technologies. Future research should continue to explore this dynamic, particularly by examining how varying levels of cognitive load and motivation interact to impact learning outcomes.

## References

- Albaret, J. M. and Aubert, E. (1996). Etalonnage 15-19 ans du test de rotation mentale de Vandenberg. *Evolutions psychomotrices*, pages 206–209.
- Albus, P., Vogt, A., and Seufert, T. (2021). Signaling in virtual reality influences learning outcome and cognitive load. *Computers & Education*, 166:104154.
- Ali, N., Ullah, S., and Khan, D. (2022). Minimization of students’ cognitive load in a virtual chemistry laboratory via contents optimization and arrow-textual aids. *Education and Information Technologies*, 27(6):7629–7652.
- Andersen, M. S. and Makransky, G. (2021). The validation and further development of a multidimensional cognitive load scale for virtual environments. *Journal of Computer Assisted Learning*, 37(1):183–196. [\\_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/jcal.12478](https://onlinelibrary.wiley.com/doi/pdf/10.1111/jcal.12478).
- Aridan, N., Bernstein-Eliav, M., Gamzo, D., Schmeidler, M., Tik, N., and Tavor, I. (2024). Neuroanatomy in virtual reality: Development and pedagogical evaluation of photogrammetry-based 3D brain models. *Anatomical Sciences Education*, 17(2):239–248. [\\_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/ase.2359](https://onlinelibrary.wiley.com/doi/pdf/10.1002/ase.2359).
- Ayres, P. and Sweller, J. (2014). The Split-Attention Principle in Multimedia Learning. In Mayer, R. E., editor, *The Cambridge Handbook of Multimedia Learning*, Cambridge Handbooks in Psychology, pages 206–226. Cambridge University Press.
- Bakdash, J. Z., Linkenauger, S. A., and Proffitt, D. (2008). Comparing Decision-Making and Control for Learning a Virtual Environment: Backseat Drivers Learn Where They are Going. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 52(27):2117–2121. Publisher: SAGE Publications Inc.
- Barteit, S., Lanfermann, L., Bärnighausen, T., Neuhann, F., and Beiersmann, C. (2021). Augmented, Mixed, and Virtual Reality-Based Head-Mounted Devices for Medical Education: Systematic Review. *JMIR Serious*

*Games*, 9(3):e29080. Company: JMIR Serious Games Distributor: JMIR Serious Games Institution: JMIR Serious Games Label: JMIR Serious Games Publisher: JMIR Publications Inc., Toronto, Canada.

- Bogomolova, K., Hierck, B. P., Looijen, A. E. M., Pilon, J. N. M., Putter, H., Wainman, B., Hovius, S. E. R., and van der Hage, J. A. (2021). Stereoscopic three-dimensional visualisation technology in anatomy learning: A meta-analysis. *Medical Education*, 55(3):317–327. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/medu.14352>.
- Bonawitz, E., Shafto, P., Gweon, H., Goodman, N. D., Spelke, E., and Schulz, L. (2011). The double-edged sword of pedagogy: Instruction limits spontaneous exploration and discovery. *Cognition*, 120(3):322–330.
- Borrelli, M., Leung, B., Morgan, M., Saxena, S., and Hunter, A. (2018). Should drawing be incorporated into the teaching of anatomy? *Journal of Contemporary Medical Education*, 6(2). Publisher: EJManager.
- Boulet, J. R., Jeffries, P. R., Hatala, R. A., Korndorffer, J. R., Feinstein, D. M., and Roche, J. P. (2011). Research regarding methods of assessing learning outcomes. *Simulation in Healthcare: Journal of the Society for Simulation in Healthcare*, 6 Suppl:S48–51.
- Buchner, J. (2023). Beyond Media Comparison: Investigating When and How Learning with Augmented and Virtual Reality Works. *Digital Psychology*, 4(1S):1–3. Number: 1S.
- Buchner, J. and Kerres, M. (2023). Media comparison studies dominate comparative research on augmented reality in education. *Computers & Education*, 195:104711.
- Carr, H. and Koch, H. (1919). The influence of extraneous controls in the learning process. *Psychological Review*, 26(4):287–293. Place: US Publisher: Psychological Review Company.
- Chen, J., Fu, Z., Liu, H., and Wang, J. (2024). Effectiveness of Virtual Reality on Learning Engagement: A Meta-Analysis. *International Journal of Web-Based Learning and Teaching Technologies (IJWLTT)*, 19(1):1–14. Publisher: IGI Global.
- Chrastil, E. R. and Warren, W. H. (2012). Active and passive contributions to spatial learning. *Psychonomic Bulletin & Review*, 19(1):1–23. Place: Germany Publisher: Springer.
- Chua, Y., Sridhar, P. K., Zhang, H., Dissanayake, V., and Nanayakkara, S. (2019). Evaluating IVR in Primary School Classrooms. In *2019 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pages 169–174.
- Clark, R. E. and Feldon, D. F. (2014). Ten common but questionable principles of multimedia learning. *The Cambridge handbook of multimedia learning*, pages 151–173. Publisher: Cambridge University Press New York, NY.
- Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. Routledge.
- Curtin, F. and Schulz, P. (1998). Multiple correlations and bonferroni’s correction. *Biological Psychiatry*, 44(8):775–777.
- Deci, E. L. and Ryan, R. M. (2012). Self-determination theory. *Handbook of theories of social psychology*, 1(20):416–436. Publisher: Sage Publications Ltd.
- Duarte, M. L., Santos, L. R., Guimarães Júnior, J. B., and Peccin, M. S. (2020). Learning anatomy by virtual reality and augmented reality. A scope review. *Morphologie*, 104(347):254–266.
- Feldon, D. F., Brockbank, R., and Litson, K. (2023). Direct effects of cognitive load on self-efficacy during instruction. *Journal of Educational Psychology*, pages No Pagination Specified–No Pagination Specified. Place: US Publisher: American Psychological Association.

- Feldon, D. F., Callan, G., Juth, S., and Jeong, S. (2019). Cognitive Load as Motivational Cost. *Educational Psychology Review*, 31(2):319–337.
- Ferguson, C., van den Broek, E. L., and van Oostendorp, H. (2020). On the role of interaction mode and story structure in virtual reality serious games. *Computers & Education*, 143:103671.
- Flake, J. K., Barron, K. E., Hulleman, C., McCoach, B. D., and Welsh, M. E. (2015). Measuring cost: The forgotten component of expectancy-value theory. *Contemporary Educational Psychology*, 41:232–244.
- Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist*, 34(10):906–911. Place: US Publisher: American Psychological Association.
- Fontaine, G., Cossette, S., Maheu-Cadotte, M.-A., Mailhot, T., Lavoie, P., Gagnon, M.-P., Dubé, V., and Côté, J. (2019). Traduction, adaptation et évaluation psychométrique préliminaire d’une mesure d’engagement et d’une mesure de charge cognitive en contexte d’apprentissage numérique. *Pédagogie Médicale*, 20(2):79–90. Number: 2 Publisher: EDP Sciences.
- Fuchs, P. (2006). *Le traité de la réalité virtuelle*, volume 2. Presses des MINES.
- García-Robles, P., Cortés-Pérez, I., Nieto-Escámez, F. A., García-López, H., Obrero-Gaitán, E., and Osuna-Pérez, M. C. (2024). Immersive virtual reality and augmented reality in anatomy education: A systematic review and meta-analysis. *Anatomical Sciences Education*, 17(3):514–528.
- Gittinger, M. and Wiesche, D. (2023). Systematic review of spatial abilities and virtual reality: The role of interaction. *Journal of Engineering Education*, n/a(n/a). \_eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/jee.20568>.
- Glaser, N. and Moore, S. (2023). Redefining Immersive Technology Research: Beyond Media Comparisons to Holistic Learning Approaches. *Digital Psychology*, 4(1S):4–8. Number: 1S.
- Großmann, N. and Wilde, M. (2020). Promoting Interest by Supporting Learner Autonomy: the Effects of Teaching Behaviour in Biology Lessons. *Research in Science Education*, 50(5):1763–1788.
- Guay, F., Vallerand, R. J., and Blanchard, C. (2000). On the Assessment of Situational Intrinsic and Extrinsic Motivation: The Situational Motivation Scale (SIMS). *Motivation and Emotion*, 24(3):175–213.
- Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, 66(1):64–74. \_eprint: [https://pubs.aip.org/aapt/ajp/article-pdf/66/1/64/12040304/64\\_1\\_online.pdf](https://pubs.aip.org/aapt/ajp/article-pdf/66/1/64/12040304/64_1_online.pdf).
- Hall, S., Stephens, J., Parton, W., Myers, M., Harrison, C., Elmansouri, A., Lowry, A., and Border, S. (2018). Identifying Medical Student Perceptions on the Difficulty of Learning Different Topics of the Undergraduate Anatomy Curriculum. *Medical Science Educator*, 28.
- Hamilton, D., McKechnie, J., Edgerton, E., and Wilson, C. (2021). Immersive virtual reality as a pedagogical tool in education: a systematic literature review of quantitative learning outcomes and experimental design. *Journal of Computers in Education*, 8(1):1–32.
- Han, J., Liu, G., and Zheng, Q. (2023). Prior knowledge as a moderator between signaling and learning performance in immersive virtual reality laboratories. *Frontiers in Psychology*, 14. Publisher: Frontiers.
- Han, J., Zheng, Q., and Ding, Y. (2021). Lost in virtual reality? Cognitive load in high immersive VR environments. *Journal of Advances in Information Technology*, 12(4).
- Harris, C. and Sun, B. (2022). Assessing the Effect of Interactivity Design In VR Based Second Language Learning Tool. In *2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 18–25. ISSN: 1554-7868.

- Hennecke, M. and Kulkarni, P. (2024). Metacognitive knowledge about self-control. *Current Opinion in Psychology*, 59:101861.
- Honebein, P. C. and Reigeluth, C. M. (2021). To prove or improve, that is the question: the resurgence of comparative, confounded research between 2010 and 2019. *Educational Technology Research and Development*, 69(2):465–496.
- Huang, W., Roscoe, R. D., Craig, S. D., and Johnson-Glenberg, M. C. (2022). Extending the Cognitive-Affective Theory of Learning with Media in Virtual Reality Learning: A Structural Equation Modeling Approach. *Journal of Educational Computing Research*, 60(4):807–842. Publisher: SAGE Publications Inc.
- Huang, W., Roscoe, R. D., Johnson-Glenberg, M. C., and Craig, S. D. (2021). Motivation, engagement, and performance across multiple virtual reality sessions and levels of immersion. *Journal of Computer Assisted Learning*, 37(3):745–758. \_eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/jcal.12520>.
- James, K. H., Humphrey, G. K., Vilis, T., Corrie, B., Baddour, R., and Goodale, M. A. (2002). “Active” and “passive” learning of three-dimensional object structure within an immersive virtual reality environment. *Behavior Research Methods, Instruments, & Computers*, 34(3):383–390.
- Jin, S., Zhong, Z., Li, K., and Kang, C. (2024). Investigating the effect of guided inquiry on learners with different prior knowledge in immersive virtual environments. *Education and Information Technologies*.
- Kalyuga, S. (2007). Expertise Reversal Effect and Its Implications for Learner-Tailored Instruction. *Educational Psychology Review*, 19(4):509–539.
- Kalyuga, S. (2011). Cognitive Load Theory: How Many Types of Load Does It Really Need? *Educational Psychology Review*, 23(1):1–19.
- Karlen, Y., Merki, K. M., and Ramseier, E. (2014). The effect of individual differences in the development of metacognitive strategy knowledge. *Instructional Science*, 42(5):777–794.
- Kourtesis, P., Linnell, J., Amir, R., Argelaguet, F., and MacPherson, S. E. (2023). Cybersickness in Virtual Reality Questionnaire (CSQ-VR): A Validation and Comparison against SSQ and VRSQ. *Virtual Worlds*, 2(1):16–35.
- Lehikko, A., Nykänen, M., Lukander, K., Uusitalo, J., and Ruokamo, H. (2024). Exploring interactivity effects on learners’ sense of agency, cognitive load, and learning outcomes in immersive virtual reality: A mixed methods study. *Computers & Education: X Reality*, 4:100066.
- Mahmoud, K., Harris, I., Yassin, H., Hurkxkens, T. J., Matar, O. K., Bhatia, N., and Kalkanis, I. (2020). Does Immersive VR Increase Learning Gain When Compared to a Non-immersive VR Learning Experience? In Zaphiris, P. and Ioannou, A., editors, *Learning and Collaboration Technologies. Human and Technology Ecosystems*, pages 480–498, Cham. Springer International Publishing.
- Makransky, G. and Lilleholt, L. (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.
- Makransky, G. and Petersen, G. B. (2021). The Cognitive Affective Model of Immersive Learning (CAMIL): a Theoretical Research-Based Model of Learning in Immersive Virtual Reality. *Educational Psychology Review*, 33(3):937–958.
- Makransky, G., Petersen, G. B., and Klingenberg, S. (2020). Can an immersive virtual reality simulation increase students’ interest and career aspirations in science? *British Journal of Educational Technology*, 51(6):2079–2097. \_eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/bjet.12954>.

- Makransky, G., Terkildsen, T. S., and Mayer, R. E. (2019). Adding immersive virtual reality to a science lab simulation causes more presence but less learning. *Learning and Instruction*, 60:225–236.
- Mayer, L., Süncksen, M., Reinhold, S., and Teistler, M. (2023). vimiLab – Teaching Medical Imaging and Radiological Anatomy in a Virtual Reality Environment. In *2023 IEEE 11th International Conference on Serious Games and Applications for Health (SeGAH)*, pages 1–5. ISSN: 2573-3060.
- Mayer, R. E. and Estrella, G. (2014). Benefits of emotional design in multimedia instruction. *Learning and Instruction*, 33:12–18.
- Meade, M. E., Meade, J. G., Sauzeon, H., and Fernandes, M. A. (2019). Active Navigation in Virtual Environments Benefits Spatial Memory in Older Adults. *Brain Sciences*, 9(3):47. Number: 3 Publisher: Multidisciplinary Digital Publishing Institute.
- Miguel-Alonso, I., Checa, D., Guillen-Sanz, H., and Bustillo, A. (2024). Evaluation of the novelty effect in immersive Virtual Reality learning experiences. *Virtual Reality*, 28(1):27.
- Moreau, G., Arnaldi, B., and Guitton, P. (2018). *Virtual Reality, Augmented Reality: myths and realities*. Computer engineering series. ISTE.
- Mulders, M. (2023). Learning about Victims of Holocaust in Virtual Reality: The Main, Mediating and Moderating Effects of Technology, Instructional Method, Flow, Presence, and Prior Knowledge. *Multimodal Technologies and Interaction*, 7(3):28. Number: 3 Publisher: Multidisciplinary Digital Publishing Institute.
- Mulders, M., Buchner, J., and Kerres, M. (2020). A Framework for the Use of Immersive Virtual Reality in Learning Environments. *International Journal of Emerging Technologies in Learning (iJET)*, 15(24):208–224. Number: 24.
- Murayama, K., FitzGibbon, L., and Sakaki, M. (2019). Process Account of Curiosity and Interest: A Reward-Learning Perspective. *Educational Psychology Review*, 31(4):875–895.
- Newman, H. J., Meyer, A. J., Wilkinson, T. J., Pather, N., and Carr, S. E. (2022). Technology enhanced neuroanatomy teaching techniques: A focused BEME systematic review of current evidence: BEME Guide No. 75. *Medical Teacher*, pages 1–12.
- Oudeyer, P. Y., Gottlieb, J., and Lopes, M. (2016). Chapter 11 - Intrinsic motivation, curiosity, and learning: Theory and applications in educational technologies. In Studer, B. and Knecht, S., editors, *Progress in Brain Research*, volume 229 of *Motivation*, pages 257–284. Elsevier.
- Paas, F., Tuovinen, J. E., van Merriënboer, J. J. G., and Aubteen Darabi, A. (2005). A motivational perspective on the relation between mental effort and performance: Optimizing learner involvement in instruction. *Educational Technology Research and Development*, 53(3):25–34.
- Parong, J. and Mayer, R. E. (2018). Learning science in immersive virtual reality. *Journal of Educational Psychology*, 110(6):785–797. Place: US Publisher: American Psychological Association.
- Parong, J. and Mayer, R. E. (2021). Cognitive and affective processes for learning science in immersive virtual reality. *Journal of Computer Assisted Learning*, 37(1):226–241. \_eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/jcal.12482>.
- Peart, D. J. (2022). Hand drawing as a tool to facilitate understanding in undergraduate human biology: a critical review of the literature and future perspectives. *Studies in Science Education*, 58(1):81–93. Publisher: Routledge \_eprint: <https://doi.org/10.1080/03057267.2021.1913321>.
- Peters, D., Calvo, R. A., and Ryan, R. M. (2018). Designing for Motivation, Engagement and Wellbeing in Digital Experience. *Frontiers in Psychology*, 9. Publisher: Frontiers.

- Petersen, G. B., Petkakis, G., and Makransky, G. (2022). A study of how immersion and interactivity drive VR learning. *Computers & Education*, 179:104429.
- Poupard, M., Larrue, F., Sauz on, H., and Tricot, A. (2024). A systematic review of immersive technologies for education: Learning performance, cognitive load and intrinsic motivation. *British Journal of Educational Technology*. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/bjet.13503>.
- Reeve, J. and Tseng, C.-M. (2011). Agency as a fourth aspect of students’ engagement during learning activities. *Contemporary Educational Psychology*, 36(4):257–267.
- Ryan, R. M. and Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contemporary Educational Psychology*, 61. Place: Netherlands Publisher: Elsevier Science.
- Sayali, C., Heling, E., and Cools, R. (2023). Learning progress mediates the link between cognitive effort and task engagement. *Cognition*, 236:105418.
- Schneider, S., Nebel, S., Beege, M., and Rey, G. D. (2018). The autonomy-enhancing effects of choice on cognitive load, motivation and learning with digital media. *Learning and Instruction*, 58:161–172.
- Schomaker, J. and Wittmann, B. C. (2021). Effects of active exploration on novelty-related declarative memory enhancement. *Neurobiology of Learning and Memory*, 179:107403.
- Serki, N. and Bolkan, S. (2024). The effect of clarity on learning: impacting motivation through cognitive load. *Communication Education*, 73(1):29–45. Publisher: Routledge eprint: <https://doi.org/10.1080/03634523.2023.2250883>.
- Seufert, T. (2018). The interplay between self-regulation in learning and cognitive load. *Educational Research Review*, 24:116–129.
- Sinha, S., DeYoung, V., Nehru, A., Brewer-Deluce, D., and Wainman, B. C. (2023). Determinants of Learning Anatomy in an Immersive Virtual Reality Environment — A Scoping Review. *Medical Science Educator*, 33(1):287–297.
- Sivashankar, Y., He, P., Sauz on, H., and Fernandes, M. A. (2024). Motoric engagement, but not decision-making, during encoding influences route memory. *Spatial Cognition & Computation*, 0(0):1–26. Publisher: Taylor & Francis eprint: <https://doi.org/10.1080/13875868.2024.2359929>.
- Skulmowski, A. and Xu, K. M. (2022). Understanding Cognitive Load in Digital and Online Learning: a New Perspective on Extraneous Cognitive Load. *Educational Psychology Review*, 34(1):171–196.
- Spitzer, M., Strittmatter, Y., Marti, M., Schumacher, A., and Bardach, L. (2024). Curiosity overpowers cognitive effort avoidance tendencies.
- Sweller, J., van Merri nboer, J. J. G., and Paas, F. (2019). Cognitive Architecture and Instructional Design: 20 Years Later. *Educational Psychology Review*, 31(2):261–292. Number: 2.
- Taylor, L., Dyer, T., Al-Azzawi, M., Smith, C., Nzeako, O., and Shah, Z. (2022). Extended reality anatomy undergraduate teaching: A literature review on an alternative method of learning. *Annals of Anatomy - Anatomischer Anzeiger*, 239:151817.
- Ten, A., Kaushik, P., Oudeyer, P.-Y., and Gottlieb, J. (2021). Humans monitor learning progress in curiosity-driven exploration. *Nature Communications*, 12(1):5972. Publisher: Nature Publishing Group.
- Vandenberg, S. G. and Kuse, A. R. (1978). Mental Rotations, a Group Test of Three-Dimensional Spatial Visualization. *Perceptual and Motor Skills*, 47(2):599–604. eprint: <https://doi.org/10.2466/pms.1978.47.2.599>.

- Wang, J., Li, W., Dun, A., Zhong, N., and Ye, Z. (2024). 3D visualization technology for Learning human anatomy among medical students and residents: a meta- and regression analysis. *BMC Medical Education*, 24(1):461.
- Wen, P., Lu, F., and Mohamad Ali, A. Z. (2024). Using attentional guidance methods in virtual reality laboratories reduces students' cognitive load and improves their academic performance. *Virtual Reality*, 28(2):110.
- Won, M., Ungu, D. A. K., Matovu, H., Treagust, D. F., Tsai, C.-C., Park, J., Mocerino, M., and Tasker, R. (2023). Diverse approaches to learning with immersive Virtual Reality identified from a systematic review. *Computers & Education*, 195:104701.
- Wong, Z. Y. and Liem, G. A. D. (2022). Student Engagement: Current State of the Construct, Conceptual Refinement, and Future Research Directions. *Educational Psychology Review*, 34(1):107–138.
- Xue, J., Liu, J., Yuan, Q., Yao, Z., Xu, J., and Pan, Z. (2024). Experimental guidance and feedback via operation intention prediction with effect analysis in chemistry labs. *Education and Information Technologies*.
- Zhang, L., Bowman, D. A., and Jones, C. N. (2019). Exploring Effects of Interactivity on Learning with Interactive Storytelling in Immersive Virtual Reality. In *2019 11th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games)*, pages 1–8. ISSN: 2474-0489.