A systematic review of immersive technologies for education: learning performance, cognitive load, and intrinsic motivation.

M. Poupard[1,2,3], F. Larrue[2], H. Sauzéon [1,3], and A. Tricot[4]

[1] Flowers Team, Inria centre at the University of Bordeaux, Talence,France

[2] Centre Technologique Nouvelle-Aquitaine du numérique (CATIE), Talence, France´

[3] University of Bordeaux, BPH Bordeaux Population Health, Bordeaux, France

[4] Université Paul Valéry Montpellier 3, Epsylon EA 4556, Montpellier, France´

Abstract

Immersive technologies are assumed to have many benefits for learning due to their potential positive impact on optimizing learners' cognitive load and fostering intrinsic motivation. However, despite promising results, the findings regarding the actual impact on learning remain inconclusive, raising questions about the determinants of efficacy.

To address these gaps, we conducted a PRISMA systematic review to investigate the contributions and limitations of Virtual Reality (VR) and Augmented Reality (AR) in learning, specifically by examining their effects on cognitive load and intrinsic motivations. Through the application of an analytical grid, we systematically classified the impact of VR/AR on the causal relationship between learning performance (i.e., objective learning improvement) and cognitive load or motivation, while respecting the fundamental assumptions of the main theories related to these factors.

Analyzing 36 studies, the findings reveal that VR, often causing extraneous load, hinders learning, particularly among novices. In contrast, AR optimizes cognitive load, proving beneficial for novice learners but demonstrating less effectiveness for intermediate learners. The effects on intrinsic motivation remain inconclusive, likely due to variations in measurement methods.

The review underscores the need for detailed, sophisticated evaluations and comprehensive frameworks that consider both cognitive load and intrinsic motivation to improve understanding of the impact of immersive technologies on learning.

Keywords

- Virtual Reality
- Augmented Reality
- Learning
- Cognitive load
- Intrinsic Motivation
- Curiosity

Practitioner Notes

What is Known

- Virtual and augmented reality show promise for education, but findings are inconsistent.
- Existing studies suggest that augmented reality optimizes learners' cognitive load.
- The literature often asserts that VR and AR are expected to enhance learning motivation.

Adding

- VR introduces unnecessary cognitive load, while AR proves effective for learning performance and cognitive load, particularly for novice learners.
- The impact of AR and VR on motivation to learn is unclear.
- Our analytical grid offers a comprehensive framework for assessing the effects of AR and VR on learning outcomes.

Implications

- AR is more suitable than VR for education concerning cognitive load.
- The cost/benefit balance of VR should be carefully considered before implementation, especially for novice learners.
- Rigorous studies on motivation to learn in AR and VR contexts are essential.

1. Introduction

Immersive technologies, specifically Virtual Reality (VR) and Augmented Reality (AR), are gaining recognition as potential solutions to address the significant educational challenges of the 21st century. Defined as systems providing high sensory richness and substantial interactivity, VR creates a completely immersive, artificial environment, while AR integrates virtual elements with our perception of the physical world, enhancing real-world experiences [6, 74].

The growing interest in VR/AR technologies arises from their novel interactions, simulations, and representations of abstract elements, offering new possibilities for learning. These immersive properties are believed to enhance intrinsic motivation and engage cognitive resources during learning activities [67]. The Cognitive-Affective Theory of Learning with Media (CATLM) suggests that the high degree of sensory richness and interactivity provided by immersive technologies can foster intrinsic motivation while reducing the mental effort required to learn [28, 54]. Additionally, the Cognitive Affective Model of Immersive Learning (CAMIL) posits that the sense of presence and agency resulting from the high degree of immersion in these technologies promotes intrinsic motivation but could also lead to higher extraneous cognitive load due to the richness and complexity of information [50]. According to these models, VR/AR technologies are considered more attractive and engaging for learners due to their ability to create immersive experiences that captivate attention and enhance learning outcomes. Immersive environments allow learners to actively engage with content through interactive simulations and representations, leading to optimal and sustainable flow experiences [73].

Despite these promising attributes, the effects of VR and AR on learning performance remain contradictory. Some systematic reviews report VR as yielding superior learning gains compared to conventional conditions [86], while others suggest its promise but highlight a lack of robust evidence for its learning effectiveness [15]. Similarly, a systematic review found that AR did not consistently lead to learning gains [58]. Additionally, although cognitive and motivational dimensions are increasingly considered in evaluating immersive technologies, there has been no specific systematic review focusing on the impact of VR on cognitive load or the effect of AR on learner motivation. Moreover, these two factors are rarely simultaneously examined in the evaluation of immersive technologies.

Therefore, the aim of our study is to conduct a systematic review of the literature on immersive technologies, cognitive load, and intrinsic motivation to determine how and to what extent VR and AR influence learning. To elucidate the relationship between these variables, we have considered the following research questions:

1. How do VR and AR technologies affect learning performance through the optimization of cognitive load due to their immersive properties?
2. How do VR and AR technologies affect learning performance through increased intrinsic motivation due to their immersive properties?
3. Do causal relationships between cognitive load and intrinsic motivation provide a better understanding of the effectiveness of immersive technologies in learning?

Previous systematic reviews demonstrated considerable diversity in the pedagogical context across studies examining immersive technologies for learning [25, 66]. Hence, for each research question, we examined the influence of pedagogical variables, including the type of knowledge taught, learners' prior knowledge, and the educational levels under consideration. Additionally, as detailed in the methodology section, this systematic review concentrates on VR and AR applications from higher education through vocational training, excluding children and the elderly.

2. Background

To assess the effectiveness of instruction, traditional measures of learning outcomes are employed, which include objective metrics (e.g., accuracy and speed in learning performance) and subjective evaluations of the learner's experience (e.g., judgment of learning, cognitive, and affective perceptions). In contemporary theoretical frameworks that underpin instructional design, both types of measures are often integrated to emphasize the dynamic aspects of the flow state during learning. Among these frameworks, Cognitive Load Theory [78] and intrinsic motivation theory [69] have been extensively studied over the past two decades.

2.1 Cognitive Load

Cognitive Load Theory (CLT) posits that learning through instruction, like many tasks, requires cognitive resources in working memory [79]. 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 intrinsic cognitive load, which is relevant to learning, and extraneous cognitive load, which is non-relevant [33].

Intrinsic cognitive load depends on the task's complexity (number and interactivity of elements) and the learner's prior knowledge, representing the cognitive resources necessary for learning. Extraneous cognitive load involves processing non-relevant information during learning, such as decorative elements on learning materials. Both types of loads are additive and represent the total cognitive load. Notably, an earlier version of CLT introduced a third type of cognitive load, germane cognitive load, related to knowledge acquisition [78].

The findings from studies on immersive technologies and cognitive load are diverse. On one hand, Augmented Reality (AR) is regarded as adhering to CLT principles, particularly spatial continuity, leading to a reduction in learners' extraneous load [9, 58]. On the other hand, although no systematic review is available, Virtual Reality (VR) has been reported to induce additional non-relevant load due to the complexity of controls and/or the richness of stimuli in 3D environments [50].

2.2 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" [69]. It represents a natural inclination associated with exploratory behaviors and spontaneous interest, often characterized by states of 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. The literature consistently demonstrates that students who are intrinsically motivated tend to learn more, achieve better academic results, exhibit improved retention rates in short-term and long-term memory, and demonstrate greater persistence when facing challenges [59].

The significance of intrinsic motivation in learning behaviors is well-explained by the model of learning progress, illustrating a positive feedback loop between intrinsic motivation and knowledge acquisition. According to this model, learners experience intrinsic rewards when acquiring new knowledge, fostering curiosity-driven learning behaviors [59, 57].

Several studies have aimed to demonstrate that immersive technologies enhance learners' intrinsic motivation [18, 28]. In a systematic review, Huang et al. [28] asserted that one of the main advantages of Augmented Reality (AR) and Virtual Reality (VR) is their ability to stimulate learners' motivation for learning. However, a more recent systematic review concluded that outcomes were mixed, contingent on instructional settings and pursued pedagogical objectives [58]. Such discrepancies in results could be attributed to methodological limitations in current empirical studies, such as small sample sizes and inconsistent use of standardized measures [15].

2.3 Interactions between Cognitive Load and Intrinsic Motivation

Several recent studies have sought to examine the links between cognitive load and learners' motivational states to better understand their mediating effects on learning performance [19, 28, 50, 73]. These studies highlight two theoretical assumptions. First, engagement and intrinsic motivation may result in a perceived reduction in cognitive load and may compensate for the cognitive demands associated with the complexity of learning technologies [73]. Second, cognitive load may influence learners' motivational beliefs. For instance, a complex task, associated with a perceived high cognitive demand, leads to decreased learner engagement. In other words, the cognitive demand of a task can be perceived as a motivational cost, leading to a decrease in cognitive effort allocated to the task [19]. However, studies on the links between these two essential ingredients for learning are limited, and thus, the connections remain unclear or poorly articulated. Nevertheless, analyzing these links is crucial to shed light on the two theoretical assumptions and provide new insights to explain the disparate effects of immersive technologies on learning performance.

3. Method

To address the questions raised in this literature review, the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) method was applied to guide the systematic examination process [61]. The protocol was registered in the International Prospective Register of Systematic Reviews (https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42022335531).

3.1 Data source and search strategy

Initial searches were conducted between April and June 2022 using the Scopus, Web of Science, and PsycInfo databases. Given the interdisciplinary nature of this systematic review, spanning psychology, human-computer interaction (HCI), and computer science, Scopus and Web of Science were chosen for their comprehensive coverage across various fields, including education, psychology, and technology. PsycInfo, recognized as a leading database in psychology, was also included to ensure comprehensive coverage of relevant studies within the field. Utilizing multiple databases enhances the depth of the review, ensuring a more thorough analysis of the literature. The following query was used consistently with the research question: "Immersive technologies AND Learning AND (Cognitive load OR Intrinsic motivation)."

Table 1: Search keywords used in the identification stage.

Categories	Research Keywords

Immersive technologies	Immersion; immersive; virtual reality; augmented reality; mixed reality; virtual environment; virtual world; digital world; virtual; head mounted display
Learning	Learning; training; schooling; student; higher education; education; teaching; instruction
Cognitive load	Cognitive load; cognitive load theory; dual task; working memory; overload; germane load; germane cognitive load; intrinsic load; intrinsic cognitive load; extraneous load; extraneous cognitive load
Intrinsic motivation	Intrinsic motivation; epistemic curiosity; motivational beliefs; interest + intrinsic motivation; curiosity + intrinsic motivation

3.2 Inclusion and exclusion criteria

We ensured that the articles included in the final analysis met all the inclusion criteria described in Table 2. The PICO framework was employed to define these eligibility criteria [71]:

- *Population*: To ensure consistency in the review of the literature, studies involving K12 students and older adults were excluded. This decision was made for several reasons. Firstly, the incomplete development of working memory [13] and the minimum age requirement of 13 imposed by VR helmet manufacturers make it challenging to compare results from studies involving children with those from other populations. Similarly, research on the use of immersive technologies among older adults has revealed user constraints and a certain heterogeneity in the ability to use these devices specific to this population [53, 70]. Excluding these populations helps to maintain coherence in the analysis and interpretation of the literature.
- *Investigated conditions*: Investigated conditions: Studies included in the review had to utilize a virtual reality and/or augmented reality system. Virtual reality immerses users in a computer-generated environment, while augmented reality overlays digital content onto the real world, enhancing the user's perception of their surroundings. In both cases, a high level of sensory richness coupled with a high degree of interactivity was required [77].

- *Comparison condition:* To ensure the use of reliable, high-quality sources of information, studies had to include a control group in addition to the experimental groups (i.e., a randomized controlled trial or a non-randomized controlled trial) [88]. As indicated by Kanyongo et al. [34], a minimum of 12 participants per group is typically required to ensure statistical power and sufficient reliability. However, to maintain inclusiveness while still ensuring the reliability of the included studies, only those with a minimum of ten participants per group were considered in the review. This adjustment allows for a broader range of studies to be included while still upholding a reasonable standard of reliability.
- *Outcomes*: This review required a measure of learning performance. Therefore, studies measuring only the learning experience (perceived learning, enjoyment, etc.) were excluded. In addition to a measure of learning, studies had to include quantitative measures of cognitive load and/or intrinsic motivation.

There were no restrictions on the publication date, but the included studies needed to be written in English and published in a peer-reviewed journal or conferences.

Table 2: Inclusion and exclusion criteria.

Inclusion Criteria	Exclusion Criteria
Must use immersive technologies (VR or AR)	Immersive technologies were not used
Must be about learning and measure learning gain	Learning was not the main goal of using immersive technologies
Must consider/measure cognitive load or motivational variables (intrinsic motivation, engagement, curiosity, etc.)	Neither cognitive load nor intrinsic motivation were measured
	Non true experimental design
Must adopt a true experimental design and more than ten participants per group	Participants were K12 or elderly
	Not journal or conference papers (e.g., books, thesis, etc.)

3.3 Identification and screening process

The initial search yielded 2800 references, to which were added two articles identified by the authors
The initial search yielded 2800 references, to which were added two articles identified by the authors
before the identification phase (see Fig. 1). The two articles added by the authors [48, 65] met all the eligibility criteria but did not appear in the database search results. After eliminating duplicates, the first screening phase, based on titles and abstracts, was conducted. Three authors carefully assessed this initial screening phase, covering 10% of the results from the selection phase. This control selection of 280 papers included a random sample of articles rejected by the first author, along with all the references included by the first author. In cases of uncertainty or disagreement among the evaluators, consensus was reached through deliberation.

The examination of the full text of the 101 articles eligible for the second phase of full-text screening led to the inclusion of 30 references (31 studies) for the systematic review phase. The PRISMA flowchart outlines the main reasons for exclusions. Notably, the end month of the sample period for publication was June 2022 in this initial search phase. To incorporate all studies from the year 2022, a second search phase was executed in November 2023. This subsequent identification phase resulted in the inclusion of five additional studies from a pool of 240 identified records (see Fig. 1). Finally, a total of 35 papers (36 studies) were included in this systematic review, covering the period from March 2016 to November 2022. Such results support the criteria relevance for the studies selection process.

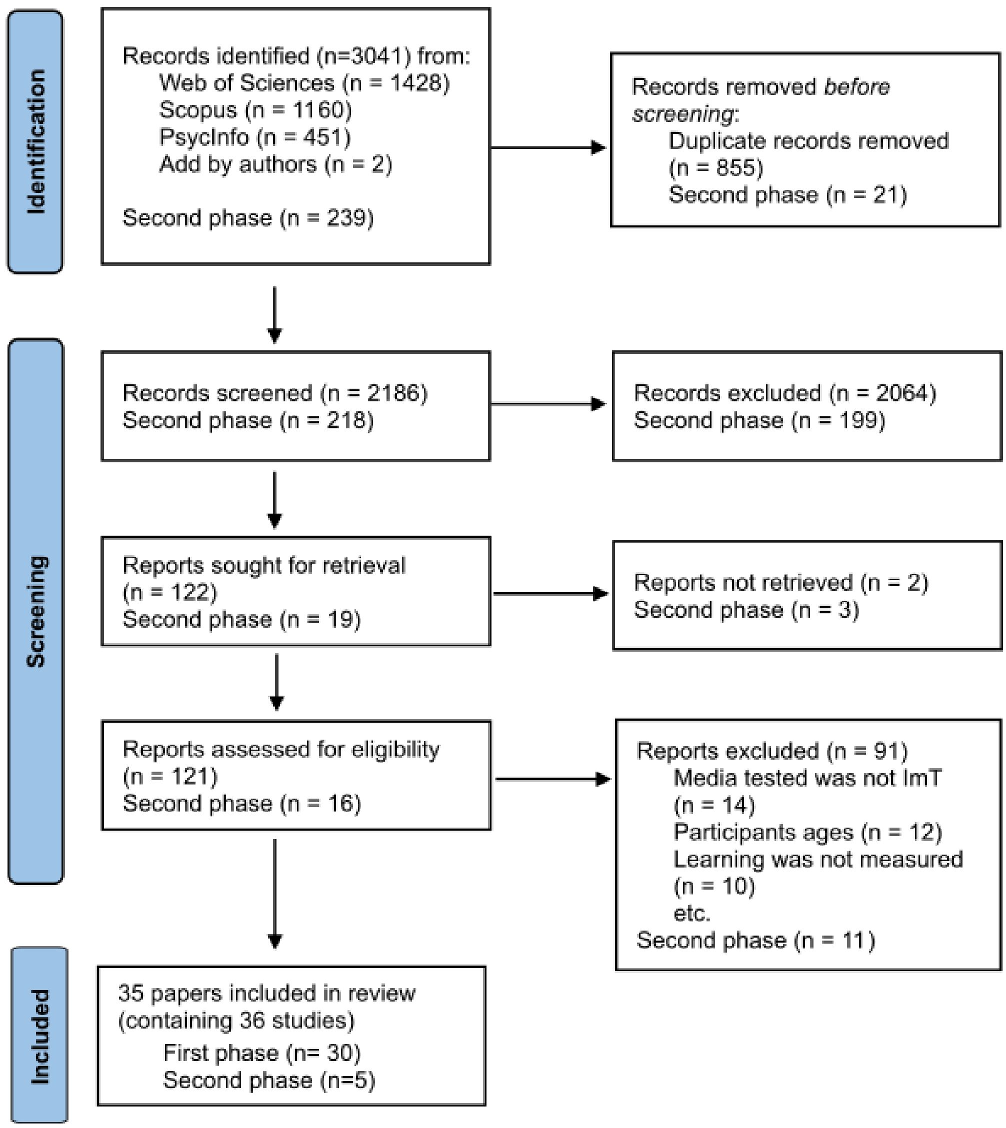

Figure 1: PRISMA flow diagram of the systematic review.

3.4 Data coding and grid analysis

The first author read all the included studies and reported information on a coding sheet presented in Table 3. The data analysis involved descriptive statistics and content analysis.

Table 3: Coding sheet used to extract and analyze results from included studies.

Category	Items
Identification	Authors, title, year of publication, type of publication
Study type	Quantitative, qualitative, randomization

Study context	Sample size, average age, number of participants per groups, learners' level of expertise
Immersive technologies	Type of technology, comparison media
Learning	Type of outcomes measured, knowledge type (based on the taxonomy of Anderson and Krathwohl (2001)), field of learning, main effect, effect size
Cognitive load and intrinsic motivation	Definition, measure, main effects, effect size

One source of confusion in immersive technology research arises from different result patterns leading authors to similar conclusions. For instance, an observed improvement in cognitive load measures due to immersive technology use is sometimes considered evidence of enhanced cognitive load, even without a corresponding improvement in learning performance. This interpretation deviates from the predictions of cognitive load theory (CLT), which posits that optimizing relevant cognitive load (ICL) and minimizing extraneous cognitive load (ECL) enhance learning performance. To address this issue and analyze the impact of immersive technologies on cognitive load, we designed and utilized an analysis grid that considers learning outcomes.

According to this grid (Figure 2.a), results can be interpreted as follows:

- Positive: An increase in learning performance accompanied by a positive effect on cognitive load (i.e., an increase in ICL and/or a decrease in ECL).
- Neutral: No discernible effect of VR/AR on learning performance and cognitive load measures.
- Negative: Immersive technology use leads to a decrease in learning performance associated with a negative effect on cognitive load (i.e., a decrease in ICL and/or an increase in ECL).

Additionally, in line with CLT, three contradictory outcome sets are recognized (Figure 2.b):

- Situations where learning gains occur alongside an increase in ECL and/or a decrease in ICL.
- Instances where there is significant cognitive load variation without changes in learning outcomes (and vice versa).
- Scenarios where immersive technology enhances ICL, decreases ECL, but paradoxically results in deteriorated learning performance.

Consistency is also crucial for intrinsic motivation-based learning models, where a positive correlation between learning performance and intrinsic motivation (IM) is expected. Applying this principle, three consistent result patterns (positive, neutral, and negative effects of technology on both learning performance and IM measures, Figure 2b) and three inconsistent patterns (Figure 2b) are distinguished. Inconsistent patterns include situations where there is a learning gain despite decreased IM, situations with no learning gain despite changes in IM, and situations where there is a learning loss despite increased IM.

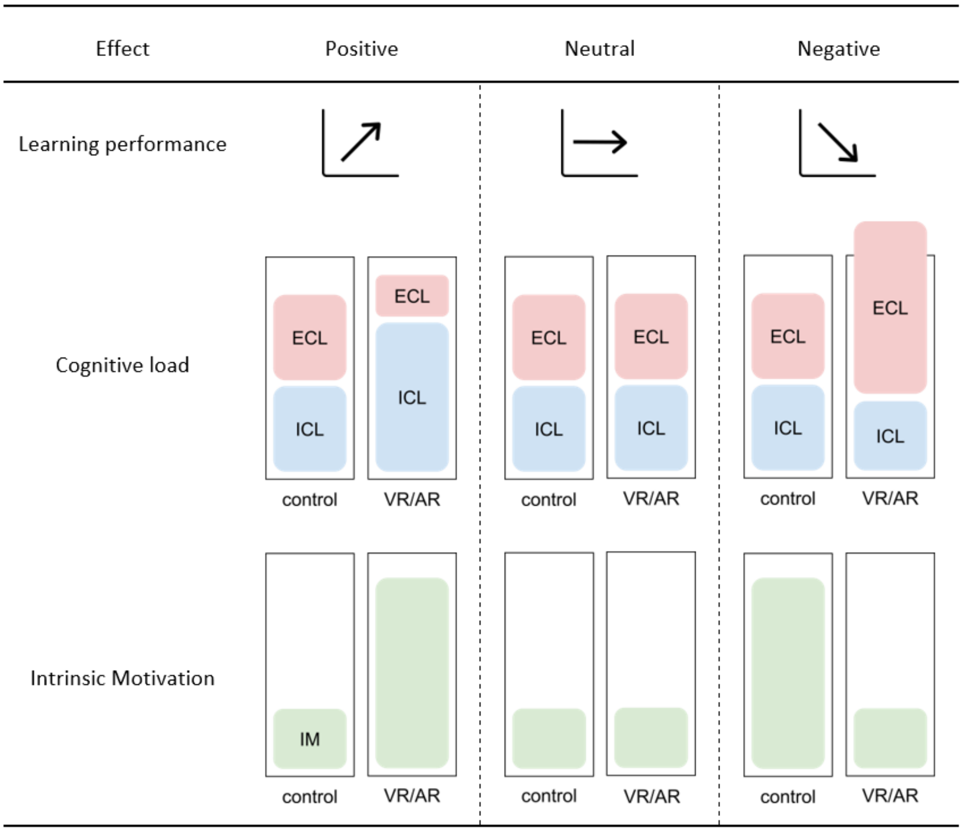

Figure 2.a: Grid for analyzing results of immersive technologies on learning considering cognitive load and intrinsic motivation results when they are consistent with the theory. ECL = extraneous cognitive load, ICL = intrinsic cognitive load, IM = intrinsic motivation

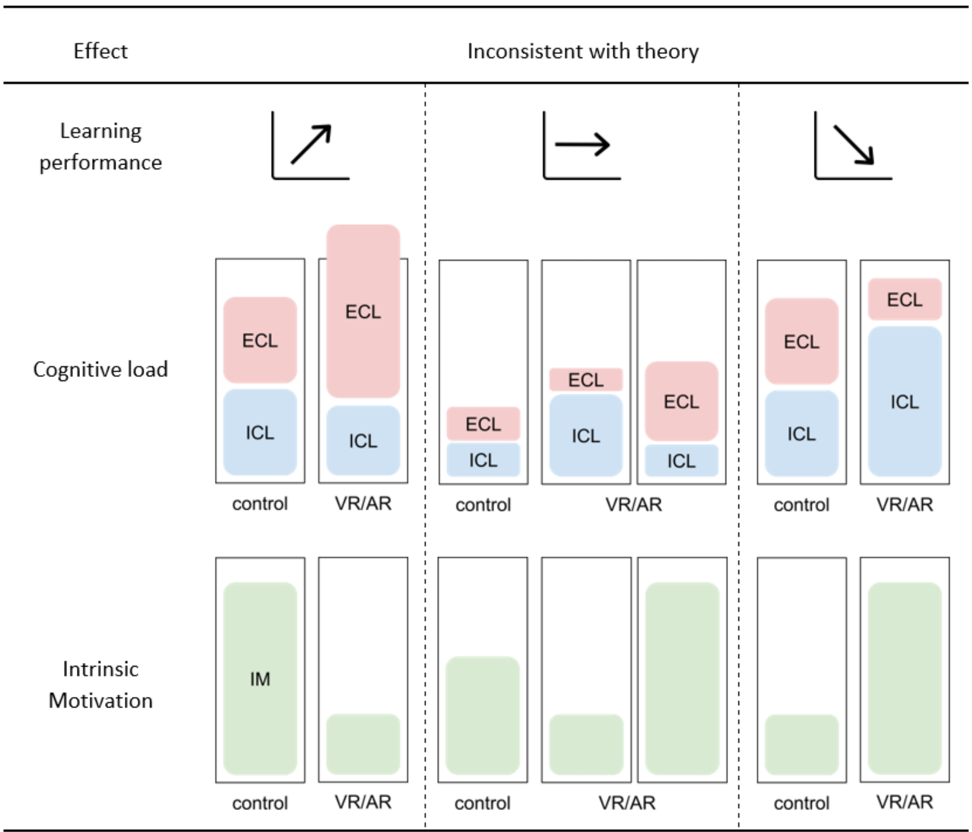

Figure 2.b: Grid for analyzing results of immersive technologies on learning considering cognitive load and intrinsic motivation results when they are inconsistent with the theory. ECL = extraneous cognitive load, ICL = intrinsic cognitive load, IM = intrinsic motivation

To interpret study results, main effects of interventions on learning, cognitive load, and motivation were reported in the grid to determine whether the observed effects were positive, neutral, negative, or inconsistent with established theoretical frameworks. To address the question “is there any evidence of an effect?”, we quantified the occurrence of studies according to the nature of their effects. Subsequently, we encapsulated the outcomes and interpretations derived from these studies, guided by the methodology proposed by Campbell et al. [11].

3.5 Risk of bias

A global risk of bias score was assigned to each study using the "RoB 2" tool from the Cochrane Collaboration [76], as shown in Table 5. This tool assesses five areas of potential bias, such as the randomization process and missing outcome data. Each domain comprises several items in the form of questions with five response options, enabling identification of potential markers of low or high risk of bias. The majority nature of the marker set determines whether the judgment of risk of bias for a dimension is low, high, or some concern. Finally, according to the five dimensions, the overall risk-of-bias judgment is set to low, high, or with some concerns. Authors used the grid to assess the overall risk of bias for each study and reported it in the results. It should be noted that none of the studies showed a high risk of bias according to this evaluation.

4. Findings and discussion

As previously mentioned, the systematic analysis encompassed a total of 35 references and 36 studies. Among these, 32 studies were experimental (randomized controlled trials), and four were quasi-experimental (controlled trials).

The primary focus of 26 studies (72%) was on investigating cognitive load, while 18 studies (50%) explored the effects of immersive technologies on intrinsic motivation. Notably, 8 studies (22%) addressed both aspects simultaneously (Figure 3). Furthermore, 22 studies (61%) utilized virtual reality (VR), 13 studies (36%) employed augmented reality (AR), and only one study compared both technologies (see Figure 3).

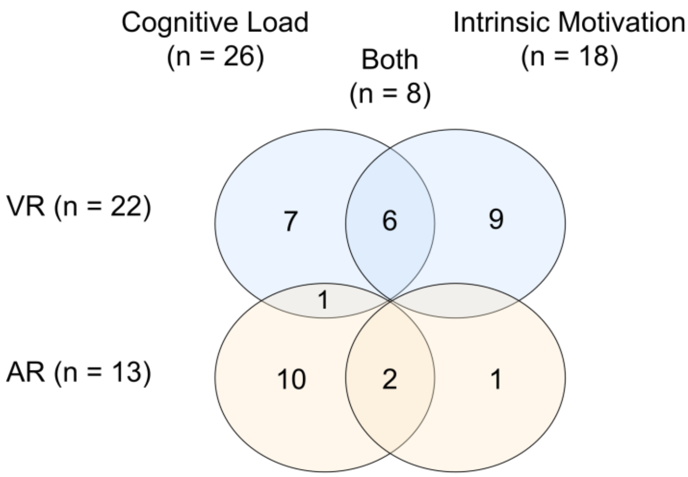

Figure 3: Distribution of studies based on cognitive load, intrinsic motivation, and the technology employed (virtual reality or augmented reality)

The sample publication period spans from March 2016 to November 2022, with the majority of studies conducted between 2020 and 2022 (Figure 4). This aligns with previous research indicating a growing scholarly interest in immersive technologies for education [66].

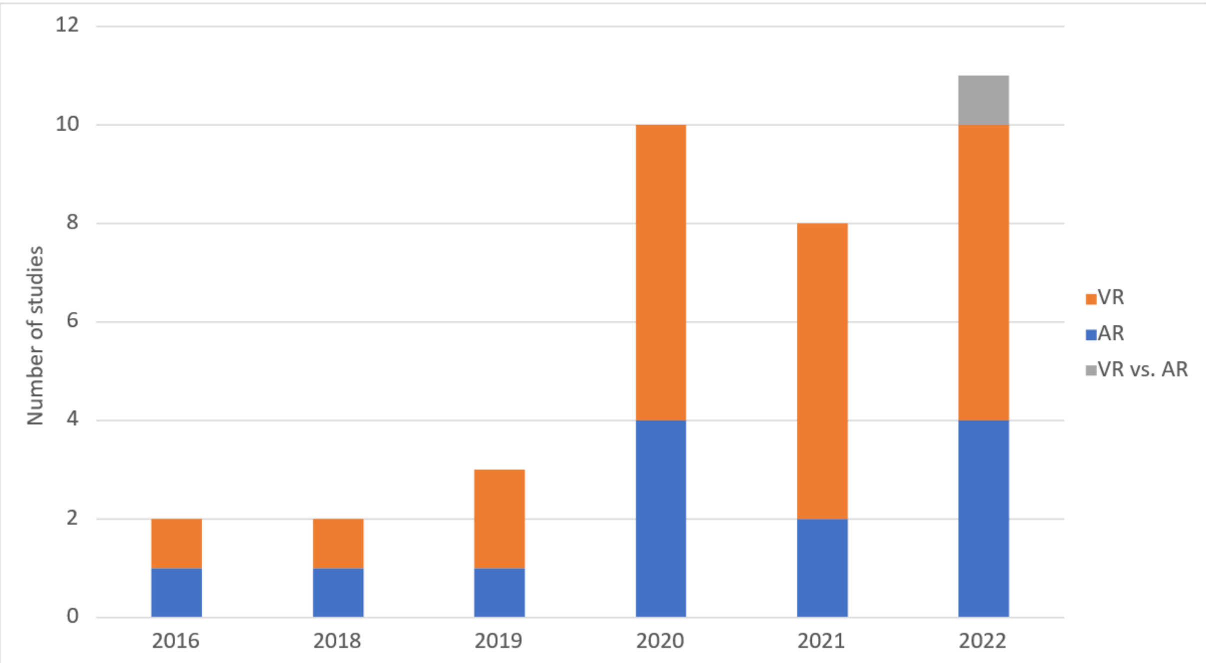

Figure 4: Distribution of publications over time and by technology

The studies involved an average of 81 participants per study (see Figure 5), predominantly consisting of university students. The participants' expertise level varied, with 16 studies including learners without prior knowledge on the subject, 11 studies including learners with intermediate-level prior knowledge, and 9 studies not specifying the learners' expertise level (see Table 4).

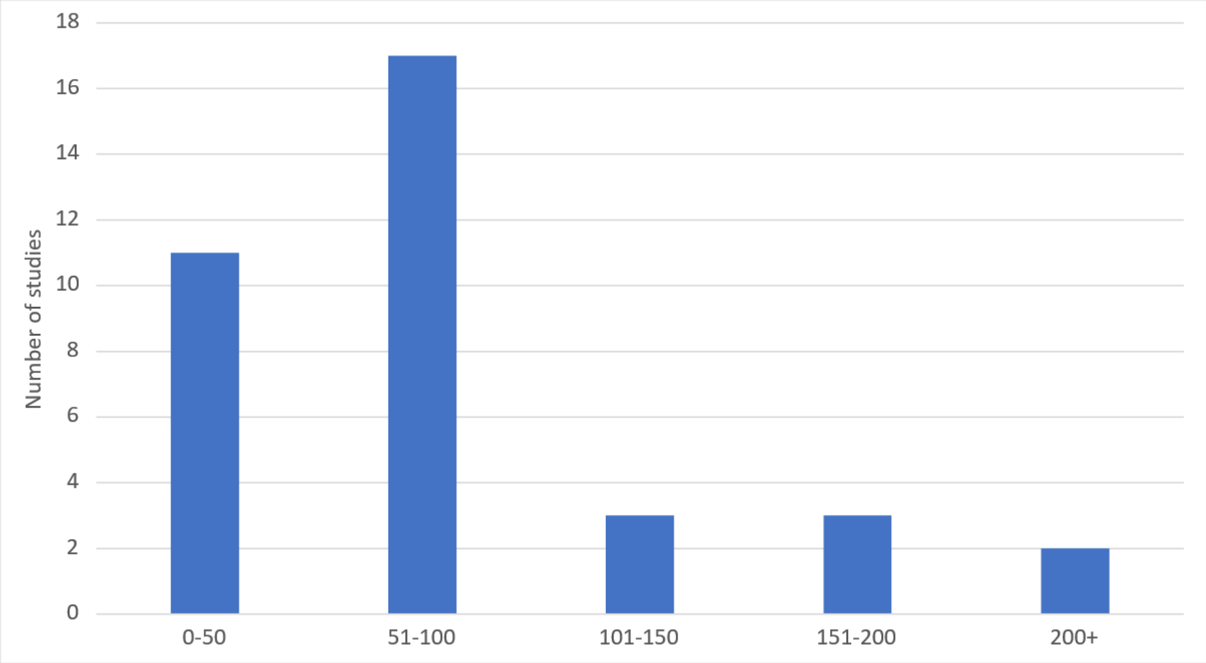

Figure 5 : Number of studies by sample size and type of experimental design

Various measures of learning were employed, encompassing diverse types of knowledge, and learning domains (Table 4). The most frequently used learning outcomes were retention (n = 29), transfer (n = 9), and skill acquisition (n = 6). The selected articles predominantly focused on declarative (factual or conceptual) and procedural knowledge. Additionally, the most studied learning domain was science.

Table 4: Distribution of studies by a) learning area b) learning outcomes c) types of knowledge and d) level of expertise of learners

Features	n	%
a) Learning domain		
science	19	53%
medicine	4	11%
art	3	8%
safety	2	6%
history	2	6%
b) Learning outcomes		
retention	29	60%
transfer	9	19%
skills acquisition	6	13%
perceived learning	2	4%
behavioral change	1	2%

Features	n	%
c) Knowledge Type		
declarative	26	68%
factual	5	13%
conceptual	6	17%
both	15	42%
procedural	9	24%
ALL	2	5%
unsure	1	3%
d) Level of expertise		
novice	16	44%
intermediate	11	31%
unsure	9	25%

4.1 What is the effect of immersive technologies on learning performance and cognitive load?

Among the 26 studies focusing on cognitive load (CL) and learning, the results exhibited significant variation. Specifically, 6 studies (23%) concluded that the use of immersive technologies positively influenced cognitive load during learning [40, 42, 47, 72, 85, 87]. In contrast, five studies (19%) reported a negative effect, suggesting that immersive technologies were less effective than other media [21, 22, 51, 64, 63]. Additionally, 6 studies (23%) found no significant impact of VR/AR on learners' cognitive load when compared to other learning methods [16, 23, 35, 84, 89, 10]. These findings underscore the diverse effects of immersive technologies on learning and cognitive load, emphasizing the need for a detailed examination of the specific conditions under which their use can be beneficial or not. In addition to the previously mentioned findings, 9 studies (35%) revealed an inconsistent effect of immersive technologies on learning and cognitive load [2, 8, 28, 36, 65, 82, 81, 80, 45].

4.1.1 Virtual reality (VR)

Among the 13 studies that investigated the impact of VR on learning and cognitive load, the majority (n = 5, 38%) reported a negative effect [21, 22, 51, 64, 63], with only two studies showing a positive effect [47, 87] (Figure 6).

According to these results, VR, particularly when an HMD is used, tends to increase cognitive load, resulting in reduced learning performance, with a medium to large effect size. Frithioff et al. [22] point out that the high complexity of the virtual environment leads to cognitive overload, resulting in impaired learning. This aligns with the observations of Frederiksen et al. [21], noting that the higher cognitive load is attributed to the increased number of elements to interact with and distractions in the environment. Makransky et al. [51] suggest that the immersive VR environment may act as a seductive detail, causing distraction and hindering cognitive assimilation. This idea resonates with the findings of Parong & Mayer [63], who argue that the perceptual richness and high-arousal emotions associated with immersive VR lead to increased cognitive distraction and, consequently, poorer learning outcomes. Parong & Mayer [64] further support this notion by arguing that high immersion in VR causes higher levels of affective processing but lower levels of cognitive processing and learning outcomes.

Collectively, these studies underscore the consistent theme that the immersive nature of VR introduces elements that contribute to extraneous cognitive load. The complexity of the VR environment, the perceived higher number of elements to interact with, and the distractions within the immersive setting all lead to increased cognitive demands, potentially hindering the learning process. Another explanation mentioned in most studies is the novelty of this technology for students. On the one hand, the fact that students must learn to use the system imposes a de facto extraneous cognitive load. On the other hand, because of the novelty effect, students may see VR as an entertainment tool rather than a learning tool compared to the traditional media they are used to [21, 51, 63, 64]. Consistent with this assumption, Frithioff et al. [22] showed a habituation effect, i.e., a reduction in cognitive load with repeated sessions.

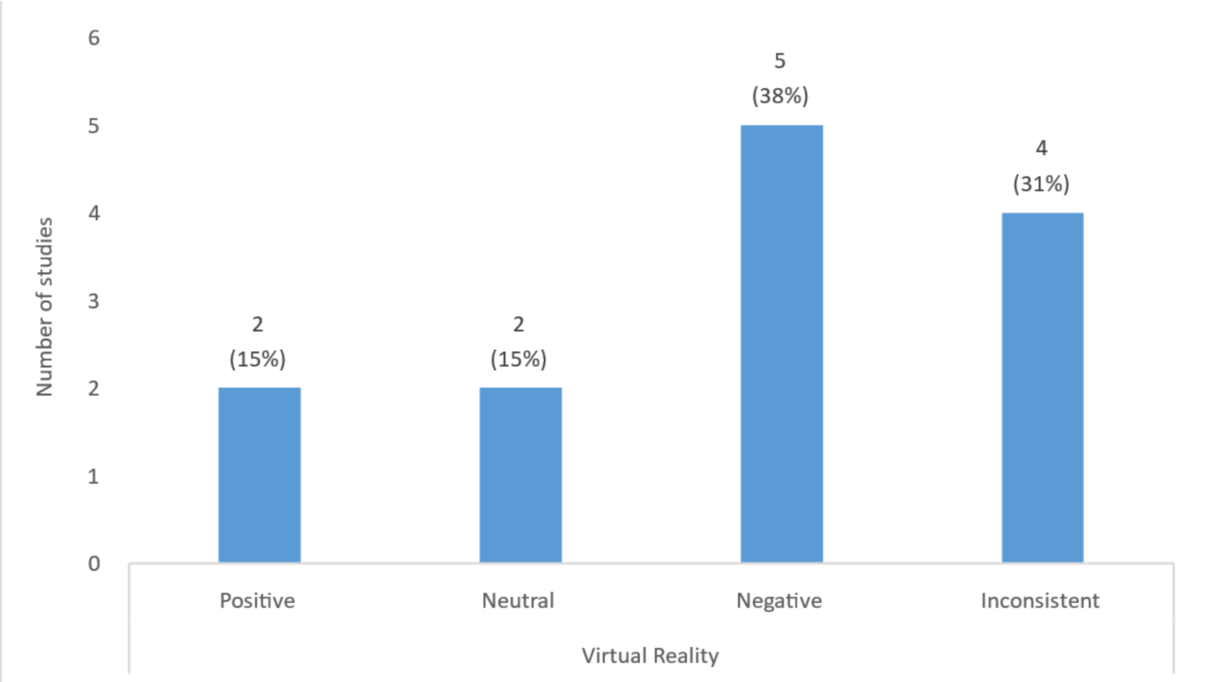

Figure 6: Effect of Virtual Reality on learning considering cognitive load.

In the five studies, it is frequently noted that the negative effect of VR on learning and cognitive load is exacerbated by the novice level of the learners. Most studies reporting a negative effect involved novice participants [21, 22, 63, 64], while all the positive studies used intermediate-level students [47, 87] (Figure 7). Specifically, VR led to better learning and reduced extraneous cognitive load for non-WEIRD professional seafarers [47], and similarly, facilitated better skill acquisition with a reduction in cognitive load for engineering students with prior basic knowledge and skills on the subject [87].

A possible explanation for the differences in effectiveness between novice and intermediate learners is that the knowledge base of intermediate learners enables them to manage the additional load imposed by VR and prevent cognitive overload. This finding can be interpreted as an expertise reversal effect [32]. This CLT effect indicates that certain cognitive load effects diminish as the learner’s expertise increases, eventually disappearing or even reversing.

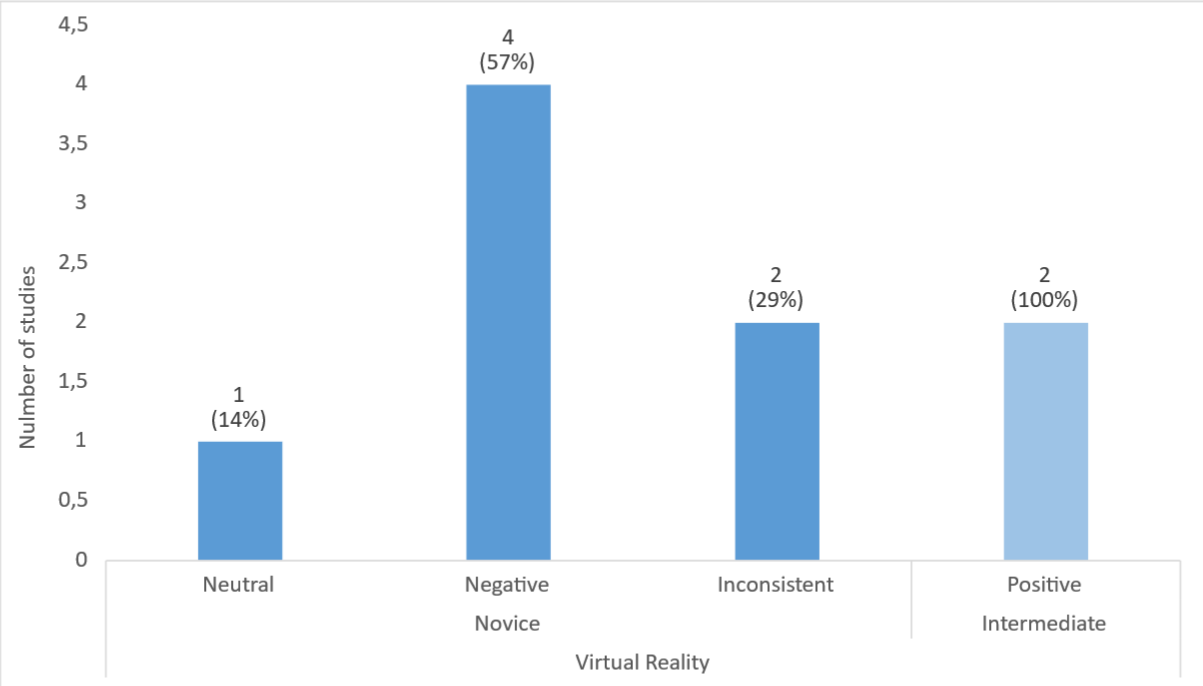

Figure 7: Effect of Virtual Reality on learning and cognitive load based on learner's prior knowledge.

This literature review reveals that some of the included studies reported surprising effects of VR on learning and cognitive load [8, 28, 65, 80]. Baceviciute et al. [8] demonstrated that learning in VR allows for better transfer than learning in real life, associated with a decrease in intrinsic cognitive load and an increase in extraneous cognitive load. Three other studies indicated that cognitive load, especially extraneous cognitive load (ECL), is not directly correlated with learning performance [65, 80, 28]. These results are inconsistent with cognitive load theory.

One explanation for these results could be the low cognitive demand of the task, leading to a "floor effect" [28, 80]. This occurs when the learning tasks are relatively easy for the participants, resulting in minimal variations in cognitive load that may not be easily detectable [79]. If the tasks lack sufficient complexity, the impact of VR on cognitive load may not be significant.

Huang et al. [28] also raised the issue of a mismatch between the activity in VR and its evaluation. The disparity between the nature of tasks conducted within the VR environment and the evaluation methods employed could introduce confounding variables that influence the relationship between cognitive load and learning outcomes.

4.1.2 Augmented reality (AR)

In contrast to virtual reality, none of the studies examining augmented reality found a negative effect of the technology on cognitive load. Instead, a slight majority of studies (n = 4) reported that AR reduced the level of cognitive load, leading to improved learning outcomes compared to less immersive media, with strong effect sizes [40, 42, 72, 85] (Figure 8). According to these studies, AR ensures spatial and temporal continuity, preventing a split-attention effect [40, 85]. This effect suggests that integrating multiple sources of information into a single source minimizes extraneous cognitive load and enhances learning, aligning with Cognitive Load Theory (CLT) and Cognitive Theory of Multimedia Learning (CTML). Lee & Hsu [42] showed that AR can reduce students' cognitive load during the learning process, allowing them to focus on important information and avoid distractions. The use of AR also actively involved students in the learning process, helping them to acquire better skills [72]. Additionally, the authors argue that AR makes previously abstract elements more concrete, contributing to the reduction in cognitive load.

However, it is worth noting that even though some studies hypothesized that their AR systems reduce extraneous load, none of them used a measurement scale to distinguish different types of cognitive loads, making it challenging to confirm whether the decrease in cognitive load corresponds specifically to a reduction in extraneous cognitive load (ECL).

Furthermore, three studies observed no significant differences between conditions for learning and cognitive load [16, 23, 35]. The reasons cited were primarily small sample sizes, but it is also noteworthy that Geng & Yamada [23] and Kapp et al. [35] reported particularly low levels of extraneous cognitive load. As a result, the observed reduction in cognitive load in these studies may not have been evident due to the already low level of unnecessary load in the traditional learning condition. According to Elford et al. [16], the expected benefits of AR on cognitive load were offset by the cognitive load induced by other elements such as gamification.

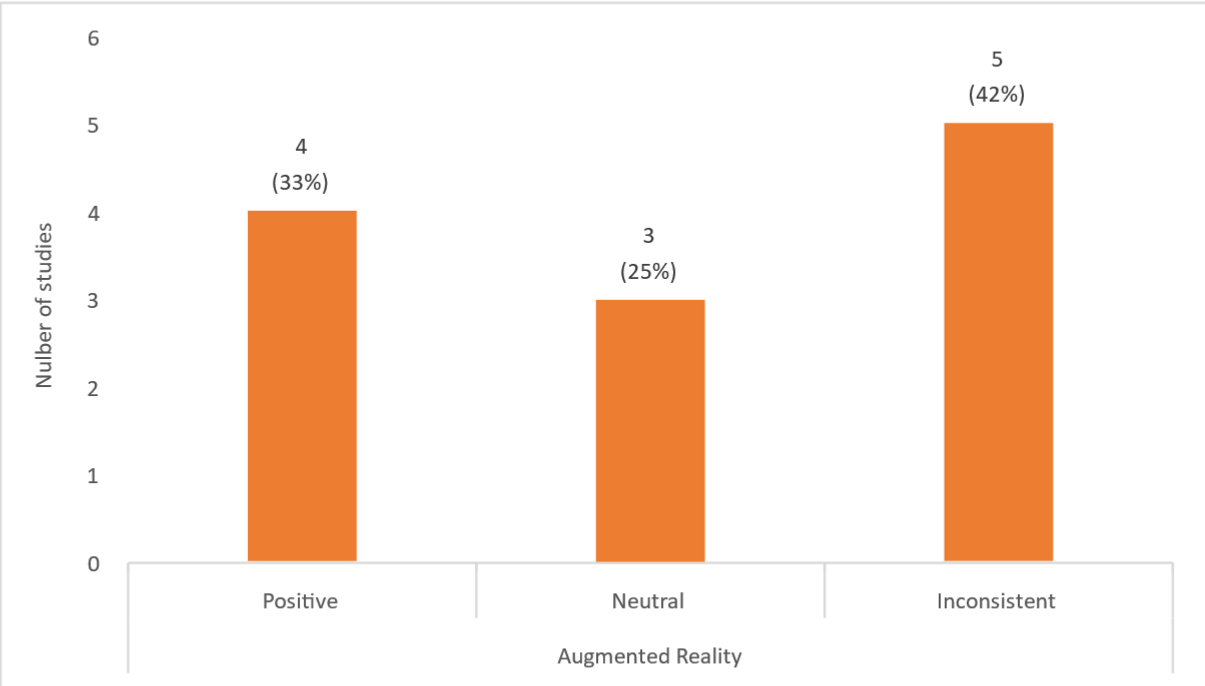

Figure 8: Effect of Augmented Reality on learning considering cognitive load.

The effectiveness of AR on learning, considering cognitive load, also appears to depend on learner prior knowledge. Most studies reporting a positive effect of AR involved novice learners [42, 72, 85]. Conversely, studies on participants with some prior knowledge of the subject seem to show that AR has no effect on their learning performance and cognitive load [23, 35] (Figure 9). Novice learners are likely to benefit more from the CLT principles ensured by AR than intermediate learners. As with VR, an expertise reversal effect seems to apply. However, more studies are needed on the effects of AR on learning and cognitive load for intermediate learners.

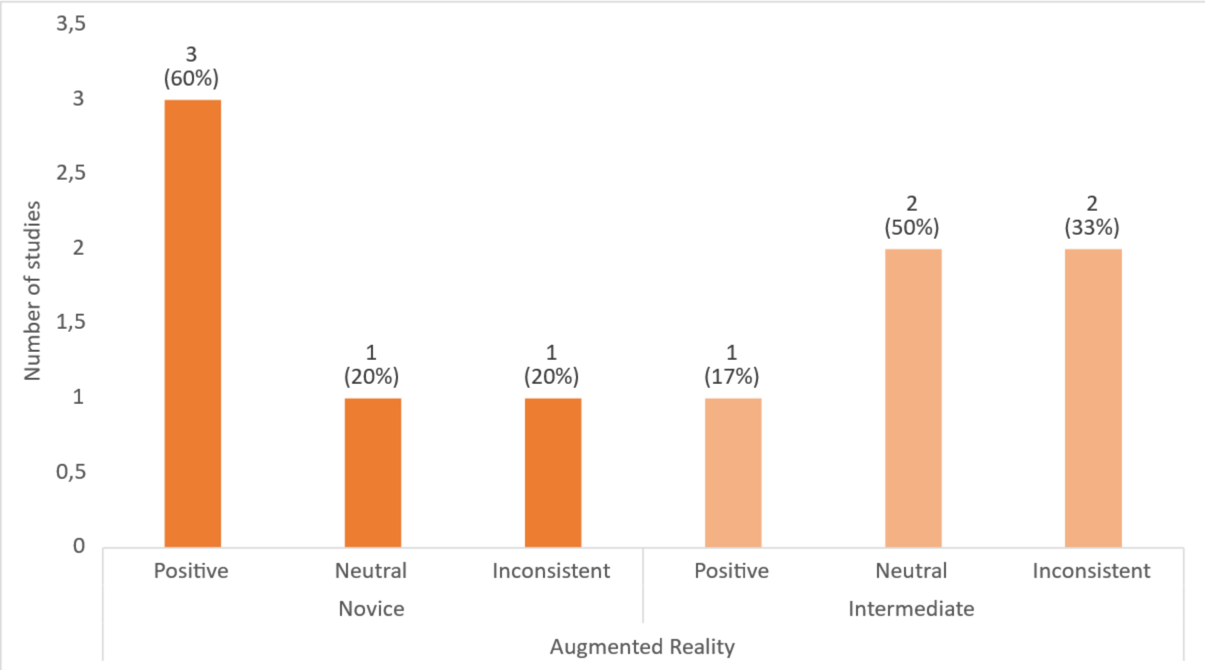

Figure 9: Effect of Augmented Reality on learning and cognitive load based on learner's prior knowledge.

Five studies (42%) demonstrated results inconsistent with CLT [2, 36, 81, 82, 45]. These inconsistencies were often associated with a lack of correlation between learning outcomes and the measured cognitive load. Several potential explanations have been proposed for these surprising results. Firstly, similar to VR, the occurrence of a "floor effect" may contribute to the inconsistency. Altmeyer et al. [2] and Thees et al. [81] reported a low level of cognitive load in each experimental condition. Another possible explanation is the hypothesis of an insufficient sample size [36]. If the number of participants in the study is not large enough, the statistical power to detect meaningful differences in cognitive load and learning outcomes may be reduced, leading to inconclusive results. Lastly, Thees et al. [82] also indicated that the measure of learning could explain these surprising results. A short-term knowledge test that is too general can obscure interesting results on more specific dimensions of learning.

4.1.3 General discussion about immersive technologies effect on learning and cognitive load

One study compared the use of AR and VR for learning about the phenomenon of "lightning" [84]. Although no significant difference was found between the two technologies in this study, the review of all included studies reveals interesting differences between VR and AR.

Overall, the results suggest that AR optimizes cognitive load and leads to learning gains [40, 42, 72, 85], while VR more often tends to overload learners cognitively, resulting in deleterious effects on learning performance [21, 22, 51, 64, 63]. This finding is consistent with previous systematic reviews on AR for learning [9, 58], but it represents a novel result for VR and its effects on cognitive load, which had not been previously subjected to a systematic review process. This difference can be attributed to the inherent technological characteristics of the two technologies. AR allows for the addition of certain virtual elements to the real learning situation, sometimes aligning with the principles of CLT and CTML [9]. In contrast, VR involves a much richer and more complex learning environment in terms of sensory rendering and interaction, which can lead to cognitive overload [21, 22, 63, 64].

This contrast in results is even more striking when the learners' level is considered. While AR is better suited to novice student profiles [42, 72, 85], reducing their irrelevant cognitive load, VR tends to overload them [21, 22, 63, 64]. On the other hand, a sufficient knowledge base in the field enables intermediate learners to take full advantage of VR's positive effects [47, 87], whereas AR offers no benefits [23, 35]. These results are similar to a well-known CLT effect, the expertise reversal effect. This effect posits that instructional techniques effective for novices may become less effective or even detrimental as learners gain expertise.

Regarding the type of knowledge taught, the studies on the effect of VR and AR on cognitive load during procedural learning were limited (n=7) and contradictory, making it challenging to draw definitive conclusions [21, 22, 51, 47, 72, 45, 87]. The effects mentioned earlier seem to mainly pertain to declarative learning (factual or conceptual). Most studies focused on information retention and, to a lesser extent, on transfer of learning. The cognitive overload effects observed with VR were evident in both retention and transfer tasks. However, for AR, studies only examined its effect on retention. Lastly, the results did not show any significant effect concerning the domain of learning.

Finally, many studies report results that are inconsistent with CLT according to the analysis grid developed for this systematic review [2, 8, 28, 36, 65, 81, 82, 80, 45]. The hypotheses to explain these results are varied but are similar for VR and AR. One idea is that the complexity of the task being taught may be too low, resulting in significantly low levels of cognitive load [2, 81, 28, 80]. This ground effect could prevent the detection of variations in cognitive load between experimental conditions. The measurement of cognitive load itself could be a contributing factor to the inconsistency. Cognitive load is often assessed using self-reported measures, which may not always be sensitive enough to capture subtle variations in cognitive load [42]. Additionally, recent studies have proposed an interesting idea that perceptually rich learning environments induce ECL that, under certain circumstances and depending on how learning is assessed, might enhance learning [73]. This suggests that the impact of cognitive load on learning outcomes is not always straightforward and may depend on the specific characteristics of the learning task and environment. Similarly, some studies have stressed the importance of using appropriate measures of learning performance [82, 28]. Measurements that are too general, too short-term, and too far removed from the task may lead to a loss of interesting results.

4.2 What is the effect of immersive technologies on learning performance and intrinsic motivation?

The results of the selection process yielded 18 (50%) studies on learning and intrinsic motivation [10, 28, 31, 38, 42, 46, 52, 47, 48, 62, 64, 63, 65, 72, 5, 17, 44, 89] (see Table 5). Two-thirds of the studies reported an inconsistent correlation between intrinsic motivation and learning [10, 28, 38, 42, 48, 64, 63, 65, 72, 5, 44, 89]. Only three studies found a positive effect [31, 46, 47], while the remaining three studies did not show any effect [52, 62, 17] (see Figure 10). These results surprisingly indicate a lack of correlation between intrinsic motivation to learn and learning performance.

4.2.1 Virtual reality (VR)

Among the fifteen studies exploring the impact of VR on learning and intrinsic motivation [10, 28, 31, 38, 46, 52, 47, 48, 62, 64, 63, 65, 5, 44, 89], only three (20%) found a positive effect with a large effect size [31, 46, 47]. These studies showed that using HMD VR increases different dimensions of learning (retention, transfer, and behavioral changes) associated with better intrinsic motivation and perceived enjoyment compared to traditional methods (e.g., manual or personal trainer). Jin et al.

[31] reported that VR made participants lose track of time and perceive the experience as a game. VR also allowed them to freely explore the environment, enhancing their curiosity, motivation to learn, and interest. This state of flow was also observed by Makransky et al. [46], who explained that VR induces positive emotions and a sense of control that elicits pleasure in learners, motivating them to actively engage with the material.

Two additional studies showed no difference in learning and motivation when using VR [52, 62]. Interestingly, Pande et al. [62] reported that VR and video control induced high levels of intrinsic motivation, and that the lack of difference could be explained by participants being accustomed to both technologies.

According to our analysis grid, two-thirds of the studies found inconsistent results between motivation and learning performance [10, 28, 38, 48, 64, 63, 65, 5, 44, 89]. The lack of consistent findings regarding the relationship between motivation to learn and learning outcomes can be explained in diverse ways. One main reason is the polysemic nature of the term "motivation," which can refer to several factors such as intrinsic motivation, motivation to learn, perceived enjoyment, or situational interest. These polysemic interpretations lead to different theoretical frameworks, resulting in studies assessing a state of motivation that is not specifically directed towards the learning activity itself. Instead, these studies may have measured motivation to explore and use the immersive technology or enjoyment in using the technology.

Similarly, the diversity of measurement scales could contribute to the inconsistent results. While the most used scale was the "interest/enjoyment" subscale of the Intrinsic Motivation Inventory (IMI) [14], six different scales were used to measure motivational states in the selected studies using VR. Furthermore, many studies (80%) adapted the original scales to their research [10, 28, 31, 38, 46, 52, 48, 62, 65, 5, 44, 89]. Although this is a widespread practice, half of these studies did not provide details on how the scales were modified [38, 46, 48, 62, 65, 89]. This leads to significant methodological differences between the measurement scales, which may explain the inconclusive results reported in the systematic review. In the case of the three studies showing a beneficial effect of VR on motivation and learning, two of them used the interest/enjoyment subscale of the IMI, as prescribed by the authors for evaluating intrinsic motivation [46, 47]. These observations suggest that the inconsistent results regarding the relationships between intrinsic motivation and learning arise from methodological issues, particularly the misuse or absence of standardized measures. This argument aligns with a previous systematic literature review that highlighted the methodological weaknesses of the field [15].

Taking a closer look at these inconsistent studies, the majority report positive effects of VR on motivation combined with no or negative effects on learning. This lack of correlation between learning and motivation may indicate that VR brings pleasure and motivation not directed towards the learning task but more towards a discovery of the system. This hypothesis supports the results of this systematic review of cognitive load (section 4.1.1), which point to a novel and distraction effect. It is possible that, due to the novelty and attractiveness of the system, users perceive it more as an entertainment tool than a learning tool.

4.2.2 Augmented reality (AR)

Only three studies have examined the effect of AR on intrinsic motivation [42, 72, 17] (see Table 5), and none of them show any effect on intrinsic motivation and learning (Figure 10). As with VR, two-thirds of the studies reported effects inconsistent with motivation theories [42, 72]. Both studies used the Index of Learning Styles (ILS) [75] to measure participants’ intrinsic motivation, which may not effectively capture intrinsic motivation to learn, potentially explaining the surprising results.

Although more studies on AR and intrinsic motivation are needed, the current review suggests that AR does not intrinsically motivate students to learn.

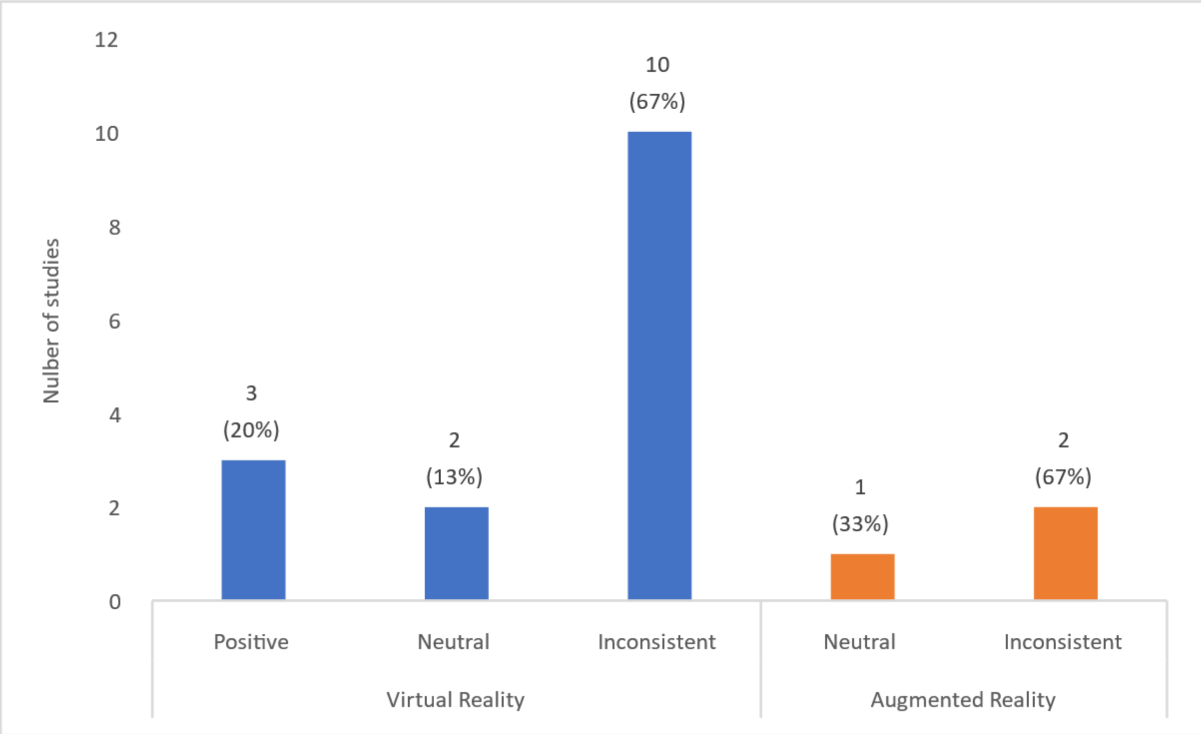

Figure 10: Effect of VR and AR on learning and intrinsic motivation.

4.2.3 General discussion about immersive technologies effect on learning and intrinsic motivation

The review of included studies about the effect of immersive technologies on learning and intrinsic motivation showed inconsistent results with motivation theories. Whether for virtual or augmented reality, this can be explained by a lack of rigor in measuring intrinsic motivation. Another hypothesis could be the influence of other factors not considered in the studies, such as cognitive load. It is possible that the positive effect of VR/AR on motivation is counterbalanced by the cognitive overload imposed by the system, leading to less efficient learning.

Focusing on AR, which has been comparatively less explored with only three identified studies, the observed differences from VR, particularly concerning cognitive load, warrant a more in-depth investigation into AR's influence on intrinsically motivated learning behaviors. Notably, the fact that two out of the three studies reported inconsistent effects supports the need for further exploration.

At present, definitively pinpointing the nature of VR and AR's impact on intrinsic motivation for learning, considering learning performance, poses a challenge. Nevertheless, it is noteworthy that none of the studies indicated negative effects. Subsequent research endeavors should delve deeper to unravel the intricacies of the relationship between intrinsic motivation and learning outcomes within immersive technologies.

4.3 What are links between cognitive load and motivational states variables in immersive learning context?

Some studies suggest that considering cognitive and motivational variables together is necessary to better understand the benefits and limitations of virtual and augmented reality [51, 80]. While several studies included measures of cognitive load and motivational states in their designs [28, 42, 47, 64, 63, 65, 72, 89], only two explored causal relationships between these two variables [28, 65]. Among the studies that did not test for causal effects, only one reported a simultaneous improvement in learning, cognitive load, and motivation with the use of VR. Specifically, learners reported being more intrinsically motivated and experiencing less extraneous cognitive load (positive effect) when using VR, leading to enhanced learning outcomes [47]. However, in the other five studies, while the effects on cognitive load and learning were similar, the effects on motivational states appeared to be entirely uncorrelated with the other two factors [42, 64, 63, 72, 89].

Two studies [28, 65] utilized structural equation modeling (SEM) to explore the relationships between cognitive load and motivation to learn in VR-based learning. In the first study, the resulting SEM model demonstrated that technical features of VR had a positive impact on motivation, mediated by psychological variables of VR (presence and agency). Furthermore, motivation to learn negatively predicted extraneous cognitive load (ECL) and positively predicted intrinsic cognitive load (ICL). The model indicated that only intrinsic cognitive load (ICL) significantly predicted learning outcomes (see Figure 11) [28].

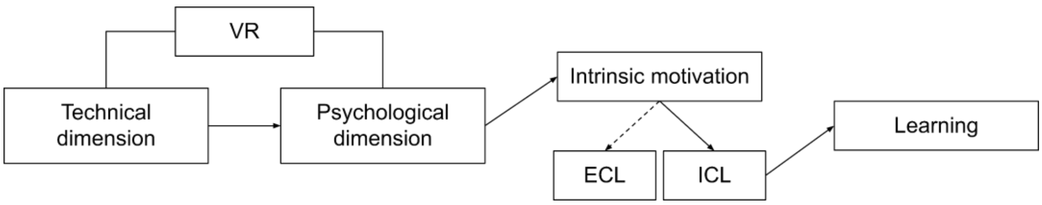

Figure 11: Huang et al. [28] structural model of the VR effect on learning performance, mediated by cognitive load and motivational states.

In the model from the study by Petersen et al. [65], situational interest directly predicted learning outcomes. Additionally, motivational variables were predicted by the sense of presence induced using VR, which also depended on the extraneous cognitive load (ECL) imposed by the system. In this context, unnecessary cognitive load (ECL) had an indirect negative effect on learners' motivation. However, the model did not integrate the concept of useful intrinsic cognitive load (ICL) (see Figure 12).

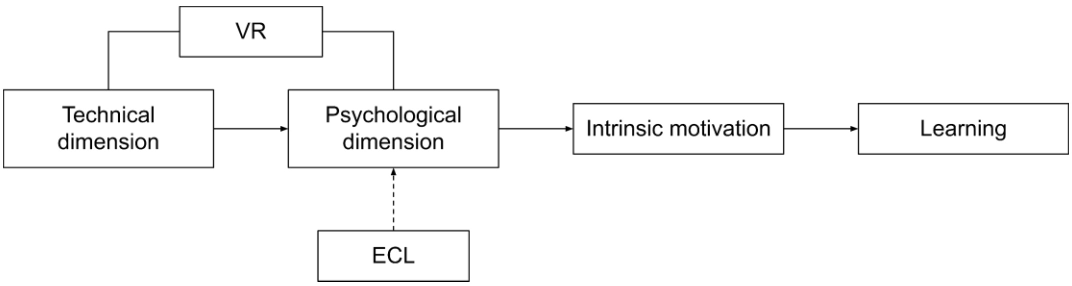

Figure 12: Petersen et al. [65] structural model of the VR effect on learning performance, mediated by cognitive load and motivational states.

The two studies indicated different causal relationships between the two factors, with intrinsic motivation determining perceived cognitive load in one case, while in the other, cognitive load influenced the sense of presence, which in turn influenced learner motivation. These contrasting results are also evident in the non-specific literature on the use of immersive technologies. On one hand, Cognitive Load Theory (CLT) assumes that a sufficient level of motivation is required for learners to invest the cognitive effort necessary for task completion [79]. Similarly, studies have shown that curiosity states enhance cognitive engagement and reduce perceived effort [55], especially by reducing perceived load, often imposed by virtual environments [73]. These results support the view that cognitive load is positively influenced by motivation. On the other hand, Feldon et al. [19] highlighted the negative effect of cognitive load on intrinsic motivation. The expected cognitive effort may diminish learners' motivational beliefs and their own motivation to learn. Thus, cognitive load could be perceived as a motivational cost. Nevertheless, the authors emphasized the lack of studies grounding this view. Moreover, unlike Huang et al. [28], the model by Petersen et al. [65] omitted the role of intrinsic cognitive load (ICL). Yet, both types of loads could influence motivation in several ways. A reduction in extraneous cognitive load (ECL) could predict higher motivation [20], while maintaining a greater sense of presence, as supported by Huang et al. [28]. On the other hand, ICL should be considered to promote better intrinsic motivation [29]. Similarly, Huang et al. [28] showed a positive effect of intrinsic motivation on generative processing (i.e., relevant load), supporting the idea of a different kind of interaction between intrinsic motivation and different types of cognitive load. This result is also consistent with recent literature [19].

Moreover, the effect of intrinsic motivation and cognitive load on learning outcomes does not appear to be solely direct but involves mediated effects between both factors. This implies that simply measuring these factors to control the effects of VR/AR is not enough. Understanding how these variables interact is necessary to design more effective learning environments. However, more studies are needed to better understand the interactions between cognitive load and intrinsic motivation.

Lastly, both models only explored the role of VR in the relationships between the two variables. It is essential to understand the role of AR in such relationships, considering the different cognitive demands imposed by VR and AR. It is likely that the mediating role of presence is more pronounced with the use of VR, as this factor is much more associated with VR than AR. Therefore, further studies, such as those conducted by Huang et al. [28] and Petersen et al. [65], are needed to better understand the role of cognitive load and motivational states in learning performances when using immersive technologies.

Table 5a: Summary table of characteristics and results of included studies using Virtual Reality. (The shaded area includes studies measuring both cognitive load and intrinsic motivation. CL = cognitive load, ICL = intrinsic cognitive load, ECL = extraneous cognitive load, IM = intrinsic motivation, ML = motivation to learn, PE = perceived enjoyment)

Authors	Participants; expertise level	Study design; immersive technology vs. comparative medium	Learning design	Risk of bias	Main results (measurements used)
Baceviciute et al. [8]	48 university students; novice	RCT; HMD VR vs. book	science; medicine; declarative; retention; transfer	Low	Reading in VR led to better transfer performance than real reading, but no significant differences were found for retention and overall cognitive load. VR was also associated with an increase of ECL [12], a decrease of ICL [7] and an increase of general EEG mental load.
Burgues et al. [10]	154 French university students; novice	RCT; HMD VR with and without control vs. tablet with and without control	art; factual; retention	Low	VR does not lead to better art learning than the tablet and does not seem to affect CL (own scale), but system interactivity, in VR or with the tablet, leads to better art retention.
Frederiksen et al. [21]	31 post-doctoral medicine students; novice	RCT; HMD VR vs. desktop VR	medicine; procedural; skills acquisition	Low	HMD VR simulation for laparoscopic surgery training induces higher CL (secondary-task reaction time) and results in worse performance than desktop VR for novices. The HMD VR environment causes additional ICL and ECL.
Frithioff et al. [22]	24 university students; intermediate	RCT; Ultra High Fidelity VR vs. desktop VR	medicine; procedural; retention	Low	Ultra-high fidelity VR simulation compared to conventional screen-based VR simulation for temporal bone surgery training results in lower comprehension and higher CL (relative reaction time) in medical learners.
Makransky et al. [51]	52 university students; unsure	RCT; HMD VR; desktop VR vs. desktop VR	science; biology; conceptual; procedural; retention; transfer	Low	Using an HMD VR science lab simulation resulted in greater presence, but less learning and greater cognitive overload (EEG), compared to desktop VR.
Tang et al. [80]	59 university students; unsure	RCT; HMD VR vs. tablet	art; factual; retention	Some concerns	Virtual reality is no better at retaining information about art than a tablet, and ECL [43] is not correlated with memory.
Yang et al. [87]	80 engineering students; intermediate	RCT; HMD VR vs. ppt; traditional	engineering; procedural; skills acquisition	Low	VR allows for better acquisition of operational skills and reduces CL (adapted from Hwang et al. [30]) compared to traditional teaching methods.
Huang et al. [28]	50 university students; unsure	RCT; HMD VR vs. lower immersive HMD VR	science; declarative; retention	Low	VR has a positive impact on LM but not on CL (both scales were revised from validated instruments). LM reduces CL and increases generative processing which increases learning.
Makransky & Klingenberg [47]	28 non-WEIRD sample of professional seafarers; intermediate	RCT; HMD VR vs. personal trainer	safety; procedural; perceived learning and behavioral change	Low	Teaching security resulted in non-WEIRED learner's higher PE [83], IM [14], perceived learning and behavioral change and lower CL [3] than learning with personal trainer (large effect size).
Parong & Mayer [63]	61 university students; novice	RCT; HMD VR vs. ppt	science; declarative; retention and transfer	Low	The VR group showed poorer transfer performance, associated with an increase in ECL (own scales). VR had little influence on IM (increased PE but not IM or interest; own scales) and that this influence did not affect learning performance.
Parong & Mayer [64]	80 university students; novice	RCT; HMD VR vs. video	history; factual; retention and transfer	Some concerns	HMD VR led to worse performance on learning outcomes than video, particularly for transfer. No significant effects were shown for emotional states (situational interest, LM and PE; authors scale), and self-reported CL (authors scale).
Petersen et al. [65]	153 university students; novice	RCT; HMD VR vs. desktop; video	science; declarative; retention	Low	The different media conditions did not affect participant retention, but the interactivity led to an increase in their ECL [3] and the sensory richness promoted their interest [39] but not IM [49]. ECL has a negative effect on situational interest and IM, mediated by sense of presence

Zhao et al. [89]	75 university students; unsure	RCT; HMD VR vs. 360° video	science; declarative; retention and transfer	Low	HMD virtual reality for biology learning has no effect on learners' CL (NASATLX [26]) and performance but increase their IM (adapted from [68]) with a large effect size.
Arayaphan et al. [5]	28 university students; novice	CT; HMD VR vs. ppt	Electronic resource management; declarative; retention	Some concerns	The VR group reported higher intrinsic motivation (short version of the IMI [14]) than the PPT group, but no significant difference in learning was shown.
Burgues et al. [10]	61 French university students; novice	RCT; HMD VR with and without control vs. tablet with and without control	art; factual; retention	Low	VR had no impact on intrinsic motivation (adapted from [14]) but better learning performance was found if the system was interactive (medium/significant effect).
Jin et al. [31]	54 Chinese university students; novice	RCT; HMD VR vs. Multi-touch table system	culture; history; factual; retention	Some concerns	Study showed that learning with HMD VR resulted in better learning retention and greater learning motivation (adapted from [27]) than learning with multiple touch tablet.
Klingenberg et al. [38]	89 first-year undergraduate students; novice	RCT; HMD VR vs. desktop VR	biochemistry; factual; conceptual; procedural; metacognitive; transfer; retention	Low	No effect difference between HMD and desktop VR on learning performance scores, IM, and perceived enjoyment (adapted from [14]) was found in the first posttest but a significant difference in favor of HMD in the second posttest for IM and PE (large effect size), indicating that the student preferred HMD VR "when they had a frame of reference after trying both media conditions."
Leung et al. [44]	217 university students; intermediate	RCT; HMD VR vs. in person; video	hospitality industry; procedural; retention	Some concerns	VR game training resulted in poorer immediate ($\eta2 = 0.08$) and long-term ($\eta2 = 0.43$) knowledge retention than video or in person training. Moreover, analysis shows that IM (SIMS from [24]) has no moderating effect on retention in VR.
Makransky et al. [51]	105 engineering students; intermediate	RCT; HMD VR vs. desktop VR; manual	safety; factual; conceptual; procedural; metacognitive; retention; transfer	Low	The VR conditions lead to a significant increase in IM (large effect; adapted from the IMI [68]) and perceived enjoyment (large effect; adapted from [83]) compared to traditional textbook learning. Although there was no effect of VR on retention score, these media were more effective on transfer than the text condition (medium effect size).
Makransky et al. [52]	189 university students; unsure	RCT; desktop VR vs. real demonstration	biology; declarative; retention; skill acquisition	Low	Using VR to prepare students for microbiology laboratory courses is no more effective, in terms of learning performance or IM (Interest/ Enjoyment Scale from [68]), than traditional face-to-face tutoring.
Makransky & Lilleholt [48]	104 European university students; unsure	RCT; HMD VR vs. desktop VR	science; declarative; procedural; perceived learning	Low	Even if students preferred using HMD rather than desktop VR with a large effect size observed for IM (adapted from [1]) and enjoyment (adapted from [83]), no differences was found for learning outcomes.
Pande et al. [62]	28 university students; unsure	RCT; HMD VR vs. video	biology; declarative; retention; skill acquisition	Low	Results showed a positive but non-significant effect of HMD VR on long-term biology learning and no effect on IM and perceived enjoyment [56] compared to video.

Table 6b: Summary table of characteristics and results of included studies using Augmented Reality. (The shaded area includes studies measuring both cognitive load and intrinsic motivation. CL = cognitive load, ICL = intrinsic cognitive load, ECL = extraneous cognitive load, IM = intrinsic motivation, ML = motivation to learn, PE = perceived enjoyment)

Authors	Participants; expertise level	Study design; immersive technology vs. comparative medium	Learning design	Risk of bias	Main results (measurements used)
Altmeyer et al. [2]	50 university students; novice	RCT; mAR vs. tablet	science; electrical circuits; conceptual; retention; transfer	Low	The application of tablet-based AR for learning electrical circuits with the addition of spatial continuity yields higher immediate conceptual knowledge gains, but no significant differences from the traditional method were found for transfer or CL (adapted from Leppink et al. [43]).

Elford et al. [16]	34 higher education students; novice	RCT; mAR vs. 2D drawings	science; chemistry; conceptual; retention	Some concerns	The results showed no significant effect between the use of AR and books on the learning of molecular structures and CL (adapted from Leppink et al. [43]).
Geng & Yamada [23]	21 Asian non-native learners; intermediate	RCT; mAR vs. images	language; declarative; retention	Low	The results showed that there was no significant difference in CL [43] between the two conditions and that perceived CL was related to learning performance and was likely to be affected by LM.
Kapp et al. [35]	56 university students; intermediate	RCT; HMD AR vs. tablet	science; laboratory course; conceptual; retention	Low	The use of HMD AR for spatial continuity compliance does not affect learning performance in laboratory of physics or the CL (adapted from Leppink et al. [43] and presented by Thees et al. [82]) of students.
Keller et al. [36]	30 university students; unsure	CT; mAR vs. text, pictures	science; declarative; retention	Some concerns	mAR for organic chemistry learning did not significantly influence learners' CL [37]. Learning performance and CL were not correlated.
Küçük et al. [40]	70 university students; intermediate	RCT; mAR vs. 2D pictures; graphs; text	medicine; neuroanatomy; declarative; retention	Low	The use of mAR for learning anatomy leads to a better performance associated with a lower CL [60] than traditional media (text, graphics, and images).
Liu et al. [45]	60 postgraduate students; unsure	CT; AR vs. Tablet, traditional training (video and real practice)	maintenance; procedural; skill acquisition	Some concerns	The results showed that using HoloLens 2 for maintenance training outperformed traditional training in terms of learning, but only for complex tasks. Similarly, cognitive load (NASA-TLX [26]) is significantly lower in AR than using a tablet.
Thees et al. [81]	107 German university students; intermediate	RCT; HMD AR vs. tablet; traditional setup	science; laboratory course; conceptual; retention; transfer	Low	The results indicated that the separate display condition (without AR) could outperform the AR condition with respect to learning gains and CL (adapted from Klepsch et al. [37])
Thees et al. [82]	74 German university students; intermediate	RCT; HMD AR vs. desktop; traditional setup	science; laboratory course; conceptual; retention	Low	No effect of spatial continuity in AR on conceptual knowledge around electronic measurement equipment and ICL (adapted from Leppink et al. [43]) was found. On the other hand, the results indicate a significant reduction of ECL in the AR condition compared to conventional teaching methods.
Turan et al. [85]	95 university students; novice	CT; mAR vs. book	geomorphology; declarative; retention	Some concerns	AR for geography learning improved performance and reduced overall self-reported CL [60] levels compared to traditional book-based learning. These results are consistent with the semi-structured interviews, in which students reported that AR increased their performance and decreased their CL levels.
Lee & Hsu [42]	70 Taiwan vocational senior high school students; novice	RCT; mAR vs. E-book	Makeup design; unsure; unsure	Low	Using AR to teach makeup resulted in better learning performance (large effect size), less mental effort (large effect size; modified from Hwang et al. [30]) but no difference in LM (adapted from [30]) compared to the e-book approach.
Singh et al. [72]	60 engineering students; novice	RCT; AR vs. traditional approach; manuals	science; procedural; skill acquisition	Low	AR allow a better laboratory skill learning (large effect) than traditional methods, associated with a decrease in CL (medium effect; adapted from Hwang et al. [30]). No effect on LM (adapted from Hwang et al. [30]) was found even if student's opinion revealed that learning in VR is more interesting, convenient and allow better understanding
Elford et al. [17]	57 university students; intermediate	RCT; mAR vs. 2D pictures	science; chemistry; declarative; retention	Some concerns	The introduction of AR did not result in significant differences in self-reported intrinsic motivation (IMI [14]) or in post-test scores on stereochemistry learning.

Table 7c: Summary table of characteristics and results of included studies using Virtual and Augmented Reality. (CL = cognitive load, IM = intrinsic motivation, ML = motivation to learn, PE = perceived enjoyment)

Authors	Participants; expertise	Study design; immersive	Learning design	Risk of	Main results (measurements used)

	level	technology vs. comparative medium		bias	
Tugtekin & Odabasi [84]	349 undergraduates' students; intermediate	RCT; desktop AR vs. HMD VR (with different multimedia principles)	Science; declarative; retention	Some concerns	The results indicate no significant differences on learning or CL (Kılıc¸ & Karadeniz [41], and secondary task reaction) between the two media conditions (AR vs. VR). In contrast, the authors demonstrate that multimedia principles affect participants' objective CL in AR and VR.

5. Recommendations

Based on the findings from this systematic review, several recommendations emerge for researchers, educators, and technology developers to advance the field of AR/VR for learning.

5.1. Use of reliable metrics for objective learning performance, cognitive load, and intrinsic motivation evaluation

Methodological improvements are imperative in researching the effects of immersive technologies on learning, cognitive load, and intrinsic motivation. Of the 2404 studies initially screened, only 35 met the eligibility criteria, indicating a deficiency in well-designed studies. Additionally, a slight majority of these included studies reported inconsistent results based on our analysis grid, revealing prevalent methodological issues, particularly concerning intrinsic motivation. Thus, our findings support those of Hamilton et al. [25], underscoring the necessity for more rigorous methodologies in evaluating the effect of immersive technologies on learning outcomes. We recommend measuring learning performance using appropriate metrics to avoid overly general, short-term measures that may not be aligned with the task at hand, thereby preventing a loss of valuable insights.

Furthermore, in assessing the effect of VR/AR on cognitive load, we advocate for tailored tasks with appropriate difficulty levels and well-designed experiments to mitigate potential floor or ceiling effects and the influence of uncontrolled variables such as gamification, as reported in some included studies. It is also essential to employ adapted and validated measurements, with greater transparency in scale modifications. Similarly, due to methodological concerns, this literature review cannot affirm an existing effect of immersive technologies on intrinsic motivation. More rigorous protocols are warranted, with clearer delineation of the motivational factors under study and the use of standardized and more adapted measurements of intrinsic motivation. A best practice may involve utilizing metrics proposed by the Center for Self-Determination Theory (https://selfdeterminationtheory.org/questionnaires).

5.2. Informing about the consistency of results between objective performance and subjective performance (CL or IM) to support interpretation of AR/VR for learning

While a reduction in cognitive load induced by a VR/AR condition may seem favorable, it is essential to ensure that this decrease is coupled with an enhancement in objective learning performance to definitively confirm the beneficial effect of AR/VR on cognitive load-related learning. The same rationale applies to the correlation between objective learning performance and intrinsic motivation.

We recommend utilizing the proposed analysis grid for evaluating immersive technologies, as it offers a unique operational approach rooted in the principle of necessary concomitance between learning and cognitive/motivational variables. This ensures that observed benefits of VR/AR on learning are credibly linked to improvements in cognitive load or intrinsic motivation, thereby enhancing the reliability and consistency of assessments and contributing to a more robust understanding of the impact of immersive technologies. Additionally, we advocate for considering both cognitive and motivational factors in VR/AR assessment.

5.3. Further exploration of the links between cognitive load and intrinsic motivation

This systematic review highlights the importance of considering cognitive load and intrinsic motivation to explain the effects of VR/AR on learning by measuring these two factors simultaneously. Moreover, links between these two variables appear to exist, as suggested by recent literature and the models of Huang and Petersen. These studies show that the interaction between cognitive load and intrinsic motivation can influence learning outcomes in VR/AR. We therefore encourage further research to explore these links, building on the work of Huang and Petersen, in order to better understand and optimize the use of immersive technologies in educational contexts, especially for AR.

5.4. Keep in mind the variable use constraints across various AR/VR devices

This systematic review highlights the distinct effects of VR and AR on learners' cognitive load. However, due to the significant variability in devices used across studies, it is essential to provide precise descriptions of the immersive technologies studied in future research endeavors.

In general, the effectiveness of immersive technologies appears to hinge upon learners' prior knowledge of the subject matter. For individuals with limited prior knowledge, particularly in the context of the subject covered in the lesson, AR emerges as a more suitable option. AR aids in attenuating irrelevant cognitive load, thereby preserving cognitive resources for knowledge acquisition. Conversely, the intricate nature and wealth of information presented in VR environments tend to overwhelm novice learners, resulting in diminished learning outcomes.

If learners already possess some prior knowledge of the course material, we recommend VR, as it appears to have positive effects on learning outcomes and cognitive load management for intermediate learners. Prior knowledge allows participants to effectively handle the cognitive load imposed by the VR system, thereby leveraging the advantages offered by VR. Conversely, augmented reality (AR) seems to be less advantageous for intermediate learners compared to those with no prior knowledge. Consistent with the expertise reversal effect, AR assistance, although beneficial for novices, may not be as necessary or impactful for learners with more advanced levels of expertise.

In a broader context, the physical and cognitive demands associated with using AR/VR warrant even greater caution, particularly for users with limited technological skills, such as K-12 students or older adults. This consideration was one of the motivations behind excluding studies involving these specific learner groups to ensure consistency in the literature review. We recommend additional studies to extend the current findings to these populations.

6. Limitations

The study has several limitations that should be acknowledged. Firstly, like any systematic review, the definition of search terms and selection criteria was subjective, potentially resulting in the exclusion of some relevant articles.

Secondly, the restriction to three databases for conducting the systematic review may have limited the identification of relevant studies. However, this limitation is somewhat mitigated by the fact that Scopus, Web of Science, and PsychInfo are commonly searched in multidisciplinary and psycho-educational contexts, and the search query was specified to ensure the replicability of the review.

Thirdly, although efforts were made to minimize selection bias through the participation of three reviewers, the control of the selection process was only conducted on a small proportion of articles (10%), which may not completely eliminate bias.

Fourthly, to ensure the inclusion of reliable, high-quality sources of information, only controlled trial designs with objective learning performance measurement were included in the analysis. However, less stringent designs and/or qualitative analyses are also valuable sources of information, particularly when considering user experience.

Fifthly, while no included studies reported any significant effects of cybersickness, it is important to acknowledge this potential limitation when considering the use of VR and AR technologies. Additionally, this systematic review did not directly address user experience and feedback, both of which play critical roles in determining the success of immersive technology interventions. However, these elements could potentially impact cognitive load and intrinsic motivation in users of immersive technologies.

Lastly, this literature review aimed to provide a broad and general overview of the effects of these technologies on learning, considering cognitive load and intrinsic motivation. However, there was considerable variability between the devices used, learning domains, and learning dimensions in the included studies. Therefore, it would be beneficial to address the questions raised in this review more specifically in future research.

7. Conclusion

The aim of this systematic review, encompassing 36 studies, was to contribute to the expanding research field on immersive technologies and learning, particularly focusing on cognitive load and intrinsic motivation, and exploring their interplay within immersive settings.

The first research question investigated the effect of immersive technologies on learning as the result of optimized cognitive load. The results indicate that this effect depends on the nature of the technology used. Specifically, regarding the first research question, the findings suggest that virtual reality tends to impose extraneous cognitive load, potentially hindering declarative retention, especially among novice learners. In contrast, augmented reality demonstrates promising results by optimizing cognitive load, making it more suitable for novice learners. Consequently, considering RQ1 about cognitive load, this literature review tends to suggest that AR appears to be better adapted for learners without prior knowledge, while VR may offer more benefits to students with prior expertise.

This literature review then examined the effect of immersive technologies on learning and intrinsic motivation. The results do not provide a definitive understanding of the impact of immersive technologies on intrinsic motivation to learn (IM). Many studies reported effects inconsistent with self-determination theory and its derived models of IM-based learning. Therefore, the need for more methodologically rigorous research, incorporating well-defined and accurately assessed intrinsic motivation, is emphasized to thoroughly address how VR and AR affect learning performance as the result of increased intrinsic motivation (RQ2).

Regarding the third research question, the results indicate potential links between cognitive load and intrinsic motivation variables. However, a comprehensive understanding of the nature of these interactions remains elusive. Consequently, further research is deemed necessary to determine if causal relationships between cognitive load and intrinsic motivation can provide a better comprehension of the effectiveness of immersive technologies in learning (RQ3).

In addition, a novel operational method was developed specifically for this literature review to analyze the effects of immersive technologies on learning, predicted by cognitive load and intrinsic motivation (Figure 2). This operational and reusable framework is based on the principle of the necessary concomitance of a real change in objective learning performance and a real change in cognitive load and/or intrinsic motivation. It ensures that the benefits of VR/AR on learning are attributed to improvements in cognitive load or intrinsic motivation. The principle excludes subjective measures of learning performance but does not exclude objective measures of cognitive load or intrinsic motivation (e.g., physiological indicators such as pupil dilation or EEG signals related to controlled attention for cognitive load and active exploration or verbal requests for curiosity). By adhering to such a principle, better categorization of effects can be achieved, aligning with the assumptions of cognitive load and motivation-focused learning theories, and minimizing the risk of drawing conclusions not supported by evidence. Therefore, the adoption of this method as a common guiding framework to analyze the effects of specific media on cognitive load and intrinsic motivation while considering learning performance is proposed.

8. Conflict of Interest

All authors declare no conflict of interest.

9. References

[1] Ai-Lim Lee, E. et al. (2010). How does desktop virtual reality enhance learning outcomes? A structural equation modeling approach. Computers & Education. https://doi.org/10.1016/j.compedu.2010.06.006.

[2] Altmeyer, K. et al. (2020). The use of augmented reality to foster conceptual knowledge acquisition in STEM laboratory courses—Theoretical background and empirical results. British Journal of Educational Technology. https://doi.org/10.1111/bjet.12900.

[3] 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. https://doi.org/10.1111/jcal.12478.

[4] Anderson et al. (2001). A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives.

[5] Arayaphan, W. et al. (2022). Enhancing Intrinsic Motivation of Librarian Students using Virtual Reality for Education in the Context of Culture Heritage Museums. TEM Journal. https://doi.org/10.18421/TEM112-16.

[6] Arnaldi, B. et al. (2018). Virtual reality and augmented reality : Myths and realities., Wiley-ISTE.

[7] Ayres, P. (2006). Using subjective measures to detect variations of intrinsic cognitive load within problems. Learning and Instruction. https://doi.org/10.1016/j.learninstruc.2006.09.001.

[8] Baceviciute, S. et al. (2021). Remediating learning from non-immersive to immersive media: Using EEG to investigate the effects of environmental embeddedness on reading in Virtual Reality. Computers & Education. https://doi.org/10.1016/j.compedu.2020.104122.

[9] Buchner, J. et al. (2022). The impact of augmented reality on cognitive load and performance: A systematic review. Journal of Computer Assisted Learning. https://doi.org/10.1111/jcal.12617.

[10] Burgues, M. et al. (2020). Immersion and Control in Learning Art Knowledge: An Example in Museum Visit. Computer Supported Education.

[11] Campbell, M. et al. (2020). Synthesis without meta-analysis (SWiM) in systematic reviews: reporting guideline. BMJ. https://doi.org/10.1136/bmj.l6890.

[12] Cierniak, G. et al. (2009). Explaining the split-attention effect: Is the reduction of extraneous cognitive load accompanied by an increase in germane cognitive load? Computers in Human Behavior. https://doi.org/10.1016/j.chb.2008.12.020.

[13] Cowan, N. (2016). Working Memory Maturation: Can We Get at the Essence of Cognitive Growth? Perspectives on Psychological Science. https://doi.org/10.1177/1745691615621279.

[14] Deci, E.L. et al. (1994). Facilitating Internalization: The Self-Determination Theory Perspective. Journal of Personality. https://doi.org/10.1111/j.1467-6494.1994.tb00797.x.

[15] Deslauriers, L. et al. (2019). Measuring actual learning versus feeling of learning in response to being actively engaged in the classroom. Proceedings of the National Academy of Sciences. https://doi.org/10.1073/pnas.1821936116.

[16] Di Natale, A.F. et al. (2020). Immersive virtual reality in K-12 and higher education: A 10-year systematic review of empirical research. British Journal of Educational Technology. https://doi.org/10.1111/bjet.13030.

[17] Elford, D. et al. (2022). Exploring the Effect of Augmented Reality on Cognitive Load, Attitude, Spatial Ability, and Stereochemical Perception. Journal of Science Education and Technology. https://doi.org/10.1007/s10956-022-09957-0.

[18] Elford, D. et al. (2022). Fostering Motivation toward Chemistry through Augmented Reality Educational Escape Activities. A Self-Determination Theory Approach. Journal of Chemical Education. https://doi.org/10.1021/acs.jchemed.2c00428.

[19] Fan, M. et al. (2020). Augmented Reality for Early Language Learning: A Systematic Review of Augmented Reality Application Design, Instructional Strategies, and Evaluation Outcomes. Journal of Educational Computing Research. https://doi.org/10.1177/0735633120927489.

[20] Feldon, D.F. et al. (2019). Cognitive Load as Motivational Cost. Educational Psychology Review. https://doi.org/10.1007/s10648-019-09464-6.

[21] Feldon, D.F. et al. (2018). Self-efficacy change associated with a cognitive load-based intervention in an undergraduate biology course. Learning and Instruction. https://doi.org/10.1016/j.learninstruc.2018.04.007.

[22] Frederiksen, J.G. et al. (2020). Cognitive load and performance in immersive virtual reality versus conventional virtual reality simulation training of laparoscopic surgery: a randomized trial. Surgical Endoscopy. https://doi.org/10.1007/s00464-019-06887-8.

[23] Frithioff, A. et al. (2020). Ultra-high-fidelity virtual reality mastoidectomy simulation training: a randomized, controlled trial. European Archives of Oto-Rhino-Laryngology. https://doi.org/10.1007/s00405-020-05858-3.

[24] Geng, X. and Yamada, M. (2020). An augmented reality learning system for Japanese compound verbs: study of learning performance and cognitive load. Smart Learning Environments. https://doi.org/10.1186/s40561-020-00137-4.

[25] Guay, F. et al. (2000). On the Assessment of Situational Intrinsic and Extrinsic Motivation: The Situational Motivation Scale (SIMS). Motivation and Emotion. https://doi.org/10.1023/A:1005614228250.

[26] Hamilton, D. et al. (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. https://doi.org/10.1007/s40692-020-00169-2.

[27] Hart, S.G. and Staveland, L.E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research, in Advances in Psychology, vol. 52, North-Holland, pp. 139–183.

[28] Huang, H.-M. et al. (2010). Investigating learners' attitudes toward virtual reality learning environments: Based on a constructivist approach. Computers & Education. https://doi.org/10.1016/j.compedu.2010.05.014.

[29] Huang, W. et al. (2022). Extending the Cognitive-Affective Theory of Learning with Media in Virtual Reality Learning: A Structural Equation Modeling Approach. Journal of Educational Computing Research. https://doi.org/10.1177/07356331211053630.

[30] Huang, X. (2017). Example-based learning: Effects of different types of examples on student performance, cognitive load and self-efficacy in a statistical learning task. Interactive Learning Environments. https://doi.org/10.1080/10494820.2015.1121154.

[31] Hwang, G.-J. et al. (2013). A concept map-embedded educational computer game for improving students' learning performance in natural science courses. Computers & Education. https://doi.org/10.1016/j.compedu.2013.07.008.

[32] Jin, S. et al. (2022). Immersive Spring Morning in the Han Palace: Learning Traditional Chinese Art Via Virtual Reality and Multi-Touch Tabletop. International Journal of Human–Computer Interaction. https://doi.org/10.1080/10447318.2021.1930389.

[33] Kalyuga, S. (2009). The Expertise Reversal Effect, in Managing Cognitive Load in Adaptive Multimedia Learning (ed.Kalyuga, S.), IGI Global, Hershey, PA, USA, pp. 58–80.

[34] Kalyuga, S. (2011). Cognitive Load Theory: How Many Types of Load Does It Really Need? Educational Psychology Review. https://doi.org/10.1007/s10648-010-9150-7.

[35] Kanyongo, G.Y. et al. (2007). Reliability and Statistical Power: How Measurement Fallibility Affects Power and Required Sample Sizes for Several Parametric and Nonparametric Statistics. Journal of Modern Applied Statistical Methods. https://doi.org/10.56801/10.56801/v6.i.284.

[36] Kapp, S. et al. (2021). Using augmented reality in an inquiry-based physics laboratory course. Computer Supported Education: 12th International Conference, CSEDU 2020, Virtual Event, May 2–4, 2020, Revised Selected Papers 12.

[37] Keller, S. et al. (2021). Cognitive Load Implications for Augmented Reality Supported Chemistry Learning. Information. https://doi.org/10.3390/info12030096.

[38] Kılıç, A.G.E. and Karadeniz, O.G.S. (2004). Hiper Ortamlarda Öğrencilerin Bilişsel Yüklenme ve Kaybolma Düzeylerinin Belirlenmesi. Kuram ve Uygulamada Eğitim Yönetimi.

[39] Klepsch, M. et al. (2017). Development and Validation of Two Instruments Measuring Intrinsic, Extraneous, and Germane Cognitive Load. Frontiers in Psychology. https://doi.org/10.3389/fpsyg.2017.01997.

[40] Klingenberg, S. et al. (2020). Investigating the effect of teaching as a generative learning strategy when learning through desktop and immersive VR: A media and methods experiment. British Journal of Educational Technology. https://doi.org/10.1111/bjet.13029.

[41] Knogler, M. et al. (2015). How situational is situational interest? Investigating the longitudinal structure of situational interest. Contemporary Educational Psychology. https://doi.org/10.1016/j.cedpsych.2015.08.004.

[42] Küçük, S. et al. (2016). Learning anatomy via mobile augmented reality: Effects on achievement and cognitive load. Anatomical Sciences Education. https://doi.org/10.1002/ase.1603.

[43] Lee, C.-J. and Hsu, Y. (2021). Sustainable Education Using Augmented Reality in Vocational Certification Courses. Sustainability. https://doi.org/10.3390/su13116434.

[44] Leppink, J. et al. (2013). Development of an instrument for measuring different types of cognitive load. Behavior Research Methods. https://doi.org/10.3758/s13428-013-0334-1.

[45] Leung, X.Y. et al. (2022). Is VR game training more effective for hospitality employees? A longitudinal experiment. Tourism Management Perspectives. https://doi.org/10.1016/j.tmp.2022.101020.

[46] Liu, X.-W. et al. (2022). Research on Training Effectiveness of Professional Maintenance Personnel Based on Virtual Reality and Augmented Reality Technology. Sustainability. https://doi.org/10.3390/su142114351.

[47] Makransky, G. et al. (2019). Motivational and cognitive benefits of training in immersive virtual reality based on multiple assessments. Journal of Computer Assisted Learning. https://doi.org/10.1111/jcal.12375.

[48] Makransky, G. and Klingenberg, S. (2022). Virtual reality enhances safety training in the maritime industry: An organizational training experiment with a non-WEIRD sample. Journal of Computer Assisted Learning. https://doi.org/10.1111/jcal.12670.

[49] 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. https://doi.org/10.1007/s11423-018-9581-2.

[50] 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. https://doi.org/10.1007/s10648-020-09586-2.

[51] Makransky, G. and Petersen, G.B. (2019). Investigating the process of learning with desktop virtual reality: A structural equation modeling approach. Computers & Education. https://doi.org/10.1016/j.compedu.2019.02.002.

[52] Makransky, G. et al. (2019). Adding immersive virtual reality to a science lab simulation causes more presence but less learning. Learning and Instruction. https://doi.org/10.1016/j.learninstruc.2017.12.007.

[53] Makransky, G. et al. (2016). Virtual Simulations as Preparation for Lab Exercises: Assessing Learning of Key Laboratory Skills in Microbiology and Improvement of Essential Non-Cognitive Skills. PLOS ONE. https://doi.org/10.1371/journal.pone.0155895.

[54] Margrett, J.A. et al. (2022). Older Adults' Use of Extended Reality: A Systematic Review. Frontiers in Virtual Reality. https://doi.org/10.3389/frvir.2021.760064.

[55] Mayer, R.E. (2020). Multimedia Learning, Cambridge University Press.

[56] Milyavskaya, M. et al. (2021). More Effort, Less Fatigue: The Role of Interest in Increasing Effort and Reducing Mental Fatigue. Frontiers in Psychology. https://doi.org/10.3389/fpsyg.2021.755858.

[57] Monteiro, V. et al. (2015). Intrinsic motivation inventory: Psychometric properties in the context of first language and mathematics learning. Psicologia: Reflexão e Crítica. https://doi.org/10.1590/1678-7153.201528302.

[58] Murayama, K. et al. (2019). Process Account of Curiosity and Interest: A Reward-Learning Perspective. Educational Psychology Review. https://doi.org/10.1007/s10648-019-09499-9.

[59] Newman, H.J. et al. (2022). Technology enhanced neuroanatomy teaching techniques: A focused BEME systematic review of current evidence: BEME Guide No. 75. Medical Teacher. https://doi.org/10.1080/0142159X.2022.2039382.

[60] Oudeyer, P.-Y. et al. (2016). Chapter 11 - Intrinsic motivation, curiosity, and learning: Theory and applications in educational technologies, in Progress in Brain Research, vol. 229, Elsevier, pp. 257–284.

[61] Paas, F.G.W.C. and Van Merriënboer, J.J.G. (1994). Variability of worked examples and transfer of geometrical problem-solving skills: A cognitive-load approach. Journal of Educational Psychology. https://doi.org/10.1037/0022-0663.86.1.122.

[62] Page, M.J. et al. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. https://doi.org/10.1136/bmj.n71.

[63] Pande, P. et al. (2021). Long-Term Effectiveness of Immersive VR Simulations in Undergraduate Science Learning: Lessons from a Media-Comparison Study. Research in Learning Technology.

[64] Parong, J. and Mayer, R.E. (2021). Cognitive and affective processes for learning science in immersive virtual reality. Journal of Computer Assisted Learning. https://doi.org/10.1111/jcal.12482.

[65] Parong, J. and Mayer, R.E. (2021). Learning about history in immersive virtual reality: does immersion facilitate learning? Educational Technology Research and Development. https://doi.org/10.1007/s11423-021-09999-y.

[66] Petersen, G.B. et al. (2022). A study of how immersion and interactivity drive VR learning. Computers & Education. https://doi.org/10.1016/j.compedu.2021.104429.

[67] Radianti, J. et al. (2020). A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. Computers & Education. https://doi.org/10.1016/j.compedu.2019.103778.

[68] Radu, I. (2014). Augmented reality in education: a meta-review and cross-media analysis. Personal and Ubiquitous Computing. https://doi.org/10.1007/s00779-013-0747-y.

[69] Reber, R. and Greifeneder, R. (2017). Processing Fluency in Education: How Metacognitive Feelings Shape Learning, Belief Formation, and Affect. Educational Psychologist. https://doi.org/10.1080/00461520.2016.1258173.

[70] Ryan, R.M. (1982). Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory. Journal of Personality and Social Psychology. https://doi.org/10.1037/0022-3514.43.3.450.

[71] Ryan, R.M. and Deci, E.L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. American Psychologist. https://doi.org/10.1037/0003-066X.55.1.68.

[72] Santos Silva, R. et al. (2019). Virtual reality for older users: a systematic literature review. International Journal of Virtual Reality. https://doi.org/10.20870/IJVR.2019.19.1.2908.

[73] Schardt, C. et al. (2007). Utilization of the PICO framework to improve searching PubMed for clinical questions. BMC Medical Informatics and Decision Making. https://doi.org/10.1186/1472-6947-7-16.

[74] Singh, G. et al. (2019). Evaluating the impact of the augmented reality learning environment on electronics laboratory skills of engineering students. Computer Applications in Engineering Education. https://doi.org/10.1002/cae.22156.

[75] Skulmowski, A. (2023). Guidelines for Choosing Cognitive Load Measures in Perceptually Rich Environments. Mind, Brain, and Education. https://doi.org/10.1111/mbe.12342.

[76] 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. https://doi.org/10.1007/s10648-021-09624-7.

[77] Slater, M. and Wilbur, S. (1997). A Framework for Immersive Virtual Environments (FIVE): Speculations on the Role of Presence in Virtual Environments. Presence: Teleoperators and Virtual Environments. https://doi.org/10.1162/pres.1997.6.6.603.

[78] Soloman, B.A. and Felder, R.M. (2005). Index of learning styles questionnaire. NC State University. Available online at: http://www. engr. ncsu. edu/learningstyles/ilsweb. html (last visited on 14.05. 2010).

[79] Sterne, J.A.C. et al. (2019). RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ. https://doi.org/10.1136/bmj.l4898.

[80] Suh, A. and Prophet, J. (2018). The state of immersive technology research: A literature analysis. Computers in Human Behavior. https://doi.org/10.1016/j.chb.2018.04.019.

[81] Sweller, J. (1988). Cognitive Load During Problem Solving: Effects on Learning. Cognitive Science. https://doi.org/10.1207/s15516709cog1202_4.

[82] Sweller, J. et al. (2019). Cognitive Architecture and Instructional Design: 20 Years Later. Educational Psychology Review. https://doi.org/10.1007/s10648-019-09465-5.

[83] Tang, Q. et al. (2022). Experiencing an art education program through immersive virtual reality or iPad: Examining the mediating effects of sense of presence and extraneous cognitive load on enjoyment, attention, and retention. Frontiers in Psychology. https://doi.org/10.3389/fpsyg.2022.957037.

[84] Thees, M. et al. (2022). Augmented Reality for Presenting Real-Time Data During Students' Laboratory Work: Comparing a Head-Mounted Display With a Separate Display. Frontiers in Psychology. https://doi.org/10.3389/fpsyg.2022.804742.

[85] Thees, M. et al. (2020). Effects of augmented reality on learning and cognitive load in university physics laboratory courses. Computers in Human Behavior. https://doi.org/10.1016/j.chb.2020.106316.

[86] Tokel, S.T. and İsler, V. (2015). Acceptance of virtual worlds as learning space. Innovations in Education and Teaching International. https://doi.org/10.1080/14703297.2013.820139.

[87] Tugtekin, U. and Odabasi, H.F. (2022). Do Interactive Learning Environments Have an Effect on Learning Outcomes, Cognitive Load and Metacognitive Judgments? Education and Information Technologies. https://doi.org/10.1007/s10639-022-10912-0.

[88] Turan, Z. et al. (2018). The impact of mobile augmented reality in geography education: achievements, cognitive loads and views of university students. Journal of Geography in Higher Education. https://doi.org/10.1080/03098265.2018.1455174.

[89] Wu, B. et al. (2020). Effectiveness of immersive virtual reality using head-mounted displays on learning performance: A meta-analysis. British Journal of Educational Technology. https://doi.org/10.1111/bjet.13023.

[90] Yang, J. et al. (2021). Effect of virtual reality technology on the teaching of urban railway vehicle engineering. Computer Applications in Engineering Education. https://doi.org/10.1002/cae.22371.

[91] Zhang, L. et al. (2022). There is an Evidence Crisis in Science Educational Policy. Educational Psychology Review. https://doi.org/10.1007/s10648-021-09646-1.

[92] Zhao, J. et al. (2020). Using the Summarizing Strategy to Engage Learners: Empirical Evidence in an Immersive Virtual Reality Environment. The Asia-Pacific Education Researcher. https://doi.org/10.1007/s40299-020-00499-w.