Computers & Education: X Reality 7 (2025) 100117

Contents lists available at ScienceDirect

Computers & Education: X Reality

journal homepage: www.journals.elsevier.com/computers-and-education-x-reality

Effects of school-based immersive virtual reality interventions on learning in the K-6 range: A systematic literature review

Josua Dubach [a,b,*], Sofia Anzeneder [c], Sebastian Tempelmann [a], Trix Cacchione [c]

[a] Bern University of Teacher Education, Fabrikstrasse 8, 3012, Bern, Switzerland
[b] Zurich University of Teacher Education, Lagerstrasse 2, 8090, Zürich, Switzerland
[c] Institute of Primary Education, University of Applied Sciences and Arts Northwestern Switzerland, Bahnhofstrasse 6, 5210, Windisch, Switzerland

ARTICLE INFO

Keywords:
Immersive virtual reality
Primary education
Affordances
Motivation
Self-efficacy
Embodiment
Cognitive load

ABSTRACT

Immersive virtual reality (IVR) is a rapidly evolving technology that has the potential to enhance learning. Recent experimental studies report higher learning gains in IVR environments compared to other VR types or analog teaching methods. However, effect sizes vary significantly, suggesting that IVR's effectiveness may be constrained by moderators related to IVR affordances (presence and agency), underlying motivational and cognitive factors, individual characteristics (age, prior knowledge of curriculum content), and IVR instructional setting. To address this issue, we conducted a PRISMA-guided systematic review of randomized controlled and quasi-experimental studies with K–6 learners across four databases (PsycINFO, PubMed, Web of Science, ERIC). Twenty-four studies were included. We analyzed IVR's effects on content-dependent knowledge acquisition and transfer, considering IVR affordances of agency and presence, underlying motivational and cognitive factors, and potential differential effects related to individual characteristics and IVR instructional setting (e.g. scaffolding measures). Results suggest that IVR can effectively enhance knowledge acquisition and transfer for K-6 learners, outperforming both analog teaching methods and non-immersive VR. These gains appear to be partially driven by increased agency and presence, as well as modulations in motivational and cognitive factors. However, contrary to previous evidence, individual characteristics such as age and prior knowledge do not consistently moderate IVR's effects on learning. Preliminary evidence further highlights the importance of instructional settings, particularly the inclusion of reflective activities and scaffolding measures. To maximize the potential of IVR in education, further research should systematically examine the individual and interactive effects of IVR affordances, cognitive and motivational factors, and instructional design.

1. Introduction

In recent years, virtual reality (VR) has assumed an increasingly significant role in learner-centred approaches, owing to its capacity to create engaging and interactive environments that enable learners to take an active role in the learning process (Makransky & Petersen, 2021). Learner-centred educational practices have been shown to enhance learners' academic achievement, engagement, and motivation (Bremner et al., 2022; Li, Y.D. et al., 2021). From a constructivist perspective, learners build their own knowledge by extracting meaning from different sources of information, particularly through sensorimotor experiences (Ertmer & Newby, 2013). This is especially relevant for children in the kindergarten through 6th grade (K-6), for whom bodily interaction with the environment constitutes a central mechanism of learning (Barsalou, 2008; Laakso, 2011).

To support such learner-centred processes, models like the 5E cycle highlight the importance of active participation across phases of engagement, exploration, explanation, elaboration, and evaluation (Putra et al., 2018; Sarac, 2018). Active engagement with sensory information is fundamental to the construction of all forms of knowledge. Even abstract concepts must be anchored in real-world experiences to foster stable and flexible understanding (Wilson & Golonka, 2013). Learning efficacy depends on the degree to which children acquire and transfer factual knowledge, conceptual understanding, procedural skills, and metacognitive knowledge through active involvement (Anderson et al., 2001). Factual knowledge, which includes discrete elements such as terminology and specific details, is more effectively retained when linked to concrete sensory experiences (Abrahamson, 2014). Conceptual

* Corresponding author. Bern University of Teacher Education, Fabrikstrasse 8, 3012, Bern, Switzerland.
E-mail address: josua.dubach@phbern.ch (J. Dubach).

https://doi.org/10.1016/j.cexr.2025.100117
Received 9 April 2025; Received in revised form 3 September 2025; Accepted 9 September 2025
Available online 25 September 2025

knowledge, encompassing categories, principles, theories, and models, deepens when abstract ideas are grounded in tangible experiences (Lindgren et al., 2022). Procedural knowledge - skills, techniques, and methods - is best developed through hands-on activities that embody the principle of learning by doing (Catrambone & Yuasa, 2006). Finally, metacognitive knowledge, including strategic, contextual, and self-knowledge, is enhanced when learners actively engage and reflect on their own thinking processes (Anderson et al., 2001). Thus, active sensorimotor engagement is essential not only for acquiring, but also for flexibly applying knowledge across all four domains.

1.1. Virtual reality systems

Virtual reality (VR) is broadly defined as a computer-generated simulation that enables users to interact with three-dimensional environments at varying levels of immersion (Di Natale et al., 2020). Its adoption in learner-centred approaches is driven by its potential to actively engage learners and enable hands-on interaction within virtual environments (Makransky & Petersen, 2021), thereby creating new opportunities to enhance educational practice (Amarin & Ghishan, 2013). VR systems are typically classified as non-immersive (e.g., desktop VR), semi-immersive (e.g., fulldome or mixed reality), and fully immersive VR (IVR). The latter provides the highest level of immersion through head-mounted displays or CAVE environments, surrounding users perceptually and enhancing presence via egocentric navigation, stereoscopic vision, and real-time sensorimotor contingencies (Dalgarno & Lee, 2010; Di Natale et al., 2020; Slater & Sanchez-Vives, 2016).

Within this spectrum, IVR has gained particular interest in educational contexts (Coban et al., 2022). It allows learners to experience 360° three-dimensional (3D) environments from a first-person perspective while interacting dynamically with visual objects via handheld controllers and a virtual body. These interactions integrate visual, auditory, and tactile feedback, providing opportunities for hands-on experimentation (Makransky & Lilleholt, 2018). Compared to less immersive systems, IVR is distinguished by its multisensory stimulation, situated immersion, real-time action-based perception, a sense of embodiment, and multimodal interaction (Di Natale et al., 2020; Ijsselsteijn & Riva, 2003). Together, these features support experiential learning, enhance the processing of multimodal (visual, auditory, and sensorimotor) information, increase learner engagement and motivation, and ultimately improve knowledge acquisition and transfer (Coban et al., 2022; Southgate et al., 2019).

1.2. Evidence from meta-analyses and systematic reviews

Several systematic reviews and meta-analyses have synthesized evidence on VR- and IVR-based learning across different populations (Coban et al., 2022; Di Natale et al., 2020; Jensen & Konradsen, 2018; Villena-Taranilla et al., 2022a; Wu et al., 2020). Taken together, these works reveal an inconsistent pattern of findings in higher education, but generally larger learning gains for primary school students when learning in IVR environments compared to other forms of VR or traditional methods (Coban et al., 2022; Conrad et al., 2024; Villena-Taranilla et al., 2022a; Wu et al., 2020). One possible explanation is that primary school children depend more heavily on active engagement during learning and therefore benefit particularly from the interactive, immersive qualities of IVR (Makransky & Petersen, 2021).

At the same time, meta-analytical evidence reveals substantial variation in reported effect sizes for IVR interventions in primary education, ranging from 0.24 to 0.57 (Coban et al., 2022; Villena-Taranilla et al., 2022a; Wu et al., 2020). This variability suggests that the overall effectiveness of IVR interventions in primary education may depend on moderating factors beyond the medium itself. Prior reviews point to specific IVR affordances (Villena-Taranilla et al., 2022a), underlying motivational and cognitive mechanisms (Makransky & Petersen, 2021; Petersen et al., 2022), individual learner characteristics such as age and prior knowledge (Coban et al., 2022; Wu et al., 2020), and instructional design features, including how IVR is implemented in educational settings (Conrad et al., 2024). Importantly, no previous review has simultaneously focused on content-related knowledge acquisition and transfer in younger learners while considering IVR affordances and cognitive-motivational mechanisms. This absence highlights the specific research gap that underlies and motivates the present study. Taken together, previous reviews demonstrate both the promise and the variability of IVR-based learning outcomes. To better understand these mixed findings and the mechanisms through which IVR exerts its effects, it is essential to consider theoretical models that explain how specific affordances shape learning processes.

1.3. Theoretical models explaining IVR learning

IVR affordances are features of the virtual environment that facilitate learning behaviors, such as selecting relevant information and integrating it into coherent knowledge structures (Di Natale et al., 2020). Among these, agency and presence are particularly important psychological affordances. They are shaped by technical attributes of IVR systems, including immersion, interaction, and the fidelity of virtual representations (Makransky & Petersen, 2021; Petersen et al., 2022). Yet, it remains unclear how different IVR configurations, such as head-mounted displays compared to mobile-based VR, differ in their immersive quality and how these differences impact agency and presence (Di Natale et al., 2020). Agency refers to the learner's sense of control, autonomy, and ability to act meaningfully within the virtual environment, whereas presence, often described as the subjective experience of "being there", remains conceptually debated (Cummings & Bailenson, 2016). Lee (2004) distinguishes between three forms of presence: social presence (interacting with social agents such as virtual instructors), self-presence (awareness of one's body within the environment), and spatial presence (the sense of being physically located in a realistic virtual space). Most empirical studies, however, focus on spatial presence (e.g., Di Natale et al., 2020). This concept is grounded in Wirth et al.'s (2007) two-stage model of presence formation: individuals first perceive the mediated environment as coherent and plausible, and subsequently experience themselves as physically situated within it.

The Cognitive Affective Model of Immersive Learning (CAMIL) explains the motivational and cognitive mechanisms through which the IVR affordances of agency and presence influence knowledge acquisition and transfer (Makransky & Petersen, 2021; Petersen et al., 2022). It identifies six key factors: 1. situational interest (i.e., focused attention and affective engagement with specific stimuli; Hidi & Renninger, 2006), 2. intrinsic motivation (i.e., engagement driven by inherent satisfaction with the activity; Ryan & Deci, 2015), 3. self-regulation (i.e., ability to manage behavior, maintain focus, and carry out relevant tasks; Schunk & DiBenedetto, 2016), 4. self-efficacy (i.e., perceived competence in learning or performing actions; Schunk & DiBenedetto, 2016), 5. cognitive load (i.e., mental effort required to process information in working memory; Sweller et al., 2011), and 6. embodiment (i.e., sensorimotor experiences during interaction with the environment that shape cognitive processes; Barsalou, 2008; Ertmer & Newby, 2013).

These factors reflect different blends of affective, motivational, and cognitive components. Situational interest and intrinsic motivation are primarily motivational, cognitive load and self-regulation more cognitive, while self-efficacy and embodiment combine both. For example, self-efficacy fosters confidence in one's abilities, while embodiment links sensorimotor experiences with cognitive engagement. CAMIL posits that IVR environments can enhance motivational and self-regulatory factors, namely, situational interest, intrinsic motivation, self-efficacy, and self-regulation, by supporting basic psychological needs for autonomy, competence, and social relatedness (Deci & Ryan, 2015). Features such as choice, acknowledgment of learners' perspectives, and feedback from virtual instructors contribute to these needs. Moreover, authentic and controllable immersive experiences can strengthen self-efficacy by reinforcing learners' sense of acting effectively within a believable virtual context (Makransky & Petersen, 2021). Empirical findings consistently support the learning benefits of these four motivational and self-regulatory factors, whereas evidence on embodiment and cognitive load is more mixed (Makransky & Petersen, 2021; Petersen et al., 2022). On the one hand, embodiment can foster self-presence, which appears to reduce cognitive load (Makransky & Petersen, 2021; Paas & Sweller, 2012). For instance, the experience of owning and controlling a virtual body can support embodied cognition by linking sensorimotor input with abstract learning content, thereby grounding cognitive processes and reducing mental effort. On the other hand, IVR may hinder learning (Lawson & Mayer, 2024; Makransky et al., 2021) when it induces extraneous cognitive load through irrelevant or overly complex elements (Sweller, 2020; Sweller et al., 2011). High levels of agency, for instance, can increase cognitive load when learners navigate complex simulations with too much autonomy, leading to poorer outcomes compared to more guided formats such as instructional videos (Makransky et al., 2021).

1.4. Learner characteristics and instructional context as moderators

In addition to IVR affordances and the associated motivational and cognitive factors, individual learner characteristics may moderate IVR's impact on learning. Meta-analyses show that younger learners (K-12) benefit more from IVR learning environments than higher education students (Coban et al., 2022). One explanation is that IVR allows learners to observe, explore, and actively manipulate objects in a simulated 3D environment, reducing the cognitive effort of transforming 2D representations into 3D context (Araiza-Alba et al., 2021). This benefit is particularly relevant for K–6 learners, whose limited working memory makes them more susceptible to intrinsic cognitive load (i.e., the mental effort required to process new information given its complexity and the learner's expertise; Sweller et al., 2011, 2020). Evidence further suggests that even within the K–6 range, younger children may benefit more from IVR than older children (Mavilidi et al., 2022). Another key moderator is prior knowledge of curriculum content. Prior knowledge reduces intrinsic cognitive load by allowing learners to more easily interpret and organize new experiences (Meyer et al., 2019). Learners with high prior knowledge can better connect virtual experiences to existing conceptual frameworks, facilitating meaningful learning. Conversely, learners with little or no prior knowledge may experience IVR mainly as engaging or entertaining due to high levels of presence and agency, but may struggle to select, organize, and integrate information into coherent knowledge structures (Meyer et al., 2019).

Lastly, the effects of IVR on knowledge acquisition and transfer may also depend on the instructional setting. School-based IVR interventions may be more effective when accompanied by didactic supports such as teacher scaffolding, peer feedback, or generative strategies that help learners actively make sense of the content. These elements reduce the cognitive load associated with IVR environments, thereby enhance learning (Conrad et al., 2024; Mayer et al., 2023).

1.5. Research gap and study aim

To the best of our knowledge, existing reviews and meta-analyses have predominantly examined IVR interventions that are detached from curriculum content (e.g., Checa & Bustillo, 2020; Jensen & Konradsen, 2018), developed for secondary and higher education context (e.g., Hamilton et al., 2021; Radianti et al., 2020), and focused primarily on the acquisition of factual, conceptual, and procedural knowledge, while largely neglecting metacognitive knowledge (e.g., Conrad et al., 2024; Jensen & Konradsen, 2018; Villena-Taranilla et al., 2022a). Moreover, no systematic literature review to date has synthesized the effects of school-based IVR interventions on the acquisition and transfer of curriculum-related knowledge among K–6 learners using a theoretical framework that integrates the affordances of agency and presence alongside their associated motivational and cognitive mechanisms. A more nuanced understanding of the conditions under which IVR environments support learning in primary education could enable educators to make informed decisions about which lessons and learning processes are likely to benefit most from immersive technologies.

The primary aim of this systematic literature review was to synthesize the effects of IVR interventions integrated into academic lessons for grades K-6 on the acquisition and transfer of curriculum-related factual, conceptual, procedural and metacognitive knowledge. In doing so, the review draws on the CAMIL framework (Makransky & Petersen, 2021; Petersen et al., 2022) by examining not only the roles of **agency** and **presence**, but also the associated **motivational and cognitive factors** that shape learning in immersive environments. Building on emerging evidence of differential effects based on individual characteristics (Coban et al., 2022; Meyer et al., 2019), the second aim was to explore whether learners' age and prior knowledge of curriculum content impact the effects of IVR environments on knowledge acquisition and transfer. A third, exploratory aim was to examine whether instructional settings, that is, the extent and nature of didactic support within school-based IVR interventions, influence learning outcomes for K–6 students.

Accordingly, the present review addresses the following research questions:

(1) How do IVR interventions integrated into academic lessons for grades K–6 affect acquisition and transfer of factual, conceptual, procedural, and metacognitive knowledge, and what role do agency, presence, and associated motivational and cognitive factors play in this process?
(2) To what extent do learners' age and prior knowledge moderate the effects of IVR on knowledge acquisition and transfer?
(3) How do instructional settings influence learning outcomes in IVR-based interventions for K–6 students?

Our hypotheses were as follows: (1) In line with the CAMIL framework, agency and presence are expected to enhance knowledge acquisition and transfer to the extent that the IVR environment fosters situational interest, intrinsic motivation, self-regulation, self-efficacy, and/or embodiment, while reducing extraneous cognitive load. (2) and (3) due to the fragmentary nature of existing findings, no a priori hypotheses were formulated regarding the differential effects based of age (Coban et al., 2022), prior knowledge (Meyer et al., 2019), and instructional setting (Conrad et al., 2024; Mayer et al., 2023).

2. Methods

The current systematic literature review was performed in conformity with the PRISMA 2020 statement (Preferred Reporting Items for Systematic Reviews and Meta-Analyses; Page et al., 2021).

2.1. Search strategy

A comprehensive search strategy was developed to identify relevant reports investigating the effects of IVR interventions with learners in the K-6 range on curriculum-content dependent factual, conceptual, procedural, and metacognitive knowledge acquisition and transfer. A total of four electronic databases, including APA PsycINFO (OvidSP), PubMed, Web of Science (All Databases), and ERIC (EBSCO) were systematically searched. The search strategy included free text terms and related truncations referring to (#1) kindergarten and primary school children, combined with terms referring to (#2) IVR, (#3) factual, conceptual, procedural, and metacognitive knowledge acquisition and transfer, and (#4) randomized controlled trials (RCT) or quasi-experimental studies (QRS; see Appendix A for detailed search strings). A manual review was performed to identify additional relevant records from the reference lists of the included records and from published systematic literature reviews on school-based IVR interventions. The systematic literature search was conducted in May 2024.

2.2. Eligibility criteria

The PICOs model was used to define reports' inclusion and exclusion criteria (Amir-Behghadami & Janati, 2020). Inclusion criteria applied in the selection process were:

- (#1) *Population*: Kindergarten and primary school children aged between 5 and 13 years without developmental disorders.
- (#2a) *Intervention*: At least one experimental group with an IVR intervention during school lessons (i.e., head-mounted displays, mobile phone-based VR, CAVE).
- (#2b) *Comparison*: Waiting list, analog teaching methods, non-immersive VR intervention (e.g., computer, laptops, Tablets) or other IVR interventions.
- (#3) *Outcome*: Curriculum-content related factual, conceptual, procedural, and metacognitive knowledge acquisition and transfer, assessed through tests or task performances.
- (#4) *Study design*: Randomized controlled trials or quasi-experimental studies published until May 2024 in peer-reviewed journals in English language. The English language represents the common denominator to avoid unbalanced inclusion of few national languages.

2.3. Selection process

The search results were imported into Zotero bibliographic software, and duplicates were removed. Titles and abstracts of the identified records were then screened using ASReview Lab, a machine learning-based tool for systematic review screening. During this process, two reviewers (JD, SA) manually labeled each title and abstract as either relevant or irrelevant, allowing the tool to continuously refine its predictions and prioritize the most relevant records. A stopping criterion of 250 consecutive irrelevant titles and abstracts was applied. Following title and abstract screening, two reviewers (JD, SA) independently conducted full-text screening to assess study eligibility. Disagreements were resolved through discussion with a third reviewer (TC). Eligible reports were included for data extraction and methodological quality assessment.

2.4. Data extraction and methodological quality assessment

To ensure the validity and reliability of data analyses, the first author (JD) extracted data from all included studies, and a research assistant independently extracted 20 % of them. Discrepancies were resolved through discussion with a third reviewer (TC). The following information was extracted from each report based on a standardized data extraction form: (a) article details (authors, publication year, country), (b) study design (QED or RCT), (c) sample information (age, number of participants), (d) lesson (subject) and learning phase in the (IVR) intervention (engagement, exploration, explanation, elaboration, or evaluation; Sarac, 2018), (e) description of the intervention [technological device, frequency and duration, level of integration into the academic lesson (embedded, isolated, unclear), social interaction in the VR environment or outside, level of presence (very high, high, mid, low and very low), level of agency (passive or active), instructional setting accompanying the IVR intervention (e.g., reflective activities, scaffolding)], (f) control condition (waiting list, analog teaching methods, non-immersive VR, other IVR intervention), (g) measured motivational and cognitive factors as described in the CAMIL framework (situational interest, intrinsic motivation, self-regulation, self-efficacy, cognitive load, and embodiment; Makransky & Petersen, 2021; Petersen et al., 2022), (h) knowledge domain assessed (factual, conceptual, procedural, and metacognitive knowledge), (i) results, and (j) limitations. In case of missing data, the corresponding author has been contacted.

One author (JD) assessed the risk of bias of the reports included, using the "Medical Education Research Study Quality Instrument" (MERSQI; Asmri et al., 2023). A second author (SA) independently assessed the risk of bias in 20 % of the reports included. Disagreements were resolved by consensus. The tool assesses risk of bias in relation to 19 items concerning study objective, sampling, blinding, type of data and measurements, data analysis, and outcome. For each applicable item, 0–2 points are awarded (2 if fully met, 1 if partially met, 0 if not met). In addition, points are awarded based on the study design and based on whether randomization was present (0–5). The summary score is calculated by dividing the total score achieved by the maximum possible score. The maximum possible score depends on the number of items that are applicable based on the type of research.

3. Results

3.1. Search results

The systematic search yielded a total of 1521 records. After removing duplicates, 1164 records were screened using ASReview Lab with the stopping criterion met after labeling 672 titles and abstracts. Of the 46 full texts reviewed, 24 reports met eligibility criteria for qualitative synthesis. The PRISMA flowchart of the literature screening process is presented in Fig. 1.

3.2. Methodological quality

The mean (SD) MERSQI summary score of reports included was 0.78 (0.06) out of 1. MERSQI summary scores for individual reports are listed in Appendix B.

3.3. Study characteristics

Main characteristics of the included studies are summarized in Appendix B. Most studies were conducted in Asia, with 38 % in China ($N = 9$), 13 % in Taiwan ($N = 3$), and 8 % each in Indonesia, Israel, and Turkey ($N = 2$). Single studies (4 % each) were conducted in Jordan, Singapore (alongside China), Korea, Australia, the United Kingdom (alongside China), Cyprus, Italy (alongside Ireland), and Spain. The 8 randomized controlled trials (RCT) and 16 quasi-experimental studies (QED) employed various types of control groups. Notably, five studies (21 %) compared the IVR intervention to multiple control conditions. Overall, 54 % ($N = 13$) contrasted the IVR intervention with non-immersive VR, 38 % ($N = 9$) with traditional analog teaching methods, 25 % ($N = 6$) with another IVR intervention, and 4 % ($N = 1$) with a waiting list control. In most studies (83 %, $N = 20$), knowledge acquisition and transfer were assessed at two time points: before and after the intervention. One of these studies additionally conducted an assessment during the intervention. In 17 % of the studies ($N = 4$), knowledge acquisition was measured only after intervention. All studies took place in the school setting. Across the 24 studies, there was a total of 2737 participants ($M = 114.04$, $SD = 96.31$), with sample sizes ranging from 28 to 370 participants.

3.4. Descriptive results of participant's characteristics and interventions

Most studies (71 %) were conducted in upper primary school (4th to 6th grade; $N = 17$), 21 % in lower primary school (1st to 3rd grade; $N = 5$) and 4 % in kindergarten ($N = 1$). One study (4 %) could not be categorized because no precise age was available (Iasha et al., 2023). Regarding school subjects, 38 % of the studies focused on language ($N = 9$), 25 % each on science and mathematics ($N = 6$), and 13 % on social sciences ($N = 3$). Based on Sarac's (2018) model, IVR interventions were categorized by learning phase as follows: Engagement (33 %, $N = 8$), exploration (67 %, $N = 16$), explanation (58 %, $N = 14$), elaboration (50

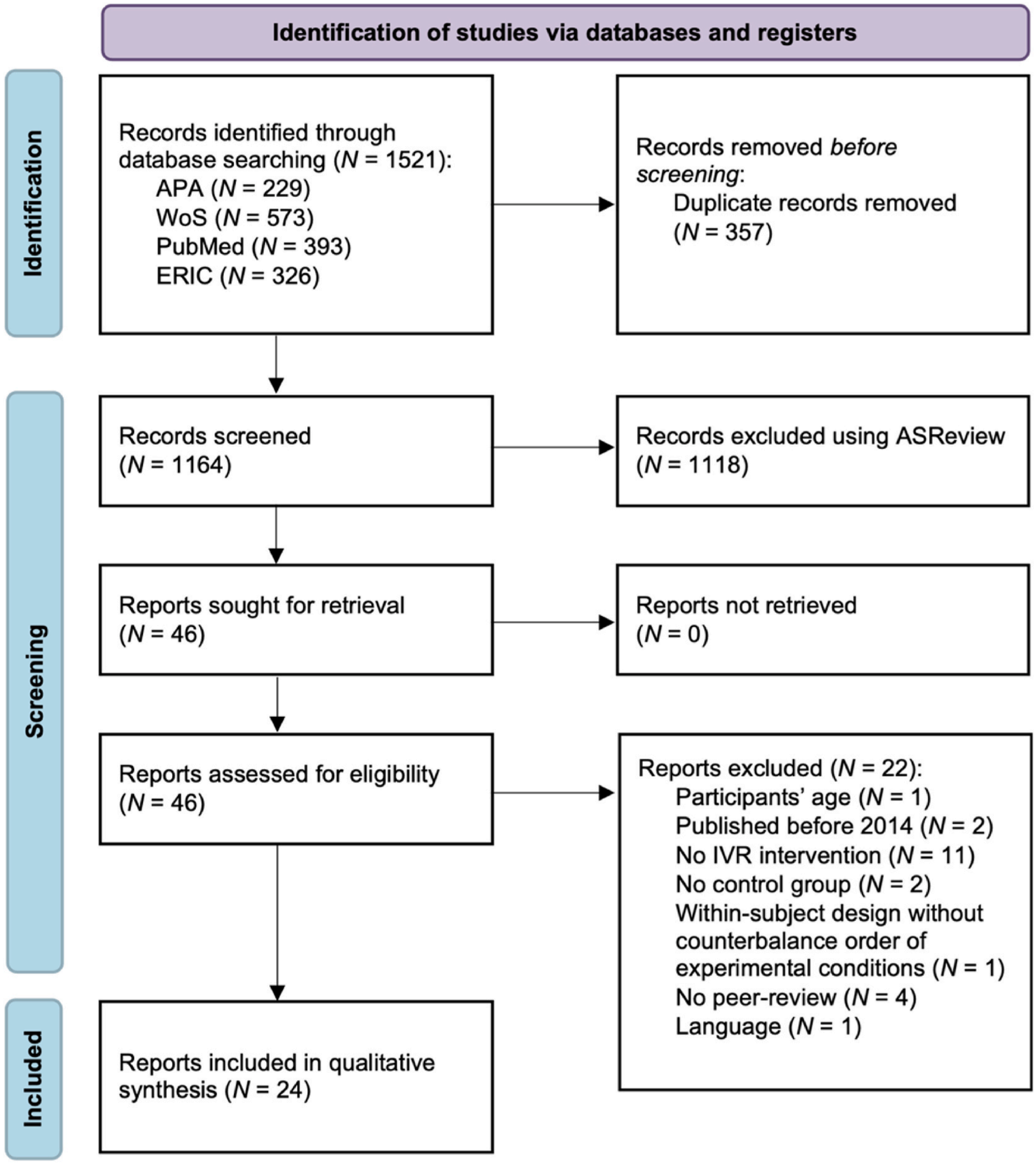

Fig. 1. PRISMA flowchart of the literature screening process.

%, $N = 12$), and evaluation (21 %, $N = 5$), with most of the interventions spanning multiple phases (63 %, $N = 15$).

A variety of technological devices were used: 79 % of the studies employed a head-mounted display ($N = 19$), 13 % a mobile phone-based VR system ($N = 3$), and 4 % 3D glasses and a 3D screen ($N = 1$). One study (4 %) did not specify the device used ($N = 1$; Iasha et al., 2023). In 63 % of the studies, the IVR intervention was embedded in the lesson ($N = 15$), while 38 % used it as a stand-alone additive learning unit ($N = 9$). The frequency and duration of the interventions varied widely. IVR session duration ranged from 4.5 to 90 min ($M = 29.35$, $SD = 24.79$), with 38 % of the studies using the device two to ten times ($N = 9$), and 54 % using it only in one session ($N = 13$). In 8 % of the studies ($N = 2$), this information was not provided. Regarding social interaction, a distinction was made, whether it took place within the (I)VR environment (e.g., interaction with another person) or outside of it (e.g., discussions about the experience). For the experimental groups, 13 % of the studies reported that participants interacted with others within the IVR environment ($N = 3$), 38 % reported that social interaction took place outside the IVR environment ($N = 9$), and 38 % reported no social interaction ($N = 9$). In 13 % of the studies this information was not available ($N = 3$). In the control conditions, 33 % of the studies ($N = 8$) reported that the social interaction took place outside the VR environment, while 54 % of the studies ($N = 13$) reported no social interaction. In 13 % of the studies, this information was not available ($N = 3$). The characteristics of the participants and the interventions are detailed in Appendix C.

3.5. Descriptive results of presence, agency and CAMIL factors

Only one study (4 %; Araiza-Alba et al., 2021) assessed the sense of spatial presence using a questionnaire [Sense of Presence Questionnaire for Children, adapted from the ITC-Sense of Presence Inventory (Lessiter et al., 2001) and the Presence Questionnaire (Witmer & Singer, 1998)] ($N = 1$). For all other studies (96 %) the type of presence was defined by the authors based on the device used and the description of the intervention. Spatial presence was categorized according to the medium used to deliver the learning content ($N = 23$). Spatial presence was rated highest when studies reported that children could interact directly with objects in the real world (very high), followed by studies in which children perceived objects in the IVR environment as real (high), non-immersive learning environments (mid), images (low) and text (very low).

No studies assessed the sense of agency using a questionnaire. The authors classified the children's involvement as either active or passive based on the device used and the description of the intervention. Active indicates that the children could interact with objects or make decisions about the course of the learning environment, while passive means that children had no opportunity to intervene in any way during the IVR learning environment. The active/passive distinction was also applied for the control conditions, where the evaluation pertained to the corresponding alternative learning environment. Participants in the experimental condition were actively involved in 71 % of the studies ($N = 17$) and passively involved in 29 % ($N = 7$). In the control conditions, participants were actively involved in 58 % ($N = 14$) and passively involved in 33 % of the studies ($N = 8$). For 11 % of the studies ($N = 3$), this information was not available for the control group. Several studies compared different control groups, which is why the total number of control groups exceeds the total number of studies.

Regarding potential underlying CAMIL factors (Makransky & Petersen, 2021; Petersen et al., 2022), most studies focused on motivational and self-regulatory factors (75 %, $N = 18$), while cognitive load and embodiment were examined in only six studies (25 %). However, not all studies measured these factors for the control condition, so comparisons across all studies were not possible. A total of 15 studies (63 %) assessed motivational and self-regulatory factors in all groups, and three studies (13 %) measured cognitive load and embodiment. Two studies (8 %) employed interviews instead of scales, enabling qualitative assessments (Weng et al., 2019; Yang et al., 2021). A detailed description of presence, agency and CAMIL factors can be found in Appendix D.

3.6. Descriptive results of knowledge domains

Regarding knowledge domains, most studies ($N = 16$, 67 %) addressed multiple domains, with 50 % focusing on factual knowledge ($N = 12$), 46 % on conceptual knowledge ($N = 11$), 42 % on procedural knowledge ($N = 10$), and 21 % on metacognitive knowledge ($N = 5$). Five studies (21 %) assessed only procedural knowledge, two studies (8 %) only factual knowledge and one (4 %) only conceptual knowledge. Knowledge acquisition and transfer were mainly assessed by paper-and-pencil performance tests (67 %, $N = 16$) and in 33 % of the studies by standardized performance assessments ($N = 8$). *Prior knowledge* was categorized into *explicit prior knowledge* of the learning content (i.e., when it was stated that the children already had knowledge of the learning content, or when a learning phase on the new content occurred before the VR intervention) and *presumed prior knowledge* (e.g., writing skills that are not entirely new but are further developed). *No prior knowledge* referred to interventions using content that had not been previously taught in the classroom. In 75 % of the studies, participants had prior knowledge of the curriculum content ($N = 18$). This was explicitly stated in 46 % of the studies ($N = 11$) and estimated in 29 % of the studies ($N = 7$). In 25 % of the studies, the participants had no prior knowledge of the learning content ($N = 6$), which was explicitly mentioned in 8 % of the studies ($N = 2$) and estimated in 17 % of the studies ($N = 4$). A detailed description of knowledge domains can be found in Appendix E. As concerns the instructional setting of IVR, among the six studies (25 %) comparing IVR interventions with different instructional elements, one study (4 %) investigated the effect of the co-construction (Chang et al., 2024), two studies (8 %) the effect of reflective activities (Chen et al., 2024; Feng et al., 2021), two studies (8 %) the effects of scaffolding measures (Li, W., et al., 2023; Yang et al., 2024), and one study (4 %) the effect of near-transfer training effects (Cuturi et al., 2023).

3.7. Primary research question

In the following section, we present results related to the primary research question concerning the effect of IVR interventions on curriculum-related knowledge in K–6 education, and the role of agency, presence, and associated motivational and cognitive factors. We categorized studies into two categories: (1) Comparisons of IVR with analog teaching methods or waiting list, and (2) comparisons of IVR with non-immersive VR. To ensure clarity and avoid conflating distinct effects, we analyze these categories separately, accounting for differences in control conditions and methodological approaches. We begin with an overview of IVR's overall impact on knowledge acquisition and transfer. The results for the primary research question are then presented in separate subsections for each study category, following this structure:

1. General effects of IVR compared to the control group in terms of knowledge gain (i.e., the number of studies showing a positive effect).
2. For IVR affordances of agency and presence, we report how many studies implemented these features in the different groups and whether learning gains compared to the control group varied by type of implementation (presence: very high, high, mid, low, very low; agency: active, passive).
3. For the CAMIL factors, we report (a) the number of studies that included each CAMIL factor in the different groups, (b) the extent to which the IVR group demonstrated more learning-supportive outcomes than the control group (e.g., higher motivation, lower cognitive load), and (c) whether studies with such enhanced CAMIL factor outcomes also showed greater learning gains than the control group.

3.8. Overall effects of IVR interventions on knowledge acquisition and transfer

As reported in Table 1, in 75 % of the studies, children benefited from the IVR intervention compared to the control group ($N = 18$), with effect sizes ranging between $\eta^2 = 0.05$ and $\eta^2 = 0.68$. 21 % of the studies

Table 1
Learning gains.

Study	Learning gains	Effect sizes
Akman and Çakır (2020)	→	n.a.
Alhajya et al. (2018)	IVR ↑	n.a.
Araiza-Alba et al. (2021)	N-I VR ↑	n.a.
	AT ↑	n.a.
Cao et al. (2024)	IVR ↑	n.a.
Chang et al. (2020)	→	n.a.
Chang et al. (2024)	IVR ↑	n.a.
Chen et al. (2023)	IVR ↑	$\eta^2 = 0.09$
Chen et al. (2024)	IVR ↑	$\eta^2 = 0.08$
Cuturi et al. (2023)	IVR ↑	n.a.
Demitriadou et al. (2020)	N-I VR →	n.a.
	AT →	n.a.
Feng et al. (2021)	IVR ↑	$\eta^2 = 0.09$
Hung et al. (2023)	N-I VR →	$\eta^2 = 0.02$
	AT →	$\eta^2 = 0.02$
Iasha et al. (2023)	IVR ↑	n.a.
Li, F., et al. (2023)	→	n.a.
Li, W., et al. (2023)	TC: ↑	$\eta^2 = 0.11$
	S: ↑	$\eta^2 = 0.05$
Liu et al. (2022)	IVR ↑	n.a.
Passig and Schwartz (2014)	IVR ↑	$\eta^2 = 0.10$
Passig et al. (2016)	N-I VR ↑	$\eta^2 = 0.12$
	AT →	$\eta^2 = 0.09$
	WL ↑	$\eta^2 = 0.68$
Sally Wu and Alan Hung (2022)	IVR ↑	n.a.
Sarioğlu and Grigin (2020)	IVR ↑	n.a.
Villena-Taranilla et al. (2022b)	IVR ↑	$\eta^2 = 0.08$
Weng et al. (2019)	IVR ↑	$\eta^2 = 0.30$
Yang et al. (2021)	IVR ↑	n.a.
Yang et al. (2024)	IVR ↑	$\eta^2 = 0.49$

Note. IVR: Immersive virtual reality, N-I VR: Non-immersive VR, AT: Analog teaching methods, WL: Waiting list, n.a.: not available, TC: Textual cues, S: Summarizing, ↑: Beneficial effects →: No significant differences between groups.

reported no significant differences between the IVR and control conditions ($N = 5$) and 4 % showed mixed results ($N = 1$). No studies reported detrimental effects of IVR interventions on knowledge acquisition or transfer. However, the effects of IVR on knowledge vary depending on the control group used for comparison.

3.9. Effects of IVR interventions compared to analog teaching methods (N = 9)

A consistent pattern emerges when comparing IVR interventions to analog teaching methods ($N = 9$): 67 % of studies reported that IVR positively impacts knowledge acquisition and transfer across various school subjects, including mathematics, language and science ($N = 6$). Effect sized ranged from $\eta^2 = 0.08$ to $\eta^2 = 0.30$. Among the 9 studies comparing IVR with analog teaching methods, IVR interventions demonstrated advantages across various knowledge domains with children showing gains in factual (three of five comparisons; Sally Wu & Alan Hung, 2022; Villena-Taranilla et al., 2022b; Weng et al., 2019), conceptual (four of six; Alhajya et al., 2018; Passig & Schwartz, 2014; Sally Wu & Alan Hung, 2022; Weng et al., 2019), and procedural knowledge (four of six; Alhajya et al., 2018; Araiza-Alba et al., 2021; Passig & Schwartz, 2014; Sally Wu & Alan Hung, 2022). The only study that assessed metacognitive knowledge showed no significant advantage of the IVR condition over analogue teaching. One study compared the IVR intervention to a waiting list (among other comparisons), reporting beneficial effects on conceptual and procedural knowledge (Passing et al., 2016). The comparison with the waiting list was not considered for further analysis, as presence or agency were not measurable in the control condition, and CAMIL factors were not investigated (Makransky & Petersen, 2021; Petersen et al., 2022).

3.10. Agency in IVR interventions compared to analog teaching methods (N = 9)

Regarding agency, 67 % involved active participant engagement in the IVR intervention ($N = 6$; Alhajya et al., 2018; Araiza-Alba et al., 2021; Demitriadou et al., 2020; Hung et al., 2023; Passig & Schwartz, 2014; Passig et al., 2016), while 33 % included passive engagement ($N = 3$; Sally Wu & Alan Hung, 2022; Villena-Taranilla et al., 2022b; Weng et al., 2019). In control groups, 44 % had active engagement ($N = 4$; Araiza-Alba et al., 2021; Hung et al., 2023; Passig & Schwartz, 2014; Passig et al., 2016), 44 % had passive engagement ($N = 4$; Demitriadou et al., 2020; Sally Wu & Alan Hung, 2022; Villena-Taranilla et al., 2022b; Weng et al., 2019), and 11 % could not be classified ($N = 1$; Alhajya et al., 2018). Only studies assessing agency in both conditions were considered. Greater learning gains for IVR were reported in two of four studies with active engagement in both groups (Araiza-Alba et al., 2021; Passig & Schwartz, 2014), and all three studies with passive engagement (Sally Wu & Alan Hung, 2022; Villena-Taranilla et al., 2022b; Weng et al., 2019). These findings suggest IVR has learning advantages in passive engagement settings, while in active participation scenarios, these advantages are inconsistent (see Fig. 2).

3.11. Presence in IVR interventions compared to analog teaching methods (N = 9)

The following section examines the relationship between spatial presence and learning in IVR. Only one study (11 %) directly assessed subjective presence in the IVR condition using a questionnaire (Araiza-Alba et al., 2021), reporting a high degree of presence for the IVR condition. All other studies ($N = 8$; 89 %) did not measure presence at all, so we estimated the level of presence experienced via the medium used. Across all studies, 100 % of the IVR learning environments showed high levels of presence ($N = 9$). In contrast, the control groups demonstrated very high levels of presence in 22 % of studies ($N = 2$) and low levels in 66 % ($N = 6$). In one study (11 %), it was not possible to estimate the level of presence in the control group (Alhajya et al., 2018). Again, we focus on studies in which spatial presence could be assessed for both the experimental and control conditions. In four of six comparisons (67 %), where the level of presence was higher in the IVR condition than in the control group, children in the IVR intervention demonstrated greater knowledge gains compared to the control group (Passig & Schwartz, 2014; Sally Wu & Alan Hung, 2022; Villena-Taranilla et al., 2022b; Weng et al., 2019). In the studies where the control condition exhibited higher presence, no consistent relationship between presence and learning was observed. These findings indicate that the IVR condition tends to show learning advantages when the level of presence is higher in the IVR condition, while learning gains are random when the control group exhibits a higher level of presence (see Fig. 3).

3.12. CAMIL factors in IVR interventions compared to analog teaching methods (N = 9)

Regarding CAMIL factors (Makransky & Petersen, 2021; Petersen

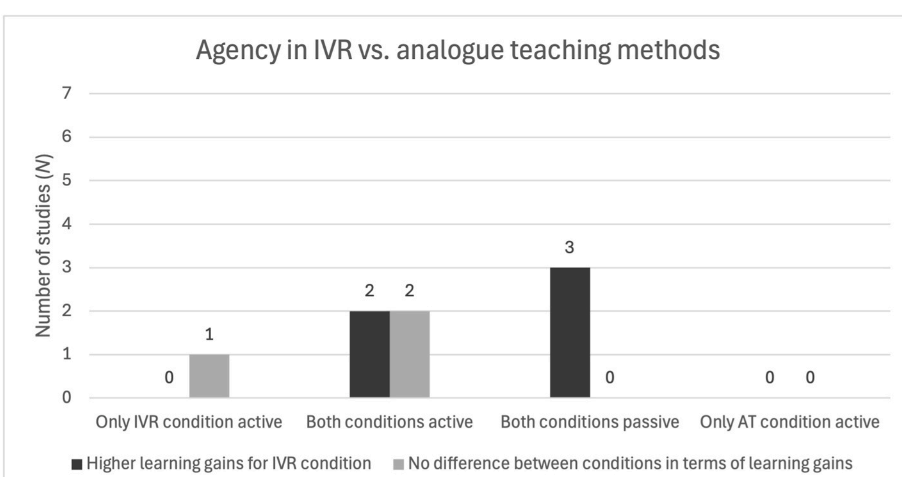

Fig. 2. Relationship between agency and learning gains in the comparison IVR vs. analogue teaching methods.
Note. IVR: Immersive virtual reality, AT: Analogue teaching. For each of the four conditions (e.g., both conditions active), the number of studies reporting higher learning gains for the IVR condition is shown in black, while the number of studies reporting no difference in learning gains between conditions is shown in gray. In one study, the level of agency could not be classified.

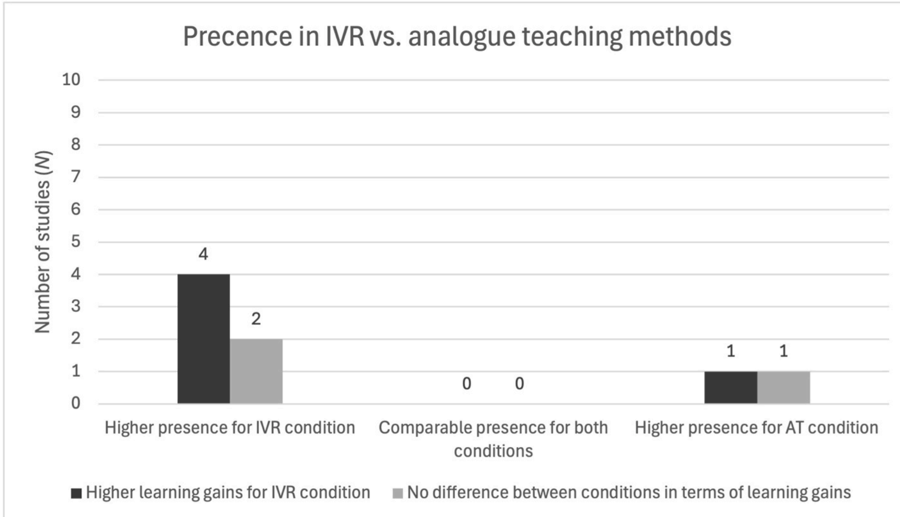

Fig. 3. Relationship between presence and learning gains in the comparison IVR vs. analogue teaching methods.
Note. IVR: Immersive virtual reality, AT: Analogue teaching. For each of the three conditions (e.g., higher presence for IVR condition), the number of studies reporting higher learning gains for the IVR condition is shown in black, while the number of studies reporting no difference in learning gains between conditions is shown in gray. In one study, the level of presence could not be classified.

et al., 2022), six studies (67 %) assessed motivational and self-regulatory factors (situational interest, intrinsic motivation, self-efficacy, self--regulation) in experimental and control groups ($N = 6$; Araiza-Alba et al., 2021; Demitriadou et al., 2020; Hung et al., 2023; Sally Wu & Alan Hung, 2022; Villena-Taranilla et al., 2022b; Weng et al., 2019). One study (11 %) examined embodiment, limited to the IVR group (Araiza-Alba et al., 2021), while none assessed cognitive load. In the following, we focus on studies assessing CAMIL factors in both experimental and control conditions to examine their relationship with learning outcomes. Among six studies measuring situational interest, five reported higher levels in the experimental group (Araiza-Alba et al., 2021; Demitriadou et al., 2020; Hung et al., 2023; Villena-Taranilla et al., 2022b; Weng et al., 2019), while one found no difference (Sally Wu & Alan Hung, 2022). Of the five studies showing increased situational interest, three also reported higher learning gains in IVR compared to analog methods (Araiza-Alba et al., 2021; Villena-Taranilla et al., 2022b; Weng et al., 2019). In the study where situational interest was equal, children still learned more in IVR (Sally Wu & Alan Hung, 2022). All four studies measuring intrinsic motivation found higher levels in the IVR group (Demitriadou et al., 2020; Hung et al., 2023; Sally Wu & Alan Hung, 2022; Villena-Taranilla et al., 2022b), and two also reported greater learning gains for IVR (Sally Wu & Alan Hung, 2022; Villena-Taranilla et al., 2022b). Of four studies measuring self-regulation, two reported higher levels in the IVR group (Demitriadou et al., 2020; Hung et al., 2023), while the others found no difference (Sally Wu & Alan Hung, 2022; Villena-Taranilla et al.,

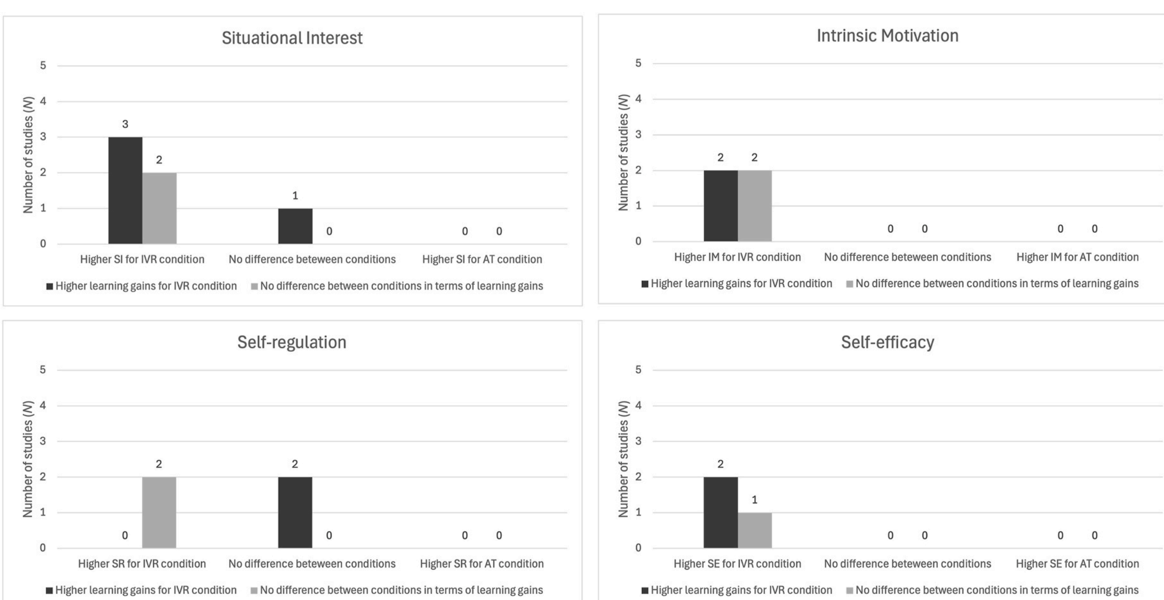

Fig. 4. Relationship between CAMIL factors and learning gains in the comparison IVR vs. analogue teaching methods.
Note. SI: Situational interest, IM: Intrinsic motivation, SR: Self-regulation, SE: Self-efficacy, IVR: Immersive virtual reality, AT: Analogue teaching. For each of the three conditions (e.g., higher SI for IVR condition), the number of studies reporting higher learning gains for the IVR condition is shown in black, while the number of studies reporting no difference in learning gains between conditions is shown in gray. Only constructs with at least one comparison available per construct are included in the graph.

2022b). None of the studies showing higher self-regulation reported greater learning gains for IVR, but in the two studies where self-regulation was equal, children learned more in the IVR condition. All three studies measuring self-efficacy found higher levels in the IVR group (Hung et al., 2023; Sally Wu & Alan Hung, 2022; Villena-Taranilla et al., 2022b), and two also reported higher learning gains for IVR (Sally Wu & Alan Hung, 2022; Villena-Taranilla et al., 2022b). In sum, learning gains in IVR appear moderately linked to situational interest and self-efficacy, while intrinsic motivation and self-regulation show no consistent trend (see Fig. 4).

3.13. Effects of IVR interventions compared to non-immersive VR interventions (N = 13)

When comparing IVR with non-immersive VR interventions, 69 % of studies (9 out of 13) report knowledge gains across various school subjects, including mathematics, language, and science following IVR interventions. Effect sized ranged from $\eta^2 = 0.09$ to $\eta^2 = 0.12$ (see Table 1 for a detailed description). Across the 13 studies comparing IVR to non-immersive VR, IVR demonstrated beneficial effects in four out of six comparisons for factual knowledge (Cao et al., 2024; Chang et al., 2024; Liu et al., 2022; Sarioğlu & Grigin, 2020), five out of six for conceptual knowledge (Cao et al., 2024; Iasha et al., 2023; Liu et al., 2022; Passig et al., 2016; Sarioğlu & Grigin, 2020), four out of seven for procedural knowledge (Araiza-Alba et al., 2021; Chen et al., 2023; Passig et al., 2016; Yang et al., 2021), and two out of three for meta-cognitive knowledge (Cao et al., 2024; Chen et al., 2023).

3.14. Agency in IVR interventions compared to non-immersive VR interventions (N = 13)

Regarding agency, 85 % of studies involved active engagement in the experimental group (*N* = 11; Akman & Çakır, 2020; Araiza-Alba et al., 2021; Cao et al., 2024; Chang et al., 2024; Demitriadou et al., 2020; Hung et al., 2023; Iasha et al., 2023; Liu et al., 2022; Passig et al., 2016; Sarioğlu & Grigin, 2020; Yang et al., 2021), while 15 % involved passive engagement (*N* = 2; Chen et al., 2023; Li, F., et al., 2023). In the control groups, 69 % of studies involved active engagement (*N* = 9; Akman & Çakır, 2020; Araiza-Alba et al., 2021; Chang et al., 2024; Demitriadou et al., 2020; Hung et al., 2023; Iasha et al., 2023; Passig et al., 2016; Sarioğlu & Grigin, 2020; Yang et al., 2021), 15 % passive engagement (*N* = 2; Chen et al., 2023; Li, F., et al., 2023), and 15 % could not be classified (*N* = 2; Cao et al., 2024; Liu et al., 2022). Considering only studies in which agency could be classified for both experimental and control conditions, greater learning gains for children in the IVR intervention compared to non-immersive VR intervention were observed in: six out of nine studies where children in both groups actively participated (Araiza-Alba et al., 2021; Chang et al., 2024; Iasha et al., 2023; Passig et al., 2016; Sarioğlu & Grigin, 2020; Yang et al., 2021); and one out of two studies where both groups had passive engagement (Chen et al., 2023). Thus, unlike the comparison between IVR and analog teaching methods, the learning advantage of IVR compared to non-immersive VR was particularly evident in active learning scenarios (see Fig. 5).

3.15. Presence in IVR interventions compared to non-immersive VR interventions (N = 13)

The following section focuses on the relationship between spatial presence and learning for IVR - non-immersive VR comparisons. Only one study directly assesed subjective presence in the IVR condition using a questionnaire (Araiza-Alba et al., 2021), reporting a high degree of presence for the IVR condition. All other studies did not measure presence at all, so we estimated the level of presence experienced via the medium used. In all 13 studies, children experienced high levels of presence in all IVR learning environments (*N* = 13) and moderate levels of presence in all control conditions (*N* = 13). In nine of these 13 studies, children showed greater knowledge gains in the IVR condition than in the non-immersive VR control group (Araiza-Alba et al., 2021; Cao et al., 2024; Chang et al., 2024; Chen et al., 2023; Iasha et al., 2023; Liu et al., 2022; Passig et al., 2016; Sarioğlu & Grigin, 2020; Yang et al., 2021), suggesting a link between higher levels of presence and improved learning outcome (see Fig. 6).

3.16. CAMIL factors in IVR interventions compared to non-immersive VR interventions (N = 13)

Focusing on CAMIL factors (Makransky & Petersen, 2021; Petersen et al., 2022), most studies (92 %) assessed motivational and self-regulatory factors (*N* = 12; Akman & Çakır, 2020; Araiza-Alba et al., 2021; Cao et al., 2024; Chang et al., 2024; Chen et al., 2023; Demitriadou et al., 2020; Hung et al., 2023; Iasha et al., 2023; Li, F., et al., 2023; Liu et al., 2022; Sarioğlu & Grigin, 2020; Yang et al., 2021), while only four (31 %) considered embodiment or cognitive load (Araiza-Alba

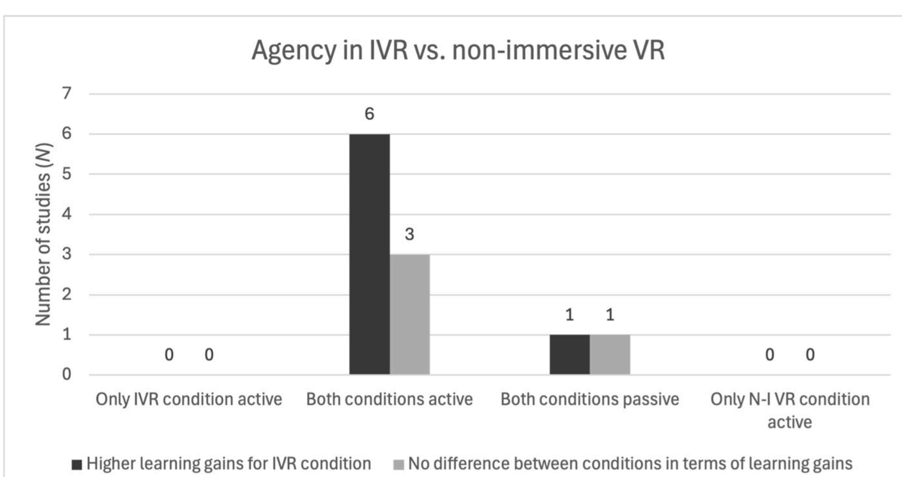

Fig. 5. Relationship between agency and learning gains in the comparison IVR vs. non-immersive VR.
Note. IVR: Immersive virtual reality, N-I VR: Non-immersive VR. For each of the four conditions (e.g., both conditions active), the number of studies reporting higher learning gains for the IVR condition is shown in black, while the number of studies reporting no difference in learning gains between conditions is shown in gray. In two studies, the level of agency could not be classified.

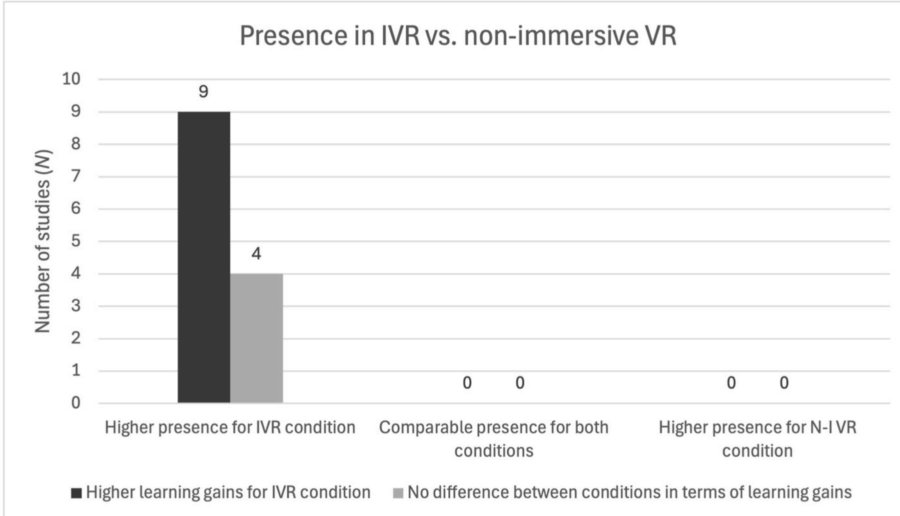

Fig. 6. Relationship between presence and learning gains in the comparison IVR vs. non-immersive VR.
Note. IVR: Immersive virtual reality, N-I VR: Non-immersive VR. For each of the three conditions (e.g., higher presence for IVR condition), the number of studies reporting higher learning gains for the IVR condition is shown in black, while the number of studies reporting no difference in learning gains between conditions is shown in gray.

et al., 2021; Chang et al., 2024; Liu et al., 2022; Yang et al., 2021). In four studies (31 %), CAMIL factors were assessed solely in the experimental IVR group, limiting between-group comparisons (Araiza-Alba et al., 2021; Chang et al., 2024; Iasha et al., 2023; Sarioğlu & Grigin, 2020).

Below, we consider only studies assessing CAMIL factors in both experimental and control conditions to examine their relationship with learning outcomes. Among seven studies measuring situational interest, four reported higher levels in the experimental group (Araiza-Alba et al., 2021; Hung et al., 2023; Liu et al., 2022; Yang et al., 2021), while three found no difference (Chen et al., 2023; Demitriadou et al., 2020; Li, F., et al., 2023). Of the four studies with increased situational interest in the experimental group, three observed higher knowledge gains in IVR compared to non-immersive VR (Araiza-Alba et al., 2021; Liu et al., 2022; Yang et al., 2021). When situational interest was equal, children in the IVR group learned more in one out of three studies. Among seven studies measuring intrinsic motivation, three reported higher levels in the experimental group (Hung et al., 2023; Liu et al., 2022; Yang et al., 2021), while four found no difference (Akman & Çakır, 2020; Chen et al., 2023; Demitriadou et al., 2020; Li, F., et al., 2023). Of the three studies with higher intrinsic motivation in the experimental group, two showed higher knowledge gains in IVR compared to non-immersive VR (Liu et al., 2022; Yang et al., 2021). When intrinsic motivation was equal, children in the IVR group learned more in one out of four studies. Of eight studies measuring self-regulation, four reported higher levels in the experimental group (Cao et al., 2024; Hung et al., 2023; Liu et al., 2022; Yang et al., 2021), three found no difference (Akman & Çakır, 2020; Chen et al., 2023; Demitriadou et al., 2020), and one reported higher self-regulation in the control group (Li, F., et al., 2023). Of the four studies with higher self-regulation in the experimental group, three observed higher learning gains in IVR (Cao et al., 2024; Liu et al., 2022; Yang et al., 2021). When self-regulation was comparable, children in the IVR group outperformed in one out of three studies. Among seven studies measuring self-efficacy, four showed higher levels in the experimental group (Cao et al., 2024; Hung et al., 2023; Liu et al., 2022; Yang et al., 2021), while three found no difference (Akman & Çakır, 2020; Chen et al., 2023; Li, F., et al., 2023). Of the four studies with higher self-efficacy in the experimental group, three reported higher learning gains in IVR (Cao et al., 2024; Liu et al., 2022; Yang et al., 2021). When self-efficacy was equal, children in the IVR group learned more in one out of three studies. Only one study assessed embodiment (Yang et al., 2021), finding higher embodiment and greater knowledge gains in IVR compared to non-immersive VR. Similarly, only one study assessed cognitive load (Liu et al., 2022), reporting lower cognitive load and higher knowledge gains in IVR.

In sum, learning gains in IVR appear moderately linked to situational interest, intrinsic motivation, self-regulation, and self-efficacy. Single data points suggest a positive relationship for embodiment and a negative relationship for cognitive load (see Fig. 7).

3.17. Exploratory research questions two and three

In the following section, we present results for exploratory questions two and three, analyzing differential effects of children's age and prior knowledge, and the impact of instructional settings in IVR interventions.

3.18. Differential effects based on children's age and prior knowledge of curriculum content

In this section, we present the results separately for each of the two study categories: Comparisons of IVR with analog teaching methods and with non-immersive VR (see Table 2).

Focusing on age, the studies comparing IVR with analog teaching methods show that the IVR group demonstrated higher learning gains in kindergarten in the only study conducted with this population (Passig & Schwartz, 2014), in two out of three studies with lower primary school children (Alhajya et al., 2018; Araiza-Alba et al., 2021), and in three out of five studies with upper primary school children (Sally Wu & Alan Hung, 2022; Villena-Taranilla et al., 2022b; Weng et al., 2019). In the comparison of IVR with non-immersive VR, no studies involved kindergarten children, but all three studies with lower primary school children reported beneficial effects on knowledge acquisition and transfer (Araiza-Alba et al., 2021; Chang et al., 2024; Passig et al., 2016). Among upper primary school children, five out of nine studies showed similar positive effects (Cao et al., 2024; Chen et al., 2023; Liu et al., 2022; Sarioğlu & Grigin, 2020; Yang et al., 2021).

Focusing on prior knowledge, in studies comparing IVR with analog methods, larger learning gains were observed in three out of five cases where children had explicitly mentioned or estimated prior knowledge (Alhajya et al., 2018; Sally Wu & Alan Hung, 2022; Villena-Taranilla et al., 2022b). Similar gains were reported in three out of four studies where prior knowledge was neither explicitly mentioned nor estimated

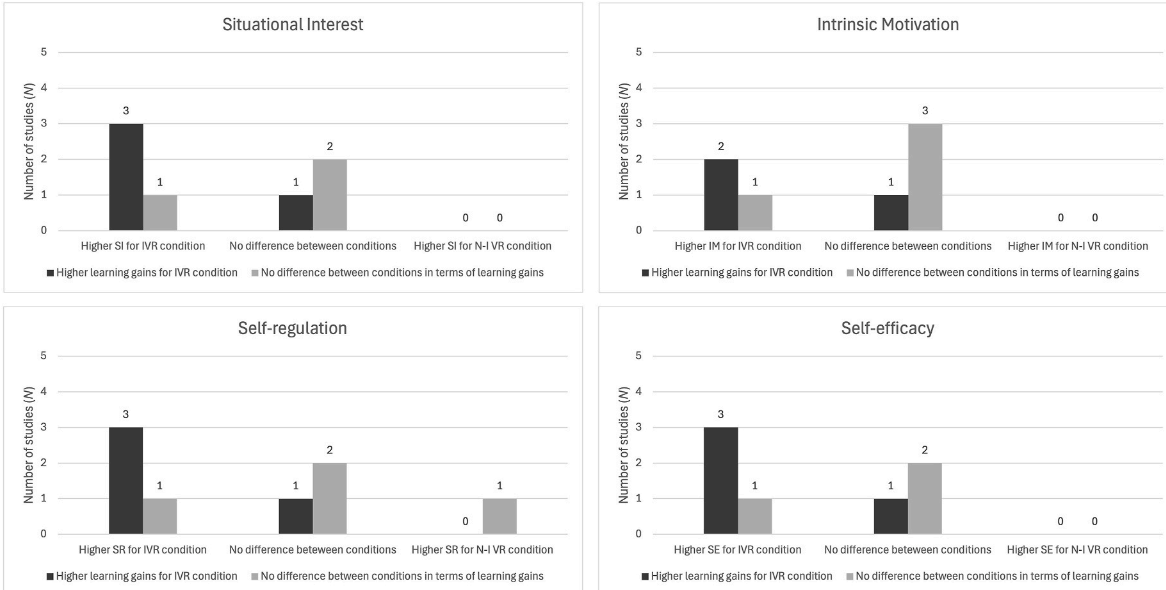

Fig. 7. Relationship between CAMIL factors and learning gains in the comparison IVR vs. non-immersive VR.
Note. SI: Situational interest, IM: Intrinsic motivation, SR: Self-regulation, SE: Self-efficacy, IVR: Immersive virtual reality, N-I VR: Non-immersive VR. For each of the three conditions (e.g., higher SI for IVR condition), the number of studies reporting higher learning gains for the IVR condition is shown in black, while the number of studies reporting no difference in learning gains between conditions is shown in gray. Only constructs with at least one comparison available per construct are included in the graph.

Table 2
Learning success of IVR conditions compared to control groups depending on children's age and prior knowledge of curriculum content.

Dimension	Group	Higher learning success for IVR condition compared to analog teaching methods	Higher learning success for IVR condition compared to non-immersive VR
Age	Kindergarten	100 % (1 out of 1)	n.a. (0 out 0)
	Lower primary School	67 % (2 out of 3)	100 % (3 out of 3)
	Upper primary School	60 % (3 out of 5)	56 % (5 out of 9)
Prior Knowledge	Prior knowledge	60 % (3 out of 5)	70 % (7 out of 10)
	No prior knowledge	75 % (3 out of 4)	67 % (2 out of 3)

(Araiza-Alba et al., 2021; Passig & Schwartz, 2014; Weng et al., 2019). In comparisons of IVR with non-immersive VR, larger learning gains were found in seven out of ten studies where children had explicitly mentioned or estimated prior knowledge (Cao et al., 2024; Chang et al., 2024; Chen et al., 2023; Liu et al., 2022; Passig et al., 2016; Sarioğlu & Grigin, 2020; Yang et al., 2021) and in two out of three studies where prior knowledge was neither explicitly mentioned nor estimated (Araiza-Alba et al., 2021; Iasha et al., 2023).

3.19. Effects of different instructional settings in IVR interventions ($N = 6$)

Six studies examined the effects of different instructional designs within IVR environments on learning outcomes (i.e., factors such as allowing children to create IVR content, providing peer feedback, offering individual debriefing, using textual or problem-solving scaffolds, or engaging in content-specific training within the IVR environment). Chang et al. (2020) compared children who constructed IVR content themselves with those using a pre-built IVR environment. Both groups experienced similar levels of spatial presence and reported similar levels of cognitive load, intrinsic motivation, and self-efficacy. No significant differences were observed in knowledge acquisition or transfer. Two of these studies focused on the impact of reflective activities (Chen et al., 2024; Feng et al., 2021). Chen et al. (2024) found that peer feedback between IVR sessions improved self-regulation, self-efficacy, and learning gains compared to no peer feedback, despite both groups experiencing passive engagement and high spatial presence. Feng et al. demonstrated that individual teacher debriefing after IVR sessions led to greater knowledge gains compared to no debriefing, though both groups were actively engaged and experienced comparably high levels of spatial presence (CAMIL factors were not assessed). Two studies implemented different scaffolds as part of the IVR interventions (Li, W., et al., 2023; Yang et al., 2024). Li, W., and colleagues (2023) compared textual cues and summary scaffolds, finding that textual cues significantly enhanced knowledge acquisition and transfer, while summary scaffolds improved only acquisition. Of the CAMIL factors, only cognitive load was measured, with no significant differences between conditions. Yang et al. (2024) observed that problem-solving scaffolds increased intrinsic motivation, self-regulation, and learning outcomes and transfer, but found no differences in situational interest or self-efficacy. Both studies reported comparable levels of spatial presence between groups; however, Yang et al.'s study involved only active engagement in both groups, whereas Li et al.'s study involved only passive engagement in all conditions. Lastly, the study by Cuturi and colleagues (2023) showed that content-specific training (Cartesian plane) within IVR improved factual and conceptual knowledge compared to a non-related IVR game. Active engagement and high spatial presence was experienced in both conditions.

4. Discussion

Overall, the results suggest that IVR interventions in primary education (K–6) substantially enhance children's learning outcomes across various school subjects, outperforming both analog teaching methods and non-immersive VR interventions. Notably, no evidence was found suggesting that IVR negatively affects learning processes. These beneficial effects appear to be partially driven by increased agency and spatial presence, as well as changes in underlying motivational and cognitive factors. However, the current evidence base remains too limited to draw definitive conclusions. Regarding individual learner characteristics, variables such as age and prior knowledge did not consistently influence IVR effects. Nevertheless, there is an emerging trend suggesting that younger learners may benefit more from IVR than from non-immersive VR. In addition, preliminary evidence points to the importance of instructional design, particularly the use of scaffolding and reflective activities, in supporting learning in IVR contexts.

The following section is structured according to our three research questions:

1. The effectiveness of IVR interventions in fostering curriculum-related factual, conceptual, procedural, and metacognitive knowledge in K–6 education, and the role of agency, presence, and associated motivational and cognitive factors.
2. The role of individual learner characteristics (age, prior knowledge).
3. The impact of instructional settings.

Regarding the first research question, IVR interventions consistently showed beneficial effects on the acquisition and transfer of factual, conceptual, procedural, and metacognitive knowledge in K–6 education. These effects were observed across a range of school subjects, including mathematics, language, and science, and emerged in comparisons with both analog teaching methods and non-immersive VR. To better understand the mechanisms underlying these learning gains, we analyzed the roles of agency and presence, focusing on studies that either measured these variables directly or allowed for classification based on methodological descriptions.

4.1. Spatial presence and agency

Spatial presence and agency are considered key affordances driving the learning effectiveness of IVR environments (Makransky & Petersen, 2021). Compared to non-immersive VR, IVR consistently elicited higher levels of presence and agency, due to its immersive, interactive features (Di Natale et al., 2020; Ijsselsteijn & Riva, 2003). In contrast, when compared to analog instruction, the advantage of IVR is more context dependent. While traditional classrooms can foster agency and presence through hands-on activities, such opportunities are not always feasible. In many traditional classroom settings, instruction relies heavily on text-based materials or static visuals, which may limit opportunities for embodied or interactive engagement. In such contexts, IVR can offer a distinct advantage by allowing learners to interact with otherwise inaccessible content, for example, abstract or microscopic phenomena in science lessons (see Martarelli et al., 2024). In these scenarios, IVR's immersive features are likely to elicit higher spatial presence than conventional methods. Across the included studies, spatial presence was consistently higher in IVR than in non-immersive VR and in three-quarters of comparisons with analog teaching. In studies where IVR elicited higher presence, learners outperformed their peers in both analog (4 out of 6 studies; Passig & Schwartz, 2014; Sally Wu & Alan Hung, 2022; Villena-Taranilla et al., 2022b; Weng et al., 2019) and non-immersive VR conditions (9 out of 13 studies; Araiza-Alba et al., 2021; Cao et al., 2024; Chang et al., 2024; Chen et al., 2023; Iasha et al., 2023; Liu et al., 2022; Passig et al., 2016; Sarioğlu & Grigin, 2020; Yang et al., 2021).

In contrast, agency did not emerge as a consistent differentiator. Most studies reported no significant differences in agency between IVR and control conditions, likely because the tasks were either active or passive across both groups. The role of agency appears to depend more strongly on the type of control condition and the learning context. Within analog instructional settings, the clearest advantage of IVR emerged in passive conditions, where its immersive features helped sustain attention and engagement. In contrast, when analog instruction involved active, hands-on tasks, the added value of IVR diminished. Compared to non-immersive VR, however, active engagement in IVR was consistently associated with improved learning outcomes, likely due to deeper cognitive processing enabled by immersive interaction. These findings align with prior research emphasizing the importance of spatial presence in immersive learning environments (Conrad et al., 2024; Mikropoulos & Natsis, 2011), as well as theoretical models that highlight the role of interactivity, constructiveness, and flow experience in learning (Chi & Wylie, 2014; Csikszentmihalyi, 1996; Makransky & Petersen, 2021; Petersen et al., 2022; Wirth et al., 2007). Consistent with the concept of flow experience (Csikszentmihalyi, 1996), IVR fosters an enjoyable sense of spatial presence (Wirth et al., 2007), which helps direct cognitive resources toward relevant learning stimuli and away from external distractions. However, the context-dependent role of agency underscores the need for more targeted research on the instructional conditions under which IVR can realize its full potential.

4.2. CAMIL factors and their role in IVR learning

The CAMIL framework (Makransky & Petersen, 2021; Petersen et al., 2022) offers a valuable lens to interpret the context-dependent effects of IVR by linking the affordances of agency and presence to motivational and cognitive learning mechanisms. Our analysis of CAMIL factors was limited to studies that systematically examined at least one CAMIL variable (situational interest, intrinsic motivation, self-regulation, self-efficacy, cognitive load, or embodiment) across all experimental conditions. Studies that assessed CAMIL factors only in the IVR group or did not include comparable data across conditions were excluded from this part of the analysis.

Comparison with analog instruction. In comparisons with analog settings, studies focused exclusively on motivational and self-regulatory factors; no reliable data were available for embodiment or cognitive load. IVR consistently elicited higher levels of situational interest, intrinsic motivation, and self-efficacy, while self-regulation showed no differences. None of these factors were found to be higher in the analog condition. However, only moderate trends suggested that situational interest and self-efficacy were associated with improved learning; no consistent link was found for intrinsic motivation or self-regulation.

Comparison with non-immersive VR. The pattern was less consistent when comparing IVR to non-immersive VR. Most studies again focused on motivational and self-regulatory factors; embodiment and cognitive load were assessed in only one study. Among factors examined in at least three studies, CAMIL variables were higher in IVR in about half of the cases, with the rest showing no differences. Notably, when CAMIL factors were elevated in IVR, learning gains also tended to be higher. In contrast, when CAMIL levels were similar between conditions, learning gains more often favored the control group. A possible explanation is that executive demands in IVR (e.g., navigation and interaction) may increase cognitive load, thus reducing learning efficiency in some cases.

In sum, these patterns reflect a moderate trend in the expected direction and are consistent with existing literature on the motivational appeal of IVR. Novel and intense sensory experiences can enhance situational interest and enjoyment, fostering short-term engagement (Makransky & Petersen, 2021) and knowledge acquisition (Hartmann et al., 2023). Contingent system feedback supports autonomy and competence, boosting intrinsic motivation, self-efficacy, and self-regulation (Ryan & Deci, 2015; Makransky & Petersen, 2021). From a constructivist and embodied cognition perspective, IVR enables learning through sensorimotor interaction with the environment

(Barsalou, 2008; Ertmer & Newby, 2013), potentially enhancing cognitive processing and working memory (Paas & Sweller, 2012). Interestingly, systematic differences emerged between the two comparison types. In analog comparisons, CAMIL mediators were generally more pronounced in IVR but not consistently linked to learning gains. In contrast, in IVR vs. non-immersive VR comparisons, elevated CAMIL factors more clearly coincided with improved learning.

While initial trends support the relevance of CAMIL factors in explaining IVR-related learning gains, the limited number of studies, especially on embodiment and cognitive load, prevents firm conclusions. To better understand the mechanisms underlying IVR effectiveness, further research should systematically investigate these and other learning-mediating variables in school-based IVR contexts.

4.3. Children's age and prior knowledge

Regarding the secondary research question, the impact of IVR on learning appeared to vary by age group, depending on the type of control condition. When compared with analog teaching methods, IVR showed beneficial effects on knowledge acquisition and transfer across all age groups. In contrast, when compared with non-immersive VR, learning benefits were more consistently observed in studies with younger primary school children.

There is tentative evidence that younger learners may benefit more from IVR, possibly because they rely more heavily on direct sensory input and hands-on interaction during learning. IVR environments support this through multimodal stimulation (visual, auditory, tactile) and interaction with virtual objects, thereby fostering embodiment (Coban et al., 2022; Villena-Taranilla et al., 2022a). This advantage is particularly evident when compared to non-immersive VR or analog environments with limited hands-on activity. However, the evidence base remains too limited to support definitive conclusions, especially for the youngest learners. Even if future research confirms that younger children benefit from IVR, this does not automatically imply that kindergarten-aged children should be targeted. At this developmental stage, learning is grounded in sensorimotor exploration, social interaction, and fine motor development. In addition, kindergarten curricula rarely involve the abstract content that IVR is well suited to support. The physical and sensory demands of IVR, combined with uncertain educational benefits, raise legitimate concerns about its suitability for early childhood settings. These issues highlight a critical gap in research regarding both the potential and limitations of IVR in early childhood education. Regarding prior knowledge, which has been hypothesized to reduce cognitive load (Meyer et al., 2019), no consistent moderating effect was observed. However, also this finding must be interpreted with caution: in most cases, prior knowledge was not assessed with objective measures but inferred from the intervention context. Future studies should examine this factor systematically, using valid and reliable instruments.

4.4. Instructional settings in IVR interventions

Regarding the third exploratory aim, comparisons between different IVR interventions provide initial insights into how instructional design can optimize IVR-based learning. Studies that incorporated scaffolding, debriefing, or peer feedback after IVR sessions suggest that these strategies can enhance learning by deepening understanding and supporting knowledge transfer. These findings align with previous research emphasizing the importance of integrating IVR into well-structured educational contexts supported by sound pedagogical strategies (Evans et al., 2024). For instance, instructional framing and the reduction of extraneous stimuli are essential to minimize cognitive load in younger learners (Han et al., 2021).

4.5. Limitations

First, due to the limited number of studies examining agency, presence, and underlying CAMIL factors, our analysis relied primarily on qualitative comparisons. Although informative, these do not permit statistical verification of IVR's learning effects. Moreover, the methodological quality of the available studies varied substantially, with many lacking standardized assessments, detailed procedural descriptions, or appropriate control conditions. In particular, agency and presence were rarely assessed using validated instruments, and CAMIL factors were often operationalized inconsistently or not examined across all experimental groups.

Second, the findings regarding individual learner characteristics, such as age and prior knowledge, must be interpreted with caution. These variables were often either not reported, inferred indirectly, or examined only in isolated conditions. As a result, the current evidence base does not allow for robust conclusions about their role as potential moderators of IVR effectiveness. Similarly, the number of studies exploring the impact of instructional support strategies (e.g., scaffolding, peer feedback, debriefing) remains small and highly heterogeneous in design, further limiting interpretability.

Third, the categorization of knowledge types (factual, conceptual, procedural, metacognitive) was based on the intended learning goals of the interventions, which were not always clearly delineated. In cases where multiple knowledge types were targeted simultaneously, classification became imprecise.

Fourth, a broader limitation of the field concerns the heterogeneity of study designs, which affects the generalizability of findings. Studies varied widely in terms of: (1) the type of IVR technology used (e.g., head-mounted displays vs. mobile-based VR), the subject matter taught (science, mathematics, language), and (2) the degree of curricular integration and instructional framing. These variations complicate cross-study comparisons and reduce the potential to isolate meaningful patterns.

4.6. Future directions and research gap

This systematic review is the first to examine the effectiveness of curriculum-integrated IVR interventions in the K–6 age range, while also considering key psychological affordances—presence and agency—and their associated motivational and cognitive mechanisms, as outlined in the CAMIL framework (Makransky & Petersen, 2021; Petersen et al., 2022). By narrowing the scope to three key dimensions, immersive VR, primary education, and curriculum-related learning, this review offers targeted insights into how IVR compares to analog and non-immersive VR environments, both of which represent standard practice in today's classrooms. However, this review also revealed several critical gaps in the current literature:

- Individual differences (e.g., age, prior knowledge) have been identified as potential moderators in previous research, but there is a lack of methodologically rigorous studies, especially in the K–6 range, that systematically assess their role in IVR-based learning.
- While motivational and self-regulatory CAMIL factors (e.g., situational interest, self-efficacy) were occasionally included, cognitive load and embodiment, central to the CAMIL model, were severely underexplored. Future studies should prioritize these factors and use standardized measures across conditions.
- The role of instructional settings remains under-investigated. More research is needed on how specific didactic elements (e.g., scaffolding types, reflective tasks, pacing) interact with immersive environments and influence learning outcomes.
- Finally, the differentiated effects of IVR technologies (e.g., head-mounted displays vs. desktop VR vs. mobile-based solutions) on presence, cognitive load, and learning remain insufficiently understood.

Addressing these gaps is essential for developing evidence-based recommendations for the pedagogical design and implementation of IVR in primary education.

5. Conclusion

School-based IVR interventions consistently lead to enhanced learning outcomes for K–6 learners across a variety of subjects. These effects appear to be partially mediated by the affordances of spatial presence and agency, as well as by motivational and cognitive factors described in the CAMIL framework. The advantage of IVR is particularly evident when compared to non-immersive VR and analog instruction that relies heavily on text and static visuals. In contrast, real-world hands-on activities can elicit similarly high levels of presence and agency, thereby reducing IVR's relative benefit.

There is tentative evidence suggesting that IVR may be especially beneficial for younger learners, who rely more on sensorimotor experiences. Moreover, instructional settings, such as the inclusion of reflective elements and scaffolding strategies, emerge as critical to optimizing learning in IVR environments.

However, the limited number of high-quality studies and the heterogeneity of existing research prevent definitive conclusions. To unlock IVR's full educational potential, future research must move beyond broad effect comparisons and instead focus on specific learning mechanisms, systematically manipulate IVR affordances, and investigate the role of instructional design. Only through such targeted inquiry can IVR be meaningfully integrated into primary education to support sustained knowledge acquisition and transfer.

CRediT authorship contribution statement

Josua Dubach: Writing – review & editing, Writing – original draft, Visualization, Project administration, Investigation, Formal analysis, Data curation, Conceptualization. **Sofia Anzeneder:** Conceptualization, Data curation, Validation, Writing – review & editing. **Sebastian Tempelmann:** Supervision, Validation, Writing – review & editing. **Trix Cacchione:** Conceptualization, Supervision, Writing – review & editing.

Funding

This study was supported by the Swiss National Science Foundation (grant number 407740_187327), which was awarded to Corinna Martarelli, Trix Cacchione, and Sebastian Tempelmann.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. – Search strategy

APA PsychInfo (06.05.2024) with filter from 2014 to current: 229 results

((“child” or “primary school” or “elementary” or “K-12″ or “K-6″) and (“virtual reality” or “immersive virtual reality” or “virtual” or “VR” or “IVR” or “head-mounted display” or “360°") and (“learning” or “academic performance” or “academic achievement” or “academic” or “science” or “math” or “reading” or “vocabulary” or “spelling” or “language” or “knowledge” or “conceptual” or “factual” or “procedural” or “transfer”) and (“randomized controlled trial” or “experimental” or “quasi-experimental” or “intervention")).af. not (“adult” or “disorder” or " clinical” or “ADHD” or “autism” or “hospital” or “review” or “meta-analysis").ab.

PubMed (06.05.2024) with filter last 10 years: 393 results

((((“child” OR “primary school” OR “elementary” OR “K-12” OR “K-6”) AND (“virtual reality” OR “immersive virtual reality” OR “virtual” OR “VR” OR “IVR” OR “head-mounted display” OR “360°”)) AND (“learning” OR “academic performance” OR “academic achievement” OR “academic” OR “science” OR “math” OR “reading” OR “vocabulary” OR “spelling” OR “language” OR “knowledge” OR “conceptual” OR “factual” OR “procedural” OR “transfer”)) AND (“randomized controlled trial” OR “experimental” OR “quasi-experimental” OR “intervention”)) NOT ((“adult”[Title/Abstract] OR “disorder”[Title/Abstract] OR “clinical”[Title/Abstract] OR “ADHD”[Title/Abstract] OR “autism”[Title/Abstract] OR “hospital”[Title/Abstract] OR “review”[Title/Abstract] OR “meta-analysis”[Title/Abstract])

ERIC (Ovid) (06.05.2024) with filters: from January 1st, 2014 to current: 326 results

((“child” or “primary school” or “elementary” or “K-12″ or “K-6″) and (“virtual reality” or “immersive virtual reality” or “virtual” or “VR” or “IVR” or “head-mounted display” or “360°") and (“learning” or “academic performance” or “academic achievement” or “academic” or “science” or “math” or “reading” or “vocabulary” or “spelling” or “language” or “knowledge” or “conceptual” or “factual” or “procedural” or “transfer”) and (“randomized controlled trial” or “experimental” or “quasi-experimental” or “intervention")).af. not (“adult” or “disorder” or " clinical” or “ADHD” or “autism” or “hospital” or “review” or “meta-analysis").ab.

Web of science (06.05.2024) in all databases with filters: Topic, last 10 years: 573 results

(((TS=(child OR primary school OR elementary OR K-12 OR K-6)) AND TS=(virtual reality OR immersive virtual reality OR virtual OR VR OR IVR OR head-mounted display OR 360°)) AND TS=(learning OR academic performance OR academic achievement OR academic OR science OR math OR reading OR vocabulary OR spelling OR language OR knowledge OR conceptual OR factual OR procedural OR transfer)) AND TS=(randomized controlled trial OR experimental OR quasi-experimental OR intervention) NOT Abstract?=(adult OR disorder OR clinical OR ADHD OR autism OR hospital OR review OR meta-analysis).

Appendix B

Table 3
Study characteristics

Study	MERSQI score	Country	Design	CG	Sample (*N*)	
					Total sample	Sample distribution
Akman and Çakır (2020)	0.83	Turkey	QED	N-I VR	64	EG = 32 CG = 32
Alhajya et al. (2018)	0.73	Jordan	QED	AT	48	EG = 24 CG = 24
Araiza-Alba et al. (2021)	0.83	Australia	RCT	N-I VR AT	120	EG = 40 N-I VR = 40 AT = 40
Cao et al. (2024)	0.73	China	QED	N-I VR	73	EG = 31 CG = 42
Chang et al. (2020)	0.73	Taiwan	QED	Other IVR	44	EG = 25 CG = 19
Chang et al. (2024)	0.85	Korea	RCT	N-I VR	300	EG = 150 CG = 150
Chen et al. (2023)	0.80	China/UK	QED	N-I VR	111	EG = 56 CG = 55
Chen et al. (2024)	0.75	China	QED	Other IVR	79	EG = 39 CG = 40
Cuturi et al. (2023)	0.71	Italy/Ireland	RCT	Other IVR	49	EG = 25 CG = 24
Demitriadou et al., 2020	0.75	Cyprus	QED	N-I VR AT	30	EG = 10 N-I VR = 10 AT = 10
Feng et al. (2021)	0.88	China	RCT	Other IVR	77	EG = 39 CG = 38
Hung et al. (2023)	0.83	Taiwan	QED	N-I VR AT	119	EG = 35 N-I VR = 45 AT = 39
Iasha et al. (2023)	0.73	Indonesia	QED	N-I VR	370	EG = 185 CG = 185
Li, F., et al. (2023)	0.75	China	QED	N-I VR	28	EG = 14 CG = 14
Li, W., et al. (2023)	0.83	China/Singapore	RCT	Other IVR textual cues (TC) Other IVR summarizing (S)	152	(TC, S) = 38 (TC, NS) = 39 (NTC, S) = 37 (NTC, NS) = 38
Liu et al. (2022)	0.78	China	QED	N-I VR	362	EG = 170 CG = 192
Passig and Schwartz (2014)	0.75	Israel	QED	AT	56	EG = 28 CG = 28
Passig et al. (2016)	0.85	Israel	RCT	N-I VR AT WL	117	EG = 36 N-I VR = 36 AT = 24 WL = 21
Sally Wu and Alan Hung (2022)	0.75	Taiwan	QED	AT	56	EG = 28 CG = 28
Sarioğlu and Grigin (2020)	0.75	Turkey	QED	N-I VR	100	EG = 50 CG = 50
Villena-Taranilla et al. (2022b)	0.73	Spain	QED	AT	98	EG = 45 CG = 53
Weng et al. (2019)	0.80	Indonesia	RCT	AT	80	EG = 40 CG = 40
Yang et al. (2021)	0.85	China	RCT	N-I VR	40	EG = 20 CG = 20
Yang et al. (2024)	0.75	China	QED	Other IVR	78	EG = 39 CG = 39

Note. MERSQI score: Medical Education Research Study Quality Instrument (MERSQI; Asmri et al., 2023), RCT: Randomized controlled trials, QED: Quasi-experimental studies, EG: Experimental IVR group, CG: Control group, N-I VR: Non-immersive VR, AT: Analog teaching methods, WL: Waiting list.

Appendix C

Table 4
Characteristics of participants and interventions

Study	Grade	Subject	Learning phase	Technological device	Frequency, duration	Level of integration IVR and instructional setting	Social interaction
Akman and Çakır (2020)	4th grade (UPS)	Mathematics	Elaboration	MPB	10–15 min 1x weekly for 4 weeks	embedded	IVR: no SoI CG: no SoI
Alhajya et al. (2018)	3rd grade (LPS)	Language	Engagement	3D glasses	n.a. 4 weeks	embedded	n.a.
Araiza-Alba et al. (2021)	2nd-3rd grade (LPS)	Mathematics	Engagement	HMD	10 min one session	additive	IVR: no SoI N-I VR: no SoI AT: no SoI
Cao et al. (2024)	4th grade (UPS)	Science	Engagement, exploration, explanation	HMD	40 min one session	embedded	IVR: no SoI CG: no SoI
Chang et al. (2020)	5th grade (UPS)	Science	Engagement, exploration, elaboration	HMD	80 min one session	embedded creative activities (hands-on design VR)	EG: SoI outside CG: no SoI
Chang et al. (2024)	2nd-3rd grade (LPS)	Language	Exploration	HMD	20 min one session	additive	IVR: no SoI CG: no SoI
Chen et al. (2023)	4th grade (UPS)	Language	Engagement	HMD	30 min 1 x weekly for 4 weeks	embedded	IVR: SoI outside CG: SoI outside
Chen et al. (2024)	4th grade (UPS)	Language	Exploration, explanation, elaboration, evaluation	HMD	50 min one session	embedded reflective activities (peer-feedback)	EG: SoI outside CG: SoI outside
Cuturi et al. (2023)	n.a. $M = 9.07$ years $SD = 1.24$ (UPS)	Mathematics	Exploration, elaboration	HMD	10 min daily for 3 days	additive thematically relevant/ irrelevant activities	EG: SoI outside CG: no SoI
(Demitriadou et al., 2020)	4th-6th grade (UPS)	Mathematics	Exploration, elaboration, evaluation	MPB	6 min one session	embedded	IVR: no SoI N-I VR: no SoI AT: no SoI
Feng et al. (2021)	2nd-3rd grade (LPS)	Social sciences	Exploration, explanation, elaboration, evaluation	HMD	5–8 min (2x) one session	additive reflective activities (debriefing)	EG: SoI outside CG: no SoI
Hung et al. (2023)	n.a. $M = 10.78$ years, $SD = 0.67$ (UPS)	Language	Exploration	HMD	20–25 min (25 min +20 min) one session	additive	IVR: SoI outside N-I VR: SoI outside AT: SoI outside
Iasha et al. (2023)	n.a.	Social Science	Engagement, exploration, explanation	n.a.	n.a.	additive	n.a.
Li, F., et al. (2023)	4th grade (UPS)	Language	Engagement	HMD	10 min 1 x weekly for 3 weeks	embedded	IVR: SoI outside CG: SoI outside
Li, W., et al. (2023)	4th-5th grade (UPS)	Science	Explanation, elaboration	HMD	4.5–5 min one session	Embedded scaffolding (textual cues, summarizing)	TC,S = no SoI TC, NS = no SoI NTC, S = no SoI NTC, NS = no SoI
Liu et al. (2022)	4th grade (UPS)	Science	Exploration, explanation, elaboration, evaluation	HMD	5–10 min (plus 35–50 min of peer observation) 1 x weekly for 3 weeks	embedded	IVR: SoI CG: SoI outside
Passig and Schwartz (2014)	KG (KG)	Mathematics	Exploration, explanation, elaboration, evaluation	HMD	15 min two times in one week	additive	IVR: SoI CG: SoI outside
Passig et al. (2016)	1st-2nd grade (LPS)	Mathematics	Exploration, explanation, elaboration	HMD	30 min one session	additive	IVR: no SoI N-I VR: no SoI AT: no SoI WL:

(continued on next page)

Table 4 (*continued*)

Study	Grade	Subject	Learning phase	Technological device	Frequency, duration	Level of integration IVR and instructional setting	Social interaction
Sally Wu and Alan Hung (2022)	6th grade (UPS)	Language	Explanation	HMD	90 min 1 x weekly for 10 weeks	embedded	IVR: SoI outside CG: SoI outside
Sarioğlu and Grigin (2020)	6th grade (UPS)	Science	Explanation	HMD	5–10 min 3x in two weeks	embedded	n.a.
Villena-Taranilla et al. (2022b)	4th grade (UPS)	Social Sciences	Exploration, explanation	HMD	45 min one session	embedded	IVR: no SoI CG: no SoI
Weng et al. (2019)	5th grade (UPS)	Science	Exploration, explanation, elaboration	MPB	30 min one session	additive	IVR: no SoI CG: no SoI
Yang et al. (2021)	4th grade (UPS)	Language	Exploration, explanation, elaboration	HMD	60 min one session	embedded	IVR: SoI CG: no SoI
Yang et al. (2024)	5th grade (UPS)	Language	Engagement, exploration, explanation	HMD	n.a.	embedded scaffolding (problem clues)	EG: SoI outside CG: SoI outside

Note. IVR: Immersive virtual reality, EG: Experimental IVR group, CG: Control group, N-I VR: Non-immersive VR, AT: Analog teaching methods, WL: Waiting list, TC: Textual cues, S: Summarizing, KG: Kindergarten, LPS: Lower primary school, UPS: Upper primary school, HMD: Head-mounted display, MPB: Mobile phone-based VR system, SoI: Social interaction, n.a.: not available.

Appendix D

Table 5
Presence, agency and CAMIL factors

Study	Presence	Agency	CAMIL factors
Akman and Çakır (2020)	IVR: high (est) CG: mid (est)	IVR: active (est) CG: active (est)	IM: IVR = CG SR: IVR = CG SE: IVR = CG
Alhajya et al. (2018)	IVR: high (est) CG: n.a.	IVR: active (est) CG: n.a.	–
Araiza-Alba et al. (2021)	IVR: high (ex) N-I VR: mid (est) AT: very high (est)	IVR: active (est) N-I VR: active (est) AT: active (est)	SI: IVR > N-I VR, IVR > AT EMB: IVR only
Cao et al. (2024)	IVR: high (est) CG: mid (est)	IVR: active (est) CG: n.a.	SR: IVR > CG SE: IVR > CG
Chang et al. (2020)	EG: high (est) CG: high (est)	EG: active (est) CG: active (est)	IM: EG = CG SE: EG = CG CL: EG only
Chang et al. (2024)	IVR: high (est) CG: mid (est)	IVR: active (est) CG: active (est)	SI: IVR only IM: IVR only SR: IVR only SE: IVR only CL: IVR only
Chen et al. (2023)	IVR: high (est) CG: mid (est)	IVR: passive (est) CG: passive (est)	SI: IVR = CG IM: IVR = CG SR: IVR = CG SE: IVR = CG
Chen et al. (2024)	EG: high (est) CG: high (est)	EG: passive (est) CG: passive (est)	SR: EG > CG SE: EG > CG
Cuturi et al. (2023)	EG: high (est) CG: n.a.	EG: active (est) CG: active (est)	–
Demitriadou et al. (2020)	IVR: high (est) N-I VR: mid (est) AT: low (est)	IVR: active (est) N-I VR: active (est) AT: passive (est)	SI: IVR = N-I VR, IVR > AT IM: IVR = N-I VR, IVR > AT SR: IVR = N-I VR, IVR > AT
Feng et al. (2021)	EG: high (est) CG: high (est)	EG: active (est) CG: active (est)	–
Hung et al. (2023)	IVR: high (est) N-I VR: mid (est)	IVR: active (est) N-I VR: active (est)	SI: IVR > N-I VR, IVR > AT IM: IVR > N-I VR, IVR > AT

(*continued on next page*)

Table 5 (*continued*)

Study	Presence	Agency	CAMIL factors
Iasha et al. (2023)	AT: low (est) IVR: high (est) CG: mid (est)	AT: active (est) IVR: active (est) CG: active (est)	SR: IVR > N-I VR, IVR > AT SE: IVR > N-I VR, IVR > AT SI: IVR only IM: IVR only SR: IVR only SE: IVR only
Li, F., et al. (2023)	IVR: high (est) CG: mid (est)	IVR: passive (est) CG: passive (est)	SI: IVR = CG IM: IVR = CG SR: IVR < CG SE: IVR = CG
Li, W., et al. (2023)	TC,S = high (est) TC, NS = high (est) NTC, S = high (est) NTC, NS = high (est)	TC,S = passive (est) TC, NS = passive (est) NTC, S = passive (est) NTC, NS = passive (est)	CL: TC,S = TC, NS = NTC, S = NTC, NS
Liu et al. (2022)	IVR: high (est) CG: mid (est)	IVR: active (est) CG: n.a.	SI: IVR > CG IM: IVR > CG SR: IVR > CG SE: IVR > CG CL: IVR < CG
Passig and Schwartz (2014)	IVR: high (est) CG: low (est)	IVR: active (est) CG: active (est)	–
Passig et al. (2016)	IVR: high (est) N-I VR: mid (est) AT: very high (est) WL:	IVR: active (est) N-I VR: active (est) AT: active (est) WL:	–
Sally Wu and Alan Hung (2022)	IVR: high (est) CG: low (est)	IVR: passive (est) CG: passive (est)	SI: IVR = CG IM: IVR > CG SR: IVR = CG SE: IVR > CG
Sarioğlu and Grigin (2020)	IVR: high (est) CG: mid (est)	IVR: active (est) CG: active (est)	SI: IVR only IM: IVR only SR: IVR only SE: IVR only
Villena-Taranilla et al. (2022b)	IVR: high (est) CG: low (est)	IVR: passive (est) CG: passive (est)	SI: IVR > CG IM: IVR > CG SR: IVR = CG SE: IVR > CG
Weng et al. (2019)	IVR: high (est) CG: low (est)	IVR: passive (est) CG: passive (est)	SI: IVR > CG (Interview)
Yang et al. (2021)	IVR: high (est) CG: mid (est)	IVR: active (est) CG: active (est)	SI: IVR > CG IM: IVR > CG SR: IVR > CG SE: IVR > CG EMB: IVR > CG (Interview)
Yang et al. (2024)	EG: high (est) CG: high (est)	EG: active (ex) CG: active (est)	SI: EG = CG IM: EG > CG SR: EG > CG SE: EG = CG

Note. IVR: Immersive virtual reality, EG: Experimental IVR group, CG: Control group, N-I VR: Non-immersive VR, AT: Analog teaching methods, WL: Waiting list, TC: Textual cues, NTC: No textual cues, S: Summarizing, NS: No summarizing, CAMIL: Cognitive affective model of immersive learning, SI: Situational interest, IM: Intrinsic motivation, SR: Self-regulation, SE: Self-efficacy, EMB: Embodiment, CL: Cognitive load, IVR > CG means that the value for IVR is significantly higher (< smaller and = no significant difference).

Appendix E

Table 6
Knowledge domains

Study	Knowledge domain	Prior knowledge
Akman and Çakır (2020)	Procedural knowledge	+ (est)
Alhajya et al. (2018)	Procedural knowledge, conceptual knowledge	+ (est)
Araiza-Alba et al. (2021)	Procedural knowledge	- (ex)
Cao et al. (2024)	Factual knowledge, conceptual knowledge, metacognitive knowledge	+ (est)

(*continued on next page*)

Table 6 (*continued*)

Study	Knowledge domain	Prior knowledge
Chang et al. (2020)	Factual knowledge, conceptual knowledge, procedural knowledge, metacognitive knowledge	+ (ex)
Chang et al. (2024)	Factual knowledge	+ (ex)
Chen et al. (2023)	Procedural knowledge, metacognitive knowledge	+ (est)
Chen et al. (2024)	Factual knowledge, procedural knowledge, metacognitive knowledge	+ (est)
Cuturi et al. (2023)	Factual knowledge, procedural knowledge	- (ex)
Demitriadou et al. (2020)	Factual knowledge, conceptual knowledge	- (est)
Feng et al. (2021)	Factual knowledge, procedural knowledge	+ (est)
Hung et al. (2023)	Factual knowledge, procedural knowledge, metacognitive knowledge	- (ex)
Iasha et al. (2023)	Conceptual knowledge	- (est)
Li, F., et al. (2023)	Procedural knowledge	+ (ex)
Li, W., et al. (2023)	Factual knowledge, conceptual knowledge	+ (ex)
Liu et al. (2022)	Factual knowledge, conceptual knowledge	+ (est)
Passig and Schwartz (2014)	Conceptual knowledge, procedural knowledge	- (est)
Passig et al. (2016)	Conceptual knowledge, procedural knowledge	+ (ex)
Sally Wu and Alan Hung (2022)	Factual knowledge, conceptual knowledge, procedural knowledge	+ (ex)
Sarioğlu and Grigin (2020)	Factual knowledge, conceptual knowledge	+ (ex)
Villena-Taranilla et al. (2022b)	Factual knowledge	+ (ex)
Weng et al. (2019)	Factual knowledge, conceptual knowledge	- (est)
Yang et al. (2021)	Procedural knowledge	+ (est)
Yang et al. (2024)	Procedural knowledge	+ (ex)

Note. Ex: Explicitly mentioned, Est: Estimated, +/−: Yes/no.

References

Abrahamson, D. (2014). Building educational activities for understanding: An elaboration on the embodied-design framework and its epistemic grounds. *International Journal of Child-Computer Interaction, 2*(1), 1–16.

Akman, E., & Çakır, R. (2020). The effect of educational virtual reality game on primary school students' achievement and engagement in mathematics. *Interactive Learning Environments*, 1–18.

Alhajya, N. M., Alzaghamim, S. S., & Arouri, Y. M. (2018). The impact of virtual trips on the development of Arabic language oral skills among third grade students in Jordan. *Journal of Technology and Science Education, 8*(1), 72–85.

Amarin, N. Z., & Ghishan, R. I. (2013). Learning with technology from a constructivist point. *International Journal of Business, Humanities and Technology, 3*(1), 52–57.

Amir-Behghadami, M., & Janati, A. (2020). Population, Intervention, Comparison, Outcomes and Study (PICOS) design as a framework to formulate eligibility criteria in systematic reviews. *Emergency Medicine Journal, 37*(6), 387.

Anderson, D. R., Krathwohl, P. W., Airasian, P., Cruikshank, K., Mayer, R., Pintrich, P., Raths, J., & Wittrock, M. (2001). *A taxonomy for learning, teaching, and assessing: A revision of bloom's taxonomy of educational objectives*. Pearson.

Araiza-Alba, P., Keane, T., Chen, W. S., & Kaufman, J. (2021). Immersive virtual reality as a tool to learn problem-solving skills. *Computers & Education, 164*, Article 104121.

Asmri, M. A., Haque, M. S., & Parle, J. (2023). A modified medical education research study quality instrument (MERSQI) developed by delphi consensus. *BMC Medical Education, 23*, 63.

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology, 59*, 617–645.

Bremner, N., Sakata, N., & Cameron, L. (2022). The outcomes of learner-centered pedagogy: A systematic review. *International Journal of Educational Development, 94*, Article 102649.

Cao, S., Chu, J., Zhang, Z., & Liu, L. (2024). The effectiveness of VR environment on primary and secondary school students' learning performance in science courses. *Interactive Learning Environments, 32*(10), 7321–7337.

Catrambone, R., & Yuasa, M. (2006). Acquisition of procedures: The effects of example elaborations and active learning exercises. *Learning and Instruction, 16*(2), 139–153.

Chang, S.-C., Hsu, T.-C., Chen, Y.-N., & Jong, M. S. (2020). The effects of spherical video-based virtual reality implementation on students' natural science learning effectiveness. *Interactive Learning Environments, 28*(7), 915–929.

Chang, H., Park, J., & Suh, J. (2024). Virtual reality as a pedagogical tool: An experimental study of English learners in lower elementary grades. *Education and Information Technologies, 29*(4), 4809–4842.

Checa, D., & Bustillo, A. (2020). A review of immersive virtual reality serious games to enhance learning and training. *Multimedia Tools and Applications, 79*(9–10), 5501–5527.

Chen, Y.-T., Li, M., & Cukurova, M. (2023). Unleashing imagination: An effective pedagogical approach to integrate into spherical video-based virtual reality to improve students' creative writing. *Education and Information Technologies, 29*, 6499–6523.

Chen, Y.-T., Li, M., Cukurova, M., & Jong, M. S.-Y. (2024). Incorporation of peer-feedback into the pedagogical use of spherical video-based virtual reality in writing education. *British Journal of Educational Technology, 55*(2), 519–540.

Chi, M., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist, 49*(4), 219–243.

Coban, M., Islam Bolat, Y., & Goksu, I. (2022). The potential of immersive virtual reality to enhance learning: A meta-analysis. *Educational Research Review, 36*, Article 100452.

Conrad, M., Kablitz, D., & Schumann, S. (2024). Learning effectiveness of immersive virtual reality in education and training: A systematic review of findings. *Computer Education: X Reality, 4*, Article 100053.

Csikszentmihalyi, M. (1996). *Creativity: Flow and psychology of discovery and invention*. Harper Collins Publication.

Cummings, J. J., & Bailenson, J. N. (2016). How immersive is enough? A meta-analysis of the effect of immersive technology on user presence. *Media Psychology, 19*(2), 272–309.

Cuturi, L. F. F., Cooney, S., Cappagli, G., Newell, F. N. N., & Gori, M. (2023). Primary schoolers' response to a multisensory serious game on cartesian plane coordinates in immersive virtual reality. *Cyberpsychology, Behavior, and Social Networking, 26*(8), 648–656.

Dalgarno, B., & Lee, M. J. W. (2010). What are the learning affordances of 3-D virtual environments? *British Journal of Educational Technology, 41*(1), 10–32.

Deci, E. L., & Ryan, R. M. (2015). Self-determination theory. In J. D. Wright (Ed.), *International encyclopedia of the social & behavioral sciences* (2nd ed., pp. 486–491). Elsevier.

Demitriadou, E., Stavroulia, K.-E., & Lanitis, A. (2020). Comparative evaluation of virtual and augmented reality for teaching mathematics in primary education. *Education and Information Technologies, 25*(1), 381–401.

Di Natale, A. F., Repetto, C., Riva, G., & Villani, D. (2020). Immersive virtual reality in K-12 and higher education: A 10-year systematic review of empirical research. *British Journal of Educational Technology, 51*, 2006–2033.

Ertmer, P. A., & Newby, T. J. (2013). Behaviorism, cognitivism, constructivism: Comparing critical features from an instructional design perspective. *Performance Improvement Quarterly, 26*(2), 43–71.

Evans, P., Vansteenkiste, M., Parker, P., Kingsford-Smith, A., & Zhou, S. (2024). Cognitive load theory and its relationship with motivation: A self-determination theory perspective. *Educational Psychology Review, 36*, 7.

Feng, Q., Luo, H., Li, W., Chen, Y., & Zhang, J. (2021). The moderating effect of debriefing on learning outcomes of IVR-based instruction: An experimental research. *Applied Sciences, 11*(21), Article 10426.

Hamilton, D., McKechnie, J., Edgerton, E., & Wilson, C. (2021). Immersive virtual reality as a pedagogical tool in education: A systematic literature review of quantitative learning outcomes and experimental design. *Journal of Computers in Education*, 1–32.

Han, J., Zheng, Q., & Ding, Y. (2021). Lost in virtual reality? Cognitive load in high immersive VR environments. *Journal of Advances in Information Technology, 12*(4), 302–310.

Hartmann, C., Orli-Idrissi, Y., Pflieger, L. C. J., & Bannert, M. (2023). Imagine & immerse yourself: Does visuospatial imagery moderate learning in virtual reality? *Computers & Education, 207*, 1–19.

Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist, 41*(2), 111–127.

Hung, C.-Y., Lin, Y.-T., Yu, S.-J., & Sun, J. C.-Y. (2023). Effects of AR- and VR-based wearables in teaching English: The application of an ARCS model-based learning design to improve elementary school students' learning motivation and performance. *Journal of Computer Assisted Learning, 39*(5), 1510–1527.

Iasha, V., Japar, M., Maksum, A., & Setiawan, B. (2023). Let's go on A virtual reality trip!: The effect on the students' literacy, interest, and satisfaction in cultural learning. *TEM Journal-Technology Education Management Informatics, 12*(4), 2488–2499.

Ijsselsteijn, W. A., & Riva, G. (2003). Being there: The experience of presence in mediated environments. In G. Riva, F. Davide, & W. A. Ijsselsteijn (Eds.), *Being there: Concepts, effects and measurement of user presence synthetic environments*. Ios Press.

Jensen, L., & Konradsen, F. (2018). A review of the use of virtual reality head-mounted displays in education and training. *Education and Information Technologies, 23*, 1515–1529.

Laakso, A. (2011). Embodiment and development in cognitive science. *Cognition, Brain, Behavior, 15*(4), 409–425.
Lawson, A. P., & Mayer, R. E. (2024). Effect of pre-training and role of working memory characteristics in learning with immersive virtual reality. *International Journal of Human-Computer Interaction*, 1–18.
Lee, K. M. (2004). Presence, explicated. *Communication Theory, 14*(1), 27–50.
Lessiter, J., Freeman, J., Keogh, E., & Davidoff, J. (2001). A Cross-Media Presence Questionnaire: The ITC-Sence of Presence Inventory. *Presence: Teleoperators and Virtual Environments, 10*(3), 282–297.
Li, F., Cheng, L., Wang, X., He, X., & Wang, Y. (2023). The effects of spherical video-based virtual reality and conventional video on students' descriptive writing achievement and motivation: A comparative study. *Sage Open, 13*(3).
Li, Y. D., Ding, G. H., & Zhang, C. Y. (2021). Effects of learner-centered education on academic achievement: A meta-analysis. *Educational Studies, 50*(3), 1–14.
Li, W., Feng, Q., Zhu, X., Yu, Q., & Wang, Q. (2023). Effect of summarizing scaffolding and textual cues on learning performance, mental model, and cognitive load in a virtual reality environment: An experimental study. *Computers & Education, 200*, 1–14.
Lindgren, R., Morphew, J. W., Kang, J., Planey, J., & Mestre, J. P. (2022). Learning and transfer effects of embodied simulations targeting crosscutting concepts in science. *Journal of Educational Psychology, 114*(3), 462–481.
Liu, R., Wang, L., Koszalka, T. A., & Wan, K. (2022). Effects of immersive virtual reality classrooms on students' academic achievement, motivation and cognitive load in science lessons. *Journal of Computer Assisted Learning, 38*(5), 1422–1433.
Makransky, G., Andreasen, N. K., Baceviciute, S., & Mayer, R. E. (2021). Immersive virtual reality increases liking but not learning with a science simulation and generative learning strategies promote learning in immersive virtual reality. *Journal of Educational Psychology, 113*(4), 719–735.
Makransky, G., & Lilleholt, L. (2018). A structural equation modeling investigation of the emotional value of immersive virtual reality in education. *Educational Technology Research & Development, 66*, 1141–1164.
Makransky, G., & Petersen, G. B. (2021). The cognitive affective model of immersive learning (CAMIL): A theoretical research-based model of learning in immersive virtual reality. *Educational Psychology Review, 33*, 937–958.
Martarelli, C. S., Dubach, J., Schelleis, N., Cacchione, T., & Tempelmann, S. (2024). Virtual reality in primary science education: improving knowledge of the water cycle. *Educational Technology Research and Development, 73*, 999–1024.
Mavilidi, M. F., Pesce, C., Benzing, V., Schmidt, M., Paas, F., Okely, A. D., & Vazou, S. (2022). Meta-analysis of movement-based interventions to aid academic and behavioral outcomes: A taxonomy of relevance and integration. *Educational Research Review, 37*, Article 100478.
Mayer, R. E., Makransky, G., & Parong, J. (2023). The promise and pitfalls of learning in immersive virtual reality. *International Journal of Human-Computer Interaction, 39* (11), 2229–2238.
Meyer, O. A., Omdahl, M. K., & Makransky, G. (2019). Investigating the effect of pre-training when learning through immersive virtual reality and video: A media and methods experiment. *Computers & Education, 140*, Article 103603.
Mikropoulos, T. A., & Natsis, A. (2011). Educational virtual environments: A ten-year review of empirical research (1999-2009). *Computers & Education, 56*, 769–780.
Paas, F., & Sweller, J. (2012). An evolutionary upgrade of cognitive load theory: Using the human motor system and collaboration to support the learning of complex cognitive tasks. *Educational Psychology Review, 24*, 27–45.
Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *British Medical Journal, 372*(71).
Passig, D., & Schwartz, T. (2014). Solving conceptual and perceptual analogies with virtual reality among kindergarten children of immigrant families. *Teachers College Record, 116*(2), 1–36.
Passig, D., Tzuriel, D., & Eshel-Kedmi, G. (2016). Improving children's cognitive modifiability by dynamic assessment in 3D Immersive Virtual Reality environments. *Computers & Education, 95*, 296–308.
Petersen, G. B., Petkakis, G., & Makransky, G. (2022). A study of how immersion and interactivity drive VR learning. *Computers & Education, 179*, Article 104429.
Putra, F., Kholifah, I. Y. N., Subali, B., & Rusilowati, A. (2018). 5e-learning cycle strategy: Increasing conceptual understanding and learning motivation. *Jurnal Ilmiah Pendidikan Fisika Al-BiRuNi, 7*(2), 171.
Radianti, J., Majchrzak, T. A., Fromm, J., & Wohlgenannt, I. (2020). A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. *Computers and Education, 147*, Article 103778.
Sally Wu, Y.-H., & Alan Hung, S.-T. (2022). The effects of virtual reality infused instruction on elementary school students' English-speaking performance, willingness to communicate, and learning autonomy. *Journal of Educational Computing Research, 60*(6), 1558–1587.
Sarac, H. (2018). The effect of learning cycle models on achievement of students: A meta-analysis study. *International Journal of Educational Methodology, 4*(1), 1–18.
Sarioğlu, S., & Grigin, S. (2020). The effect of using virtual reality in 6th grade science course the cell topic on students academic achievements and attitudes towards the course. *Turkish Journal of Science Education, 17*(1), 109–125.
Schunk, D. H., & DiBenedetto, M. K. (2016). Self-efficacy theory in education. In K. R. Wentzel, & D. B. Miele (Eds.), *Handbook of motivation at school* (pp. 34–54). Routledge.
Slater, M., & Sanchez-Vives, M. V. (2016). Enhancing our lives with immersive virtual reality. *Frontiers in Robotics and AI, 3*, 74.
Southgate, E., Smith, S. P., Cividino, C., Saxby, S., Kilham, J., Eather, G., Scevak, J., Summerville, D., Buchanan, R., & Bergin, C. (2019). Embedding immersive virtual reality in classrooms: Ethical, organisational and educational lessons in bridging research and practice. *International Journal of Child-Computer Interaction, 19*, 19–29.
Sweller, J. (2020). Cognitive load theory and educational technology. *Educational Technology Research & Development, 68*, 1–16.
Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. Springer.
Villena-Taranilla, R., Cózar-Gutiérrez, R., González-Calero, J. A., & López Cirugeda, I. (2022b). Strolling through a city of the Roman Empire: An analysis of the potential of virtual reality to teach history in Primary Education. *Interactive Learning Environments, 30*(4), 608–618.
Villena-Taranilla, R., Tirado-Olivares, S., Cozar-Gutierrez, R., & Gonzalez-Calero, J. A. (2022a). Effects of virtual reality on learning outcomes in K-6 education: A meta-analysis. *Educational Research Review, 35*, Article 100434.
Weng, C., Rathinasabapathi, A., Weng, A., & Zagita, C. (2019). Mixed reality in science education as a learning support: A revitalized science book. *Journal of Educational Computing Research, 57*(3), 777–807.
Wilson, A. D., & Golonka, S. (2013). Embodied cognition is not what you think it is. *Frontiers in Psychology, 4*, 58.
Wirth, W., Hartmann, T., Böcking, S., Vorderer, P., & Jäncke, P. (2007). A process model of the formation of spatial presence experiences. *Media Psychology, 3*(9), 493–525.
Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments, 7*(3), 225–240.
Wu, B., Yu, X., & Gu, X. (2020). Effectiveness of immersive virtual reality using head-mounted displays on learning performance: A meta-analysis. *British Journal of Educational Technology, 51*(6), 1991–2005.
Yang, G., Chen, Y., Zheng, X., & Hwang, G. (2021). From experiencing to expressing: A virtual reality approach to facilitating pupils' descriptive paper writing performance and learning behavior engagement. *British Journal of Educational Technology, 52*(2), 807–823.
Yang, G., Zhou, W., Rong, Y.-D., Xu, Y.-J., Zeng, Q.-F., & Tu, Y.-F. (2024). Designing a second-order progressive problem-based scaffold strategy to promote students' writing performance in an SVVR environment. *Education and Information Technologies, 29*, 14591–14620.