Conrad, Matthias; Kablitz, David; Schumann, Stephan

Learning effectiveness of immersive virtual reality in education and training. A systematic review of findings

Computers & education: X reality 4 (2024), 11 S.

Computers & Education: X Reality 4 (2024) 100053

Contents lists available at ScienceDirect

Computers & Education: X Reality

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

Learning effectiveness of immersive virtual reality in education and training: A systematic review of findings

Matthias Conrad *, David Kablitz, Stephan Schumann

University of Konstanz Universitaetsstrasse 10 78464 Konstanz Germany

A R T I C L E I N F O

Keywords:
Virtual reality
Education
Learning
Engagement

A B S T R A C T

The use of immersive virtual reality (IVR) offers a variety of design possibilities for action-oriented teaching and learning that enables the promotion of specific knowledge and skills. In order to use IVR applications as an effective teaching and learning medium, it is important to understand the potential advantages and disadvantages of this technology compared to other media. This raises the question of what type of learning environment is most effective in promoting specific knowledge and skills. To answer these questions, a systematic review of research on learning with IVR in an educational context was conducted using the PRISMA method (Liberati et al., 2009; Page et al., 2021). The study analyzed 30 relevant research articles to compare the relative effectiveness of IVR-based learning and its impact on learner engagement, as outlined in the ICAP framework (Chi & Wylie, 2014). The results indicate that IVR has a positive impact on learning compared to other types of media. The study suggests that IVR technology is suitable for learning environments that prioritize active learner engagement and practical application, such as active manipulation and constructive creation. In summary, the results offer more insights into the advantages of using IVR to accomplish particular learning objectives.

1. Introduction

The use of immersive virtual reality (IVR) enables complex instructional arrangements that can promote the achievement of specific learning goals. In the area of education and training, particular potential is seen in the promotion of procedural knowledge required to solve complex problems in novel or unfamiliar work contexts (Hamilton et al., 2021). Research on learning with IVR in secondary and higher education suggests several conclusions about the effectiveness of learning in promoting different forms of knowledge, with some evidence that IVR tends to outperform non-immersive technologies in the acquisition of conceptual information and procedural knowledge (ibid.). Under certain conditions, the use of IVR has also been found to be conducive to the acquisition of declarative knowledge (Hamilton et al., 2021; Jensen & Konradsen, 2018). Other studies find an advantage in achieving cognitive learning goals when using IVR, as well as greater learner engagement (Concannon et al., 2019). This paper compares the benefits of this technology in terms of learning outcomes with the use of other types of media. It also explores what type of learning environment is most effective in promoting specific knowledge and skills when using IVR technologies. Understanding the benefits of using this technology in terms of learning outcomes, especially in comparison to alternative media, is of paramount importance in the field of computer assisted learning and instruction. In addition, it is imperative for educators and educational institutions to understand the particular domain of knowledge that is best advanced through the integration of IVR in the field of Computer Assisted Learning and Instruction. A nuanced understanding of the specific benefits of IVR enables educators to tailor instructional strategies to create immersive and engaging learning experiences that are tailored to the preferences of the learner. It helps ensure that instructional content is seamlessly aligned with the inherent strengths of this technology. Such precision in approach is critical to increasing the overall impact on learning outcomes. By identifying the specific types of learning environments in which IVR is superior, educators gain the ability to create compelling and memorable learning experiences that increase student engagement and retention. Maintaining an awareness of the distinct advantages offered by IVR is critical to keeping educational institutions abreast of technological advancements in this rapidly evolving landscape. Such awareness not only ensures technological relevance, but also serves as a catalyst for continued research and

* Corresponding author.
E-mail addresses: matthias.conrad@uni-konstanz.de (M. Conrad), david.kablitz@uni-konstanz.de (D. Kablitz), stephan.schumann@uni-konstanz.de (S. Schumann).

https://doi.org/10.1016/j.cexr.2024.100053
Received 22 September 2023; Received in revised form 17 December 2023; Accepted 20 January 2024
Available online 13 February 2024

innovation in the field of educational technology. This, in turn, facilitates the development and refinement of novel IVR applications and learning experiences. The main added value of this study lies in the study-based identification of the relative superiority of IVR-based learning environments in terms of specific forms of learning, with special consideration of the associated learner engagement, as assumed in the ICAP framework (Chi & Wylie, 2014). The results can provide additional perspectives on the beneficial use of this technology, particularly with regard to the achievement of specific learning objectives.

2. Theory

In order to approach the topic of IVR-based learning from a theoretical perspective, it is important to first consider the basic paradigms of relevant theories. Learning is often characterized as a lasting change in behavior resulting from practice or experience, but it is important to note that learning does not always result in behavioral change (Lachman, 1997). In this regard, the most important goal of educational processes is the acquisition of declarative and procedural knowledge that individuals need to solve specific problems (Anderson, 1976, 2005; Renkl, 2015). Both procedural and declarative knowledge play an important role in how we understand and interact with our environment. Although declarative and procedural knowledge are presented as two separate domains, they are interrelated (Anderson & Krathwohl, 2001). In cognitive psychology, declarative knowledge refers to knowledge of facts and events, knowledge of complex relationships, knowledge of rules and constraining conditions, and knowledge of semantic relationships (Renkl, 2015). Procedural knowledge, on the other hand, focuses on application, such as calculating a mathematical problem or writing a textual argument. Procedural knowledge does not need to be reconstructed explicitly and consciously (Anderson & Krathwohl, 2001). Thus, procedural knowledge represents the ability to perform certain actions purposefully and effectively or to reproduce sequences of actions in a verbalized form. Almost all learning and memory phenomena are procedural. Purely declarative knowledge is the exception rather than the rule (ten Berge & van Hezewijk, 1999). It can be assumed that procedural knowledge builds on declarative knowledge (ibid.). Higher levels of knowledge are usually characterized by an increasing degree of proceduralization, so that in the course of this process knowledge turns into ability. Therefore, procedural knowledge can also be identified with skill (Renkl, 2015). Both types of knowledge can refer to technical, domain-specific knowledge and also have interdisciplinary components. In general, it should be noted that knowledge acquisition is not only about the quantitative acquisition of individual knowledge elements, but also about the quality of knowledge and the networking of knowledge structures (ibid.). Due to the possibility of computer-based simulation and interaction, IVR is considered to promote specific knowledge and practical skills (Radianti et al., 2020).

According to the *Cognitive Affective Model of Immersive Learning* (CAMIL), the user's sense of presence and agency are critical affordances for learning in IVR, with immersion, control factors, and representational fidelity able to facilitate these affordances (Makransky & Petersen, 2021). In particular, the importance of *agency* in the context of IVR-based learning cannot be overstated. It refers to the learner's sense of control, autonomy, and ability to take meaningful action within the virtual environment. The perception of agency can increase engagement in the virtual environment, contributing to a heightened sense of presence and a deeper understanding of the subject matter. In this context, a suitable system for categorising IVR-based learning in terms of different levels of engagement is Chi and Wylie's (2014) ICAP framework, which describes four different modes of learner engagement: *Passive Receiving*, *Active Manipulating*, *Constructive Generating* and *Interactive Dialoguing*. At the passive receiving level, learners are mere recipients of information, engaging in activities such as watching lectures or reading without active participation. Passive receiving refers to a low-level engagement where learners are passive recipients of information. In this mode, learners do not actively process or interact with the content but rather absorb it without significant cognitive involvement. Moving up the hierarchy, at the active manipulating level, learners process information and participate in basic interactions, fostering a more engaged learning experience. Active manipulating involves a higher level of cognitive engagement, where learners actively manipulate information, apply concepts, or perform hands-on activities. This level requires learners to go beyond passive reception and actively interact with the learning materials. Constructive generating emphasizes the active construction of knowledge, encouraging learners to integrate concepts and develop problem-solving skills. In this mode of engagement, learners actively create knowledge by integrating and synthesizing information to form a deeper understanding. It goes beyond basic interaction and requires learners to generate new insights or solutions. Group work, projects, and simulations are examples of constructive activities. Finally, the interactive dialoging level represents the pinnacle, where learners not only construct knowledge, but also share and discuss it with others through group discussions, cooperative learning, or online interactions. This mode represents the highest level of engagement in the ICAP model. It involves not only constructing knowledge but also sharing, discussing, and refining it through interactions with peers. Interactive dialoging emphasizes collaborative learning and the social construction of knowledge.

The ICAP framework emphasizes the critical link between interactivity and constructiveness in learning activities and highlights their central role in enhancing cognitive processing and comprehension. It provides a solid foundation for designing learning environments that actively engage learners and enhance their cognitive skills. The corresponding hypothesis predicts that the more engaged learners are in the learning environment, the more they will learn (Chi & Wylie, 2014). Following Chi and Wylie's (ibid) category system, we have simplified the basic modes of IVR-based learning environments to distinguish passive forms of engagement from more active forms. The resulting dichotomous distinction is between *IVR environments with passive learner engagement* (e.g. viewing a digital artefact) and *IVR environments with active learner engagement* (e.g. bidirectional interaction with an IVR environment). Given this context, scenarios surpassing passive observation can be categorized as learning environments with active learner engagement. This is due to the necessity for proactive participation from the learner, thereby stimulating heightened cognitive processes. The resulting classification of IVR-based learning environments can be used to design learning scenarios that meet specific learning objectives. In the present study, we will use these dichotomous categories to compare basic types of learning environments according to the corresponding

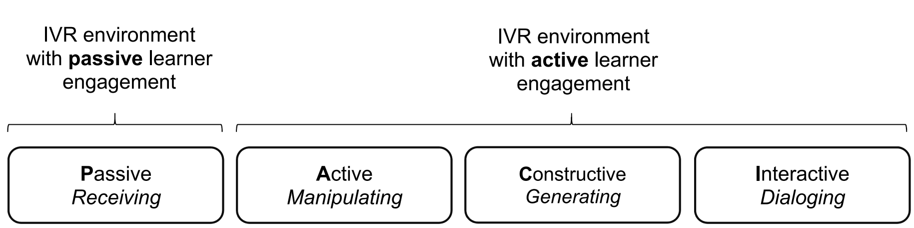

Fig. 1. Types of IVR-based learning environments along the ICAP framework (original graphic).

learner engagement (Fig. 1).

2.1. Reviews on teaching and learning with IVR

To obtain a general overview of previous systematic literature reviews on the topic of “IVR in education”, a systematic search for systematic literature reviews on the learning-related effectiveness of IVR technologies was conducted using the educational and psychological databases *APA PsycArticles*, *APA PsycInfo*, *Education Resources Information Center* (ERIC), and *Education Resources Information Center*. To ensure sufficient comparability of study results, only reviews from 2010 onwards that were embedded in formal educational contexts and based on the use of *head-mounted displays* (HMDs) were included, as this technology is the most common form of IVR use. Based on this specification, five relevant reviews were identified (Table 1).

The identified reviews focus on different aspects of learning and different sectors of secondary and tertiary education (e.g., high school, university, vocational training) (Concannon et al., 2019; Hamilton et al., 2021; Radianti et al., 2020). Although the varying focus of the studies does not allow for a definitive conclusion on the effects of learning with IVR, a number of findings can be derived regarding the effectiveness of learning in terms of achieving relevant learning goals.

Overall, the use of IVR technologies tends to be superior to non-immersive technologies in terms of recall of abstract and conceptual information and acquisition of procedural knowledge (Hamilton et al., 2021). However, the effects of these studies have reported mixed results (Concannon et al., 2019; Hamilton et al., 2021; Jensen & Konradsen, 2018). Jensen and Konradsen (2018) note inconsistent findings regarding the development of cognitive skills when using IVR for education and training. Some of the studies reviewed here find an increase in cognitive skills, while other studies find no positive effects of IVR-based learning, especially when compared to desktop-based training and other learning methods or media. However, the use of IVR is beneficial when the focus is on visual and spatial perception, such as in safety or medical training. For the acquisition of psychomotor and practical skills, IVR appears to have an advantage over the use of non-immersive technologies such as desktop computers (ibid). Hamilton et al. (2021) categorize the effects of IVR use into cognitive, procedural, and affective learning, and identify positive effects on the achievement of affective learning goals (e.g., attitudes and values) and partially on the acquisition of procedural knowledge. For the majority of studies in post-secondary education, Concannon et al. (2019) identify an advantage in motivational aspects and cognitive learning goals of using IVR technology compared to other types of media, as well as greater learner engagement. Despite these findings, it is crucial to emphasize that while there are theoretical categories for grouping learning environments in IVR-related research, there has not been a systematic differentiation concerning learner engagement within these environments. However, addressing this gap is important to offer further guidance on the significance of this aspect.

Table 1
Selected reviews on learning with IVR technologies as of 2010.

Authors	Number of studies included	Study focus
Billingsley et al. (2019)	7	Teacher training: Effects of IVR on understanding, knowledge, skills and attitudes.
Concannon et al. (2019)	119	Areas of application, learning outcomes, and conceptual and theoretical rationale for the use of IVR in post-secondary education.
Hamilton et al. (2021)	29	Comparison of learning outcomes between IVR users and control groups (including desktop VR, PowerPoint, video).
Jensen and Konradsen (2018)	21	Acquisition of affective, cognitive and psychomotor skills in the field of education.
Radianti et al. (2020)	38	Learning content and application areas of IVR in higher education.

3. Research questions

Given this context, it seems obvious that IVR technologies can have some benefit for specific learning. However, the specific learning environments in which VR applications truly capitalize on their added value remain largely unclear. In this study, we use Chi and Wylie’s (2014) ICAP framework as a theoretical foundation to classify learning environments as either *active* or *passive*. Building on this foundation, the first investigation explored the extent to which IVR-supported teaching and learning environments provide benefits in terms of different types of knowledge acquisition. Subsequently, the study asks whether passive or active IVR learning environments (based on the ICAP framework) demonstrate superiority over traditional teaching approaches. In order to gain more information about the effectiveness of IVR compared to other media, the following research question was formulated.

1) Is IVR more effective than other media in achieving specific learning outcomes?

The second research question relates to the relative advantage of specific types of learning environments in terms of learning outcomes. In this context, the general suitability of this technology in terms of learner engagement will be examined more closely in relation to the assumptions of the ICAP framework (Chi & Wylie, 2014).

2) What type of learning environment is most effective when using IVR?

The answers to these research questions can enhance the intentionality and effectiveness of IVR-based learning environments by considering the learner’s underlying activity in relation to the intended learning objectives. In this way, the results can provide further insight into the beneficial use of this technology, including the achievement of specific learning goals associated with it.

4. Methods

To answer these questions, a systematic review was conducted using the PRISMA method (Liberati et al., 2009; Page et al., 2021). This method primarily concentrates on reporting reviews assessing the effects of experimental interventions but can also serve as a foundation for reporting systematic reviews with objectives extending beyond the evaluation of interventions. A systematic review, as defined by Liberati et al. (2009), is characterized by clear objectives with explicit, reproducible methods. It involves a systematic search to identify all studies that meet predefined selection criteria. It also includes an assessment of the validity of the results of the included studies. The process is completed by a systematic presentation and synthesis of the characteristics and results of the included studies. These features serve to create a transparent, comprehensible and reproducible research process.

4.1. Definition of the search principle and the search components

For the present study, a sensitive search principle was used to identify the highest possible number of relevant hits. Compared to the specific search principle, a sensitive search requires more effort because a high proportion of publications appear in the hit list. At the same time, a sensitive search minimizes the probability of skipping relevant hits. The purpose of defining search components is to translate the research question into a searchable format. This step determines which elements of the research question will be included in the search strategy. With regard to the research question, the following search components were derived: (1) *HMD-based*, (2) *Immersive Virtual Reality* (3) *learning outcome*. The limitation to these three search components was intended to provide as many hits as possible without overlooking potentially suitable studies. The same applies to the specification of possible comparison media, which could not be comprehensively narrowed down in advance.

4.2. Selection of the databases to be searched

A comprehensive literature search requires a search of relevant specialist databases. The selection of databases is based on the extent to which they cover publications relevant to the research topic. The following databases specializing in educational or psychological aspects were included in the systematic literature search.

- APA PsycArticles
- APA PsycInfo
- Education Resources Information Center (ERIC)
- Teacher Reference Center

These databases were selected because they specialize in educational and/or psychological subject areas and thus seemed appropriate for thematic searches related to learning outcomes associated with the use of technology.

4.3. Identification of synonymous keywords

The next step was to derive appropriate synonymous keywords from the identified search components. The keywords identified and used in this work are listed in Table 2. A very sensitive approach was chosen, especially for the concept of learning outcome, as it can appear in different forms.

4.4. Development of the syntax

The search algorithm set up as part of the database search allowed for high relevance and quality of search results. The search mode was based on a Boolean algorithm and thus on a logical combination of the previously defined terms. In this study, the three Boolean operators AND, OR, and NOT were used to specify the search. In order to appear in the hit list, a keyword of the search components *Immersive Virtual Reality* and *learning outcome* had to be mentioned in the title or abstract of the article. For the search component *HMD-based*, however, only one of the keywords had to appear in the text, as preliminary tests showed that otherwise the number of hits would be significantly lower. However, results containing the keywords "therapy" or "rehabilitation" in the title or abstract were excluded in order to sort out a large number of hits that focus on the learning progress of patients as part of a medical-rehabilitative measure. The source selection was confined to scientific journals to guarantee that the identified studies originated from reputable, peer-reviewed publications. To maintain relevance to the research context, the publication date criterion for the identified articles was set to January 2010, given the constant technological evolution of IVR technologies. All searches were conducted in September 2020.

Table 2
Derivation of keywords from the search components.

Search components	Keywords	
HMD-based	- cardboard	- HMD
	- goggles	- HTC Vive
	- head mounted display	- immersive
	- head mounted helmet	- Oculus Rift
	- head-mounted display	- Samsung Gear
	- head-mounted helmet	- spectacles
	- headset	
Immersive Virtual Reality (IVR)	- IVE (Immersive Virtual Environment)	- virtual reality
		- virtual simulation
	- IVLE (Immersive Virtual Learning Environment)	- virtual world
		- VLE (Virtual Learning Environment)
	- VE (Virtual Environment)	
	- virtual environment	- VR
	- virtual learning environment	
Learning outcome	- cognition	- learning outcome
	- cognitive	- learning performance
	- competence	- learning result
	- educational achievement	- learning success
	- educational effectiveness	- outcome in learning
	- educational outcome	- outcome of learning
	- educational success	- performance
	- expertise	- skill
	- know	- solve
	- know-how	- study achievement
	- knowledge	- study effectiveness
	- learn	- study outcome
	- learning achievement	- study result
	- learning effectiveness	- study success
	- learning improvement	- understand

4.5. Inclusion and exclusion criteria

All included studies measured learning outcomes as a dependent variable and related to either primary, secondary or tertiary education. Moreover, the inclusion criteria covered studies related to continuing education, incorporating specific educational and training initiatives (e.g., road safety education or safety training). Furthermore, supplementary content and quality criteria were defined. An article was included in the review if.

➢ it was published in a scientific journal
➢ it had successfully undergone a peer review process
➢ it was an empirical study
➢ learning success was measured as a dependent variable within the study
➢ an IVR-based HMD device or HMD-like model was used
➢ the article was published in an included database between 2010 and 2020
➢ at least one other group besides the HMD group participated in the study, in order to include only empirical intervention studies with an experimental control group design, as this allows a direct comparison with alternative scenarios

However, an article was excluded if.

➢ it was a project report, article, edited volume, monograph (e.g., in the form of a dissertation), conference paper, or technology report, as these forms of publication may not have sufficient quality control with respect to the information they contain
➢ discussed only the theories, concepts, frameworks, or proposals without following up with an experiment or case study to validate them
➢ learning was based on a subjective assessment of the subjects, as these do not allow for objective measurement
➢ the dependent variables were based solely on body- or brain-specific measures (e.g., activity of the cerebral cortex), as no conclusions can be drawn about the acquisition of relevant knowledge
➢ other forms of IVR (e.g., desktop IVR or CAVE) were used instead of HMD
➢ the individuals in the sample were undergoing medical rehabilitation as part of the study, or if the individuals in the sample were diagnosed with a physical or mental impairment or disease (e.g., autism, ADHD), as these participants may have impaired perceptual and/or cognitive performance

4.6. Selection process

The selection process consisted of a two-step filtering procedure. The first step was the semi-automatic filtering described above, using the exclusion and inclusion criteria formulated in the syntax. This was followed by a manual filtering process, which can be divided into three distinct steps.

1. Read title and abstract

The database search identified 923 articles. After removing 96 duplicates, 827 records were analyzed. In this way, 643 records were subsequently excluded.

2. Reading of the content

From a total of 144 records, a further 114 could be excluded after checking the content for the above criteria. Exclusions were mainly due to the fact that no HMD technology was used, no comparison with other media was made or the learning success was based on the subjective assessment of the participating subjects.

3. Backward search

The bibliographies of the remaining 30 studies were searched backwards for additional potential matches. No suitable articles were identified that met the inclusion criteria.

Fig. 2 shows a flowchart of the selection process.

4.7. Categorization of relevant media

The different types of reference media were categorized according to the media identified in the database search, with no restrictions on acceptable media types. This was done to include all forms of relevant reference media. The resulting classification distinguishes between the use of *electronic* and *analog media*. On the one hand, the use of analog media includes verbal or textual instruction as well as textbooks, texts, and practical exercises. On the other hand, electronic media includes computer-based technologies (e.g., tablet computers, laptops, desktop computers, AR technologies) and other electronic media such as digital slide presentations, videos, and digital simulations. This approach should allow for a specific comparison between different types of media and the use of an IVR-based HMD.

4.8. Categorization of learning outcomes

The distinction between *declarative knowledge* and *procedural knowledge* used in this study is based on the considerations of Anderson (1976; 2005; Anderson & Krathwohl, 2001). Based on this concept, the following rule was applied to categorize the respective types of knowledge: If the study aimed at reproduction or reorganization in a task as a requirement, the learning outcome was classified as *declarative knowledge*. If the task involved the transfer or solution of a specific problem, the corresponding learning outcome was classified as *procedural knowledge*. A quantitative measure of learning requires a comparison between an initial and a final state, which can be accomplished by a pre-post test comparison (Brogan & Kutner, 1980). The difference

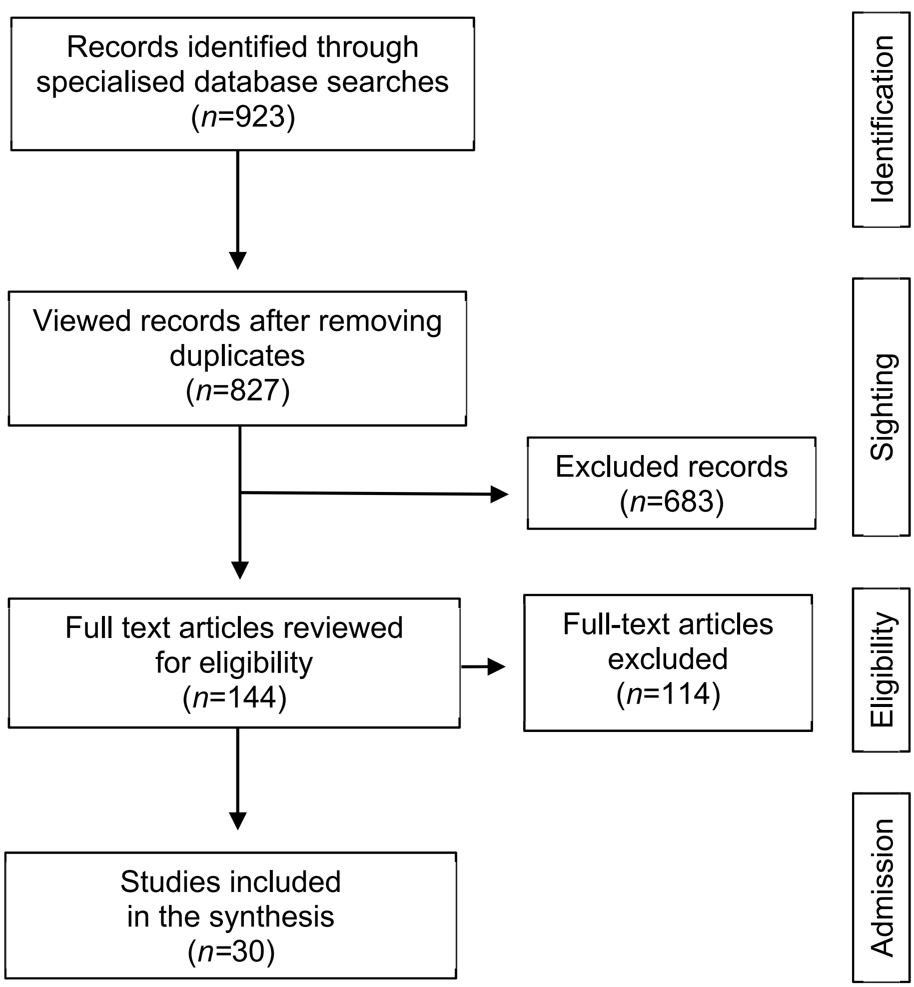

Fig. 2. Selection process using the PRISMA flow chart (original graphic).

between these states was interpreted as learning outcome.

4.9. Categorization of learning environments

The underlying categorization of learning environments was based on Chi and Wylie's (2014) ICAP framework. In contrast to Chi and Wylie's model, which distinguishes four different types of learner engagement, the present study only distinguished between IVR environments with *passive learner engagement* and those with *active learner engagement*. If the IVR-based learning environment was used only to perceive or explore through passive observation (e.g., viewing a chemical molecule), it was classified as *passive*. Learning environments that went beyond passive observation or exploration were classified as *active*, such as action-oriented simulations or active construction and design. This dichotomous distinction was intended to allow for fundamental differences between more passive and more active learner roles.

4.10. Evaluation of studies in terms of learning effects

The studies were analyzed in terms of their reported effects on learning. Where the necessary information was reported, effect sizes were calculated manually using *Cohen's d* formula. According to Cohen's (1988) convention for interpreting effect sizes, values below 0.2 are negligible, values from 0.2 to below 0.5 represent a small effect, values from 0.5 to below 0.8 represent a medium effect, and values above 0.8 represent a large effect.

5. Results

A detailed breakdown of the identified studies, including authors, HMD technologies used, countries, and the context in which each study was conducted, is presented in Table 3. The 30 studies included a total of 2404 subjects. Of these, 1110 were female, 809 were male, and one was diverse. Nine studies failed to distinguish between gender, making it impossible to ascertain the gender of 485 participants. The study samples were predominantly composed of tertiary students (36 %) and primary and secondary students (28 %). The remaining part of the sample consisted of vocational training attendees and individuals undergoing job-related further training.

Regarding the hardware employed in the identified studies, Oculus-branded devices were the most frequently utilized, followed by Samsung, HTC, and Sony. Seven studies involved models from other manufacturers, while five studies used HMDs without providing brand details. Additionally, three studies employed more than one HMD model.

The media comparison studies identified in the review are reported below (Table 4). The overview categorises the results according to the reference medium (e.g. video) and provides further information on the authors, the year of publication, the types of knowledge examined (declarative and/or procedural) and the associated findings.

5.1. Impact on declarative and procedural knowledge

To address the first research question regarding the effectiveness of IVR for specific learning outcomes compared to other types of media, we analyzed the identified studies regarding the reference media and the resulting effect sizes. A comprehensive analysis was conducted to assess whether HMD-based IVR technologies are superior or inferior to other electronic and analog media.

A total of 45 learning outcome measures were identified in 30

Table 3
Descriptive overview of the included studies.

Authors	HMD Technology	Country	Sample	*N*	m	f	d
Butt et al. (2018)	Oculus Rift	USA	University Students	20	n.s.	n.s.	n.s.
Buttussi and Chittaro (2018)	Sony HMZ-T1 HMD, Oculus Rift DK2	Italy	Training Attendees	96	55	41	0
Chang et al. (2019)	VR Box 2nd Generation	Taiwan	University Students	64	n.s.	n.s.	n.s.
Chittaro and Buttussi (2015)	Sony HMZ-T1	Italy	Training Attendees	48	26	22	0
Innocenti et al. (2019)	Not specified	Italy	Students in school	36	20	16	0
Ferguson et al. (2020)	Sony PlayStation VR Headset (CUH-ZVR1)	Netherlands	Students in school	25	12	13	0
Gutiérrez-Maldonado et al. (2015)	Oculus Rift DK1	Spain	University Students	52	n.s.	n.s.	n.s.
Huang et al. (2019)	Oculus Rift	USA	University Students	109	28	81	0
Kozhevnikov et al. (2013)	nVisor SX60	USA	University Students	37	19	18	0
LaFortune and Macuga (2018)	Sensics zSight	USA	Training Attendees	56	18	38	0
Lai et al. (2019)	HMD	Taiwan	University Students	90	n.s.	n.s.	n.s.
Makransky et al. (2020)	Experiment 1: Samsung Gear VR	USA	University Students	131	47	84	0
	Experiment 2: Samsung Gear VR	USA	Students in school	165	54	111	0
Makransky, Terkildsen & Mayer (2019)	Samsung Gear VR	Denmark	University Students	52	22	30	0
Makransky, Borre-Gude & Mayer (2019)	Samsung Gear VR	Denmark	University Students	105	49	56	0
Meyer et al. (2019)	Samsung Gear VR	Denmark	University Students	118	63	55	0
Moro, Štromberga, and Stirling (2017)	Oculus Rift, Samsung Gear VR	Australia	University Students	59	28	31	0
Negro Cousa et al. (2019)	VR i7	USA	University Students	42	15	27	0
Oh et al. (2019)	Oculus Rift	USA	University Students	129	64	65	0
Parong and Mayer (2018)	HTC Vive	USA	University Students	55	17	38	0
Parong and Mayer (2020)	HTC Vive	USA	University Students	61	20	40	1
Passig et al. (2016)	HMD	USA	University Students	117	61	56	0
Rupp et al. (2019)	Oculus Rift DK2, Oculus Rift CV1, Google Cardboard	USA	University Students	136	70	66	0
Sportillo et al. (2018)	HTC Vive	France	Training Attendees	60	30	30	0
Sundar et al. (2017)	Cardboard	USA	University Students	129	29	100	0
Tai et al. (2020)	Samsung Gear VR	Taiwan	Sudents in school	49	27	22	0
Ventura et al. (2019)	HMD	Spain	University Students	42	25	17	0
Villena Taranilla et al. (2019)	Netway Vita	Spain	Sudents in school	98	52	46	0
Webster (2016)	Sony HMZ T1	USA	Professionals	140	136	4	0
Yang et al. (2018)	HTC Vive	China	University Students	60	26	34	0
Ye et al. (2019)	HMD	China	University Students	62	12	50	0

N = Number of participants; m = male; f = female; d = diverse; n.s. = not specified.

Table 4
Overview of the studies by comparative media characteristics.

Reference Medium	Authors	Dependent Learning Outcome	Findings
Video	Sundar et al. (2017)	Declarative knowledge	No significant difference between IVR groups in newspaper story recall performance.
	Tai et al. (2020)	Declarative knowledge	Significantly greater vocabulary learning success on both the immediate and delayed posttest for the IVR group ($d = 0.89$).
	Rupp et al. (2019)	Declarative knowledge	Significantly higher recall performance when using immersive IVR devices to view 360-degree video.
	LaFortune and Macuga (2018)	Declarative knowledge	No significant differences in dance movement reproduction between video and IVR groups.
	Meyer et al. (2019)	Declarative and procedural knowledge	With pre-training: IVR group slightly better in recall and transfer test (not significant); without pre-training: video group significantly better in recall test, no significant difference in transfer test.
	Chang et al. (2019)	Declarative and procedural knowledge	Significantly increased learning of the IVR group in medical obstetrics ($d = 0.22$).
	Makransky et al. (2020)	Experiment 1: Declarative and procedural knowledge	No significant difference in declarative and procedural knowledge gain in forensic DNA analysis.
		Experiment 2: Declarative and procedural knowledge	Significantly greater increase in declarative knowledge for participants in the video group ($d = 0.23$); no significant difference between groups in procedural knowledge increase and transfer task.
Computer Desktop	Oh et al. (2019)	Declarative knowledge	Significantly better recall of information from a conversation by the laptop group ($d = 0.56$).
	Buttussi and Chittaro (2018)	Declarative knowledge	No superiority of one group in the subsequent and delayed posttest on the correct behavior in problem and accident scenarios during a flight.
	Ferguson et al. (2020)	Declarative and procedural knowledge	Significantly better spatial recall in the IVR group ($d = 0.93$); but no significant difference in knowledge.
	Makransky, Borre-Gude, and Mayer (2019) Makransky, Borre-Gude, and Mayer (2019)	Declarative and procedural knowledge	No significant difference in laboratory safety knowledge test; desktop group scored higher on first transfer task, but IVR group scored higher on average on second transfer task.
	Makransky, Terkildsen, and Mayer (2019)	Declarative and procedural knowledge	The desktop group had significantly more knowledge in terms of laboratory activities ($d = 1.30$); there was no difference in the transfer outcome.
	Kozhevnikov et al. (2013)	Procedural knowledge	IVR group significantly better at applying the relative velocity concept than the desktop IVR-group ($d = 0.15$).
	Passig et al. (2016)	Procedural knowledge	IVR group with significantly greater learning gains in geometric figure arrangement; also better performance on transfer test.
	Gutiérrez-Maldonado et al. (2015)	Procedural knowledge	Both groups with similar diagnostic skills when interviewing eating disorder patients.
Tablet PC	Moro, Štromberga, and Stirling (2017)	Declarative knowledge	No significant difference in anatomical tests between groups.
	Ventura et al. (2019)	Declarative knowledge	No significant difference between media in terms of recall performance for image items; however, significantly higher recall for items from the IVR environment.
	Negro Cousa et al. (2019)	Declarative knowledge	No significant difference between immersive and non-immersive group in recall of specific items in an image.
AR/Simulator	Huang et al. (2019)	Declarative knowledge	IVR users significantly better at recalling visual information when viewing the solar system ($d = 0.39$); AR group significantly better at recalling information previously presented auditorily ($d = 0.43$).
	Moro, Štromberga, Raikos, and Stirling (2017)	Declarative knowledge	No significant difference in anatomical tests between groups.
	Sportillo et al. (2018)	Procedural knowledge	No technology superiority in taking over the steering function in conditionally automated driving; no significant difference in taking evasive action.
Slide Presentation	Parong and Mayer (2020)	Declarative and procedural knowledge	Slideshow group with better reproductive performance (not significant) and significantly better transfer performance.
	Parong and Mayer (2018)	Declarative and procedural knowledge	Significantly higher scores on the declarative ($d = 1.12$) but not the conceptual knowledge test for slideshow learners.
	Sportillo et al. (2018)	Procedural knowledge	In the first two of three rounds, the IVR group took over the steering function significantly faster during conditionally automated driving ($d = 1.12$); no significant difference in the completion of evasive maneuvers.
	Lai et al. (2019)	Procedural knowledge	VR group significantly better at programming or developing algorithms (computational thinking) ($d = 1.74$).
Textbook	Sundar et al. (2017)	Declarative knowledge	Text group with better recall of two newspaper stories (statistically significant for story 2) ($d = 3.75$).
	Makransky, Borre-Gude, and Mayer (2019) Makransky, Borre-Gude, and Mayer (2019)	Declarative and procedural knowledge	No significant differences in the knowledge test on laboratory safety; significantly better transfer performance of the VR group.
	Chittaro and Buttussi (2015)	Declarative knowledge	After Air Traveler Safety Training: No significant difference on the immediate post-test, but the IVR group scored significantly better on the delayed post-test one week later.
	Villena Taranilla et al. (2019)	Declarative knowledge	Significantly greater learning success in a history lesson for the IVR group compared to the traditional textbook group ($d = 0.56$).
Teacher-based Instruction/ Lecture	Webster (2016)	Declarative knowledge	IVR group with significantly higher test score on erosion damage to aircraft than lecture group ($d = 0.61$).

(*continued on next page*)

Table 4 (*continued*)

Reference Medium	Authors	Dependent Learning Outcome	Findings
	Innocenti et al. (2019)	Declarative knowledge	Significantly better performance of the IVR group compared to teacher-based music instruction ($d = 2.09$).
Practical Exercise	Passig et al. (2016)	Procedural knowledge	No significant difference in learning outcomes between the groups. HMD users performed better on the transfer test.
	Yang et al. (2018)	Procedural knowledge	IVR group scored significantly higher on creative drawing than paper and pencil participants ($d = 0.69$).
	Butt et al. (2018)	Procedural knowledge	For medical catheterization, same results in posttest; however, significantly more practice time in pretraining in the IVR group.
	Ye et al. (2019)	Procedural knowledge	IVR group significantly better at classroom management than microteaching group in terms of error detection ($d = 0.86$) and handling of misbehavior ($d = 0.78$); no significant difference in response time to error detection.

Notice: Effect sizes were calculated and reported if the necessary information was available.

studies. Two measures could not be precisely assigned. Another study contained controversial findings and was therefore not included in the comparison. The results were categorized according to the acquisition of either declarative or procedural knowledge. At first glance, Table 4 presents an overall heterogeneous picture of the effectiveness of IVR as a teaching and learning tool. Both the direction of the effects and the effect sizes presented in this table indicate mixed results. While some of the identified studies did not provide information on effect sizes or relevant parameters for their calculation, effect sizes ranging from $d = 0.15$ to $d = 3.75$ can be reported for a total of 17 studies (see Table 4).

In the following, the collected results are systematically organized according to their relative advantages and disadvantages, based on the number of studies identified. Fig. 3 illustrates that the use of IVR applications tends to outperform analog media or traditional instructional methods, particularly in the acquisition of declarative knowledge. A closer examination of the results presented in Table 4 reveals a clear advantage of IVR, especially when compared to teacher-based instructional approaches, particularly in the acquisition of declarative knowledge (Innocenti et al., 2019; Webster, 2016). Furthermore, when compared to practical exercises, there are overall tentative advantages of IVR, particularly for conveying procedural knowledge, with studies reporting medium to high effects ($0.69 < d < 0.86$; Passing et al., 2016; Yang et al., 2018; Ye et al., 2019). Comparison with textbook use also shows predominantly positive effects of IVR use in both procedural and declarative knowledge domains, accompanied by medium to high effect sizes (Makransky, Borre-Gude, & Mayer, 2019; Chittaro & Buttussi, 2015; Villena Taranilla et al., 2019).

The overall picture of the comparison between the use of HMD-based IVR and other types of electronic media (e.g., video, tablet PC, AR), as shown in Fig. 4, does not reveal a clear difference in terms of declarative and procedural knowledge acquisition. Excluding studies that show no difference, there is a discernible trend toward the superiority of IVR over the selected reference media for procedural knowledge acquisition. The majority of studies in this research area describe this form of knowledge as predominantly advantageous, while only a few cases, measured by the number of studies, are classified as relatively disadvantageous. A closer look reveals that none of the reference media stand out as particularly advantageous or disadvantageous compared to IVR. Instead, a rather mixed picture emerges within the respective reference media groups. Given this overall picture, neither a relative advantage nor a relative disadvantage of IVR compared to other types of electronic media can be determined.

To answer the second research question, further analyses were conducted to compare the effects of HMD-based IVR in terms of relative superiority or inferiority between learning environments with *active* and *passive* learner engagement. For this purpose, a dichotomous distinction was made according to the assumptions of the ICAP framework (Chi & Wylie, 2014). This comparison shows that the majority of studies report no difference in the learning success achieved with IVR compared to other types of media (see Fig. 5). On the one hand, some studies report a relative superiority when using IVR. On the other hand, some studies report a relative inferiority. Concerning the differentiation between declarative and procedural knowledge, a greater number of studies indicate a disadvantage in acquiring declarative knowledge when utilizing HMD-based IVR compared to other media. In summary, the overall depiction of this comparison is somewhat varied, revealing no distinct tendency in terms of relative advantage or disadvantage.

Looking at the number of studies that have examined the effects of IVR on learning, it is clear that the majority of research has focused on learning scenarios with *passive* learner engagement. While these scenarios also show a somewhat mixed picture, the use of IVR in learning environments with *active* learner engagement is on par with or superior to other types of media (see Fig. 6). However, when analyzing for the acquisition of procedural knowledge, most studies show no difference in learning outcomes compared to other types of media. Nevertheless, this type of learning environment shows a clear trend in favor of IVR.

6. Discussion

The analysis conducted in this study aims to provide data-based answers to two research questions related to the use of immersive virtual reality (IVR) in an educational context. On the one hand, the focus is on identifying evidence regarding the relative effectiveness of the use of IVR for the acquisition of declarative and procedural knowledge. On the other hand, the focus is on identifying IVR learning environments that are particularly conducive to learning in comparison.

The results suggest that the use of an IVR-based HMD may offer advantages over other media. Upon closer examination, there is a tendency for IVR to be more advantageous in terms of knowledge acquisition, especially when compared to the use of analog media. Specifically, when compared to teacher-based instruction or the use of textbooks, the results show a clear advantage for IVR, particularly in the acquisition of declarative knowledge. In addition, when compared to hands-on practice, there are overall preliminary advantages of IVR, especially in the acquisition of procedural knowledge, with studies reporting moderate to high effects. However, when IVR usage is compared to other electronic media, the results are more ambiguous. The results indicate that none of the reference media stand out as particularly advantageous or disadvantageous compared to IVR.

In terms of a dichotomous distinction according to the assumptions of the ICAP framework (Chi & Wylie, 2014), there is no clear advantage or disadvantage in the use of IVR for the type of learning environment that is more passive in relation to the role of the learner, especially since the majority of studies report no difference in the learning success achieved with IVR compared to other types of media. The situation is different when we consider learning environments with a more active involvement of the learner, which, compared to the use of other media, is mainly advantageous for the acquisition of both declarative and procedural knowledge. In light of these findings, the use of this technology seems particularly appropriate for action-oriented settings that focus on practical action (i.e., manipulation, generation, dialog). In this respect, the results are in line with the findings of Jensen and Konradsen (2018) and Hamilton et al. (2021).

When interpreting the results, it should be noted that the studies reviewed here were conducted in a wide variety of contexts and subject areas. The associated effect sizes have a high variance and range from negligible to very strong effects. Given the large variance in the magnitude of the effects observed here, the question arises as to why these effects are so different. Possible reasons include the didactic fit and methodological implementation of each setting, as well as the comparison medium used. In addition, the studies reviewed here are conducted with different IVR hardware and software, which show significant differences in their effectiveness in promoting learning. Considering the heterogeneity of the associated didactic settings, substantial variations in the observed effects may be attributed to these differences.

Considering the assumptions proposed by Chi and Wylie (2014) regarding the various forms of learner engagement, the results of this study can also be analyzed using this model. Initially, the classification of learning environments into those with more active or passive engagement is a simplified approach to categorize their general design based on the learners' activity. Nevertheless, the results provide evidence that a rather active learner engagement can enhance knowledge acquisition, including both declarative and procedural knowledge. In light of this observation, the results suggest that IVR scenarios that actively engage learners and thus trigger more complex cognitive processes tend to be more conducive to learning. In this regard the findings are in line with Chi and Wylie's ICAP hypothesis where a higher level of engagement leads to a higher level of cognitive processing and therefore leads to a higher learning outcome. With respect to the importance of the perception of agency in learning with IVR, as described by Makransky and Petersen (2021), the findings are also consistent with the specific assumptions of the CAMIL framework. On the other hand, no general statement can be made about the suitability of this technology for use in more passive scenarios. In this context, this observation is consistent with Chi and Wylie's ICAP hypothesis, where lower engagement is associated with lower cognitive processing, as well as with the assumptions of Makransky and Petersen's model, where a lower perception of agency leads to reduced learning. The findings underline that active learner engagement in IVR-based learning environments tend to be advantageous for knowledge acquisition.

While this study provides insights into the relative advantages and disadvantages of IVR use, there are a number of limitations to this research. Initially, the fundamental categorization into electronic and analog media warrants a critical examination. This is mainly due to the existence of hybrid, mixed forms of media use and the fact that in some cases the underlying methodology cannot be fully separated from the media use. In addition, each use of media tends to follow a specific methodology, a facet not explored in depth in the current study. In this review, effect sizes were manually computed whenever the relevant information was available. However, certain studies included in the analysis do not provide information on effect sizes or the necessary data for such calculations. Consequently, conducting a meaningful assessment of these studies and, by extension, the reported effects proves to be challenging. It is noteworthy that the effect sizes exhibit considerable variance, precluding a comprehensive statement on the overall learning benefits of IVR use. Furthermore, the categorical distinction between declarative and procedural knowledge was based on the intended learning goals of the reported interventions. Given that multiple learning goals were occasionally targeted, this classification may not always be entirely straightforward. Additionally, the simplistic reduction of the ICAP model categories into *active* and *passive* environments represents a further limitation in this context, potentially introducing ambiguities.

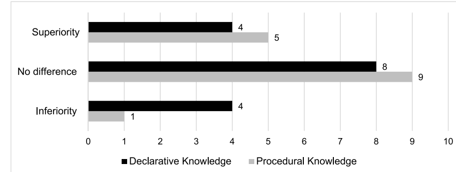

Fig. 4. Effectiveness of IVR on declarative and procedural knowledge compared to other types of electronic media [represented by number of studies].

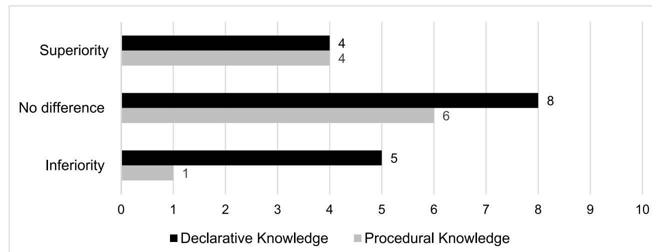

Fig. 5. Effectiveness of IVR learning environments with *passive learner engagement* compared to other types of media [represented by the number of studies].

7. Conclusion

Given the current state of research, the overall aim of this study was to analyse different types of IVR-based learning environments in terms of promoting declarative and procedural knowledge, with particular attention to the associated learner engagement (i.e. passive observation, active manipulation, constructive creation), as envisaged in the ICAP framework (Chi & Wylie, 2014). A systematic review was conducted using the PRISMA method, which identified 30 relevant research papers. The results show positive learning effects with HMD-based IVR compared to other media types. The decision to use IVR in the context of education and training appears to be particularly advantageous compared to other instructional settings that rely on analog media such as text or hands-on exercises, as well as traditional forms of instruction. In addition, the results suggest that IVR is particularly well suited to

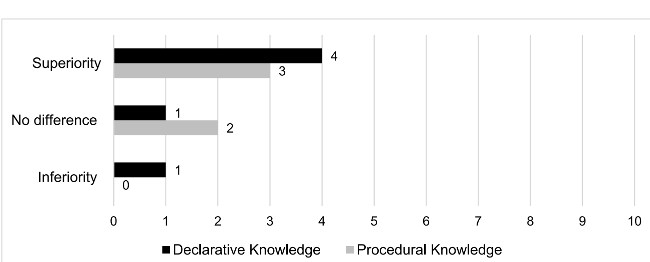

Fig. 3. Effectiveness of IVR on declarative and procedural knowledge compared to analog media [represented by the number of studies].

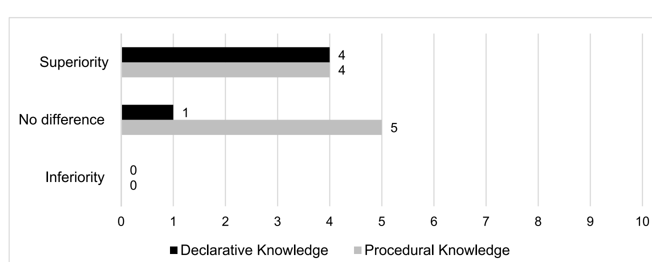

Fig. 6. Effectiveness of IVR learning environments with *active learner engagement* compared to other types of media [represented by the number of studies].

action-oriented environments that prioritize higher levels of learner engagement. It is important to note that the results do not establish IVR as inherently superior to other media or instructional methods. Rather, they highlight the importance of considering both the reference medium chosen and the learning objectives to be achieved.

In summary, the present study suggests that IVR can offer advantages over the use of other media, especially when the learner plays a more active role. These findings have practical implications for teaching and learning practice. When designing appropriate instructional interventions, it is worthwhile to base didactic planning on established models, such as the ICAP framework. In addition, empirical evidence such as Makransky and Petersen's CAMIL should be taken into account. This will increase the likelihood of a fit between the use of IVR and the intended instructional goals.

In reviewing the results, it is clear that the conditions under which IVR provides a learning benefit are still questionable. Further research is needed to identify evidence-based instructional design options for promoting specific learning goals in specific forms of IVR environments. By addressing these considerations, the use of this technology to enable and promote learning can be optimised while overcoming the instructional challenges associated with its implementation. It is essential to acknowledge that IVR is merely a digital tool, and the achievement of learning outcomes depends not only on the medium itself but also on effective instructional design and comprehensive methodological implementation.

CRediT authorship contribution statement

Matthias Conrad: Writing – original draft. **David Kablitz:** Writing – original draft. **Stephan Schumann:** Writing – review & editing.

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.

References

Anderson, J. R. (1976). *Language, memory, and thought*. Lawrence Erlbaum Associates.
Anderson, J. R. (2005). *Cognitive psychology and its implications* (6th ed.) Worth.
Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). *A Taxonomy for learning, teaching and assessing: A revision of bloom's taxonomy of educational objectives*. Longman.
Billingsley, G., Smith, S., Smith, S., & Meritt, J. (2019). A systematic literature review of using immersive virtual reality technology in teacher education. *Journal of Interactive Learning Research, 30*(1), 65–90.
Brogan, D. R., & Kutner, M. H. (1980). Comparative analyses of pretest-posttest research designs. *The American Statistician, 34*(4), 229–232.
Butt, A. L., Kardong-Edgren, S., & Ellertson, A. (2018). Using game-based virtual reality with haptics for skill acquisition. *Clinical Simulation in Nursing, 16*, 25–32.
Buttussi, F., & Chittaro, L. (2018). Effects of different types of virtual reality display on presence and learning in a safety training scenario. *IEEE Transactions on Visualization and Computer Graphics, 24*(2), 1063–1076.
Chang, C.-Y., Sung, H.-Y., Guo, J.-L., Chang, B.-Y., & Kuo, F.-R. (2019). Effects of spherical video-based virtual reality on nursing students' learning performance in childbirth education training. *Interactive Learning Environments*, 1–17.
Chi, M., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist, 49*(4), 219–243. https://doi.org/10.1080/00461520.2014.965823
Chittaro, L., & Buttussi, F. (2015). Assessing knowledge retention of an immersive serious game vs. a traditional education method in aviation safety. *IEEE Transactions on Visualization and Computer Graphics, 21*(4), 529–538.
Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Erlbaum Associates, Publishers.
Concannon, B. J., Esmail, S., & Roduta Roberts, M. (2019). Head-mounted display virtual reality in post-secondary education and skill training: A systematic review. *Frontiers in education, 4*(80), 1–23.
Ferguson, C., van den Broek, E. L., van Oostendorp, H., de Redelijkheid, S., & Giezeman, G.-J. (2020). Virtual reality aids game navigation: Evidence from the hypertext lostness measure. *Cyberpsychology, Behavior, and Social Networking, 23*(9), 635–641.
Gutiérrez-Maldonado, J., Ferrer-García, M., Plasanjuanelo, J., Andrés-Pueyo, A., & Talarn-Caparrós, A. (2015). Virtual reality to train diagnostic skills in eating disorders. Comparison of two low cost systems. In *Annual review of cybertherapy and telemedicine 2015: Virtual reality in healthcare: Medical simulation and experiential interface* (Vol. 13, pp. 75–81).
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. https://doi.org/10.1007/s40692-020-00169-2
Huang, K.-T., Ball, C., Francis, J., Ratan, R., Boumis, J., & Fordham, J. (2019). Augmented versus virtual reality in education: An exploratory study examining science knowledge retention when using augmented reality/virtual reality mobile applications. *Cyberpsychology, Behavior, and Social Networking, 22*(2), 105–110.
Innocenti, E. D., Geronazzo, M., Vescovi, D., Nordahl, R., Serafin, S., Ludovico, L. A., & Avanzini, F. (2019). Mobile virtual reality for musical genre learning in primary education. *Computers & Education, 139*, Article 102117.
Jensen, L., & Konradsen, F. (2018). A review of the use of virtual reality headmounted displays in education and training. *Education and Information Technologies, 23*(4), 1515–1529.
Kozhevnikov, M., Gurlitt, J., & Kozhevnikov, M. (2013). Learning relative motion concepts in immersive and non-immersive virtual environments. *Journal of Science Education and Technology, 22*(6), 952–962.
Lachman, S. J. (1997). Learning is a process: Toward an improved definition of learning. *Journal of Psychology, 131*(5), 477–480.
LaFortune, J., & Macuga, K. L. (2018). Learning movements from a virtual instructor: Effects of spatial orientation, immersion, and expertise. *Journal of Experimental Psychology: Applied, 24*(4), 521.
Lai, Y.-H., Chen, S.-Y., Lai, C.-F., Chang, Y.-C., & Su, Y.-S. (2019). Study on enhancing AIoT computational thinking skills by plot image-based VR. *Interactive Learning Environments*, 1–14.
Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P. A., Clarke, M., Devereaux, P. J., Kleijnen, J., & Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration. *PLoS Medicine, 6*(7), 1–28.
Makransky, G., Andreasen, N. K., Baceviciute, S., & Mayer, R. E. (2020). 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*, 1–17. https://doi.org/10.1037/edu0000473. Advance Online Publication.
Makransky, G., Borre-Gude, S., & Mayer, R. E. (2019). Motivational and cognitive benefits of training in immersive virtual reality based on multiple assessments. *Journal of Computer Assisted Learning, 35*(6), 691–707.
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*(3), 937–958.
Makransky, G., Terkildsen, T. S., & Mayer, R. E. (2019). Adding immersive virtual reality to a science lab simulation causes more presence but less learning. *Learning and Instruction, 60*, 225–236.
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*, 1–17. Article 103603.
Moro, C., Štromberga, Z., Raikos, A., & Stirling, A. (2017). The effectiveness of virtual and augmented reality in health sciences and medical anatomy. *Anatomical Sciences Education, 10*(6), 549–559.
Moro, C., Štromberga, Z., & Stirling, A. (2017). Virtualisation devices for student learning: Comparison between desktop-based (Oculus Rift) and mobile-based (Gear VR) virtual reality in medical and health science education. *Australasian Journal of Educational Technology, 33*(6), 1–10.
Negro Cousa, E., Brivio, E., Serino, S., Heboyan, V., Riva, G., & Leo, G. de (2019). New frontiers for cognitive assessment: An exploratory study of the potentiality of 360 technologies for memory evaluation. *Cyberpsychology, Behavior, and Social Networking, 22*(1), 76–81.
Oh, C., Herrera, F., & Bailenson, J. (2019). The effects of immersion and real-world distractions on virtual social interactions. *Cyberpsychology, Behavior, and Social Networking, 22*(6), 365–372.
Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *International Journal of Surgery, 88*, Article 105906. https://doi.org/10.1136/bmj.n71
Parong, J., & Mayer, R. E. (2018). Learning science in immersive virtual reality. *Journal of Educational Psychology, 110*(6), 785–797.
Parong, J., & Mayer, R. E. (2020). Cognitive and affective processes for learning science in immersive virtual reality. *Journal of Computer Assisted Learning*, 1–16.
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.
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 & Education, 147*, 1–29.
Renkl, A. (2015). Wissenserwerb [knowledge acquisition]. In E. Wild, & J. Möller (Eds.), *Pädagogische psychlogie [pedagogical psychlogy]* (2nd ed., pp. 3–26). Springer.
Rupp, M. A., Odette, K. L., Kozachuk, J., Michaelis, J. R., Smither, J. A., & McConnell, D. S. (2019). Investigating learning outcomes and subjective experiences in 360-degree videos. *Computers & Education, 128*, 256–268.
Sportillo, D., Paljic, A., & Ojeda, L. (2018). Get ready for automated driving using virtual reality. *Accident Analysis & Prevention, 118*, 102–113.
Sundar, S. S., Kang, J., & Oprean, D. (2017). Being there in the midst of the story: How immersive journalism affects our perceptions and cognitions. *Cyberpsychology, Behavior, and Social Networking, 20*(11), 672–682.

Tai, T.-Y., Chen, H. H.-J., & Todd, G. (2020). The impact of a virtual reality app on adolescent EFL learners' vocabulary learning. *Computer Assisted Language Learning*, 1-26. ten Berge, T., & van Hezewijk, R. (1999). Procedural and Declarative Knowledge: An Evolutionary Perspective. *Theory & Psychology, 9*(5), 605–624.

Ventura, S., Brivio, E., Riva, G., & Baños, R. M. (2019). Immersive versus non-immersive experience: Exploring the feasibility of memory assessment through 360° technology. *Frontiers in Psychology, 10*, 1–7. https://doi.org/10.3389/fpsyg.2019.02509

Villena Taranilla, R., Cózar-Gutiérrez, R., González-Calero, J. A., & López Cirugeda, I. (2019). 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*, 1–11.

Webster, R. (2016). Declarative knowledge acquisition in immersive virtual learning environments. *Interactive Learning Environments, 24*(6), 1319–1333.

Yang, X.[X.]., Lin, L., Cheng, P.-Y., Yang, X.[X.]., Ren, Y., & Huang, Y.-M. (2018). Examining creativity through a virtual reality support system. *Educational Technology Research & Development, 66*(5), 1231–1254.

Ye, X., Liu, P.-F., Lee, X.-Z., Zhang, Y.-Q., & Chiu, C.-K. (2019). Classroom misbehaviour management: An SVVR-based training system for preservice teachers. *Interactive Learning Environments*, 1–18.