Lost in Virtual Reality? Cognitive Load in High Immersive VR Environments

Jining Han, Qiyu Zheng, and Yunan Ding
Faculty of Education, Southwest University, Chongqing, China
Email: jininghan@swu.edu.cn, {zheng73, dyn2001}@email.swu.edu.cn

***Abstract*—With the rapid development of high Immersive Virtual Reality (IVR), cognitive load computing has been much studied in its relationship to the design and implementation of high IVR. This study presents a systematic review of the literature on cognitive load in high IVR. A total of 46 empirical studies are selected based on selection criteria. The study reviews the studies and extracts information to answer research questions. The study reports trends in publication, including dates and originating countries of publication, participant selection and sample size, research areas, high IVR devices used, measurement channels and methods, and main results. Findings identify research gaps and suggest reconsidering the relationship between cognitive load and outcomes in high IVR. Furthermore, this study points to several future directions researchers could take to help designers and practitioners to apply high IVR more effectively.**

***Index Terms*—immersive high virtual reality, cognitive load, research gaps**

I. Introduction

With the improvement of Virtual Reality (VR) technology, the development of VR products, and the formation of a relatively new technology, Immersive Virtual Reality (IVR) [1], there are increasing calls to discover the potential implementations of IVR. VR can be described as "the sum of the hardware and software systems that seek to perfect an all-inclusive, sensory illusion of being present in another environment" [2]. IVR integrates immersion as a significant factor into VR, endowing VR with immersion function and allowing users to establish deeper connections with the 3D environment [3]. Some scholars further classify IVR into low IVR and high IVR [4]. Low IVR is based on lower immersion devices, such as 2D computer screens or conventional media, and high IVR usually needs to be based on higher immersion devices, such as Head-Mounted Displays (HMD) and the Cave Automatic Virtual Environment (CAVE). High IVR has three important benefits [3]: one is high physical immersion, with the users employing the HMD or the CAVE to enter the VR environment. The large amount of sensory input in high IVR allows users to visualize and feel integrated into the real environment [5]. The second benefit is high mental immersion, allowing users to generate various emotions through deep participation in the high IVR. High IVR provides users with a real, complex, and rich immersion environment [6], in which they can experience high mental immersion. The third one is high interaction and high imagination: the multiple sensors in high IVR devices can detect the users' input, such as users' eye movements and gestures, and the devices can provide instant responses. Because high IVR could produce many scenes that humans cannot experience personally, such as diving into the deep sea or into space, it can also trigger users' imagination and creativity [7]. High IVR has been widely used in medicine [8], training [9], education [10], simulation of important historical events [1], tourism and other fields. High IVR has its benefits and development potential, but there are some challenges that have concerned scholars: one type of research discusses the shortcomings of high IVR from the perspective of Cognitive Load (CL).

Cognitive Load Theory (CLT) establishes that "all novel information first is processed by a capacity and duration limited working memory and then stored in an unlimited long-term memory for later use" [11]. Our limited working memory can only handle a limited amount of input at a time. When CLT is applied to the high IVR field, it can be understood that a bunch of new information elements are provided in the high IVR environment, which increases users' CL and has a negative impact on users' experiences of high IVR [1], [12]. Some scholars have conducted empirical studies to investigate the impact of high IVR on users' CL and on users' outcomes. A model of four components in learning with media has been established to reveal that CL is directly related to outcomes (see Fig. 1). However, in regard to the impact of high IVR on CL or the impact of high IVR on outcomes, the research conclusions are still inconsistent.

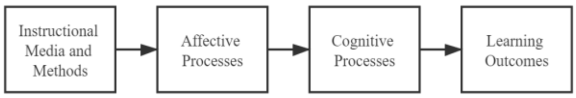

Figure 1. Four components in learning with media [1].

From the perspective of high IVR future development, CLT will play a significant role in users' applications of high IVR devices. Many CLT researchers have also called for studies to explore CL in emerging environments [11]. Therefore, this is a critical synthesis

Manuscript received October 4, 2021; revised October 27, 2021.

doi: 10.12720/jait.12.4.302-310

of the existing studies on CL and learning outcomes in high IVR environments, in order to explore the impact of high IVR on CL and learning outcomes. The study aims to identify research gaps and find future research directions, with the research results providing implications for the design, development, and implementation of high IVR.

II. Objectives and Methodology

This article aims to report a systematic review of eligible empirical studies that examine CL in a high IVR environment. This study summarizes existing evidence concerning CL in high IVR with regard to the trends, benefits, drawbacks, challenges, measurement channels and methods, factors, and outcomes. Based on the results, we identify existing gaps and suggest future research directions.

The following research questions were raised:

1) What is the landscape of research publications on CL in high IVR in the Web of Science and Scopus?
2) What are the main measurement channels and methods used in the selected papers?
3) What are the main results addressed in the selected papers?

A. *Search Strategy*

In this study, we found empirical research articles from peer-reviewed journals published till September 2021. We systematically searched the following electronic databases: Web of Science and Scopus. Search keywords for publications on CL in high IVR include the following groups of words: (a) immersive, immersion; (b) virtual reality, VR; and (c) cognitive load. The three groups of search keywords were combined by means of Boolean AND. Based on the criteria, selected articles were retrieved for full-text review.

B. *Inclusion and Exclusion Criteria*

Before selecting articles to include in the systematic review, we established the following inclusion and exclusion criteria. Articles that fell into the following criteria were excluded: (a) non-English publications, (b) non-peer reviewed articles, and (c) conference proceedings or presentations. Furthermore, the following criteria were applied in the full-text screening process: (e) research must use high IVR devices, (d) research must focus on CL in high IVR settings, (f) research must be empirical studies with reported data and human participants.

C. *Data Analysis*

This study applied the PRISMA guidelines (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) to analyze the data. Fig. 2 presents the selection procedures in terms of the PRISMA statement [13]. 78 articles were retrieved from the two databases, but 17 were duplicated. 61 articles were subject to the full text screening. Each article was assessed and summarized by two researchers. 12 articles were removed because the studies were conducted in low IVR environments, and 3 articles were removed because the studies were without data or human participants. In sum, 46 studies were used to answer the three research questions.

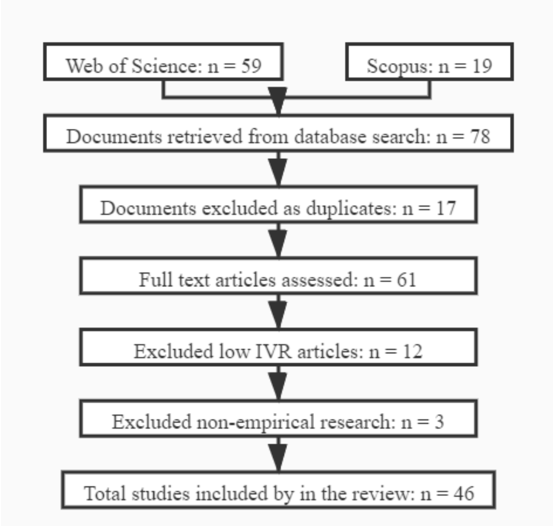

Figure 2. Overview of the search conducted in September 2021 based on PRISMA statement.

III. Results

A. *(RQ1) The Landscape of CL in High IVR Research Publications*

1) Places of study

CL in high IVR environments has been studied in many countries around the world. The selected 46 articles examined CL in high IVR environment in 13 countries (see Table I). USA was the most prolific country with 11 articles, and noticeably all the articles were conducted in university settings. Germany was the second most prolific country. China and Denmark tied as the third most prolific countries, each with six articles. Three articles in China were conducted in mainland China [14]-[16], and three articles took place in China Taiwan [17]-[19]. Other countries contributing to the research of CL in high IVR included Australia, Canada, Croatia, France, New Zealand, Spain, Switzerland, Turkey, and UK. Interestingly, Europe, which contributed 25 articles from 8 countries, was well represented in the publications on CL in high IVR. Most of the studies were conducted in economically developed countries, due to the high cost of high IVR devices, few articles study the usage of high IVR in underdeveloped areas.

TABLE I. A Summary of Countries and Participants of Papers of CL in High IVR

Place of study	Article	Participants
Australia (n=1)	Birbara & Pather (2021) [20]	37 junior and senior university students
Canada (n=2)	1. Carr, Pichora-Fuller, Li, & Campos (2020) [21]	55 participants were 60 years of age or older.
	2. Carr, Pichora-Fuller, Li, Phillips, & Campos (2019) [22]	55 participants were 60 years of age or older.
China (n=6)	1. Han, Diao, Yin, Jin, Kangwa, & Ebohon (2021) [14]	43 undergraduate or graduate students from the subjects of civil

		engineering or construction management.
	2. Liu, Tang, & Wang (2021) [15]	120 college students aged 18-25 years.
	3. Lu, Wu, Cheng, & Lou (2018) (Taiwan) [18]	104 students from the 12th grade of a high school.
	4. Wu, Hsu, Yang, & Chen, Jiang-Jie (2021) (Taiwan) [19]	140 first-year students majoring in landscape design at a university.
	5. Zhao, Lin, Sun, &Liao (2020) [16]	75 college students, mean age is 21.08.
	6. Huang, Luo, Yang, Lu, & Chen (2019) (Taiwan) [17]	77 grade 11 students in a senior high school, aged 15-18 years.
Croatia (n=1)	Lukacevic, Skec, Perisic, Horvat, & Storga (2020) [23]	40 participates aged from 19 to 31 years majored in engineering.
Denmark (n=6)	1. Baceviciute, Terkildsen, and Makransky (2021) [24]	51 participants aged 18-34 years.
	2. Baceviciute, Lucas, Terkildsen & Makransky (2021) [25]	73 participants aged 18-41 years.
	3. Makransky, Terkildsen, & Mayer (2019) [10]	52 university students aged 19-42 years.
	4. Frederiksen, Sorensen, Konge, Svendsen, Sondergaard; Nobel-Jorgensen, Bjerrum, & Andersen (2019) [8]	31 first-year residents without previous laparoscopic experience.
	5. Meyer, Omdahl, & Makransky (2019) [4]	118 participant students at a large European University, aged 18-25 years.
	6. Petersen, Klingenberg, Mayer, & Makransky (2020) [26]	102 seventh and eighth grade students.
France (n=3)	1. Armougum, Gaston-Bellegarde, Joie-La Marle, & Piolino (2020) [27]	124 undergraduate students, of mean age (M=24, SD=1.4) years old.
	2.Ros, Neuwirth, Ng, Debien, Molinari, Gatto, & Lonjon (2021) [28]	89 4th-year medical students aged 19-36 years.
	3.Wang, Chardonnet, & Merienne (2021) [29]	15 subjects including engineers and students (mean age = 23.1) from the university.
Germany (n=7)	1.Schatzschneider, Bruder, & Steinicke (2016) [30]	21 students or members of the department of informatics at a university, aged 18-42 years.
	2. Gloy, Weyhe, Nerenz, Kaluschke, Uslar, Zachmann & Weyhe (2021) [31]	16 high school students.
	3. Ariali & Zinn (2021) [32]	102 participants ages 18-47.
	4. Lerner, Mohr, Schild, Goering, & Luiz (2020) [33]	18 participating emergency physicians, the mean age was 36.6.
	5. Pletz & Zinn (2020) [34]	13 participates aged from 21 to 53 years.
	6. Wechsler, Drescher, Janouch, Haeger, Voelcker-Rehage, & Bock (2018) [35]	124 participants: 63 young (age 20–30 years) and 61 older (age 65–75 years).
	7. Ariali & Zinn (2021) [32]	102 participants, mean age was 27.74.
New Zealand (n=1)	Sridhar, Chan, Chua, Quin, & Nanayakkara (2019) [36]	54 participants. (18 Kindergartners, 36 children between 11–13 years).
Spain (n=1)	Lorenzo-Alvarez, Rudolphi-Solero, Ruiz-Gomez, & Sendra-Portero (2019) [37]	90 radiology students, mean age was 22.6.
Switzerland (n=4)	1. Schlegel, Geering, & Weber (2021) [38]	32 certified nursing professionals, average age is 31.78 years.
	2. Wenk, Penalver-Andres, Buetler, Nef, Muri, & Marchal-Crespo (2021) [39]	20 participates aged from 19 to 42
	3. Moser, Chiquet, Strahm, Mast, & Bergamin (2020) [40]	174 students from a secondary school, a vocational college, and a university, mean age was 18.42.
	4. Wenk, Penalver-Andres, Palma, Buetler, Muri, Nef, & Marchal-Crespo (2019) [41]	20 participates aged 19 -42.
Turkey (n=1)	Mahmoudzadeh, Afacan, & Adi (2021) [42]	30 graduate or Ph.D. students aged 18-38 years.
UK (n=2)	1. Pan, Collingwoode-Williams, Antley, Brenton, Congdon, Drewett, Gillies, Swapp, Pleasence, Fertleman, & Delacroix (2018) [43]	64 general practitioners.
	2. Cassarino, Maisto, Esposito, Guerrero, Chan, & Setti (2019) [44]	38 undergraduate and graduate students at University College Cork, mean age was 22.1.
USA (n=11)	1. Parong & Mayer (2021) [45]	61 participants aged 18-38 years.
	2. Parong, Pollard, Files, Oiknine, Sinatra, Moss, Passaro, & Khooshabeh (2020) [46]	73 university students and community members, aged 19-33 years.
	3. [47], Rottigni, Cavallo, Bailey, Patton, & Marai (2019)	30 university students aged 18-32 years.
	4. Lok, Naik, Whitton, & Brooks (2004) [48]	40 undergraduate students in computer science classes.
	5. Madden, Pandita, Schuldt, Kim, Won, & Holmes (2020) [49]	172 undergraduate students aged 18-24 years.
	6. Bozgeyikli, Bozgeyikli, Raij, Alqasemi, Katkoori, & Dubey (2017) [50]	18 college-aged participants
	7. Parong & Mayer (2021) [1]	80 undergraduate students aged 17-23 years.
	8. Holdnack & Brennan (2021) [51]	84 healthy participants (aged 18-75 years).
	9. Parong, Pollard, Files, Oiknine, Sinatra, Moss, Passaro, & Khooshabeh (2020) [46]	61 university students and community members, aged 19-29 years.
	10. Rao, Smalt, Rodriguez, Wright, Mehta, Brattain, Edwards, Harvey M., Lammert, Heaton, & Quatieri (2020) [52]	8 participants aged 19-37 years.
	11. Shi, Du, & Worthy (2020) [53]	120 undergraduate and graduate students in engineering, ages ranged from 18 to 45 years.

2) Trend of publications

The distribution of the number of articles by year of publication from 2004 to 2021 is shown in Fig. 3. The first article examining CL in high IVR was published back in 2004. This pioneering study used a HMD with a 3rdTech HiBall optical tracker to develop a purely VR environment aimed to examine the effects of handling real objects and self-avatar fidelity on cognitive task performance [48]. After that, few studies paid attention to CL in high IVR until 2016, when one article was published on CL in high IVR. The experiment was conducted in 2015: the study applied Oculus Rift Developers Kit DK2 HMD in the study to examine the effects of CL [30]. It may relate to the release of the Oculus Rift Developers Kit DK2 in July 2014. Fig. 3

shows that the number of CL in high IVR studies gradually increased from 1 paper in 2016 to 16 papers in the first 9 months of 2021, which reflects an increasing interest in investigating CL in high IVR.

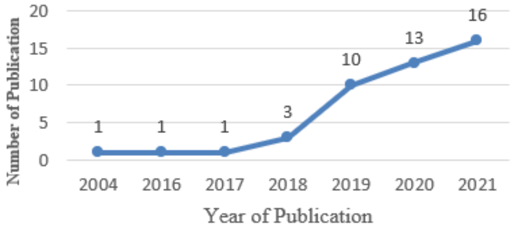

Figure 3. Number of papers on CL in high IVR.

3) Participants and sample size

The sample sizes in Table I ranged from 8 [52] to 174 [40], with an overall total of 3066 participants; the mean number of participants is 67, and the median number is 58. The age ranges of participants in Fig. 4 varied greatly, ranging from 3 to 60 above. The 46 studies involved participants from k-12 schools, colleges, hospitals, and communities. Remarkably, the majority of participants (n=2542, 83%) were aged 18 to 50 years. 18 participants (0.6%) aged 3-5 years, 36 participants (1.2%) aged 10-12 years, 102 participants (3%) aged 12 to 15 years, 197 participants (6%) aged 15-18 years, and 171 participants (5.6%) aged over 60 years.

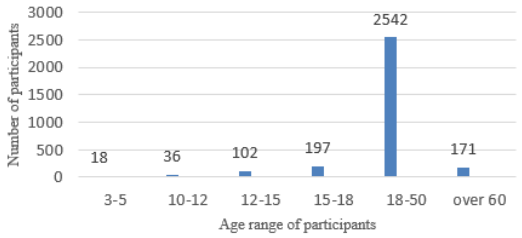

Figure 4. Number of papers on CL in high IVR.

4) Research areas and high IVR devices

CL in high IVR was examined in a variety of contexts. Based on the categories of WOS and Scopus, 34 studies were situated in Technology and Computer Science, 35 studies were in Social Sciences, and 18 studies were in Life Sciences Biomedicine. Medical and Psychology tied as the largest number of studies, each with 13 articles (28%), high IVR technology was used to do these things (simulate medical procedures so medical students could practice the procedures and assess patients' spatial abilities so doctors could diagnose impaired patients) and that the studies aimed to find out how effectively the technology met those goals. Other fields that produced research studies on CL in high IVR included the following: Biology (n=6, 13%), Engineering (n=4, 9%), Architecture (n=2, 4%), Astronomy and Geography (n=2, 4%), Driving (n=2, 4%), Reading (n=1, 2%), Lab (n=1, 2%), History (n=1, 2%), and Game (n=1, 2%) (See Table II). 31 studies applied Head-mounted display (HMD) with computers, of which 24 used HTC Vive, 5 used Oculus Rift, and 2 used HMD with optical trackers (See Table III).

TABLE II. CONTEXT OF USING IVR

Context of using IVR	n	%	Example
Medicine	13	28%	e.g., medical knowledge, 3D skull model, emergency physicians, autism spectrum disorder, CPR, lumbar puncture procedure, anesthesia.
Psychology	13	28%	e.g., spatial ability, head movements, gazing.
Biology	6	13%	e.g., biology education, human cell, the body VR: Cell.
Engineering	4	9%	e.g., lighting system, CAD, pipe maintenance experiment
Architecture	2	4%	e.g., construction site hazard
Astronomy and Geography	2	4%	e.g., moon phases, climate Change
Driving	2	4%	e.g., simulated car driving
Reading	1	2%	e.g., reading texts
Lab	1	2%	e.g., science lab
History	1	2%	e.g., historical events
Game	1	2%	e.g., video games

TABLE III. IVR DEVICES

VR medium	Frequent	Examples
HMD with computers	31	e.g., HTC Vive (24), Oculus Rift (5), HMD with optical tracker (2).
HMD with phones	5	e.g., Google cardboard (2), Samsung Gear (1)
CAVE	7	e.g., virtual theater environment, AVIE 360-degree stereoscopic immersive interactive visualization system, CEAL, Caren
Others	2	e.g., Simulated Car driving

B. (RQ2) Measurement Channels and Methods

1) Textual channel

The textual measurement channel was the most commonly used measurement channel for CL in high IVR (See Fig 5). 39 studies used questionaries, self-reports, or observations to collect data. For example, existing questionnaires such as the NASA TASK Load Index (Nasa-TLX), Multi-dimensional rating-scales, and subjective Workload Assessment Technique (SWAT) were widely used for self-measurement of CL [14], [16], [27], [31], [42].

2) Test channel

Conducting a test was also a major measurement method for assessing CL in high IVR. The test types used in the research on CL in high IVR varied greatly depending on different research purposes, disciplines, research settings and so forth. For example, the Rey Auditory Verbal Learning Test (RAVLT), the Trail Making Test (TMT), the Early Treatment Diabetic Retinopathy Study (ETDRS), and the Timed Up and Go (TUG) test trials were used to evaluate older adults' multi-tasking performance. In addition, comparing the percentage of correct answers [44], pre-tests and post-tests scores [46], memory tests [27], mental rotations test (MRT) tests [23], and response time [36] were also used to evaluate CL in high IVR.

3) Physiological channel

21 studies used physiological assessments to measure participants' CL in high IVR. Those researchers tracked participants' EEG, ECG, EDA, DDT, and eye tracking into their research. Single-channel EEG headsets were

used to measure participants' brainwave activity and mental workload [20]. EDA data could reflect physiological and emotional arousal by calculating average skin conductance level [45]. ECG signal indicates heart rate and Heart Rate Variability (HRV), which reflected physiological arousal associated with emotions [46]. Eye trackers were used to track eyeball movement and pupil size [53] indicating the levels of users' attention.

4) Visual channel

10 studies made visual observations of participants in the high IVR environments. Head movements were captured to evaluate the effects of user perspective on user performance [47]. Participants' movements were tracked by a movement tracking device that shows the pitch, yaw, and roll rotation and the X, Y, and Z position of participants' hands and head: researchers analyzed this movement to assess engagement and outcomes. For example, researchers inferred students' level of engagement from the volume of their movement, including the volume of their button-pressing [49]. Eye gaze and facial expression were commonly used measures to sense affect; for example, a combination of various response, such as particular facial and bodily gestures, can indicate the emotion of anger [36]. A structured qualitative video analysis was able to investigate knowledge transfer, learning sequence, learning challenges, and users' responses [34].

5) Multimodal channel

Fig 5 shows that using multimodal channels was a major measurement method in the studies of CL in high IVR: 31 studies (67%) used multimodal channels, and 15 studies (33%) used a single channel. The integration of textual and test channels was the most widely used multimodal channel in CL in high IVR studies. 12 studies used questionnaires or self-reports combined with tests to measure CL. Physiological measurement shows an increasing trend, with 16 studies applying physiological measurements. In addition, it is worth noting that 10 studies combined three or four measurement channels to evaluate CL.

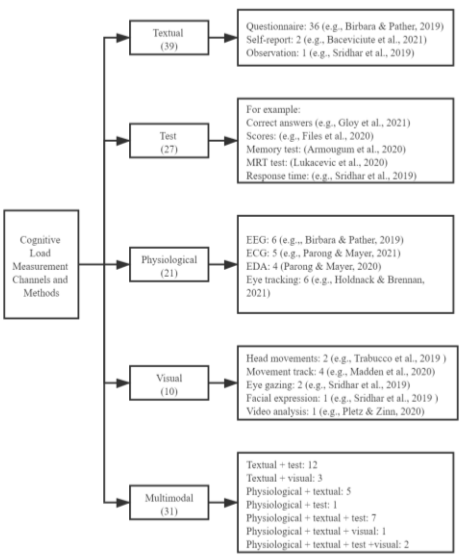

Figure 5. A summary of CL measurement channels and methods.

C. (RQ3) The Main Results Addressed in the Selected Papers

1) Results mixed

As to whether high IVR could increase users' CL, the reported results of the 46 studies are mixed and complicated (See Table IV). 14 studies clearly pointed out that users have higher CL in the high IVR environments and that high IVR reduced users' cognitive engagement resulting in decreased attention to the high IVR environment.

In particular, when texts were displayed in the high IVR, it was difficult for users to read texts in the high IVR environments leading to extremely high CL [24]. However, the results of other studies paint a different picture. Seven studies pointed out that the CL of users was not increased in high IVR because the high IVR environment a) was closer to the real space [48], b) enhanced users' interests and motivations [49], c) increased users' presence: the high IVR environment gave users a sense that they were truly, physically present in virtual environment [46], and d) promoted positive attitude [33], which can reduce distraction and CL [50].

Regarding the subsequent outcomes of CL in high IVR, some studies point to negative effects and outcomes of using high IVR. Through the comparison of test scores or task performances between high IVR environments and other traditional environments, some studies found that users' test scores were lower [45], and the task performances [8] were worse in the high IVR environments. Scholars concluded that high IVR is more suitable for entertainment and should not be used for serious tasks [42]. However, other studies have pointed out that high IVR can improve knowledge transfer, learning efficiency, and learning effects, especially for increasing the proficiency and accuracy of operational skills [23], [28], [50]. In addition, it helps users to enhance technology acceptances [34], self-regulations and attitudes [19].

2) Factors influencing CL in high IVR

Based on the analysis of the 46 studies, we found that research settings are varied, which may lead to differences in research results. We concluded that there are four major factors that influence participants' CL in high IVR: individual differences, prior experiences, task design, and IVR design. Individual differences included users' perspectives and attitudes, users' health conditions [47], gender [32], age [21], usage of working memory [53], and learning styles [17]. Prior experiences involved participants' engagement in pre-assessment tasks designed to familiarize them with the technology [47], pre-training [4], and participants' proficiency levels in using HMD [52]. Task design consisted of task numbers [21], task complexity [52], task types [46], and task time [14]. IVR design included instructional design [34], real environments [44], learning and teaching strategies [16], information formats [53], high fidelity environment [18], and numbers of visual components [21].

TABLE IV. PROS AND CONS OF IVR TOWARD CL AND OUTCOMES, AND THE FACTORS INFLUENCING THE EFFECTS OF IVR

	CL and affective		Outcomes
IVR Negative	1) increase extraneous CL, decrease cognitive engagement (e.g., Birbara & Pather, 2019; Parong & Mayer 2020). (Frequency: 14). 2) need more time to read texts (e.g., Baceviciute *et al.*, 2021; Makransky , Terkildsen, & Mayer, 2019). 3) Distracting (e.g., Parong & Mayer 2020). 4) Latency (e.g., Lok, Naik, Whitton, & Brooks, 2004). 5) improve excessive positive emotions (e.g., Parong & Mayer, 2021).	**IVR Negative**	1) IVR is for enjoyable, not serious tasks (e.g., Mahmoudzadeh, Afacan, & Adi, 2021). 2) No differences in sense of presence (e.g., Lok, Naik, Whitton, & Brooks, 2004). 3) learn less in IVR (e.g., Makransky , Terkildsen, & Mayer, 2019). 4) influence pick-up concerns (Pan *et al.*, 2018). 5) low grades (Parong & Mayer, 2021). 6) worse performances in surgical training (e.g., Frederiksen *et al.*, 2019).
IVR Positive	1) less cognitive processing (e.g., Baceviciute *et al.*, 2021; Moser *et al.*, 2020; Schlegel, Geering, & Weber, 2021). (Frequency: 7). 2) closer to real space (e.g., Lok, Naik, Whitton, & Brooks, 2004). 3) No distraction (e.g., Bozgeyikli *et al.*, 2017). 4) Students preferred the IVR (e.g., Madden *et al.*, 2020). 5) increase motivation and interests (e.g., Wenk *et al.*, 2020). 6) decrease LF/HF ratio indicating lesser mental effort (e.g., Sridhar *et al.*, 2019).	**IVR Positive**	1) promote knowledge transfer (e.g., Baceviciute et al., 2021; Liu, Tang, & Wang, 2021). 2) improve learning efficiency and effects (e.g., Gloy *et al.*, 2021; Lu *et al.*, 2018). 3) better for long-term retention of knowledge (e.g., Gloy *et al.*, 2021; Liu, Tang, & Wang, 2021). 4) promote mental rotation (e.g., Ariali & Zinn, 2021). 5) six vocational skills improved (e.g., Bozgeyikli *et al.*, 2017). 6) skill accuracy (e.g., Lukacevic *et al.*, 2020; Ros *et al.*, 2021). 7) comprehension (e.g., Ros *et al.*, 2021). 8) enhance learning attitudes (e.g., Wu *et al.*, 2021). 9) increase self-regulation (e.g., Wu *et al.*, 2021). 10) increase technology acceptance (e.g., Pletz & Zinn, 2020). 11) CL and learning outcomes may not be positively correlated (Huang *et al.*, 2019).
Factors	**1) Individual differences (e.g., Sridhar *et al.*, 2019):** users' perspectives and attitudes (e.g., Trabucco *et al.*, 2019). users' conditions (e.g., Trabucco *et al.*, 2019). gender (e.g., Ariali & Zinn, 2021). age (e.g., Carr *et al.*, 2020; Huang *et al.*, 2019). usage of working memory (e.g., Patrong *et al.*, 2020; Shi, Du, & Worthy, 2020). learning styles (e.g., Huang *et al.*, 2019). listening and speech-in-noise detection (e.g, Carr *et al.*, 2019). **2) Prior experiences (e.g., e.g., Birbara & Pather, 2019):** exercise requirements (e.g., Trabucco *et al.*, 2019). pre-training (e.g., Meyer, Omdahl, & Makransky, 2019). the level of experimentally (e.g., Armougum *et al.*, 2021; Rao *et al.*, 2020). **3) Tasks design (10, e.g., Pletz & Zinn, 2020):** task numbers (e.g., Carr *et al.*, 2021). task complexity (e.g., Rao *et al.*, 2020). task types (e.g., Files *et al.*, 2020). task time (e.g., Han *et al.*, 2021). **4) IVR design:** instructional design (e.g., Pletz & Zinn, 2020). real environments (e.g., Cassarino *et al.*, 2019). learning and teaching strategies (e.g., Zhao *et al.*, 2020). information formats (e.g., Shi, Du, & Worthy, 2020). high fidelity environment (e.g., Lu *et al.*, 2018). numbers of visual components (e.g., Carr *et al.*, 2020).		

IV. DISCUSSION

A. Research Gaps in Investigating CL in High IVR

Participants' age range: 83% of participants were between 18 and 60 years old. With the development of high IVR, users of all ages should be able to use high IVR products. However, only 0.6% participants in current studies were aged 3-5 years old, 10.2% participants aged 10 to 18 years old, and 5.6% participants aged over 60 years old. For users under the age of 18, their CL and outcomes in high IVR are still not deeply investigated. These participants may have different characteristics: I-generation and digital natives have grown up in technology environment, and they preferred accessing pictures and visuals to reading texts [17].

Places of study and IVR devices: most of the empirical studies were conducted in developed countries. High IVR devices are expensive, so it only makes sense that wealthier countries have greater access to them. It also seems that wealthier countries have larger potential markets for such devices and thus have a greater incentive to do research on them. Students in developed regions and students in under-developed regions may have different technology abilities, technology acceptances, and prior knowledge of using technology devices. The CL in high IVR of the users in underdeveloped areas are still not being widely examined. 31 studies applied HMDs connected with computers into their empirical studies, 24 used HTC Vive, 5 used Oculus Rift, and 2 used HMDs with optical tracker. Only seven studies used CAVE in their research. Future research may need to integrate new and updated high IVR devices into empirical studies. In addition, only 5 studies used HMDs connected with phones (e.g., Google Cardboard and Samsung Gear), which are low-cost high IVR devices. These low-cost devices are more conducive to large-scale use, especially for users in under-developed regions. No research focused on All-in-One high IVR device, such as Oculus Go and HTC Vive Focus; thus, future research may also consider using ALL-in-One devices in the study.

Measurement channels and methods: 67% of the studies used multimodal channels and methods to measure users' CL in high IVR. 33% of the studies used single channel. Single measurement channel and method is not accurate enough: multimodal channels could overcome the limitations of each single channel and further improve the accuracy of the measurement [54]. When using multimodal channels to measure CL in high IVR, studies might face challenges in managing huge amount of data, ethical issues, technical difficulties, interpreting data, and costs.

B. Reconsider the Relationships between CL and Outcomes

The model of four components in learning with media [1] establishes that CL is a mediator between instructional media and outcomes. The current study believes that CL is not necessarily positively correlated with outcomes in high IVR: increased CL does not directly lead to negative outcomes, and it is not one size fits all. This finding resonates with some studies, which suggested that CL may not positively correlate to learning outcomes, as shown in Fig. 6. Future studies need to reconsider the relationships among high IVR, CL, and outcomes, taking into account factors including individual differences, prior experiences, task design, and high IVR design.

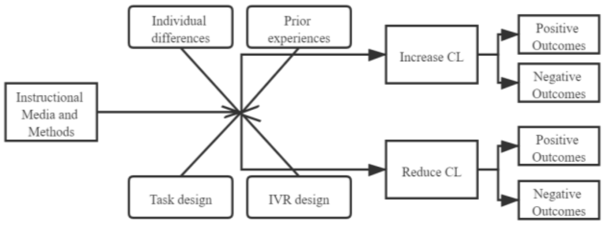

Figure 6. A revised model of variables in learning with media.

C. Considering the Influencing Factors When Designing Tasks and Components in High IVR

Through adjusting tasks, designing components, simulation effects, pre-training, and teaching strategies, CL can be effectively changed, and further positive outcomes can be produced. For example, when developing high IVR products, designers need to reduce the number of tasks, reduce the complexity of a single task, and set an appropriate task time [14], [21], [52]. At the same time, reducing the number of texts, pictures, and videos that appear at one time can also influence users' CL. Enhancing the effect of simulation, making the high IVR environment more realistic and authentic can promote users' interest, motivation, and engagement [18], [21], [44]. In addition, the instructional strategies are also crucial. Through clear and reasonable instructions, users can better involve themselves in high IVR [34]. Pre-training is an effective method, which is consistent with the Pre-training Principle in Multimedia Learning [12]; especially when users encounter a new high IVR device or face a novel task in high IVR, pre-training can help users more quickly adapt to the new tasks and environments [4].

V. Conclusion

This study has reported an overview of cognitive load computing in high IVR. The review covers the relevant studies indexed in Web of Sciences and Scopus. The final 46 papers are reviewed to answer three research questions. Eight findings are reported (1) most of the studies were conducted in economically developed countries; (2) the number of studies investigating CL in high IVR has significantly increased in the most recent three years; (3) the majority of research participants were aged 18 to 50 years, and the sample sizes ranged from 8 to 174; (4) most of the studies applied HMD with computers; (5) most of the studies were conducted in Medical and Psychology; (6) using multimodal channels was a major measurement method; (7) CL may not positively correlate to learning outcomes in high IVR environments; (8) designers and developers need to considering the influencing factors when designing tasks and components in high IVR. The study hopes that the findings can provide researchers, designers, developers, and practitioners new insights into the more effective developments and applications of high IVR.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Jining Han, Qiyu Zheng, and Yunan Ding conducted the research and collected data; all authors analyzed the data and wrote the paper; all authors had approved the version.

Acknowledgement

This study was financially supported by the "the Fundamental Research Funds for the Central Universities" under Grant SWU1909764 and the "Southwest University Education and Teaching Reform Research Project" under Grant 2020JY038.

References

[1] J. Parong and R. E. Mayer, "Learning about history in immersive virtual reality: Does immersion facilitate learning?" *Educ. Technol. Res. Dev.*, vol. 69, no. 3, pp. 1433-1451, Jun. 2021.

[2] F. Biocca and B. Delaney, *Immersive Virtual Reality Technology*, Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc, 1995.

[3] M. Mulders, J. Buchner, and M. Kerres, "A framework for the use of immersive virtual reality in learning environments," *Int. J. Emerg. Technol. Learn.*, vol. 15, no. 24, pp. 208-224, 2020.

[4] O. A. Meyer, M. K. Omdahl, and G. Makransky, "Investigating the effect of pre-training when learning through immersive virtual reality and video: A media and methods experiment," *Comput. Educ.*, vol. 140, Oct. 2019.

[5] W. Winn, H. Hoffman, A. Hollander, K. Osberg, and H. Rose, "The effect of student construction of virtual environments on the performance of high- and low-ability students," in *Proc. Annual Meeting of the American Educational Research Association*, Jan. 1997.

[6] J. Martín-Gutiérrez, C. E. Mora, B. Añorbe-Díaz, and A. González-Marrero, "Virtual technologies trends in education," *Eurasia J. Math. Sci. Technol. Educ.*, vol. 13, no. 2, pp. 469-486, 2017.

[7] D. Jonassen, "Transforming learning with technology: Beyond modernism and post-modernism or whoever controls the technology creates the reality," *Educ. Technol.*, vol. 40, Jan. 2000.

[8] J. G. Frederiksen, *et al.*, "Cognitive load and performance in immersive virtual reality versus conventional virtual reality simulation training of laparoscopic surgery: A randomized trial," *Surg. Endosc.*, vol. 34, no. 3, pp. 1244-1252, Mar. 2020.

[9] A. D. Kaplan, J. Cruit, M. Endsley, S. M. Beers, B. D. Sawyer, and P. A. Hancock, "The effects of virtual reality, augmented reality, and mixed reality as training enhancement methods: A meta-analysis," *Hum. Factors*, vol. 63, no. 4, pp. 706-726, Jun. 2021.

[10] G. Makransky, T. S. Terkildsen, and R. E. Mayer, "Adding immersive virtual reality to a science lab simulation causes more

presence but less learning," *Learn. Instr.*, vol. 60, pp. 225-236, Apr. 2019.

[11] J. Sweller, J. J. G. V. Merriënboer, and F. Paas, "Cognitive architecture and instructional design: 20 years later," *Educ. Psychol. Rev.*, vol. 31, no. 2, pp. 261-292, 2019.

[12] R. E. Mayer, *Multimedia Learning*, 3rd ed., Cambridge: Cambridge University Press, 2020.

[13] D. Moher, A. Liberati, J. Tetzlaff, and D. G. Altman, "Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement," *BMJ*, vol. 339, p. b2535, Jul. 2009.

[14] Y. Han, Y. Diao, Z. Yin, R. Jin, J. Kangwa, and O. J. Ebohon, "Immersive technology-driven investigations on influence factors of cognitive load incurred in construction site hazard recognition, analysis and decision making," *Adv. Eng. Informatics*, vol. 48, Apr. 2021.

[15] Q. Liu, Q. Tang, and Y. Wang, "The effects of pretraining intervention in immersive embodied virtual reality cardiopulmonary resuscitation training," *Behav. Inf. Technol.*, 2021.

[16] J. Zhao, L. Lin, J. Sun, and Y. Liao, "Using the summarizing strategy to engage learners: Empirical evidence in an immersive virtual reality environment," *Asia-Pacific Educ. Res.*, vol. 29, no. 5, pp. 473-482, Oct. 2020.

[17] C. L. Huang, Y. F. Luo, S. C. Yang, C. M. Lu, and A. S. Chen, "Influence of students' learning style, sense of presence, and cognitive load on learning outcomes in an immersive virtual reality learning environment," *J. Educ. Comput. Res.*, vol. 58, no. 3, pp. 596-615, Jun. 2020.

[18] C. M. Lu, P. L. Wu, Y. M. Cheng, and S. J. Lou, "Effects on patterns of learning-support design in immersive virtual reality system," *J. Inf. Hiding Multim. Signal Process*., vol. 9, no. 5, 2018.

[19] W. L. Wu, Y. Hsu, Q. F. Yang, and J. J. Chen, "A spherical video-based immersive virtual reality learning system to support landscape architecture students' learning performance during the COVID-19 era," *Land*, vol. 10, no. 6, 2021.

[20] N. S. Birbara and N. Pather, "Instructional design of virtual *learning* resources for anatomy education," in *Instructional Design of Virtual Learning Resources for Anatomy Education*, Springer International Publishing, 2021, pp. 75-110.

[21] S. Carr, M. K. Pichora-Fuller, K. Z. H. Li, and J. L. Campos, "Effects of age on listening and postural control during realistic multi-tasking conditions," *Hum. Mov. Sci.*, vol. 73, Oct. 2020.

[22] S. Carr, M. K. Pichora-Fuller, K. Z. H. Li, N. Phillips, and J. L. Campos, "Multisensory, multi-tasking performance of older adults with and without subjective cognitive decline," *Multisens. Res.*, vol. 32, no. 8, pp. 797-829, 2019.

[23] F. Lukačević, S. Škec, M. M. Perišić, N. Horvat, and M. Štorga, "Spatial perception of 3D CAD model dimensions and affordances in virtual environments," *IEEE Access*, vol. 8, pp. 174587-174604, 2020.

[24] S. Baceviciute, T. Terkildsen, and G. Makransky, "Remediating learning from non-immersive to immersive media: Using EEG to investigate the effects of environmental embeddedness on reading in Virtual Reality," *Comput. Educ.*, vol. 164, Apr. 2021.

[25] S. Baceviciute, G. Lucas, T. Terkildsen, and G. Makransky, "Investigating the redundancy principle in immersive virtual reality environments: An eye-tracking and EEG study," *J. Comput. Assist. Learn.*, Aug. 2021.

[26] G. B. Petersen, S. Klingenberg, R. E. Mayer, and G. Makransky, "The virtual field trip: Investigating how to optimize immersive virtual learning in climate change education," *Br. J. Educ. Technol.*, vol. 51, no. 6, pp. 2098-2114, Nov. 2020.

[27] A. Armougum, A. Gaston-Bellegarde, C. J. L. Marle, and P. Piolino, "Expertise reversal effect: Cost of generating new schemas," *Comput. Human Behav.*, vol. 111, Oct. 2020.

[28] M. Ros, *et al.*, "The effects of an Immersive Virtual Reality Application in First Person Point-of-View (IVRA-FPV) on the learning and generalized performance of a lumbar puncture medical procedure," *Educ. Technol. Res. Dev.*, vol. 69, no. 3, pp. 1529-1556, Jun. 2021.

[29] Y. Wang, J. R. Chardonnet, and F. Merienne, "Enhanced cognitive workload evaluation in 3D immersive environments with TOPSIS model," *International Journal of Human-Computer Studies*, vol. 147, p. 102572, 2021.

[30] C. Schatzschneider, G. Bruder, and F. Steinicke, "Who turned the clock? Effects of manipulated zeitgebers, cognitive load and immersion on time estimation," *IEEE Trans. Vis. Comput. Graph.*, vol. 22, no. 4, pp. 1387-1395, Apr. 2016.

[31] K. Gloy, *et al.*, "Immersive anatomy atlas: Learning factual medical knowledge in a virtual reality environment," *Anat. Sci. Educ.*, Jun. 2021.

[32] S. Ariali and B. Zinn, "Adaptive training of the mental rotation ability in an immersive virtual environment," *Int. J. Emerg. Technol. Learn.*, vol. 16, no. 9, pp. 20-39, 2021.

[33] D. Lerner, S. Mohr, J. Schild, M. Göring, and T. Luiz, "An immersive multi-user virtual reality for emergency simulation training: Usability study," *JMIR Serious Games*, vol. 8, no. 3, Jul. 2020.

[34] C. Pletz and B. Zinn, "Evaluation of an immersive virtual learning environment for operator training in mechanical and plant engineering using video analysis," *Br. J. Educ. Technol.*, vol. 51, no. 6, pp. 2159-2179, Nov. 2020.

[35] K. Wechsler, U. Drescher, C. Janouch, M. Haeger, C. Voelcker-Rehage, and O. Bock, "Multitasking during simulated car driving: A comparison of young and older persons," *Front. Psychol.*, vol. 9, Jun. 2018.

[36] P. K. Sridhar, S. W. T. Chan, Y. Chua, Y. W. Quin, and S. Nanayakkara, "Going beyond performance scores: Understanding cognitive-affective states in Kindergarteners and application of framework in classrooms," *Int. J. Child-Computer Interact.*, vol. 21, pp. 37-53, Sep. 2019.

[37] R. Lorenzo-Alvarez, T. Rudolphi-Solero, M. J. Ruiz-Gomez, and F. Sendra-Portero, "Game-Based learning in virtual worlds: A multiuser online game for medical undergraduate radiology education within second life," *Anat. Sci. Educ.*, vol. 13, no. 5, pp. 602-617, Sep. 2020.

[38] C. Schlegel, A. Geering, and U. Weber, "Learning in virtual space: An intergenerational pilot project," *GMS Z. Med. Ausbild.*, vol. 38, pp. 1-14, Jan. 2021.

[39] N. Wenk, J. Penalver-Andres, K. A. Buetler, T. Nef, R. M. Müri, and L. Marchal-Crespo, "Effect of immersive visualization technologies on cognitive load, motivation, usability, and embodiment," *Virtual Real.*, 2021.

[40] I. Moser, S. Chiquet, S. K. Strahm, F. W. Mast, and P. Bergamin, "Group decision-making in multi-user immersive virtual reality," *Cyberpsychology, Behav. Soc. Netw.*, vol. 23, no. 12, pp. 846-853, Dec. 2020.

[41] N. Wenk, *et al.*, "Reaching in several realities: Motor and cognitive benefits of different visualization technologies," in *Proc. IEEE 16th International Conference on Rehabilitation Robotics*, 2019.

[42] P. Mahmoudzadeh, Y. Afacan, and M. N. Adi, "Analyzing occupants' control over lighting systems in office settings using immersive virtual environments," *Build. Environ.*, vol. 196, Jun. 2021.

[43] X. Pan, *et al.*, "A study of professional awareness using immersive virtual reality: The responses of general practitioners to child safeguarding concerns," *Front. Robot. AI*, vol. 5, 2018.

[44] M. Cassarino, M. Maisto, Y. Esposito, D. Guerrero, J. S. Chan, and A. Setti, "Testing attention restoration in a virtual reality driving simulator," *Front. Psychol.*, vol. 10, Feb. 2019.

[45] J. Parong and R. E. Mayer, "Cognitive and affective processes for learning science in immersive virtual reality," *J. Comput. Assist. Learn.*, vol. 37, no. 1, pp. 226-241, Feb. 2021.

[46] J. Parong, *et al.*, "The mediating role of presence differs across types of spatial learning in immersive technologies," *Comput. Human Behav.*, vol. 107, Jun. 2020.

[47] J. T. Trabucco, A. Rottigni, M. Cavallo, D. Bailey, J. Patton, and G. E. Marai, "User perspective and higher cognitive task-loads influence movement and performance in immersive training environments," *BMC Biomed. Eng.*, vol. 1, no. 1, Dec. 2019.

[48] M. Whitton and F. P. Brooks. (2003). Effects of handling real objects and self-avatar fidelity on cognitive task performance and sense of presence in virtual environments. [Online]. Available: http://direct.mit.edu/pvar/article-pdf/12/6/615/1624015/105474603322955914.pdf

[49] J. Madden, S. Pandita, J. P. Schuldt, B. Kim, A. S. Won, and N. G. Holmes, "Ready student one: Exploring the predictors of student learning in virtual reality," *PLoS One*, vol. 15, no. 3, 2020.

[50] L. Bozgeyikli, E. Bozgeyikli, A. Raij, R. Alqasemi, S. Katkoori, and R. Dubey, "Vocational rehabilitation of individuals with

autism spectrum disorder with virtual reality," *ACM Trans. Access. Comput.*, vol. 10, no. 2, Apr. 2017.
[51] J. Holdnack and P. Brennan, "Usability and effectiveness of immersive virtual grocery shopping to assess cognitive fatigue in healthy controls: Protocol for a randomized control trial (Preprint)," *JMIR Res. Protoc.*, vol. 10, Feb. 2021.
[52] H. M. Rao, *et al.*, "Predicting cognitive load and operational performance in a simulated marksmanship task," *Front. Hum. Neurosci.*, vol. 14, Jul. 2020.
[53] Y. Shi, J. Du, and D. A. Worthy, "The impact of engineering information formats on learning and execution of construction operations: A virtual reality pipe maintenance experiment," *Autom. Constr.*, vol. 119, Nov. 2020.
[54] E. Yadegaridehkordi, N. F. B. M. Noor, M. N. B. Ayub, H. B. Affal, and N. B. Hussin, "Affective computing in education: A systematic review and future research," *Comput. Educ.*, vol. 142, p. 103649, 2019.

Jining Han is an assistant professor of educational technology in the Faculty of Education at Southwest University. He obtained his Ph.D. in second language acquisition and instructional technology at the University of South Florida, and he obtained his M.A. in pedagogy at Arizona State University. His main research interests include computer-assisted learning, online education, and STEM education.

Qiyu Zheng is a graduate student in the Faculty of Education at Southwest University. She serves as a research assistant at the Research Center of Smart Education Project at Chongqing Universities. Her main research interests include virtual reality in education and computer-assisted learning.

Yunan Ding is an undergraduate student in the Faculty of Education at Southwest University. She serves as a research assistant at the research project of "Smart Education: Students' Feedback, Views, and Attitudes". Her main research interests include learning management system and deep learning.