IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 4 October 2024, accepted 26 October 2024, date of publication 1 November 2024, date of current version 23 December 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3489233

SURVEY

A Systematic Review of Immersive Virtual Reality in STEM Education: Advantages and Disadvantages on Learning and User Experience

PEDRO ACEVEDO[1], ALEJANDRA J. MAGANA[2], BEDRICH BENES[3], (Senior Member, IEEE), AND CHRISTOS MOUSAS[1], (Member, IEEE)
[1]Department of Computer Graphics Technology, Purdue University, West Lafayette, IN 47907, USA
[2]Department of Computer and Information Technology, Purdue University, West Lafayette, IN 47907, USA
[3]Department of Computer Science, Purdue University, West Lafayette, IN 47907, USA

Corresponding author: Pedro Acevedo (paceved@purdue.edu)

This work was supported in part by U.S. National Science Foundation under Award 2309564, Award 2417510, and Award 2201019.

ABSTRACT Immersive virtual reality (IVR) is widely recognized as an innovative technology that enhances learning experiences. IVR affordances could serve as a learning medium, and instructors and designers should consider using available tools to teach practical STEM (science, technology, engineering, and mathematics) subjects, incorporating the latest practices and recommendations. Due to innovations in virtual reality hardware and affordability in recent years, there is a higher chance that IVR will be adopted in educational settings. Thus, designers and developers must stay informed about current IVR trends and the aspects addressed and recommended for instructing STEM concepts. To target this issue, we systematically reviewed the design and development of customized IVR experiences and their multiple effects on STEM learning settings. In this systematic review, we identified several advantages and disadvantages reported in 30 papers based on user studies in higher education scenarios. We proposed a conceptual framework to categorize the design and features of the IVR tools discussed in the papers, based on their levels of embodiment, immersion, and the type of learning they facilitate. IVR can provide an intuitive and practical immersive learning experience, reporting multiple results on motivation, engagement, usability, and learning performance. However, we should consider the features that could directly affect the user experience and learning outcomes. We recommend that designers and developers explore developing customized IVR experiences tailored for STEM learning, especially for invisible and complex subjects where 3D visualization could benefit students.

INDEX TERMS Immersive virtual reality, STEM, learning design, learning outcomes, user experience.

I. INTRODUCTION

Science, technology, engineering, and mathematics (STEM) education has been adopted as a priority for students worldwide [1]. The National Science Foundation report on science and engineering indicators reflects the growth of the STEM labor force and education [2]. The report shows that the investment of the United States in STEM research and development has increased significantly in recent years. Many other countries are taking steps to ensure their students have access to high-quality STEM education. For example, in 2015, Australia launched its *“National STEM School Education Strategy 2016-2026,”* focusing on foundation skills, developing mathematical, scientific, and digital literacy, and promoting problem-solving. This program has reported promising outcomes around the proposed initiatives [3]. Likewise, Finland launched the *“National STEM Strategy and Action Plan in 2021”* to boost STEM education, research,

The associate editor coordinating the review of this manuscript and approving it for publication was Tai-Hoon Kim.

and careers [4]. With the increasing demand, different trainers and professionals should be able to instruct all these concepts and materials at multiple educational levels. Then, a need emerged for qualified professionals to design and teach materials on STEM effectively in different educational institutions [5]. Teaching methods must be tailored to engage young learners constructively. With technological advancements and the widespread use of mobile devices, there are challenges and opportunities to innovate traditional learning methods. Among these, virtual reality (VR) has emerged as a promising educational tool to enhance learning in STEM disciplines [6].

Immersive virtual reality (IVR) can provide the possibility to explore different concepts and manipulate reality within a simulated environment. Understanding affordance as a characteristic of the environment that, when perceived, affords an agent the opportunity for action based on the agent's capabilities [7], IVR has two main affordances: the sense of presence and agency. The sense of presence, defined as the *"sense of being there"* [8], immerses users in a 3D simulated world. Agency establishes the sense of ownership over one's actions within the environment and enables potential interactive learning scenarios [9]. Through IVR, users can be immersed in a 3D simulation, for example, being teleported on a historical architecture to explore unknown places [10] or in front of a hazardous procedure in a construction environment [11]. Modern VR head-mounted displays (HMDs), such as HTC VIVE or Meta Quest, allow users to experience high immersion. Unlike mobile VR (e.g., Google Cardboard or Samsung Gear VR), HTC VIVE or Meta Quest can be considered high-end HMDs due to the inclusion of various accessories and features designed to control the user's immersion. High-end HMDs enable users to move and interact with the virtual environment in six degrees of freedom (6 DOF), consisting of three translational movements (left/right, up/down, and forward/backward) and three rotational movements (roll, pitch, and yaw) [12]. By leveraging these hardware capabilities, designed experiences across different levels of immersion can transform IVR into a truly unique learning experience, especially for immersive learning. Immersive learning occurs when a student experiences a technological, narrative, and challenge-based state of deep mental involvement within a simulated reality isolated from the real world [13].

IVR benefits in education and STEM have been explored, and the conclusions are not ready to be confirmed. Findings about whether the usage of IVR in education is practical or necessary have been contradictory. Researchers have highlighted the need to investigate IVR's potential advantages in STEM, particularly in higher education, as previous analyses have shown more significant effects on learning outcomes in K-12 scenarios compared to higher education settings [14], [15]. Despite this, a possible adoption of the technology could happen, and instructors must develop competencies to integrate IVR into the educational curriculum effectively. Designers and developers must understand how IVR can be implemented, including developer tools, recommended frameworks, possible learning approaches, devices, and expected learning outcomes. Moreover, unlike available third-party VR applications, customized IVR solutions can enable instructors to create experiences tailored to the specific needs of their students and educational contexts. Therefore, it is crucial for designers and developers to be aware of current trends in IVR and the best practices for implementing these experiences to teach STEM concepts effectively. To address this issue, we systematically reviewed the design and development of IVR experiences and their different effects (advantages and disadvantages) on learning and user experience in higher education.

We divide this review paper into the following sections. In Section II, we discuss related reviews, findings, limitations, and opportunities. In Section III, we introduce the conceptual framework for categorizing and understanding the reviewed papers and their findings. In Section IV, we present the methodology used in the conducted review. In Section V, we compile all results and metadata from the systematic review and categorize the papers based on the proposed conceptual framework. In Section VI, we delimit the discussion around the findings and address the research questions. Finally, we present our conclusions in Section VII.

II. RELATED WORKS

Researchers have recently shown interest in using IVR for education and training. They have conducted several reviews and surveys to establish the current state and identify opportunities and gaps [16], [17], [18]. In education, these surveys have disclosed various relationships and recommendations regarding IVR usage in the classroom. IVR has shown advancements in learning, evidenced by students' positive attitudes, engagement, learning outcomes, and performance across different STEM fields [17]. However, few authors have grounded their IVR designs or activities in theoretical learning frameworks or evaluated knowledge acquisition and skill development [14], [16]. Radianti et al. [16] examined VR in education concerning learning content, VR design elements, and learning theories as a foundation for successful VR-based learning. However, their study focused solely on papers published between 2016 and 2018, suggesting that any conclusions regarding the adoption and utilization of IVR may have changed since then. Notably, their findings showed that most papers (68%) did not include learning theories as the foundation of their VR design. Won et al. [19] classified the design elements used and their level of integration for IVR in education, identifying patterns in the use of VR affordances in these studies. They reviewed 219 studies, categorizing design features based on learning tasks and context. However, the authors did not provide details on how the reviewed experiences were implemented or whether the authors of those papers developed a customized IVR application. Lui et al. [14] mapped different design approaches for IVR experiences in higher education based on learning theories. They outlined various strategies for designing IVR educational experiences, considering factors affecting learning outcomes and cognitive load. However, their review did not provide details on how the IVR experiences reviewed were leveraged or which methodologies those studies followed. Additionally, their review focused narrowly on science-related topics, disregarding other STEM concepts.

This review focuses on the design and development of IVR learning experiences, particularly their implemented features (e.g., haptic feedback, realistic hands, virtual avatars) and their effects, including advantages and disadvantages, on learning outcomes (e.g., cognitive load, motivation) and user experience (e.g., usability, presence) in higher education STEM-related concepts. We focused this review exclusively on STEM concepts due to their relevance to technological advancements. Additionally, technologies such as VR have the potential to enhance the instruction and learning of STEM topics due to their affordances. Lui et al. [20] noted mixed results on the impact of IVR on learning outcomes in higher education, contrasting with the benefits observed in K-12 and high school education. Therefore, this review centers on higher education to provide insights into how customized IVR experiences can enhance student learning in these settings. Previous reviews have reported findings around the design choices and the delimited VR features to enhance learning in science fields [14], [16], as well as other topics, including training and health [19]. Some reviews have also reported incorporating learning theories [14] and the focus of the state-of-the-art literature on learning through VR. However, an analysis of the development of these applications, specifically focusing on high-end HMDs, has not been detailed. Finally, the advantages and disadvantages of the designed IVR experiences for STEM concepts in learning and usability warrant further discussion.

III. CONCEPTUAL FRAMEWORK: EMBODIMENT LEVEL, IMMERSION, AND FORMATIVE ASSESSMENT

Three taxonomies or frameworks inform our conceptual framework: taxonomy of embodiment [21], framework for immersion [22], [23], and framework for learning assessment [24]. We summarized the framework in Figure 1.

Johnson-Glenberg and Megowan-Romanowicz [21] proposed a taxonomy of embodiment based on three factors: *sensorimotor engagement* (SE), *gestural congruency* (GC), and *immersion* (IM). *Sensorimotor engagement* measures physical involvement in learning, *gestural congruency* assesses how well gestures match learning content, and *immersion* refers to the learner's feeling of being inside the experience. They propose four degrees of embodiment:

- **First-degree:** Little to no sensorimotor engagement, gestural congruency, and immersion;
- **Second-degree:** Low to moderate sensorimotor engagement, some gestural congruency, and moderate immersion;
- **Third-degree:** Moderate to high sensorimotor engagement, high gestural congruency, and high immersion; and
- **Fourth-degree:** High sensorimotor engagement, high gestural congruency, and high immersion.

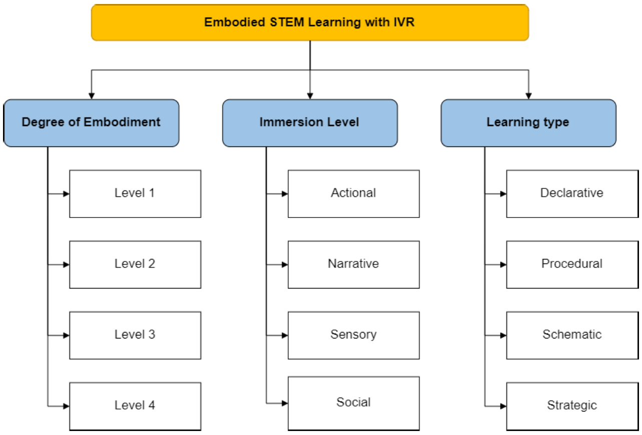

FIGURE 1. Conceptual framework to guide the analysis of the paper composed by the degree of embodiment [21], type of learning [24], and type of immersion [22].

Embodied learning is grounded in the theory that using bodily actions and interactions in VR can enhance learning [25]. The degree of embodiment has been used to clarify how the designed IVR environment is composed in relation to the instructional topic and the integration of VR affordances. Johnson-Glenberg et al. [26] evaluated different degrees of embodiment by comparing PC and VR with varying interaction levels. They found that the low-embodied VR group performed significantly worse than the high-embodied VR group, primarily due to the lower agency offered in the low-embodied VR. Chatain et al. [27] conducted a study on varying degrees of embodiment in a math lesson based on the taxonomy but did not find differences in learning outcomes across different degrees of embodiment. These contrasting results provide an opportunity to evaluate how embodied IVR applications can be designed to promote active learning.

Shavelson et al. [24] outlined a conceptual framework for the embedded formative assessment. The authors define an embedded assessment as a formative evaluation to diagnose students' understanding and actions during a lesson, thereby enhancing their learning achievement, motivation, and conceptual change. The framework is divided into four types of knowledge and reasoning necessary for achievements in learning: declarative (knowing facts and concepts), procedural (knowing steps to accomplish tasks), schematic (connecting and explaining knowledge), and strategic (knowing when and where to apply knowledge). The relevance of formative assessment in STEM has been explored [28], [29]. When designing a lesson, regardless of the medium, different objectives and expected outcomes are defined. The lesson should provide learners with sufficient tools and scaffolding to construct new knowledge, making assessment a central element of a learning environment aimed at improving students' learning by tracking their progress [30].

Formative assessment assists the learning process and is often referred to as *"assessment for learning."* It involves seeking and interpreting evidence to help learners and their teachers determine the learners' current level of understanding, identify their learning goals, and decide on the best strategies to achieve those goals [29]. This approach can also be applied to lessons using immersive technology, such as VR [31], [32]. Given the importance of formative assessment in learning, we consider it necessary to include an overview of how the authors of customized IVR experiences implemented their assessments and the objectives they aimed to achieve.

Dede et al. [22], [23] described the VR scenario's main characteristics of immersion for learning, including *sensory*, *actional*, *narrative*, and *social* features. These features refer to how the design of immersive experiences is managed to leverage practical learning applications. They provide capabilities to explore novel actions (actional), trigger semantic associations through symbolisms (narrative), enhance the sense of presence through immersive devices (sensory), and include the degree of collaboration and work with pairs (social). However, Dede et al. [22] stated that immersive experiences could potentially provide learning through constructivist approaches solely on materials that require 3D to be explained (e.g., understanding of the solar system's ellipsis) or where embodied cognition can be applied (e.g., empathy through first-view experiences). Other mediums, such as 2D simulations, non-immersive environments, and traditional non-digital elements, could be practical or even more efficient than IVR, depending on the concept of the instruction. Therefore, an overview of how immersion is integrated into customized learning experiences can provide insights into the enhancements authors aim to achieve in their designed lessons. Dede's immersive interface design framework can be adapted to identify the technological and pedagogical features used in these immersive learning experiences.

We adapted these frameworks to understand the composition and enhancement of IVR experiences. The IVR affordances and their integration into the reviewed design examples provide guidelines on effective ways to develop and customize IVR for learning. With our conceptual framework, we aim to demonstrate the level of integration achieved by authors in their designs and developments, specifically examining patterns in these developments and their potential benefits or drawbacks for learning. Based on the previously discussed frameworks and objectives, this review investigates the following research questions:

- **RQ1:** What are the development features and methodological approaches for designing and evaluating IVR learning experiences?
- **RQ2:** How the customized IVR applications are classified in terms of degree of embodiment, immersion, and type of learning?
- **RQ3:** How do the targeted STEM topics in the customized IVR experiences effectively improve learning outcomes?
- **RQ4:** What are the reported advantages and disadvantages of learning outcomes and user experience when using customized IVR experiences?

IV. METHODOLOGY

A. SEARCH STRATEGY

We aimed to explore the current trends in the design of IVR in STEM. We searched recent publications and studies on multidisciplinary databases according to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [33]. We have reviewed seven scientific databases: IEEE Xplore, ACM, Wiley, Springer, Elsevier, Taylor & Francis, and Web of Science. The database platform provides different methods and searches advanced tools to access the relevant literature according to the delimited keywords. The refined search was done, including special Boolean operators (e.g., "`AND`," "`OR`," and "`NOT`"), to narrow down the possible results coming from the consulted databases. The key search used to perform the search on each database is based on the main keywords that align with targeted publications and intended revision. The key search used is stated as follows:

> ("immersive virtual reality" `OR` "virtual reality") `AND` "higher education" `AND` (design `OR` development `OR` implementation) `AND` (learning `OR` experience) `AND` (advantage `OR` benefit `OR` limitation) `NOT` (medical `OR` medicine)

The keys selected are related to our intended objective around the article's focus and content. The term "higher education" is presented as immovable due to our leading population group of interest, similar to the terms "immersive virtual reality" and "virtual reality." The key "virtual reality" is often used in different contexts and has meant slightly different technologies over the years, such as computer-based interactions, 3D simulations (e.g., second life), and high-end HMD experiences. For the possible VR misconception, we opted to include the Boolean operator "`OR`" with the inclusion of the "immersive" component to try to point to papers that discuss certain levels of immersion in their implementation. Additionally, we included "design" and "development" terms to include articles that present the outline and development process of the IVR experience in their methodology. We have omitted STEM-related terms to prevent the exclusion of subjects not explicitly addressed by the STEM acronym. However, since we have not specified any topic, we discarded articles focusing on health and medicine by including the Boolean operator "`NOT`," which disregards entries containing the mentioned terms. Searches were performed considering most of the metadata related to the articles, but mainly title, abstract, and keywords.

B. INCLUSION AND EXCLUSION CRITERIA

Consumer-grade VR HMDs, such as HTC VIVE, Sony PlayStation VR, and Oculus Rift CV1, began to be announced in 2016, significantly boosting interest in VR trends [34]. By this time, HMDs had extended the visualization aspects of VR through a more interactive experience (e.g., isolation through HMD and controllers) to enhance immersion potentially. Additionally, in 2016, the Horizon expert panel [35] anticipated widespread adoption of virtual and augmented reality in higher education in two years; however, such a range was extended to five years with the emergence of the mixed reality paradigm. We considered peer-reviewed studies published from January 1, 2016, to August 31, 2023. The literature review presents the current state of designed and developed IVR systems and their advantages and disadvantages in terms of learning outcomes and user experience. We delimited the inclusion criteria that fit our research objectives as follows:

- Papers published between 01/2016-08/2023;
- Paper with full text available;
- Peer-reviewed papers;
- Samples related to higher education at any level (e.g., undergraduate or graduate students);
- STEM-related IVR lesson:
  - The authors aim to teach/train around a concept; and
- Authors present the development of an instructional IVR experience, including:
  - Evaluation of one or multiple components of the developed tool with conditions or comparison with traditional methods; and
  - Discuss the advantages/disadvantages of using VR and the designed features (e.g., haptic feedback, pedagogical agents).

C. EXCLUSION PROCESS AND SCREENING OF THE PAPERS

From the listed strategies, we identified 2175 studies for consideration. After eliminating duplicates and entries with incomplete metadata (e.g., those lacking authors or with one-word titles), we were left with 2012 papers. Subsequently, we applied filters to these papers using the relevant keywords outlined in the specified categories (see Table 1) through an automated Python script. In this script, we scored each paper based on the occurrence of words from all categories on the paper metadata (e.g., title, keywords, abstract, journal, series, and others). We discarded papers with a resulting zero score. Furthermore, we implemented a string match filter to retain only papers containing at least one specified word, such as VR, virtual reality, immer*, headset, head mount, head-mount, HMD, mixed reality, extended reality, and XR. Following this process, we screened the titles and the abstracts of 713 filtered papers to validate their eligibility. We excluded irrelevant entries by manually reading full texts and excluding papers.

TABLE 1. Categories keywords.

Categories	Keywords
Virtual reality	virtual reality, immersive virtual reality, immersion, 3D environment, head mount display, headset, high-definition graphics, immersive learning, immersive experience, desktop
Learning context	STEM, STEAM, science, technology, engineering, math, laboratory, higher education, university, campus, student, learning, design, lesson, development, implementation, methodology, framework, guideline, theory, scaffold, instruction
Evaluation metrics	engagement, cognitive load, understanding, knowledge, transfer, advantages, disadvantages, haptics, embodiment, spatial, agency, peer, distraction, discomfort, benefits, limitation, comparison, effort, frustration, performance (pre and post) test, user study

Two of the authors examined the 625 papers based on the exclusion criteria (see Table 2) by reading the titles and abstracts. From the 625 papers, the authors, after a voting process followed by a discussion phase, decided to read the chosen 88 papers to assess their eligibility for inclusion in the systematic review. Exclusion criteria comprised the absence of conducted user studies, lack of focus on learning outcomes or STEM topics, non-use of high-end HMDs, absence of details about the design and development of the VR tool, and usage of third-party VR solutions for their studies. We compiled information regarding the learning design, virtual reality device, prominent features developed for the IVR experience, and targeted population, addressed STEM topics and findings, and discussed the advantages/disadvantages of the developed tools. The selection and filtering of the studies included in this systematic review are delimited in Figure 2.

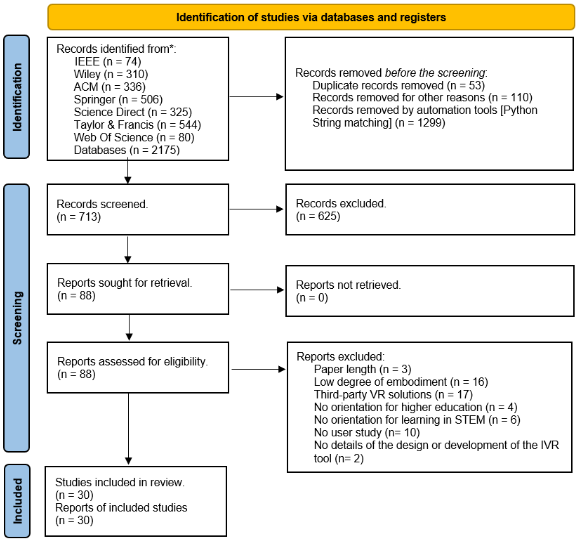

FIGURE 2. The literature identification and screening process flow chart is based on the PRISMA guidelines.

D. RISK OF BIAS ASSESSMENT

We conducted a thorough risk of bias assessment by crosschecking all the authors' choices. We adopted the

TABLE 2. Exclusion criteria for abstract and complete text revision.

Exclusion criteria	Description
Paper length.	The paper is expected to be more than five pages. We considered this to indicate that the authors provided enough information about the designed IVR tool.
The usage of a low degree of embodiment.	The authors aimed to explore VR through experience with a low level of embodiment, such as 360-degree videos, Desktop VR, or no high-end HMD.
Third-party VR solutions.	The authors did not develop any IVR experience but used tools from other sources. We discarded the paper because using already developed tools does not provide details of their design for an immersive experience.
No orientation for higher education.	The sample population for the study is not related to higher education. For example, the paper focuses on implementing for high or vocational schools.
No orientation for learning in STEM.	The paper explores different aspects of learning or user experience but not in STEM. A learning topic can justify the objectives of the author's implementation, but their current design does not focus on learning outcomes.
No user studies.	The authors should have validated their designed framework or tool with a user study in the paper. The authors did not test the proposed features and recommendations in real settings.
No details of the design or development of the IVR tool.	The authors discussed aspects outside the design and the development of the tool; even though VR could be employed, details about the tool are not provided.

Cochrane Collaboration's ROB-2 (Risk of Bias version 2) tool for randomized studies to assess the "intention-to-treat" effect [36]. Bias was categorized into five domains: (1) bias arising from the randomization process, (2) bias due to deviations from intended interventions, (3) bias due to missing outcome data, (4) bias in measurement of the outcome, and (5) bias in the selection of the reported result. Each study was scored with an indicator of risk: "low risk," "some concerns," or "high risk." For non-randomized studies that focused on validating interventions rather than conducting comparison studies, we used the ROBINS-E (Risk of Bias in Non-randomized Studies of Exposures) tool [37]. In these cases, bias was categorized into seven domains: (1) bias due to confounding, (2) bias in the selection of participants, (3) bias in classification of exposures, (4) bias due to deviations from intended exposures, (5) bias due to missing data, (6) bias in measurement of outcomes, and (7) bias in the selection of the reported result. Following the guidelines of ROB-2 and ROBINS-E, we answered signaling questions and used algorithms to estimate the level of risk for each domain and the overall risk. The risk of bias assessment for all studies is provided in the supplementary material.

E. CATEGORIZATION OF THE PAPERS

1) OVERVIEW

We rigorously identified and categorized studies aligned to suggested content analysis guidelines [38]. We recognized essential characteristics of the papers' content based on each research question's objectives. Encoded categories guided the review's findings, focusing on the papers' design and development description of IVR experiences (**RQ1**). The design should include decisions that align with the author's reasoning for choosing IVR and fit the learning content. We categorized the papers according to the proposed conceptual framework, considering the degree of embodiment, level of immersion, and learning type targeted in the designed IVR experiences (**RQ2**). Other observations included diverse STEM topics addressed in the developed IVR tool and its effectiveness regarding learning outcomes (**RQ3**). When considering design, authors had to choose topics to evaluate or emphasize, such as complex invisible phenomena or challenging-to-access materials to be replicated in 3D. Additionally, we critically classified the discussed advantages and disadvantages of the implemented IVR design. The authors drew on different arguments from previous literature, highlighting various benefits (e.g., VR increases motivation) and limitations (e.g., possible cognitive overhead) of IVR for learning. Consequently, we compiled observed advantages and disadvantages of the developed IVR tool resulting from the measured student's perception through self-reported metrics (quantitative methods) or interviews and think-out-loud (qualitative approaches) (**RQ4**).

2) CONCEPTUAL FRAMEWORK CATEGORIZATION

Based on the proposed conceptual framework (see Figure 1), we classified the reported papers according to the delimited constructs. This classification relies on the reported design, development, and results from each paper, including elements such as the selected point of view (POV), environment layout, interactions (with controllers or hand tracking), and multi-modal feedback (e.g., haptic or audio feedback), assessments used and user feedback and self-reported ratings. For the embodiment degree, we considered factors such as physical involvement, required gestures or bodily movements, and user engagement with the content (e.g., 360-degree environments). We categorized the papers by the degree (low, medium, or high) of *sensorimotor engagement*, *gestural congruency*, and *immersion*, assigning them to one of the four levels described in Section III. For immersion, we assessed integration based on *sensory*, *actional*, *narrative*, and *social* features. Sensory immersion was classified by the level (high to low) of representational fidelity, graphics quality, and multimodal feedback integration. Actional immersion was evaluated by the extent of user interaction with the environment, from passive viewing to active control and modification. Narrative immersion was assessed by including context or storylines, such as assigned roles, missions, achievements, and difficulty variations. Social immersion was determined by the presence of peer interaction and collaboration components. For learning type, we examined the knowledge assessment described in the user studies. Papers were categorized based on the knowledge types outlined in Section III, and each paper was labeled according to one or more of these categories.

V. RESULTS

In this section, we compiled the results from the systematic review according to the delimited research questions and objectives. We have included 30 papers. Of the papers, 56.67% were submitted to scientific journals, 36.67% were published as conference papers, and 6.67% were included as book chapters. In terms of publication, the "British Journal of Educational Technology (BJTE)" was the one with the highest frequency (3 articles) around the included papers, followed by "The Journal of Computer-Assisted Learning (JCAL)" (2 articles). Considering the databases we used to retrieve the papers, the distribution from the included publications is shown in Figure 3.

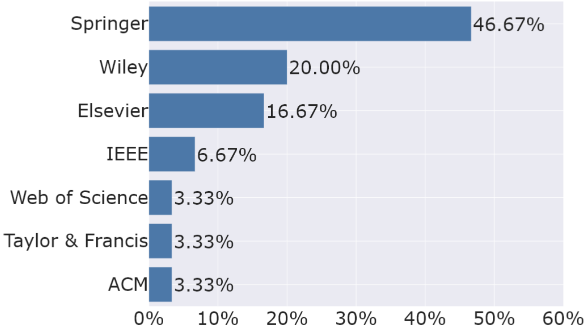

FIGURE 3. The database frequency of the included papers.

We found that the papers were published in different institutions in the USA (23.33%), Germany (13.33%), China (10.00%), Spain (6.67%), Taiwan (6.67%), Canada (6.67%), Australia (6.67%), Czech Republic, Belgium, Malaysia, New Zealand, Italy, Austria, Denmark, and Thailand (all with 3.33%). The country data is taken from the authors' affiliation, and we have considered it the most frequent country among the papers' listed authors. The year distribution of the papers is summarized in Figure 4. The years 2021 and 2023 could be considered the years with the most published papers, with ten and nine, respectively. The data reflects the trends of using high-end HMD for these learning activities in current years. The data shows that this review did not include publications published around 2022 from the revised papers.

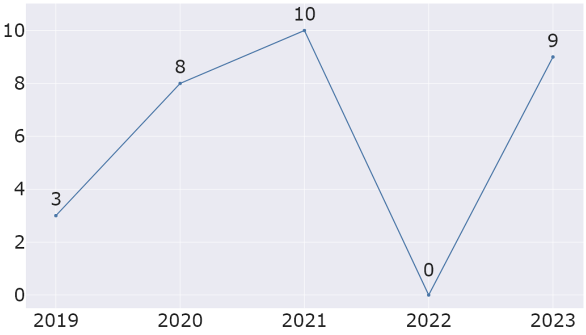

FIGURE 4. The included papers' year of publication.

A. METHODOLOGICAL APPROACHES AND METRICS (RQ1)

In terms of research methods and how the authors explored IVR experiences and their effects, the results show that authors tend to use mixed (46.67%), quantitative (40%), or qualitative (13.33%) methodologies. Only a few paper publications explicitly stated the research methodology as classified (8 out of 30). However, we inferred their intended research methodology based on the described procedures in their user studies. In Table 3, the selected papers are classified based on their intended research methods.

TABLE 3. The research methods used by the selected papers.

Research method	Frequency	Papers
Qualitative	4	[20], [39]–[41]
Quantitative	12	[11], [42]–[51]
Mixed	14	[26], [52]–[64]

As we reviewed the research paper that includes user studies, the authors aimed to answer specific research questions. To achieve this, they utilized various metrics to assess the usage of IVR and its potential impact on the students' experience and performance. In Table 4, we categorized and described the used metrics. The authors focused on previously implemented surveys to measure users' self-perceived aspects such as presence, immersion, engagement, simulation sickness, and usability [52], [53], [54], [62]. Instead, other authors preferred to design questionnaires to target the expected measurements that fit their user studies [39], [52], [53], [54], [58]. The more straightforward way that authors used to report the analysis of certain behaviors or emotions, such as engagement, was through a single question such as "How engaging did you find the game?" [26], where the authors quantified the self-perceived users' experiences through standardized scales (e.g., Likert scale).

The authors aimed to measure learning performance in various ways, often using their own designed surveys tailored to the targeted learning content [55], [57], [59], as presented in Table 4. Other standardized surveys that address learning from a specific topic are not reported, such as the Geologic Block Cross-sectioning Test (GBCT) used to evaluate knowledge around earthquakes [59]. Bagher et al. [59] explored the IVR's use to enhance student's learning experience and performance in drawing earthquake location cross-sections in 3D, focusing on learning performance in geosciences. The self-developed knowledge questionnaire and the semi-structured interview were also a way to assess learning.

Among the less used but promising metrics are multimodal measurements for student performance, including video screen recording [40], performance video recording [20], [40], [43], eye-tracking [20], and physiological measurements [60]. Video recordings allow capturing participants' activities in the virtual space on the HMD through video streams. Additionally, they allow the analysis from different perspectives [91]. Regarding eye tracking, the authors can

TABLE 4. Metrics to address IVR experiences in user studies.

Category	Questionnaire	Paper
User experience	System usability scale (SUS) [65], [66]	[43], [50], [52], [53], [59]
	User experience questionnaire (UEQ) [67]	[53], [56]
	Technology acceptance [68]	[56]
	Simulator sickness questionnaire [69], [70]	[53], [57], [60]
	Self-made questions and semi-structured interviews	[20], [26], [39]–[41], [45]–[47], [52]–[54], [58], [64]
Presence and immersion	Steed-Usoh-Slater (SUS) presence questionnaire [71]	[54]
	Bailenson's co-presence questionnaire [72]	[54]
	Immersion and presence	[55], [73]
	MEC (Measurement, Effects, and Conditions) spatial presence questionnaire [74]	[51], [59]
Cognitive load	Technology usage inventory (TUI) [75]	[49], [61]
	Nasa-Task load index (NASA-TLX) [76]	[53], [54]
	Intrinsic and extraneous cognitive load [77]	[48]
Engagement	Engagement questionnaire (GEQ) [78]	[62]
	User engagement scale (UES) [79]	[51]
	Perceived enjoyment [80]	[48]
	Attentional control scale (ACS) [81]	[59]
Learning	Factual and conceptual knowledge [82]	[48]
	Perceived learning effectiveness [83]	[59]
	Confidence [84]	[50]
	Technology-Mediated learning (TML) [85]	[44]
	Self-efficacy and intrinsic value [86]	[55]
	Spatial ability [87]	[60]
	Verbal ability [88]	[49], [61]
	Anxiety [89]	[50]
	Self-developed knowledge questionnaire and semi-structured interviews	[11], [55]–[57], [63], [90]

track what the participants are focused on, for example, to help students who struggle with specific parts of the 3D simulation [20]. The authors aimed to register user actions to understand the interactions with the implemented features [40], user attention [20], and sensor data to identify any cognitive workload during the intervention [20]. This reflects one advantage of developing applications tailored to specific research objectives: direct access to the source code allows designers to implement mechanisms for tracking learners' performance, such as embedding data in the IVR application interaction. An example is reported in Santos-Torres et al. [52] paper, which includes aspects such as the number of errors and the overall time of the whole task when using the HMD.

B. DEVELOPMENT PLATFORMS, TOOLKITS, AND PIPELINES (RQ1)

In discussing the design of IVR experiences, the authors provided details about their developed experiences, including features and the utilized developer resources. For development toolkits, authors primarily relied on existing game engines that facilitated the implementation of graphical interfaces specifically for VR. Authors used the Unity game engine the most, with 76.67%, while the Unreal game engine was utilized for one of 30 papers (3.33%). The remaining papers (20%) did not provide details about the development tools used for the IVR platforms. Regarding implementation details, a few authors discussed how the tool was developed and described the pipeline they followed for the developed IVR experience.

Checa et al. [45] developed a multiplatform (VR and Desktop) serious game experience to teach undergraduate computer hardware assembly concepts. The authors provided details around the used pipeline as (1) creation of 3D model using Blender software and imported 3D models, obtained under Creative Commons (CC) license, from different sources; (2) integration of these models in the Unreal game engine, which offers a high capacity to create photorealistic environments and its visual scripting system; (3) development of the 3D virtual environments; (4) creation of the VR learning experience; and (5) adaptations for VR and desktop applications. Try et al. [64] described their development pipeline in three main stages: (1) draw, referring to the collection of the 3D object models used in designing the VR application such as the laboratory building (made with Sketch–Up and Autodesk AutoCAD) and two nondestructive testing equipment (modeled with Blender), (2) build, enclosing to the coding phase on the Unity engine, using C# programming language as well the exported format (.exe), and (3) test, involving validation test with students.

The authors detailed the use of HMDs in their studies. The listed HMDs include the HTC Vive (50%), Meta Quest (16.7%), Oculus Rift (10%), Meta Quest 2 (10%), Oculus Go (6.7%), and 6.7% authors who did not specify the HMD they used in their papers. The HTC Vive was the most used HMD, with half of the papers relying on the device's capabilities. The authors in various papers justified their choice of HMD for their IVR experiences. For instance, Franzluebbers et al. [54] emphasized the Meta Quest's affordability and features like its high-resolution display and tracking capabilities. Qian et al. [55] detailed the HTC Vive's hardware specifications and tracking systems. Some studies integrated additional hardware, such as Tobii's eye-tracking technology [20], [60]. Qian et al. [55] developed an IVR experience specifically for Meta Quest devices. At the same time, other authors noted the Meta Quest 2's limitations in rendering highly realistic graphics due to its computing power constraints [53].

TABLE 5. Proposed IVR experiences STEM topic, immersion level, and features.

STEM topic	Citation/ (Tool name)	Immersion	Features
Science			
Geographical Information Systems (GIS)	[52]	Sensory: medium; Actional: high; Narrative: low; Social: low	Learners interacted with a map interface using either hand-tracking or controllers with ray casting. They could move the virtual map, zoom in/out, and place markers. The virtual map was presented as a large wall that occupies the user's field of view.
Computational fluid dynamics (CFD)	The Virtual Garage [53]	Sensory: high; Actional: high; Narrative: medium; Social: low	Learners visualized pre-computed CFD simulation data and interacted with the different parameters to complete the activities. They also were able to move around through different stations that included short presentations, explanations with audio, quizzes, and instructions to learn about the evaluated procedure (https://www.youtube.com/watch?v=Lv__-Hl0eMY).
Scanning Electron Microscopy (SEM)	[42]	Sensory: high; Actional: high; Narrative: medium; Social: low	Learners interacted with a training space by following the correct operations of the SEM, watching and controlling video instructions, and completing a formative assessment embedded in the virtual environment.
Oxygen preparation experiment.	[55]	Sensory: high; Actional: high; Narrative: medium; Social: low	Learners can grasp, move, and put in contact with the interaction of virtual chemical apparatuses. They also were able to recreate and visualize simulated chemical reactions. The application allows the tracking of real equipment matching the virtual objects visualized on the HMD.
Cells and the human body	The Body VR [44]	Sensory: high; Actional: high; Narrative: high; Social: low	Learners could look around 360 degrees at a simulated and enlarged human body system. They could also touch, rotate, and explore virtual objects such as red blood cells. The application Included instructional narration and animations.
DNA (lac operon)	MolgenVR [20], [60]	Sensory: high; Actional: high; Narrative: medium; Social: medium	Learners could interact with simulated complex gene regulation models by assembling elements in 3D via drag-and-drop and controlling simulation settings to generate multiple outputs, such as molecules or proteins. They were immersed in the IVR environment with the help of a facilitator outside for questions.
Earthquakes	[59]	Sensory: high; Actional: high; Narrative: medium; Social: low	Learners could visualize and interact (scale) global phenomena such as earthquakes, volcanoes, and plate boundaries on a world map. They can also walk and interact with the in-world menu through a laser pointer, which allows them to hear the information related to each data point. The data is represented on top of a 2D world map and displayed as point clouds.

TABLE 5. *(Continued.)* **Proposed IVR experiences STEM topic, immersion level, and features.**

Molecular compounds	[46]	Sensory: high; Actional: high; Narrative: high; Social: medium	Learners interacted with molecular structures using controllers, selecting atoms and bonds to combine into molecules. The applications include an in-world menu that provides instructions for component manipulations. The virtual environment changes cycled through different weather conditions and times of day.
Electromagnetic and electrostatic	MaroonVR [62]	Sensory: high; Actional: high; Narrative: medium; Social: high	Learners interacted with a simulation of electromagnetic and electrostatic physics concepts. Movement can be tracked, and actions can be performed through controllers. The simulations also include control and assessment features. The application provides a networked environment, but voice is not included (https://www.youtube.com/watch?v=LnhZmVWycx4).
Molecular interactions	[40]	Sensory: high; Actional: high; Narrative: medium; Social: high	Learners can control and visualize molecules and enzymes and build new elements by using bonds and atoms and changing the representation model (e.g., ball-and-stick or cartoon). The application includes In-world menu instructions and collaborative learning.
Virology	VR museum [48]	Sensory: high; Actional: high; Narrative: high; Social: medium	Learners are placed in a virtual museum exhibition on viruses. A pedagogical agent with narration guides them through it. They can also walk or teleport, interact with 2D animated presentations, and respond to integrated knowledge assessments. The application includes ambient music.
Natural selection and mimicry	Catch a Mimic [26]	Sensory: medium; Actional: high; Narrative: high; Social: low	Learners are placed in a rainforest and asked to catch virtual butterflies using controllers. The catch procedure includes visual and auditory feedback. The application also includes In-world menu instructions and assessment. The headset used limited the interaction to 3 DOF (https://www.youtube.com/watch?v=TQs1VkCL6Nc).
Technology			
Robotics	[49], [61]	Sensory: high; Actional: low; Narrative: high; Social: medium	Learners explored a virtual kitchen environment guided by a robot assistant while executing its service task via audio narration. Using controllers, they moved around the environment to check different perspectives. However, the learners were not able to manipulate the robot.
Computer hardware assembly	[45]	Sensory: high; Actional: high; Narrative: high; Social: medium	Learners can assemble step-by-step a computer model from modular components with the help of assistant robot guides. The application includes an in-world menu that provides instructions and information for each component (https://www.youtube.com/watch?v=lfXhAMymjW8).

TABLE 5. ***(Continued.)*** **Proposed IVR experiences STEM topic, immersion level, and features.**

Robot operations	[47]	Sensory: high; Actional: high; Narrative: high; Social: medium	Learners are placed at the construction site, where they can remote-control demolition robots. The application includes safety procedures and a simulation of failures and the consequences of poor strategies. They also interact with the environment using an Omni VR treadmill to walk/run in the virtual environment and a physical controller for the robot to operate the different actions remotely.
Hardware assembly	[50]	Sensory: high; Actional: medium; Narrative: low; Social: low	Learners could grab and place components and program them to assemble computer hardware.
Photovoltaics PV-arrays	[41]	Sensory: high; Actional: high; Narrative: medium; Social: high	Learners, represented as avatars, are placed in a realistic environment surrounded by photovoltaic arrays. They can visualize PV arrays and historical data in real-time. The application includes a networking environment with integrated voice.
Engineering			
Land Surveying	ENGREDUVR [54]	Sensory: high; Actional: high; Narrative: high; Social: high	Learners, represented through colored avatars, could collaborate and manipulate simulated machinery, such as the station and prism-rod device required for surveying. They could also share a whiteboard to take notes, communicate via voice, and teleport through the environment.
Power engineering	V-Turbine [43]	Sensory: high; Actional: high; Narrative: medium; Social: low	Learners were placed in a workshop environment where they could interact and assemble an aero-derivative turbine step-by-step. They also watched 3D animations, teleported through the environment, and in-world menu instructions that guided the activity. The environment also includes visual and audio feedback.
Machine operator	[56]	Sensory: high; Actional: high; Narrative: medium; Social: low	Learners followed a step-by-step guide (text and audio) to interact and complete a virtual machinery operation task, including in-word menu interactions and tool operations such as loosening screws. The application allowed learners to teleport through the environment and an ‘‘eye button’’ that provides more hints about the current step.
Industrial system design (cooling water)	SDVR [90]	Sensory: high; Actional: high; Narrative: medium; Social: low	Learners were placed in a factory environment to assemble a cooling water system. They used an in-world menu to select the required parts of the machine and the needed instructions. They manipulated the elements (position and rotation) using controllers and teleported through the virtual environment (https://www.youtube.com/watch?v=hErpeptKm9Y).
Product development	[58]	Sensory: high; Actional: high; Narrative: medium; Social: low	Learners visualized and manipulated 3D CAD models through realistic hands-free interactions and voice commands. They could use commands such as zoom, select, restore, and highlight matching pieces. The application includes information and context about the selected models.

TABLE 5. ***(Continued.)*** **Proposed IVR experiences STEM topic, immersion level, and features.**

Bridge construction	[11]	Sensory: high; Actional: high; Narrative: medium; Social: low	Learners can visualize and assemble 3D bridges. The bridge assembly can also be simulated and visualized as a 3D animation. They can also toggle between modes and visualize specific information about each of the components that compose the bridges. The application included in-world menu instructions and laser-pointing interactions (https://www.youtube.com/watch?v=c5gnH7xithw).
Safety in design (SiD)	[63]	Sensory: high; Actional: high; Narrative: medium; Social: low	Learners could interact and manipulate a vehicle loading crane (VLC) operation. The crane is manipulated using controllers, and real-time feedback is provided to indicate safety failures.
Building constructions principles	Immersive Building Simulation [51]	Sensory: high; Actional: high; Narrative: medium; Social: low	Learners could construct a virtual bus stop by interacting through various mediums, such as controllers, body motion tracking, and teleportation.
Non-destructive testing instruments	VRAL [64]	Sensory: high; Actional: high; Narrative: medium; Social: low	Learners are placed in a laboratory building with material testing equipment, an interactive UI, 2D textual information, and voice instructions to guide their activity. The application includes 3D animation effects that show the utilization process of non-destructive testing and interactivity through object manipulation.
Mathematics			
Polygonal mesh and geometry	[39]	Sensory: high; Actional: medium; Narrative: medium; Social: high	Learners watched displayed PDF presentations and interacted with polygonal meshes loaded from OBJ format files. They could also draw over the models enabled in a networked environment with an instructor and other students connected remotely. The experience did not feature avatars of connected users and no voice chatting.
4D-space	[57]	Sensory: medium; Actional: medium; Narrative: low; Social: low	Learners visualized a hypercube in 3D adapted from 4D to 3D through orthographic projection. They were also able to rotate it through the different axes using controllers. The application includes yes/no questions to assess spatial reasoning.

In a specific case discussed by Arntz et al. [41], their developed VR application serves as a substitute for accessing real photovoltaics (PV)-arrays, with the simulation output resembling actual machinery. This application requires providing real-time data in the virtual representation. The authors integrated a back-end solution using Modbus data into a MariaDB storage. Similarly, Qian et al. [55] detailed the integration of VR hardware to facilitate user interactions, utilizing the tracking system to calculate the coordinate transformation of the HTC HMD and controllers through the lighthouse base station and OptiTrack cameras. They provided equations necessary to retrieve the device's positions. Another aspect to consider when developing such applications is the expected duration of the IVR experience. The authors reported this information, which we presented in different ranges: less than 10 minutes (3.33%), between 10-20 minutes (23.33%), between 20-30 minutes (13.33%), between 30-40 minutes (13.33%) and 40 or more minutes (13.33%). Notably, 47% of the studies did not report the duration.

C. DEGREE OF EMBODIMENT AND IMMERSION (RQ2)

In Table 6, we classified the developed IVR experiences according to the proposed conceptual framework. The majority of the papers focused on providing a third-degree embodiment (46.43%), followed by the fourth-degree (42.86%) and the second-degree (10.71%). No first-degree embodiment was discussed due to the nature and filtering established for this review, which is aimed at discussing immersive learning experiences. Among the fourth-degree examples is MaroonVR [62], an interactive IVR physics laboratory where students engage with simulations related to electromagnetism and electrostatics. This IVR experience is adaptable to various platforms, including desktop and mobile. Regarding embodiment properties, MaroonVR features a laboratory setting with gesture congruency for interaction and room-scale movement within the 3D environment. Similarly, Tang et al. [40] designed an IVR experience for visualizing molecular interactions, highlighting five key multimodal affordances: viewing, scaling, sequencing, modeling, and manipulating. Their design enhances these affordances through interactive features, such as editing stages of the visualization and engaging in shared immersive environments.

For a third-degree embodiment experiences, Checa et al. [45] created a serious VR game for teaching computer hardware assembly. Their experience includes a step-by-step tutorial, continuous feedback from an assistant robot, and a menu with component information. Gesture congruency is evident in hands-on activities where students assemble computer parts. However, limited movement and multimodal effects restrict this IVR experience to a third-degree embodiment. In the second degree, Wang et al. [50] investigated the impact of emotion on engagement and learning in a VR STEM activity but provided minimal details on tools or interactions. Based on assembly tasks, the assumed grabbing option results in a low sensorimotor degree. Additionally, 37% of the papers named their implemented tools.

Regarding immersion level, the designed IVR experiences were primarily considered to have high sensory immersion (89.29%), followed by medium sensory immersion (10.71%), with no instances of low sensory immersion due to the use of high-end devices. For actional immersion, high levels were most common (85.71%), followed by medium (10.71%) and low (3.57%). In terms of narrative immersion, the majority of approaches were medium (60.71%) and low (10.71%), with fewer papers implementing high narratives (28.57%). Similarly, for social immersion, most papers did not focus on this aspect, with low integration being the most common (60.71%), followed by medium (21.43%) and high (17.86%).

D. STEM TOPICS AND TYPES OF LEARNING ASSESSED (RQ2 AND RQ3)

Papers conducted user studies on higher education scenarios, including sample students related to the subjects or targeted population in their designed research. The studies focused on different samples of higher education level and other stakeholders, such as undergraduate students (56.66%), graduates or postgraduate level (13.33%), or either both groups (13.33%); also, reported papers have samples with a variety of groups including students, professors, experts or workers (16.66%).

The selected subjects to instruct through IVR lessons are relevant points to discuss in this review. In Table 5, we listed the learning topics of the designed IVR experiences and classified them based on the STEM focus. From the selected topics, we found papers oriented through science concept (43.33%) for the explorations of topics such as geography [52], [59], biology [20], [26], [44], chemistry [46], [55], and physics [62]; Technology (20%) in terms of robotics [47], [49], [61], hardware assembly [45], [50], and solar panel experimentation [41]; Engineering (30%) delimited by industry safety operation [63], simulated field work [54], machinery assembly [43], [56], [90], and construction [11], [51]; and Mathematics (6.67%) for geometry [39], and 4D spaces [57]. The distribution of the papers' topics by STEM field is summarized in Figure 5.

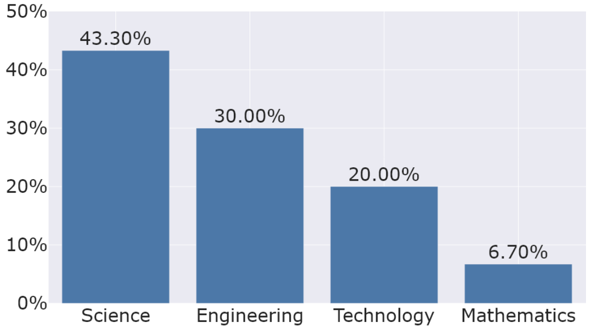

FIGURE 5. The STEM learning topics used for the design of the IVR experiences.

TABLE 6. Proposed IVR experiences degree embodiment, learning type, and findings.

Citation	Degree embodiment	Type of learning	Sample size	Findings
[39]	Third (SE: high, GC: high, IM: high)	No assessment	4	Tested students found the integration of VR in their class setting useful for visualizing 3D meshes and interpreting geometric concepts. They commented about not being able to take notes during the VR lectures.
[52]	Third (SE: medium, GC: high, IM: high)	No assessment	32	Device-based interaction performs significantly better for controlling immersive VR maps than body-based interaction, as it takes less time to complete. The users perceived the former technique as more usable. The authors conducted a within-subject study.
[53]	Third (SE: high, GC: low, IM: high)	No assessment	24	The authors compared two modules within the designed Virtual Garage. Students found both useful, while module #2 found them more cumbersome, especially for novice users. Module #1 was composed differently for VR training interactions and theory sections. Module #2 presented the pre-computed CFD data simulation. In the study, all participants experienced the proposed modules.
[54]	Fourth (SE: high due to the inclusion of voice support in the networked environment, GC: high, IM: high)	No assessment	85	The students' self-reported scores regarding task load and presence favored male ($n = 61$) participants over female ($n = 24$) ones. The students reported difficulties reading text on VR.
[42]	Third (SE: low, GC: high, IM: high)	No assessment	12	Students found the system usable and reported lower issues when interacting with it. Some participants reported a lack of guidance prompts in the IVR environment.
[55]	Fourth (SE: high multimodal sensory feedback by the touch of natural objects, GC: high, IM: high)	Procedural	42	The authors compared the usage of VR-fusion, which blends real-world objects with virtual ones, VR, and desktop learning activities in balanced groups ($n = 14$). They reported no significant differences in learning results and motivation between conditions. The VR-fusion condition has positive effects on presence and immersion but is not compared with traditional VR, even reducing the sense of presence compared to VR. Students in both VR conditions reported improvements in self-efficacy.

TABLE 6. *(Continued.)* Proposed IVR experiences degree embodiment, learning type, and findings.

[43]	Fourth (SE: high, GC: high, IM: high)	No assessment	26	Students indicated confidence and excitement regarding the IVR application's usability. The authors did not find significant differences in usability scores by gender ($M = 14$, $F = 12$) and gaming experience.
[56]	Third (SE: low, GC: high, IM: high)	No assessment	13	The authors found that steps learned to remove the construction cylinder in the IVR environment were successfully transferred to the real machine without any issues. Also, participants showed a positive score regarding self-perceived measures such as technology acceptance, motivation, and learning success. The authors found evident difficulties in the required high-fidelity representation of the procedure, specifically the steps requiring haptic feedback.
[90]	Third (SE: low, GC: high, IM: high)	Schematic	33	The authors compared interactive ($n = 17$) and design groups ($n = 16$) in two IVR modules. In one, users are asked to explore and read; in the other, they are asked to design and build the system. The authors reported that constructive generative engagement through IVR yielded a significantly elevated level of Casual Model notion with a medium-large effect size.
[57]	Second (SE: low a single position interaction, no movement is required, GC: low simple gesture interactions, IM: high)	Declarative	70	The authors compared experts ($n = 22$) and non-expert ($n = 48$) participants. They found that both groups benefit from the immersive interaction with hypercubes. Surprisingly, experts with prior theoretical knowledge benefit more regarding assessment scores.
[58]	Third (SE: low, GC: highest due to the nature of the included gestures, such as uniformly scaling, IM: high)	No assessment	16	Students found the designed IVR system useful. The proposed gestures and voice command features were used and considered intuitive, except for some cases with the rotation and scaling operations. Also, participants reported delay issues and information overlapping in the environment.
[44]	Fourth (SE: high for narration, GC: high, IM: high for the immersive 360-degree visualization inside the human body)	Not enough details	60	The authors assessed the influence of IVR on learning outcomes through affective and cognitive factors. The recorded self-perceived measurements showed that the immersion provided by the IVR experience should increase motivation before positively affecting learning effectiveness.

TABLE 6. *(Continued.)* **Proposed IVR experiences degree embodiment, learning type, and findings.**

[20], [60]	Fourth (SE: high, GC: high, IM: high due to visualizing a simulated and invisible model in a surrounding perspective)	Strategic	34 and 36	The authors explored the use of IVR simulation for complex topics, incorporating a facilitator to enhance model-based learning. The facilitator's guidance helped students develop strategies to understand the model's structures. In their second study, they compared control and experience modes (seated vs. standing) using an IVR lesson in a regular course. They found that the seated IVR group demonstrated significantly higher conceptual understanding at the course's end than the control group (normal course lectures). In contrast, the standing IVR group had scores similar to those of the control group. Additionally, the authors discovered that learning gains are influenced by prior knowledge and the mode of IVR experience.
[59]	Second (SE: low due to usage of a single point of view, GC: low simple gesture interactions, IM: high)	Declarative	21	The authors did not find a significant difference between drawing cross-sections in 2D maps ($n = 10$) and IVR conditions ($n = 11$), but IVR provided a significantly better experience for creating a mental model.
[11]	Third (SE: low, GC: high by grabbing and assembly interactions, IM: high by immersive visualizations and different scene perspectives)	No assessment	44	The authors proposed an automatic generation method of bridge scenes with spatial constraints for virtual teaching with IVR. The participants believed that the IVR bridge teaching system positively improved the interactivity and engagement of teaching.
[49], [61]	Third (SE: high, GC: low; the participants are limited in visualizing the experience, IM: high)	Declarative and strategic	81 and 64	The authors compared the effects of sequencing order and prompting on learning outcomes; they proposed four experimental conditions with prompt and text first ($n = 20$) with prompt and VR animation ($n = 20$), without prompt a text first ($n = 19$), and without prompt and VR animation ($n = 22$). They found that presenting a VR animation first fosters learning of basic concepts and definitions at the knowledge level, and using an elaboration prompt before learning in VR enables the learners to build connections between the concepts and to process the content semantically.

TABLE 6. ***(Continued.)*** **Proposed IVR experiences degree embodiment, learning type, and findings.**

[45]	Third (SE: low, GC: High inclusion of an assistant and different stations to interact with during the assembly process, IM: high)	Declarative and strategic	77	The authors compare serious VR games ($n = 40$) vs. Desktop ($n = 19$) vs. Webcam ($n = 18$) instruction. They show that VR serious games positively affect learning theoretical knowledge and provide significant advantages compared to other conditions for visual recognition. Students rated their satisfaction with the VR and Desktop groups significantly higher than traditional teaching methods.
[46]	Third (SE: low, GC: high, IM: high due to the higher control of the environment by background settings being affected)	All	223	The authors included VR activities during a whole-semester class, comparing students who opted to use VR ($n = 111$) to the ones who did not participate in those activities as a control group ($n = 113$). Their results showed a trend towards improved course grades and final exam scores in the section offered IVR activities, particularly for first-generation college students.
[62]	Fourth (SE: high, GC: high, IM: high)	No details	17 and 19	The authors summarize previously conducted studies comparing the proposed IVR design with traditional methods. They found that immersion and engagement are key elements for supporting learners. They recommend enhancing interactions with the in-world environment and objects to enhance the high level of immersion.
[40]	Fourth (SE: high, GC: high, IM: high)	No assessment	22	The authors conducted a session with participants to analyze the affordances used in their proposed designed IVR experience. They listed IVR's key affordances (viewing, sequencing, modeling, scaling, manipulating) that facilitate learning through science animation and recommend exploiting this affordance to differentiate IVR from other educational technologies.
[47]	Fourth (SE: high inclusion of external devices to provide more realistic experiences with multiple feedback (e.g., walking), GC: high, IM: high)	Declarative	50	The authors compared the effectiveness of their designed tool among students and construction workers in a balanced comparison ($n = 25$). Their results indicated that VR-based training can lead to a significantly larger increase in knowledge acquisition for construction students than for workers. In contrast, VR-based training improved trust in the robot and robot operation self-efficacy significantly more for construction workers than students.

TABLE 6. *(Continued.)* Proposed IVR experiences degree embodiment, learning type, and findings.

[48]	Fourth (SE: high narration of the guided tour, GC: high, IM: high the user is teleported to a natural museum setting)	Declarative	162	The authors compared the inclusion of a pedagogical agent in VR with different appearances and behaviors. They found that including a pedagogical agent, especially a realistic one, leads to lower factual knowledge acquisition than only including a narration. Differently, including the agent may aid the learning of conceptual information.
[26]	Fourth (SE: high auditory feedback, GC: high the used gestures congruent relationship between the performed on the VR experience by moving the hands and arms, IM: high even with the limited interaction of the used device (Oculus Go with 3 DOF); the authors provide a 360-degree immersive environment)	Declarative	217	The authors compared immersion by used platform and PC vs VR and level of embodiment. Their results show that more active and agentic conditions had a higher post-play content knowledge. However, there are no significant main effects for the used platform.
[50]	Second (SE: low, GC: low; no details are provided about the interactions used, IM: high)	Declarative	17	The authors compared the IVR learning environment ($n = 9$) with a control group with slides ($n = 8$). Their results showed that learners in the experimental group performed better, had lower anxiety, and were more confident when realizing the task than the control group.
[63]	Third (SE: low, GC: high, IM: high)	Procedural	143	The authors validated that their designed tools were grounded in multiple learning activities, such as collaborative, problem-based learning embedded in authentic engineering processes. They conclude that IVR is an effective learning tool but may benefit from supplementation with other learning approaches to provide students with diverse learning opportunities and achieve learning objectives.

TABLE 6. *(Continued.)* Proposed IVR experiences degree embodiment, learning type, and findings.

[51]	Fourth (SE: high, GC: high, IM: high)	Procedural	54	The authors compared the integrated level of immersion and interactivity in the designed IVR experience. They found significant main effects regarding interactivity, resulting in a greater perception of possible actions and rewards that could enhance learning experiences. However, there were no significant results regarding immersion; the authors mentioned that a limited Field of View could be a cause.
[64]	Fourth (SE: high, GC: high, IM: high)	No assessment	21	The authors compared three learning approaches, VR, VR-based, instructor-based, and video-based, between different students' levels (12 undergraduates and nine graduates). The result indicated that for both the undergraduate and graduate student groups, VR condition was considered more desirable and beneficial than video in terms of interactivity, cognitive interest, ease of understanding, and support for learning metrics. The graduate students viewed VR condition as the most preferred learning approach. In contrast, the undergraduate students preferred the instructor base the most.
[41]	Third (SE: high, GC: high, IM: high)	No assessment	7	Participants perceived the VR application as a motivating tool for learning. They praised the VR application's interaction and direct visual feedback, which enabled them to contextualize their actions better, experience direct consequences for operational errors, and recognize the high fidelity of the replicated equipment.

The authors highlighted several reasons for designing an IVR experience through high-end HMDs for their targeted STEM topic. They mentioned the trend of affordable devices extending the usage of immersive experiences in the classroom [26], [40], [54], [60]. Papers also point out that VR enhances visualization and understanding by allowing the visualization of three-dimensional objects and multi-dimensional information [49], [58] with a greater field of vision, such as the visualization of CFD data in an IVR environment [53]. Additionally, the authors supported that VR enables students to interact with and control their learning environments, providing hands-on and interactive experiences that are more engaging than traditional methods, such as virtual laboratories [55], [62]. VR also provides a controlled and safe environment for exploring and practicing different STEM concepts, including training in complex and dangerous situations, such as robot operations [47], which are unfeasible or expensive to simulate in real life [54], [90]. Several studies discuss how VR enhances embodied learning and cognition, allowing students to interact with learning material and improving understanding and retention physically. Research shows that VR engages and motivates students due to its interactive nature, leading to better learning outcomes [11], [59], [61]. Additionally, IVR allows students to learn at their own pace, make decisions about their learning path and methods, and engage with the material in a visually attractive way [43], [44], [45], [50], [56]. In response to public health concerns in recent years, VR provided an alternative for remote learning during the COVID-19 pandemic [48], [64].

In the sense of the used evaluation metrics and how the authors assessed learning, we provide a category of metrics in Table 7. Half of the reviewed papers (50%) did not specify their assessment method or evaluated learning directly. Conversely, the authors indicated how these assessments were constructed and what knowledge participants could acquire through their designed IVR experiences. Additionally, some papers mentioned learning in their metrics even though the questionnaires and questions related more to the perceived learning experience rather than assessing learning outcomes [56], [57], [62]. Most authors assessed learning on the declarative (nine papers), followed by procedural (four papers) and strategic (four papers) learning categories. They employed self-designed questionnaires with multiple-choice and open-ended questions covering biology, robotics, computer assembly, and electronics. Furthermore, some authors focused on learning transfer by reflecting on how the IVR activity translated into real-world settings [55]. Fewer papers focused on schematic (two papers) learning, either using pre-existing surveys or considering evaluating all course content as a metric [46], [60].

Considering the effectiveness of the designed IVR experiences, the reported findings (as presented in Table 7) highlight several positive outcomes. The authors noted that participants found their IVR tools usable and accepted the instruction methods for STEM content learning [39], [40], [42], [43], [52], [53], [58]. The studies showed positive ratings and results for self-reported measurements such as self-efficacy, motivation, and engagement, which are closely related to learning effectiveness [11], [44], [55], [56], [62], [64], [90]. Comparisons between different groups and mediums revealed that female participants experienced a higher workload and lower presence in a Land surveying IVR experience [54]. At the same time, other studies found no significant gender differences in usability measures [43]. Novice participants reported higher difficulty levels with complex interactions and visualizations compared to experts [53], although experts with prior theoretical knowledge benefitted more in terms of assessment scores [47], [51], [57].

In comparisons between IVR-designed tools and other instruction mediums, authors found IVR could reduce anxiety and increase confidence compared to desktop counterparts [50], and it is associated with higher post-play content knowledge due to its interactivity and agentic

TABLE 7. Proposed IVR experiences and learning assessment.

Citation	Learning assessment
[55]	Reproduce the experiment (virtual chemical experiment) by manipulating fundamental apparatuses.
[90]	The mental model assessment (six open-ended questions) on various aspects of the cooling water in the industrial cooling water system (CWVR)
[44]	The biology test (ten multiple-choice questions) about the information introduced in the learning application.
[59]	Two knowledge assessments: • knowledge test (five multiple-choice questions) about subduction zones, and, • draw cross-sections plotting the earthquakes' depth with distance from a subduction zone trench by hand.
[60]	Lac Operon Concept Inventory [92] questionnaire to assess knowledge about the lac operon.
[49], [61]	Knowledge assessment (seven open-ended questions) on technical concepts for robotics.
[57]	Hypercube testing on VR (Yes or No questions)
[45]	Knowledge assessment with multiple-choice type questions and open-ended questions about hardware assembly. The test was designed to respond to a multi-level assessment by focusing on retention), transfer, and understanding.
[46]	Course letter grade (A, B, C, D, F, and withdraw) and final exam grade in the course (entire course performance, so all knowledge is assumed).
[47]	Pre-posttest on knowledge acquisition.
[48]	Factual and conceptual knowledge.
[26]	Content knowledge assessment (mix of open-ended and multiple-choice questions) on natural selection, mimicry, and intentionality.
[50]	A questionnaire designed to assess learners' prior knowledge of electronics science and programming.
[63]	CHAIR is a conceptual design review tool for identifying hazards and risks early in the design [93].
[51]	Knowledge test: • multiple-choice questions on 3D rendering of the bus shelter from multiple perspectives, • identify all the tools used to construct the bus shelter, and • an open-ended question asked the participants to draw exact details of the bus shelter by hand.

nature [26]. IVR experiences resulted in significantly higher conceptual understanding at the end of courses compared to traditional lectures [20], successful learning transfer in training scenarios [56], and significant advantages in visual recognition for theoretical knowledge when compared to IVR serious games with desktop and webcam [45]. VR animations followed by prompts helped participants build connections between concepts and process content semantically [61]. Some authors reported a trend towards improved course grades and final exam scores, particularly among first-generation college students, when integrating IVR activities throughout the semester [46]. Additionally, including a pedagogical agent, especially a realistic one, led to lower factual knowledge acquisition than narration alone, but it aided the learning of conceptual information [48]. Overall, customized IVR experiences provided students with a practical and engaging way to interact with learning materials, significantly impacting learning outcomes and performance.

E. REPORTED ADVANTAGES/DISADVANTAGES IN CUSTOMIZED IVR EXPERIENCES (RQ4)

1) LEARNING OUTCOMES

The authors have different hypotheses and research questions that outline their expected outcomes from the designed and developed IVR lesson. Considering the used metrics and the proposed implementations, the advantages and disadvantages offered by the IVR interventions were compiled for each paper and then coded to be presented as a classification. In the sense of learning, authors provided different results of their assessment reflective advantages through different metrics and measurements as presented in Table 8 and reported disadvantages as listed in Table 9.

2) USER EXPERIENCE

As we delimited, the included papers should discuss the development of their own IVR lesson, adding to the expected activity and the learning enhancement. A software tool was designed so the authors could validate the prototype's usability. As described in the discussed metrics (see Table 4), the authors validate aspects of human-computer interaction and the user perception of the developed IVR. Similarly, the advantages/disadvantages of learning these are also classified from the user experience perspective and summarized in Table 10 and Table 11, respectively.

VI. DISCUSSION

Designing and developing customized IVR can be considered a complex task, primarily due to using less common input systems such as HMDs and integrated motion controllers. Various tools and resources, such as 3D modeling software or popular game engines like Unity or Unreal, are available to leverage and facilitate the creation of those VR scenarios. Computer-based experience can be integrated as a part of the learning module to provide different perspectives and views of the learning content, specifically, in these cases, through immersive learning. Our systematic review arranges information on customized IVR experiences, relationships, and possible connections for future implemented immersive lessons. Considering the importance of agency and sense of presence for immersive learning in IVR experiences, the authors, through high-end HMDs, enhanced immersion and embodiment by incorporating features like voice commands, sensory feedback, hand gestures, VR animations, 360-degree visualization, and audio effects. Consequently, 89.29% of studies demonstrated high embodiment implementation (third or fourth degree), often categorized as fourth-degree due to multimodal feedback. This aligns with [22] statement about using immersive experiences for learning content that requires it, indicating congruence between the chosen STEM topics and the included affordances in the discussed papers. Moreover, by exploring learning content, designers should create content that exploits IVR's main features [72]. Common features include interaction with virtual objects using hand-tracking or controllers, ambient environments (e.g., laboratories, factories, or museums) with free movement, and 3D visualization of complex models and data. Additionally, in-world menus and instructions, real-time feedback, audio narration, embedded assessments, collaborative environments, and control-based input interactions such as laser pointing, manipulation, and navigation.

Immersion, a predominant construct, was achieved through diverse 3D environments like museums, factories, and virtual classrooms, focusing more on actional and sensory immersion. However, social elements and networked environments were limited due to development complexities, reflecting a gap in "pedagogical features" promoting social interactions, as noted by [19]. Instructors and designers should consider how IVR can enhance the exploration and understanding of STEM content. Insights from various papers show that IVR effectively visualizes abstract science concepts [62], facilitates hands-on activities, and replicates complex or risky procedures [48]. Science topics are the most selected STEM topics for IVR, leveraging 3D perspectives to explore different abstract phenomena [14], [94]. IVR allows interaction with virtual artifacts and feedback, supporting constructivism learning principles [95]. However, as reported in previous reviews, only 50% of studies focused on learning outcomes, with others only prioritizing usability aspects of their designed tool [16]. Despite some limitations, IVR holds promise for specific STEM subjects, warranting further research. Evaluating examples and aligning IVR with learning theories can help determine its suitability for various educational purposes [26]. Other aspects not explored in the reviewed papers relate to the challenges and technical difficulties associated with implementing and using these high-end HMD devices in educational settings, so reports on these are relevant for practitioners to contextualize all involved factors when using VR.

Furthermore, the authors explored how customized IVR experiences enhance various learning outcomes such as

TABLE 8. Advantages on learning performance of the IVR designed experiences as reported in the respective papers.

Advantages	Description	Papers
Embodied experience enhances understanding of complex procedures/concepts.	The use of the body and the hands-on activities (e.g., learning by doing) benefit learners.	[20], [45]–[47], [57], [59], [90]
Construction of mental models.	IVR experience allows one to build mental models, forming meaningful connections between the concepts by manipulating and interacting with the system.	[20], [49], [59], [90]
Increase engagement.	The designed interactions (e.g., game-based activities) improve participants' engagement.	[11], [20], [51], [53], [62], [63], [90]
Increase motivation.	Users report and recognize the potential of the IVR experience as a motivation for learning STEM-related concepts.	[11], [41], [50], [56]
Immersion and presence effects on learning performance.	Immersion can give the students a spectacle in the teaching that textbooks cannot provide and isolate the learner in the virtual environment to focus on learning.	[11], [44], [45], [48], [55], [59], [62]
Designed learning IVR environment cognitive effort.	The designed learning/training IVR experience is within an acceptable range of cognitive demands.	[56]
Higher learning outcomes.	Participants obtained significantly better learning assessment results than the other less immersive mediums.	[45], [46], [50], [60]–[62]

TABLE 9. Disadvantages in learning performance of the IVR designed experiences as reported in the respective papers.

Disadvantages	Description	Papers
Learning cues for novice users.	Interactive simulations of complex topics require practice and a learning curve to be usable for a novice.	[53], [60]
Main learning task design complexity can induce cognitive load (e.g., high mental workload).	How the learning content is delivered can cause difficulties in the student's learning experience.	[20], [44], [48], [49], [53], [54], [64]
Lack of multimodal feedback (e.g., haptic).	The lack of realistic designed IVR for technical training could limit the learning transfer from the virtual environment to real-world scenarios.	[56], [64]
Included affordances limited learning transfers (e.g., asking for help in a training procedure).	Possible overuse of affordances in the IVR experience could reflect in lower learning transfer.	[26], [56]
IVR intervention has effects similar to those of less immersive mediums on learning.	Assessment results are found without a significant difference in favor of IVR-designed experiences.	[26], [49], [51], [59], [61], [63]

TABLE 10. Reported advantages on user performance of the IVR-designed experiences.

Advantages	Description	Papers
Visualization of 3D content is usually taught in 2D (e.g., textbooks and slides) or complex to replicate in real-world scenarios.	The authors developed 3D models' visualization to provide multiple perspectives of the subject.	[11], [39], [43], [45], [57], [62], [63], [90]
Control-based input interaction is an intuitive way to explore the IVR environment.	IVR developed experience opted to use control-based input that is easy to use and more intuitive to understand than body-based input.	[52], [62]
Usable and easy-to-understand designed IVR environment.	The participants understood the proposed affordances and the expected interactions.	[45], [50], [53]–[55]
UI menus and instructions information.	The information included can be perceived correctly and usable by the participants.	[53], [58]
Integrated feedback.	Participants can contextualize their actions and experience direct consequences from errors or interactions.	[41]
Simulation realism.	The virtual representation of the matched participants' real experiences accurately portrayed the expected results.	[41], [56]

collaboration, critical thinking, and mental model development. Studies like Hácha et al. [39] and Franzluebbers et al. [54] demonstrated the collaborative potential of IVR by enabling students to interact with 3D models and virtual equipment alongside peers and instructors. However, limitations like lack of voice chat affected engagement. Pirker et al. [62] highlighted IVR's ability to foster critical thinking through immersive simulations of complex physics concepts, emphasizing the importance of interaction for enhancing learning transfer. Regarding mental models, Slezaka et al. [90] and Bagher et al [59] showed how interactive tasks in IVR can deepen understanding, with students achieving higher cognitive engagement and constructing more effective mental models in immersive environments.

TABLE 11. Reported disadvantages in the user experience of the IVR designed experiences.

Disadvantages	Description	Papers
Integrated voice communication in a networked environment.	Authors implemented the option to collaborate in the IVR environment, even though voice communication was not supported.	[39]
Lectures in VR suppress note-taking.	IVR lectures did not allow taking notes, so the students were limited to visualizing and following up on the professor's actions.	[39], [62]
Movement.	Participants reported confusion and difficulties moving in the virtual environment using teleportation.	[54]
Reading text on VR.	Longer text or smaller labels included in the IVR environment could be challenging for the participants to read—overlays between the defined content.	[54], [58], [64]
Lack of guidance labels and instructions.	The IVR experience simulated procedures need to be guided more so that they can be comprehended during the initial interactions for the students. Special attention is needed for novice VR users.	[43], [50], [53]
Complex interactions (e.g., gestures) can be poorly used.	Implemented features can be avoided due to the complexity; there is a noticed preference for easy actions.	[58], [64]

As highlighted in Table 8, multiple studies reported higher learning outcomes with IVR tools compared to non-immersive mediums, demonstrating the effectiveness of their designs. For instance, M. Lui et al. [60] developed an IVR application for understanding a gene regulation system that included interactive assembly tasks, teleportation, and dynamic animations. Vogt et al. [61] designed an IVR experience featuring a robot assistant with audio narration, scene teleportation, and passive viewing. Similarly, Checa et al. [45] created an IVR tool for computer hardware learning that incorporated an assembly task, guided instruction, and a virtual instructor. Miller et al. [46] explored organic chemistry learning, integrating assembly tasks, guided instructions, and immersive ambient scenes. Pirker et al. [62] developed a laboratory-based electromagnetism simulation, which included parameter control, teleportation, and a networked environment. Wang et al. [50] focused on computer hardware learning, featuring assembly tasks and interactive components. Across these customized IVR experiences, including embodied interactions, such as assembly tasks requiring hand movements and 3D spatial recognition, proved critical. Engaging with hardware equipment, molecular structures, and other elements in the virtual environment fostered a stronger sense of agency and presence, resulting in significantly higher learning outcomes than non-immersive solutions like desktop applications and slideshow presentations.

This review highlights the main advantages and disadvantages of IVR experiences in user studies. Key advantages include an enhanced understanding of complex procedures, the development of mental models, and increased engagement and motivation, often leading to higher learning outcomes in immersive conditions. Disadvantages include the need for guidance for novice users, cognitive overload, and a possible need for multimodal feedback, such as haptic or audio elements [80], [96]. Usability benefits include intuitive control-based input interactions and realistic simulations, while challenges involve problems when reading text on VR, issues with notetaking, lack of enough guidance, and motion sickness [53]. A comprehensive analysis from learning and usability perspectives provides guidelines for future IVR designs, emphasizing adopting mixed methods for constructive feedback and enhanced virtual learning environments [45], [97].

A. PRACTICAL IMPLICATIONS

Based on methodologies, development pipelines, design features, and user studies of customized IVR experiences for STEM learning, we propose the following steps to consider when designing and developing these experiences using high-end HMD capabilities.

- **Step #1—Assessing the need for IVR:** Before creating IVR experiences, analyzing why the lesson requires an immersive experience using HMD is essential. While customized IVR experiences offer unique advantages (see Section V-E), alternative solutions such as desktops or third-party tools should also be considered. The decision to use IVR should be justified by the student's needs and potential benefits, such as simulating complex procedures, promoting mental models, or enhancing engagement, motivation, and sense of presence.
- **Step #2—Delimiting developer tools and expertise:** Once the decision to use VR is made, appropriate resources must be selected. The choice of HMD should align with the lesson's objectives, teaching methods, and classroom sizes, with devices like Meta Quest or HTC Vive as preferable for delivering high immersion levels. Other considerations, such as game engines (e.g., Unity or Unreal) and resources like 3D models, textures, and UI designs, are crucial for creating immersive environments. Depending on the lesson's complexity, programming skills may be required, particularly for interactive or simulated STEM concepts. Collaborative work across disciplines, especially in software engineering, is recommended to design and develop adaptable IVR experiences. An example pipeline is presented by [53], who used a Unity game engine with several built-in and external packages and included CFD pre-computed data for the visualization and targeted to deploy on Meta Quest 2 HMD.

- **Step #3—Instructional design and expected embodiment degree:** Planning students' activities in the IVR environment is crucial, especially for self-instructed tasks. We suggest that our proposed conceptual framework (see Figure 1) can help to outline the different considerations for the customized IVR experience. To enhance embodied cognition, the embodiment framework ensures high sensory motor sensations, adequate immersion, and gestural congruency. Most reviewed examples show positive effects on user experiences and learning effectiveness, suggesting a focus on the third and fourth degrees of embodiment. However, including complex interactions, multi-modal feedback, or social immersion increases development complexity, so features should align with instructional goals and development scope. Providing clear instructions, tutorial scenes, and explanations of expected interactions is essential, especially for novice users.
- **Step #4—Implementing and using the IVR experience:** The protocol for delivering IVR lessons is essential to this instruction. Designers and instructors should understand equipment configuration, student support during the intervention, space requirements (e.g., seated vs. standing experiences [20]), and enough guidance for device actions like controllers' buttons, teleportation, or complex gestures.
- **Step #5—Assessments and experience validations:** Customized IVR experiences can always be improved. Including assessments to validate lesson effectiveness, usability metrics, and user experience perceptions can help refine the tool. Assessments should be considered in any target learning type, such as declarative or procedural learning, to evaluate if embodied interactions aid students. Mixed methods can provide deeper insights into students' experiences and performance. Application logs, such as points of attention, element interactions, and task completion times, are recommended to monitor the lesson's length and feature relevance.

B. RESEARCH GAPS AND FUTURE DIRECTIONS

The reviewed literature highlights several limitations and offers insights for improving customized IVR tools and their application in education. Authors have emphasized aspects such as immersion, sense of presence, engagement, cognitive load, and learning outcomes to validate and demonstrate the effectiveness of their designs. However, several areas remain open for future exploration:

- **Addressing the Reported Disadvantages:** IVR has limitations, as with any educational tool. Future studies should address the drawbacks highlighted in Tables 9 and 11, such as challenges with note-taking in large VR-based lectures or difficulties in reading tasks and text formatting in immersive environments. Researchers should focus on redesigning these elements to mitigate their impact or conduct further studies to determine whether these issues are common across various IVR learning scenarios.
- **Increasing Sample Sizes:** Many authors have cited small sample sizes as a limitation, which poses challenges for achieving statistically significant results. Future research should involve larger samples to validate findings and ensure broader applicability. For example, unlike the small sample size in [39], which only involved four students, more extensive studies in immersive classrooms could provide more robust evidence of IVR's educational benefits.
- **Testing Customized Tools on a Variety of Learning Outcomes:** Although many IVR tools have been designed to target specific learning outcomes, their features and visualizations may be adaptable to broader educational objectives. Future research could explore how these tools perform across different learning domains, whether for declarative knowledge, procedural skills, or conceptual understanding. For instance, Pirker et al. [62] applied their MaroonVR tool in multiple contexts, showing its potential to address varied learning needs.
- **Focusing on Learning as the Primary Objective:** While many studies emphasize usability and user experience, future research should prioritize learning outcomes as the central goal of IVR design for STEM education. Evaluating knowledge acquisition, skill development, and the transfer of learning should take precedence, ensuring that IVR contributes to meaningful educational results. Many reviewed papers (50%) focus on usability without fully exploring how these tools impact learning. Addressing this gap is crucial for practitioners integrating IVR into STEM curricula.

C. LIMITATIONS

The review has some limitations. Exploring trends from 2016 to 2023, we encountered the exclusion of papers before 2018. We delimited strict criteria that may result in the exclusion of papers before 2018. However, recommendations suggest using individual year searches on databases to include more papers per specified year. Another limitation is exclusively selecting publications discussing their development and design of IVR experiences. Expecting authors to develop IVR applications was ambitious, leading to the elimination of publications evaluating IVR in learning due to using third-party solutions. However, these solutions may be justified as an option for designers/authors without programming experience.

Another criterion was using high-end HMDs in papers, excluding publications attempting an immersive environment with other HMD types like mobile alternatives (e.g., Google Cardboard or Samsung Gear VR), which were considered more accessible for educational settings. Regardless, in this review, we aimed to analyze VR experiences exploiting the potential of current VR technology (advanced devices).

Notably, well-known VR conferences like IEEE VR and IEEE ISMAR were underrepresented, possibly due to keywords or conference topics not aligning with education. However, when we look at the IEEE VR 2024 call for papers (https://ieeevr.org/2024/contribute/papers/), there are no topics on teaching or learning. In this sense, the VR community might focus on advances in hardware and software innovations and perceptual studies rather than understanding how VR can be used for educational purposes. That scope differs from the more frequent journals (BJTE and JCAL) between the included articles that aim to examine the use of technologies to support learning, teaching, instructional design, and development and to demonstrate whether and how applications lead to improvements in formal and non-formal education at all levels.

As we are aware, the development of IVR experiences is a challenging task. However, the discussions around the development paths used allowed an understanding of the complexity of developing this type of application. The findings can serve as a starting point for new designers and developers. Finally, as a limitation in the sense of the results and conclusions presented, these should be carefully analyzed due to the existence of multiple papers in which the authors reported among their limitations that their findings come from a small sample. We have included and combined the findings with other proposed designs (with a larger sample size). Nevertheless, we observed patterns in both articles, so it is possible to generalize the results regarding adopting IVR lessons for STEM learning.

VII. CONCLUSION

The rise of consumer VR devices (e.g., HTC Vive, Meta Quest) offers alternatives to embedded IVR technology in education, which is increasingly adopted as a classroom learning tool. In recent years, as reported by the reviewed papers from 2019-2023, IVR-designed environments have become a trend in leveraging immersive experiences for STEM learning concepts, particularly in science. High-end HMDs offer key affordances, enabling teaching topics traditionally conveyed through less immersive media like videos or slides—papers assessing learning outcomes predominantly focused on declarative and procedural learning approaches. The review highlighted advantages in learning and usability, enhancing learning outcomes, motivation, engagement, mental model building, and usable design, with intuitive interactions and congruence in included gestures. Experiences with higher embodiment levels (fourth and third degree) have shown several advantages in enhancing student learning performance, indicating that offering immersive interactions and HMD isolations can be an effective learning method. Regarding development, the authors primarily utilized the Unity game engine due to its capabilities and frameworks for developing IVR experiences, including graphical capabilities and hardware integration support. The authors discussed their development procedures and the software toolkits for implementing IVR tools. Although IVR development is not a straightforward process, we encourage researchers to explore the possibility of developing customized IVR experiences tailored to learning topics, considering learners' needs and the possibilities VR offers. We suggest using the proposed conceptual frameworks to guide the creation of customized IVR experiences for future work. However, this is not limited to the dimensions considered; other aspects, such as interactivity, cognitive load, and specific learning objectives, should be considered for more refined designs.

ACKNOWLEDGMENT

Any opinions, findings, conclusions, or recommendations expressed in this material are those of the author(s) and do not necessarily reflect those of the National Science Foundation. The publication of this article was funded in part by the Purdue University Libraries Open Access Publishing Fund.

REFERENCES

[1] B. Freeman, S. Marginson, and R. Tytler, "An international view of STEM education," in *STEM Education 2.0*. Leiden, The Netherlands: Brill, 2019, pp. 350–363.

[2] Nat. Sci. Board. (2022). *Science and Engineering Indicators 2022: The State of U.S. Science and Engineering*. Accessed: Oct. 18, 2023. [Online]. Available: https://ncses.nsf.gov/pubs/nsb20221

[3] Australian Government. (2022). *National Stem School Education Strategy 2016–2026*. Accessed: Oct. 18, 2023. [Online]. Available: https://www.education.gov.au/australian-curriculum/support-science-technology-engineering-and-mathematics-stem/national-stem-school-education-strategy-2016-2026

[4] Ministry Edu. Culture, Finland. (2023). *Finnish National Stem Strategy and Action Plan: Experts in Natural Sciences, Technology and Mathematics in Support of Society's Welfare and Growth*. Accessed: Oct. 18, 2023. [Online]. Available: https://julkaisut.valtioneuvosto.fi/handle/10024/164953

[5] *Education at a Glance 2023*, OECD, Paris, France, 2023.

[6] Q. Jin, Y. Liu, S. Yarosh, B. Han, and F. Qian, "How will VR enter university classrooms? Multi-stakeholders investigation of VR in higher education," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, New York, NY, USA, Apr. 2022, pp. 1–17.

[7] J. J. Gibson, *The Ecological Approach to Visual Perception*. London, U.K.: Psychology Press, Nov. 2014.

[8] C. Heeter, "Being there: The subjective experience of presence," *Presence: Teleoperators Virtual Environments*, vol. 1, no. 2, pp. 262–271, Jan. 1992.

[9] M. C. Johnson-Glenberg, "The necessary nine: Design principles for embodied VR and active stem education," in *Proc. Learning in a Digital World*. Singapore: Springer, 2019, pp. 83–112.

[10] D. Checa and A. Bustillo, "Advantages and limits of virtual reality in learning processes: Briviesca in the fifteenth century," *Virtual Reality*, vol. 24, no. 1, pp. 151–161, Mar. 2020.

[11] W. Li, J. Zhu, P. Dang, J. Wu, J. Zhang, L. Fu, and Q. Zhu, "Immersive virtual reality as a tool to improve bridge teaching communication," *Expert Syst. Appl.*, vol. 217, May 2023, Art. no. 119502.

[12] J. Jerald, *The VR Book: Human-Centered Design for Virtual Reality*. San Rafael, CA, USA: Morgan & Claypool, 2015.

[13] D. Beck, L. Morgado, and P. O'Shea, "Educational practices and strategies with immersive learning environments: Mapping of reviews for using the metaverse," *IEEE Trans. Learn. Technol.*, vol. 17, pp. 319–341, 2024, doi: 10.1109/TLT.2023.3243946.

[14] A. L. C. Lui, C. Not, and G. K. W. Wong, "Theory-based learning design with immersive virtual reality in science education: A systematic review," *J. Sci. Educ. Technol.*, vol. 32, no. 3, pp. 390–432, Jun. 2023.

[15] M. Coban, Y. I. Bolat, and I. Goksu, "The potential of immersive virtual reality to enhance learning: A meta-analysis," *Educ. Res. Rev.*, vol. 36, Jun. 2022, Art. no. 100452.

[16] J. Radianti, T. A. Majchrzak, J. Fromm, and I. Wohlgenannt, ''A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda,'' *Comput. Educ.*, vol. 147, Apr. 2020, Art. no. 103778.

[17] N. Pellas, A. Dengel, and A. Christopoulos, ''A scoping review of immersive virtual reality in STEM education,'' *IEEE Trans. Learn. Technol.*, vol. 13, no. 4, pp. 748–761, Oct. 2020.

[18] S. Kavanagh, A. Luxton-Reilly, B. Wuensche, and B. Plimmer, ''A systematic review of virtual reality in education,'' *Themes Sci. Technol. Educ.*, vol. 10, no. 2, pp. 85–119, 2017.

[19] M. Won, D. A. K. Ungu, H. Matovu, D. F. Treagust, C.-C. Tsai, J. Park, M. Mocerino, and R. Tasker, ''Diverse approaches to learning with immersive virtual reality identified from a systematic review,'' *Comput. Educ.*, vol. 195, Apr. 2023, Art. no. 104701.

[20] M. Lui, K.-Y.-A. Chong, M. Mullally, and R. McEwen, ''Facilitated model-based reasoning in immersive virtual reality: Meaning-making and embodied interactions with dynamic processes,'' *Int. J. Comput.-Supported Collaborative Learn.*, vol. 18, no. 2, pp. 203–230, Jun. 2023.

[21] M. C. Johnson-Glenberg and C. Megowan-Romanowicz, ''Embodied science and mixed reality: How gesture and motion capture affect physics education,'' *Cognit. Res., Princ. Implications*, vol. 2, no. 1, p. 4, May 2017, doi: 10.1186/s41235-017-0060-9.

[22] C. J. Dede, J. Jacobson, and J. Richards, ''Introduction: Virtual, augmented, and mixed realities in education,'' in *Smart Computing and Intelligence*. Singapore: Springer, 2017, pp. 1–16.

[23] C. Dede, ''Immersive interfaces for engagement and learning,'' *Science*, vol. 323, no. 5910, pp. 66–69, Jan. 2009.

[24] R. J. Shavelson, D. B. Young, C. C. Ayala, P. R. Brandon, E. M. Furtak, M. A. Ruiz-Primo, M. K. Tomita, and Y. Yin, ''On the impact of curriculum-embedded formative assessment on learning: A collaboration between curriculum and assessment developers,'' *Appl. Meas. Educ.*, vol. 21, no. 4, pp. 295–314, Sep. 2008.

[25] J. Chatain, M. Kapur, and R. W. Sumner, ''Three perspectives on embodied learning in virtual reality: Opportunities for interaction design,'' in *Proc. Extended Abstr. CHI Conf. Hum. Factors Comput. Syst.*, Apr. 2023, pp. 1–8.

[26] M. C. Johnson-Glenberg, H. Bartolomea, and E. Kalina, ''Platform is not destiny: Embodied learning effects comparing 2D desktop to 3D virtual reality STEM experiences,'' *J. Comput. Assist. Learn.*, vol. 37, no. 5, pp. 1263–1284, Oct. 2021.

[27] J. Chatain, V. Ramp, V. Gashaj, V. Fayolle, M. Kapur, R. W. Sumner, and S. Magnenat, ''Grasping derivatives: Teaching mathematics through embodied interactions using tablets and virtual reality,'' in *Proc. Interact. Design Children*, vol. 10, Jun. 2022, pp. 98–108.

[28] J. Stull, S. J. Varnum, J. Ducette, and J. Schiller, ''The many faces of formative assessment.,'' *Int. J. Teaching Learn. Higher Educ.*, vol. 23, no. 1, pp. 30–39, 2011.

[29] M. Grangeat, C. Harrison, and J. Dolin, ''Exploring assessment in STEM inquiry learning classrooms,'' *Int. J. Sci. Educ.*, vol. 43, no. 3, pp. 345–361, Feb. 2021.

[30] C. Redecker and O. Johannessen, ''Changing assessment—Towards a new assessment paradigm using ict,'' *Eur. J. Educ.*, vol. 48, pp. 79–96, Feb. 2013.

[31] K. K. Bhagat, W.-K. Liou, J. M. Spector, and C.-Y. Chang, ''To use augmented reality or not in formative assessment: A comparative study,'' *Interact. Learn. Environ.*, vol. 27, nos. 5–6, pp. 830–840, Jun. 2018.

[32] A. S. Oyelere, F. J. Agbo, and S. S. Oyelere, ''Formative evaluation of immersive virtual reality expedition mini-games to facilitate computational thinking,'' *Comput. Educ., X Reality*, vol. 2, 2023, Art. no. 100016.

[33] M. J. Page et al., ''The PRISMA 2020 statement: An updated guideline for reporting systematic reviews,'' *BMJ*, vol. 372, Mar. 2021, Art. no. 372: n71.

[34] W. R. Sherman and A. B. Craig, *Understanding Virtual Reality*. Amsterdam, The Netherlands: Elsevier, 2019.

[35] B. Alexander, K. Ashford-Rowe, N. Barajas-Murph, G. Dobbin, J. Knott, M. McCormack, J. Pomerantz, R. Seilhamer, and N. Weber, Eds., *EDUCAUSE Horizon Report: 2019 Higher Education Edition*. Louisville, CO, USA, 2019. [Online]. Available: https://library.educause.edu/-/media/files/library/2019/4/2019horizonreport.pdf?la=en&hash=C8E8D444AF372E705FA1BF9D4FF0DD4CC6F0FDD1

[36] J. A. C. Sterne et al., ''RoB 2: A revised tool for assessing risk of bias in randomised trials,'' *BMJ*, vol. 366, p. l4898, Aug. 2019.

[37] J. P. T. Higgins et al., ''A tool to assess risk of bias in non-randomized follow-up studies of exposure effects (ROBINS-E),'' *Environ. Int.*, vol. 186, Apr. 2024, Art. no. 108602.

[38] O. R. Holsti, *Content Analysis for the Social Sciences and Humanities* (content analysis). Reading, MA, USA: Addison-Wesley, 1969.

[39] F. Hácha, P. Vanecek, and L. Vása, ''A virtual reality platform for immersive education in computer graphics,'' in *Eurographics 2021—Education Papers*, S. B. Santos and G. Domik, Eds. The Eurographics Association, 2021, doi: 10.2312/eged.20211001.

[40] K.-S. Tang, M. Won, M. Mocerino, D. F. Treagust, and R. Tasker, ''Multimodal affordances of immersive virtual reality for visualising and learning molecular interactions,'' in *Learning From Animations in Science Education*. Cham, Switzerland: Springer, 2020, pp. 75–100.

[41] A. Arntz, S. C. Eimler, D. Keßler, J. Thomas, A. Helgert, M. Rehm, E. Graf, S. Wientzek, and B. Budur, ''Walking on the bright sight: Evaluating a photovoltaics virtual reality education application,'' in *Proc. IEEE Int. Conf. Artif. Intell. Virtual Reality (AIVR)*, Nov. 2021, pp. 295–301.

[42] C.-C. Lin, B.-Y. Cheng, and R.-C. Shih, ''Developing an immersive virtual reality-assisted learning system to support scanning electron microscopy learning activities,'' in *Proc. Int. Conf. Innov. Technol. Learn.*, in Lecture Notes in Computer Science. Cham, Switzerland: Springer, 2023, pp. 494–500.

[43] H. Sulaiman, S. N. Apandi, A. M. Yusof, R. A. Latif, and N. Ibrahim, ''Evaluation of a virtual reality VR learning tool for fundamental turbine engineering concepts,'' in *Advances in Visual Informatics*. Cham, Switzerland: Springer, 2019, pp. 48–59.

[44] Y. Lin, G. Wang, and A. Suh, ''Exploring the effects of immersive virtual reality on learning outcomes: A two-path model,'' in *Augmented Cognition. Human Cognition and Behavior*. Cham, Switzerland: Springer, 2020, pp. 86–105.

[45] D. Checa, I. Miguel-Alonso, and A. Bustillo, ''Immersive virtual-reality computer-assembly serious game to enhance autonomous learning,'' *Virtual Reality*, vol. 27, no. 4, pp. 3301–3318, Dec. 2023.

[46] M. D. Miller, G. Castillo, N. Medoff, and A. Hardy, ''Immersive VR for organic chemistry: Impacts on performance and grades for first-generation and continuing-generation university students,'' *Innov. Higher Educ.*, vol. 46, no. 5, pp. 565–589, Oct. 2021.

[47] P. Adami, R. Singh, P. Borges Rodrigues, B. Becerik-Gerber, L. Soibelman, Y. Copur-Gencturk, and G. Lucas, ''Participants matter: Effectiveness of VR-based training on the knowledge, trust in the robot, and self-efficacy of construction workers and university students,'' *Adv. Eng. Informat.*, vol. 55, Jan. 2023, Art. no. 101837.

[48] G. B. Petersen, A. Mottelson, and G. Makransky, ''Pedagogical agents in educational VR: An in the wild study,'' in *Proc. CHI*, New York, NY, USA, 2021.

[49] A. Vogt, F. Babel, P. Hock, M. Baumann, and T. Seufert, ''Prompting in-depth learning in immersive virtual reality: Impact of an elaboration prompt on developing a mental model,'' *Comput. Educ.*, vol. 171, Oct. 2021, Art. no. 104235.

[50] W.-S. Wang, M. Pedaste, and Y.-M. Huang, ''The influence of emotion in stem activity based on virtual reality learning environment,'' in *Proc. Int. Conf. Innov. Technol. Learn.*, in Lecture Notes in Computer Science. Cham, Switzerland: Springer, 2023, pp. 484–493.

[51] S. Q. Lima, B. Balakrishnan, and J. B. Kim, ''Virtual construction simulation: Evaluating impact of immersion and interactivity on novice designers,'' in *Proc. Int. Conf. Hum.-Comput. Interact.*, in Lecture Notes in Computer Science. Cham, Switzerland: Springer, 2023, pp. 678–690.

[52] A. Santos-Torres, T. Zarraonandia, P. Díaz, T. Onorati, and I. Aedo, ''An empirical comparison of interaction styles for map interfaces in immersive virtual environments,'' *Multimedia Tools Appl.*, vol. 79, nos. 47–48, pp. 35717–35738, Dec. 2020.

[53] S. Solmaz, L. Kester, and T. Van Gerven, ''An immersive virtual reality learning environment with CFD simulations: Unveiling the virtual garage concept,'' *Educ. Inf. Technol.*, vol. 29, no. 2, pp. 1455–1488, Feb. 2024.

[54] A. Franzluebbers, A. J. Tuttle, K. Johnsen, S. Durham, and R. Baffour, ''Collaborative virtual reality training experience for engineering land surveying,'' in *Cross Reality and Data Science in Engineering*. Cham, Switzerland: Springer, 2021, pp. 411–426.

[55] J. Qian, Y. Ma, Z. Pan, and X. Yang, ''Effects of virtual-real fusion on immersion, presence, and learning performance in laboratory education,'' *Virtual Reality Intell. Hardw.*, vol. 2, no. 6, pp. 569–584, Dec. 2020.

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

[57] J. Collins, H. Regenbrecht, and T. Langlotz, "Expertise and experience in VR-supported learning: Achieving a deep non-verbal comprehension of four-dimensional space," *Int. J. Hum.-Comput. Stud.*, vol. 152, Aug. 2021, Art. no. 102649.

[58] K. Lupinetti, B. Bonino, F. Giannini, and M. Monti, "Exploring the benefits of the virtual reality technologies for assembly retrieval applications," in *Augmented Reality, Virtual Reality, and Computer Graphics* (Lecture Notes in Computer Science). Cham, Switzerland: Springer, 2019, pp. 43–59.

[59] M. M. Bagher, P. Sajjadi, J. Carr, P. La Femina, and A. Klippel, "Fostering penetrative thinking in geosciences through immersive experiences: A case study in visualizing earthquake locations in 3D," in *Proc. 6th Int. Conf. Immersive Learn. Res. Netw. (iLRN)*, Jun. 2020, pp. 132–139.

[60] M. Lui, R. McEwen, and M. Mullally, "Immersive virtual reality for supporting complex scientific knowledge: Augmenting our understanding with physiological monitoring," *Brit. J. Educ. Technol.*, vol. 51, no. 6, pp. 2181–2199, Nov. 2020.

[61] A. Vogt, F. Babel, P. Hock, M. Baumann, and T. Seufert, "Immersive virtual reality or auditory text first? Effects of adequate sequencing and prompting on learning outcome," *Brit. J. Educ. Technol.*, vol. 52, no. 5, pp. 2058–2076, Sep. 2021.

[62] J. Pirker, M. Holly, I. Lesjak, J. Kopf, and C. Gütl, "MaroonVR—An interactive and immersive virtual reality physics laboratory," in *Smart Computing and Intelligence*. Singapore: Springer, 2019, pp. 213–238.

[63] M. Marinelli, S. A. Male, A. Valentine, A. Guzzomi, T. van der Veen, and G. M. Hassan, "Using VR to teach safety in design: What and how do engineering students learn?" *Eur. J. Eng. Educ.*, vol. 48, no. 3, pp. 538–558, May 2023.

[64] S. Try, K. Panuwatwanich, G. Tanapornraweekit, and M. Kaewmoracharoen, "Virtual reality application to aid civil engineering laboratory course: A multicriteria comparative study," *Comput. Appl. Eng. Educ.*, vol. 29, no. 6, pp. 1771–1792, Nov. 2021.

[65] J. Brooke, "SUS-A quick and dirty usability scale," *Usability Eval. Ind.*, vol. 189, no. 194, pp. 4–7, 1996.

[66] J. R. Lewis, "Item benchmarks for the system usability scale," *J. Usability Stud.*, vol. 13, no. 3, pp. 158–167, 2018.

[67] M. Schrepp, A. Hinderks, and J. Thomaschewski, "Construction of a benchmark for the user experience questionnaire (UEQ)," *Int. J. Interact. Multimedia Artif. Intell.*, vol. 4, no. 4, p. 40, 2017.

[68] C. Pletz and B. Zinn, "Technology acceptance of virtual learning and working environments in technical domains," *J. Tech. Educ.*, vol. 6, no. 4, pp. 86–105, 2018.

[69] P. Caserman, A. Garcia-Agundez, A. Gámez Zerban, and S. Göbel, "Cybersickness in current-generation virtual reality head-mounted displays: Systematic review and outlook," *Virtual Reality*, vol. 25, no. 4, pp. 1153–1170, Dec. 2021.

[70] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal, "Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness," *Int. J. Aviation Psychol.*, vol. 3, no. 3, pp. 203–220, Jul. 1993.

[71] M. Usoh, E. Catena, S. Arman, and M. Slater, "Using presence questionnaires in reality," *Presence*, vol. 9, no. 5, pp. 497–503, Oct. 2000.

[72] J. N. Bailenson, K. Swinth, C. Hoyt, S. Persky, A. Dimov, and J. Blascovich, "The independent and interactive effects of embodied-agent appearance and behavior on self-report, cognitive, and behavioral markers of copresence in immersive virtual environments," *Presence*, vol. 14, no. 4, pp. 379–393, Aug. 2005.

[73] C. Jennett, A. L. Cox, P. Cairns, S. Dhoparee, A. Epps, T. Tijs, and A. Walton, "Measuring and defining the experience of immersion in games," *Int. J. Hum.-Comput. Stud.*, vol. 66, no. 9, pp. 641–661, Sep. 2008.

[74] P. Vorderer, W. Wirth, F. R. Gouveia, F. Biocca, T. Saari, L. Jäncke, S. Böcking, H. Schramm, A. Gysbers, and T. Hartmann, "Mec spatial presence questionnaire," *Retrieved Sep.*, vol. 18, p. 6, Jan. 2004.

[75] O. D. Kothgassner, A. Felnhofer, N. Hauk, E. Kastenhofer, J. Gomm, and I. Kryspin-Exner, *TUI: Technology Usage Inventory Manual*. FFG Forschungsförderungsgesellschaft, 2012.

[76] S. G. Hart and L. E. Staveland, "Development of NASA-TLX (task load index): Results of empirical and theoretical research," in *Human Mental Workload* (Advances in Psychology), vol. 52, P. A. Hancock and N. Meshkati, Eds., Amsterdam, The Netherlands: North Holland, 1988, pp. 139–183.

[77] G. Cierniak, K. Scheiter, and P. Gerjets, "Explaining the split-attention effect: Is the reduction of extraneous cognitive load accompanied by an increase in germane cognitive load?" *Comput. Hum. Behav.*, vol. 25, no. 2, pp. 315–324, Mar. 2009.

[78] J. H. Brockmyer, C. M. Fox, K. A. Curtiss, E. McBroom, K. M. Burkhart, and J. N. Pidruzny, "The development of the game engagement questionnaire: A measure of engagement in video game-playing," *J. Experim. Social Psychol.*, vol. 45, no. 4, pp. 624–634, Jul. 2009.

[79] H. L. O'Brien, P. Cairns, and M. Hall, "A practical approach to measuring user engagement with the refined user engagement scale (UES) and new UES short form," *Int. J. Hum.-Comput. Stud.*, vol. 112, pp. 28–39, Apr. 2018.

[80] G. Makransky and L. Lilleholt, "A structural equation modeling investigation of the emotional value of immersive virtual reality in education," *Educ. Technol. Res. Develop.*, vol. 66, no. 5, pp. 1141–1164, Oct. 2018.

[81] D. Derryberry and M. A. Reed, "Anxiety-related attentional biases and their regulation by attentional control.," *J. Abnormal Psychol.*, vol. 111, no. 2, pp. 225–236, 2002.

[82] L. W. Anderson and D. R. Krathwohl, *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives: Complete Edition*. Reading, MA, USA: Addison-Wesley, 2001.

[83] E. Ai-Lim Lee, K. W. Wong, and C. C. Fung, "How does desktop virtual reality enhance learning outcomes? A structural equation modeling approach," *Comput. Educ.*, vol. 55, no. 4, pp. 1424–1442, Dec. 2010.

[84] F. C. Breslin, L. C. Sobell, M. B. Sobell, and S. Agrawal, "A comparison of a brief and long version of the situational confidence questionnaire," *Behav. Res. Therapy*, vol. 38, no. 12, pp. 1211–1220, Dec. 2000.

[85] G. Makransky and G. B. Petersen, "Investigating the process of learning with desktop virtual reality: A structural equation modeling approach," *Comput. Educ.*, vol. 134, pp. 15–30, Jun. 2019.

[86] P. R. Pintrich, "A manual for the use of the motivated strategies for learning questionnaire (MSLQ)," Nat. Center Res. Improve Postsecondary Teach. Learn., Office Educ. Res. Improvement (ED), Washington, DC, USA, Tech. Rep., 1991.

[87] R. B. Ekstrom, *Kit of Factor-Referenced Cognitive Tests*. Princeton, NJ, USA: Educational Testing Service, 1976.

[88] D. Liepmann, *Intelligenz-Struktur-Test 2000 R: IST 2000 R*. Göttingen, Germany: Hogrefe, 2007.

[89] D. K. May, "Mathematics self-efficacy and anxiety questionnaire," Ph.D. thesis, Univ. Georgia, 2009.

[90] R. J. Slezaka, N. Keren, S. B. Gilbert, M. E. Harvey, S. J. Ryan, and A. J. Wiley, "Examining virtual reality as a platform for developing mental models of industrial systems," *J. Comput. Assist. Learn.*, vol. 39, no. 1, pp. 113–124, Feb. 2023.

[91] S. J. Derry, R. D. Pea, B. Barron, R. A. Engle, F. Erickson, R. Goldman, R. Hall, T. Koschmann, J. L. Lemke, M. G. Sherin, and B. L. Sherin, "Conducting video research in the learning sciences: Guidance on selection, analysis, technology, and ethics," *J. Learn. Sci.*, vol. 19, no. 1, pp. 3–53, Jan. 2010.

[92] K. M. Stefanski, G. E. Gardner, and R. L. Seipelt-Thiemann, "Development of aLacOperon concept inventory (LOCI)," *CBE—Life Sci. Educ.*, vol. 15, no. 2, p. ar24, Jun. 2016.

[93] *Safe Design for Engineering Students. Barton: Commonwealth of Australia*, Australian Government-Dept. Educ. Employment Workplace Relations, 2006.

[94] H. Matovu, D. A. K. Ungu, M. Won, C.-C. Tsai, D. F. Treagust, M. Mocerino, and R. Tasker, "Immersive virtual reality for science learning: Design, implementation, and evaluation," *Stud. Sci. Educ.*, vol. 59, no. 2, pp. 205–244, Jul. 2023.

[95] T. M. Duffy, "Constructivism: Implications for the design and delivery of instruction," in *Handbook of Research on Educational Communications and Technology*. New York, NY, USA: Simon & Schuster, 1996, pp. 170–198.

[96] R. Moreno and R. E. Mayer, "Personalized messages that promote science learning in virtual environments," *J. Educ. Psychol.*, vol. 96, no. 1, pp. 165–173, Mar. 2004.

[97] K. Bucher, T. Blome, S. Rudolph, and S. von Mammen, "VReanimate II: Training first aid and reanimation in virtual reality," *J. Comput. Educ.*, vol. 6, no. 1, pp. 53–78, Mar. 2019.

PEDRO ACEVEDO received the B.S. and M.S. degrees in software engineering from Universidad del Norte, Colombia, in 2020 and 2021, respectively. He is currently pursuing the Ph.D. degree with the Department of Computer Graphics Technology, Purdue University. His research interests include virtual reality, human–computer interaction, procedural content generation, games, and wireless sensor networks.

ALEJANDRA J. MAGANA received the Ph.D. degree in engineering education from Purdue University, West Laffayette, IN, USA, in 2009. She is currently the W.C. Furnas Professor in enterprise excellence with the Department of Computer and Information Technology and a Professor with the School of Engineering Education, Purdue University. Her research program investigates how model-based cognition in STEM can be better supported by computational disciplinary practices, such as computation, modeling and simulation, data science practices, and artificial intelligence. In 2016, she was conferred the status of Purdue Faculty Scholar for being on an accelerated path toward academic distinction. In 2022, she was inducted into the Purdue University Teaching Academy, recognizing her excellence in teaching. She received the Faculty Early Career Development (CAREER) Award and the National Science Foundation's Most Prestigious Award for Faculty who exemplify the role of Teacher-Scholar, in 2015.

BEDRICH BENES (Senior Member, IEEE) received the M.S. and Ph.D. degrees in computer science from Czech Technical University in Prague, Prague, Czechia, in 1991 and 1998, respectively. He is currently a Professor of computer graphics technology with Purdue University and computer science (by courtesy), West Lafayette, IN, USA. He is also the Director of the High Performance Computer Graphics Laboratory, Purdue University. He is known for his work in geometric modeling, procedural modeling, scientific visualization, and software optimizations in digital manufacturing. He was a member of numerous program committees of various conferences, including ACM Siggraph and Eurographics and he was the Paper Chair of Eurographics, in 2017. He is an Editor of *Computer Graphics Forum* and an Associate Editor of *Computers and Graphics*.

CHRISTOS MOUSAS (Member, IEEE) received the five-year integrated master's (B.Sc. and M.Sc.) degrees in audiovisual science and art from the Department of Audio and Visual Arts, Ionian University, Greece, in 2009, and the M.Sc. degree in multimedia applications and virtual environments and the Ph.D. degree in informatics from the Department of Informatics, School of Engineering and Informatics, University of Sussex, in 2011 and 2014, respectively. From 2015 to 2016, he was a Postdoctoral Researcher with the Department of Computer Science, Dartmouth College. He is currently an Associate Professor and the Director of the Virtual Reality Laboratory, Department of Computer Graphics Technology, and a Core Faculty Member of the Applied AI Research Center, Polytechnic Institute, Purdue University. His research interests include virtual reality, virtual humans, computer graphics and animation, intelligent systems, and human–computer interaction. He is a member of ACM. He serves as an Associate Editor for the *Computer Animation and Virtual Worlds* and *Frontiers in Virtual Reality* journals, as well as on the organizing and program committees of numerous conferences in the virtual reality, computer graphics/animation, and human–computer interaction fields. He received the Best Paper Awards from IEEE ISMAR, in 2024, ACM TiiS, in 2023, and ACM SIGGRAPH VRCAI, in 2022, along with Honorable Mention Awards (Top 5%) from ACM CHI, in 2022 and ACM CHI PLAY, in 2021.

• • •