

Review

# The Effects of Immersive Virtual Reality Applications on Enhancing the Learning Outcomes of Undergraduate Health Care Students: Systematic Review With Meta-synthesis

Justina Yat Wa Liu<sup>1,2</sup>, PhD; Yue-Heng Yin<sup>3</sup>, PhD; Patrick Pui Kin Kor<sup>1</sup>, PhD; Daphne Sze Ki Cheung<sup>1</sup>, PhD; Ivy Yan Zhao<sup>1</sup>, PhD; Shanshan Wang<sup>1</sup>, PhD; Jing Jing Su<sup>1</sup>, PhD; Martin Christensen<sup>1</sup>, PhD; Stefanos Tyrovolas<sup>1</sup>, PhD; Angela Y M Leung<sup>1,2</sup>, PhD

<sup>1</sup>School of Nursing, The Hong Kong Polytechnic University, Hong Kong, Hong Kong

<sup>2</sup>Research Institute for Smart Ageing, The Hong Kong Polytechnic University, Hong Kong, Hong Kong

<sup>3</sup>School of Nursing, Nanjing Medical University, Nanjing, China

**Corresponding Author:**

Justina Yat Wa Liu, PhD

School of Nursing

The Hong Kong Polytechnic University

No 11 Yuk Choi Road

Hong Kong

Hong Kong

Phone: 852 2766 4097

Email: [justina.liu@polyu.edu.hk](mailto:justina.liu@polyu.edu.hk)

## Abstract

**Background:** Immersive virtual reality (IVR) applications are gaining popularity in health care education. They provide an uninterrupted, scaled environment capable of simulating the full magnitude of sensory stimuli present in busy health care settings and increase students' competence and confidence by providing them with accessible and repeatable learning opportunities in a fail-safe environment.

**Objective:** This systematic review aimed to evaluate the effects of IVR teaching on the learning outcomes and experiences of undergraduate health care students compared with other teaching methods.

**Methods:** MEDLINE, Embase, PubMed, and Scopus were searched (last search on May 2022) for randomized controlled trials (RCTs) or quasi-experimental studies published in English between January 2000 and March 2022. The inclusion criteria were studies involving undergraduate students majoring in health care, IVR teaching, and evaluations of students' learning outcomes and experiences. The methodological validity of the studies was examined using the Joanna Briggs Institute standard critical appraisal instruments for RCTs or quasi-experimental studies. The findings were synthesized without a meta-analysis using vote counting as the synthesis metric. A binomial test with  $P < .05$  was used to test for statistical significance using SPSS (version 28; IBM Corp). The overall quality of evidence was evaluated using the Grading of Recommendations Assessment, Development, and Evaluation tool.

**Results:** A total of 17 articles from 16 studies totaling 1787 participants conducted between 2007 and 2021 were included. The undergraduate students in the studies majored in medicine, nursing, rehabilitation, pharmacy, biomedicine, radiography, audiology, or stomatology. The IVR teaching domains included procedural training (13/16, 81%), anatomical knowledge (2/16, 12%), and orientation to the operating room setting (1/16, 6%). The quality of the 75% (12/16) of RCT studies was poor, with unclear descriptions of randomization, allocation concealment, and outcome assessor blinding procedures. The overall risk of bias was relatively low in the 25% (4/16) of quasi-experimental studies. A vote count showed that 60% (9/15; 95% CI 16.3%-67.7%;  $P = .61$ ) of the studies identified similar learning outcomes between IVR teaching and other teaching approaches regardless of teaching domains. The vote count showed that 62% (8/13) of the studies favored using IVR as a teaching medium. The results of the binomial test (95% CI 34.9%-90%;  $P = .59$ ) did not show a statistically significant difference. Low-level evidence was identified based on the Grading of Recommendations Assessment, Development, and Evaluation tool.

**Conclusions:** This review found that undergraduate students had positive learning outcomes and experiences after engaging with IVR teaching, although the effects may be similar to those of other forms of virtual reality or conventional teaching methods.

Given the identification of risk of bias and low level of the overall evidence, more studies with a larger sample size and robust study design are required to evaluate the effects of IVR teaching.

**Trial Registration:** International prospective register of systematic reviews (PROSPERO) CRD42022313706; [https://www.crd.york.ac.uk/prospero/display\\_record.php?RecordID=313706](https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=313706)

(*J Med Internet Res* 2023;25:e39989) doi: [10.2196/39989](https://doi.org/10.2196/39989)

## KEYWORDS

immersive virtual reality; virtual reality; effects; undergraduate health care education; systematic review; meta-synthesis

## Introduction

### Background

Clinical competency is essential for all health care professionals. Health care students, particularly undergraduate students with limited clinical experience, must incorporate theories into their skills with a good professional attitude to attain clinical competency. In addition to conventional classroom teaching for theoretical knowledge inputs and laboratory skill practice, the clinical practicum is crucial in undergraduate health care education to develop the clinical competency of students [1]. However, clinically based degree programs such as medicine, nursing, and physiotherapy are facing faculty shortages and increasing demands on clinical venues for student clinical placements [2]. In addition, clinical practicums for undergraduate health care students have been suspended in many countries because of the COVID-19 pandemic, further diminishing learning opportunities for students [3]. Therefore, there is an ongoing need for accessible, cost-effective, and high-quality methods of education to overcome these resource limitations [2].

Simulation-based learning is useful for teaching in all domains (ie, knowledge, skills, and attitudes) relevant to the education of health professionals [4]. It is a well-applied teaching and learning strategy [5] for increasing training opportunities and enhancing learning efficiency [6]. Simulation-based learning facilitates a deeper understanding of theoretical knowledge, relationships between different concepts, advanced inquiry, problem-solving, and decision-making [7]. It allows health care students to practice procedural skills without compromising patient safety and improves the quality of their patient care [8]. Many studies have shown that simulation-based learning effectively advances students' competencies in acquiring diagnostic and psychomotor skills [9,10]. However, it is not aimed at replacing the clinical practicum but at better engaging and preparing students for it.

Immersive virtual reality (IVR) applications for simulation-based training are gaining popularity in health care professional education [2,11]. They provide an immersive learning experience with a first-person viewpoint in a 3D virtual environment supported by head-mounted displays or Cave Automatic Virtual Environments (room-sized cube virtual reality [VR] environments) [12,13]. They provide direct sensory feedback or reactions to users based on their physical actions.

Compared with nonimmersive VR (such as computer-based simulation games) or other forms of traditional simulation-based training (such as skill laboratories using mannequins and

simulators), only IVR provides learners with the perception of being physically present in a synthetic world [14]. It provides an uninterrupted, scaled environment capable of simulating the full magnitude of sensory stimuli present in busy health care settings such as operating rooms, emergency departments, and surgical or medical wards [11]. Owing to the uniquely high level of immersion that IVR offers, students can practice a particular procedure or rehearse a specific action with a high level of physical and psychological fidelity, with immediate and standardized feedback provided corresponding to the users' actions. IVR can increase the competence and confidence of students by providing them with accessible and repeatable learning opportunities in a fail-safe environment [2].

The COVID-19 pandemic has led to a paradigm shift in all clinically based undergraduate programs, prompting adaptations through better use of educational technology, in which IVR is a promising alternative [15]. It is believed that the use of IVR-based simulation learning in health care professional training will persist beyond the pandemic [16]. A small number of systematic reviews have reported the efficacy of VR-based simulation learning in improving the knowledge and skill competence of students compared with conventional simulation-based training [1,17,18]. However, the learning approaches included in these reviews were mainly non-IVR approaches. Recently, a systematic review of 17 studies involving 307 surgical trainees reported significantly improved procedural completion times and greater postintervention scores on procedural checklists compared with conventional teaching methods such as watching standardized surgical training videos [11]. As previously mentioned, simulation-based learning, including the use of IVR, should be useful for learning in different domains (knowledge, skills, and attitudes) [4]. However, the extent to which IVR applications are used for teaching in different learning domains and their effectiveness in undergraduate health care professional education remain unknown [2].

### Objectives

Therefore, the aim of this review was to evaluate the effects of IVR applications in improving the learning outcomes and experiences of undergraduate health care students compared with other teaching methods. Two review questions were devised: (1) What are the effects of IVR applications on improving students' learning outcomes in different teaching domains (such as knowledge, skills, and attitudes) compared with other teaching methods? (2) What are the effects of IVR applications on enhancing the learning experiences of students (such as their level of satisfaction, perception of IVR innovation,

and self-perceived competence and confidence) compared with other teaching methods?

## Methods

This systematic review was conducted following the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses; [Multimedia Appendix 1](#)) [19] and SWiM

### Textbox 1. Inclusion and exclusion criteria.

- Inclusion criteria
  - Population: the participants were undergraduate students majoring in professional health care fields such as medicine, nursing, and physiotherapy who were included without regard to sex or race.
  - Intervention: the teaching methods used in the studies focused on immersive virtual reality (IVR) applications using either fully immersive 360° IVR, such as head-mounted display products, or Cave Automatic Virtual Environment systems, such as HTC VIVE, Samsung Oculus Rift, or other similar programs.
  - Comparison: we aimed to determine whether other teaching methods would have similar effects to IVR applications in professional health care education. For the purpose of comparison, we classified the different teaching methods into either active or passive comparison groups. The active comparison groups included other forms of simulation-based training such as nonimmersive virtual reality (VR), computer screen-based simulators, and real patient simulation. The passive control groups included mainly conventional teaching methods such as face-to-face lectures, tutorial classes, role-play, and reading materials.
  - Outcome: we determined the effects of IVR applications on the learning outcomes (ie, the primary outcome) and experiences (ie, the secondary outcome) of students. The primary outcome refers to the change in the students' theoretical knowledge and procedural skills, as reflected by any form of written examination or clinical skills such as communication skills, completion time, and error rate in a specific clinical procedure. The secondary outcome was the students' learning experience, as assessed by their satisfaction levels, perception of IVR innovations, and self-perceived competence and confidence after receiving IVR.
  - Study design: this review included randomized controlled trials and quasi-experimental studies published in English between January 2000 and March 2022 as IVR technology has developed enough to be used in health education since 2000 [17].
- Exclusion criteria
  - Studies were excluded if their target population was postgraduate health care students or students of other subjects unrelated to health care. Studies that involved only other forms of VR, such as nonimmersive VR, or other types of simulation-based technologies, such as augmented reality and manikins, were excluded. Pilot and feasibility studies were also excluded.

## Information Sources

A combination of Medical Subject Headings and free-text terms was used to search through 4 databases, namely, MEDLINE, Embase, PubMed, and Scopus, for potentially relevant abstracts. In addition, hand searches were conducted by reviewing the reference lists of all papers selected for inclusion in this review from the electronic databases, Google Scholar, and hard copies in university libraries to identify any articles missed by the database search. An alert for updated articles in each database was set to avoid missing potential up-to-date studies.

## Search Strategy

Search strategies were developed according to the 2 primary concepts of this review: the use of IVR applications in health professional undergraduate education and their effectiveness in enhancing the learning outcomes and experiences of students. To identify studies that used IVR applications, we used search terms such as “Immersive Virtual Reality” or “Simulated environment” and “Simulation” and “Healthcare” and “Students” or “Undergraduates” or “Trainees” ([Multimedia Appendix 2](#)). These terms were revised appropriately for different databases.

(Synthesis Without Meta-analysis) [20] guidelines. The review protocol was registered in PROSPERO (CRD42022313706).

## Eligibility Criteria

We used the Population, Intervention, Comparison, Outcome, and Study Design (PICOS) structure to define the eligibility criteria ([Textbox 1](#)).

## Selection Process

The search results were imported into the EndNote bibliographic software (version 20; Clarivate Analytics), and duplicate studies were removed. The titles and abstracts of all the identified studies were screened independently by 2 researchers (JYWL and YHY) to identify potentially relevant papers based on the review criteria. Both researchers compared the preliminary results of the review to reach an agreement. Full-text articles were then obtained and screened independently. Any disagreements between the reviewers were then discussed among the members of the research team to reach a consensus.

## Data Collection Process and Data Items

A specific data extraction matrix was created to collect information from each included study, including the author, year, title, country of origin, demographic data of the participants (age, sex, and type of health profession), methodological data (aims of the study, sample size, study design, educational innovation, and comparison groups), and outcome data (primary outcome, eg, examination scores; secondary outcome, eg, self-perceived competence and confidence). If any original study had been published in more than one paper, the information was extracted as 1 study based on the study protocol number. The authors of the primary studies

were contacted when clarifications were required or if any information was missing. Data extraction was conducted independently by the same 2 researchers. Once complete, the results were compared, and discrepancies were resolved through discussion. A final extraction table was developed.

### Study Risk-of-Bias Assessment

The studies selected for this review were independently assessed for methodological validity by the reviewers involved in their selection using the Joanna Briggs Institute (JBI; University of Adelaide, Australia) standard critical appraisal instruments for randomized controlled trials (RCTs) or quasi-experimental studies.

We assessed the included RCTs based on the following 12 appraisal items: methods of randomization; treatment allocation and concealment; similarity of characteristics between groups at baseline; blinding procedures for participants, interventionists, and outcome assessors; whether the comparison groups were treated identically other than in the intervention of interest (ie, IVR applications); completeness of the follow-up (ie, if there was any bias because of missing data); consistency and reliability of the outcome measurements; and appropriateness of the statistical analysis and trial design (JBI). For example, to determine if there were any biases because of missing data, we checked whether there were differences between groups with regard to the loss to follow-up (numbers or proportions, reasons, any analysis of patterns of loss to follow-up, and their impact on the internal validity of the study). We assessed the included quasi-experimental studies based on the following 9 appraisal items: clarity of dependent and independent variables, similarity of characteristics between groups at baseline, whether comparison groups were treated identically other than in the intervention of interest (ie, IVR applications), any comparison group, any pre- and postoutcome measurements, completeness of the follow-up (ie, whether there was any bias because of missing data), consistency and reliability of the outcome measurements, and appropriateness of the statistical analysis.

Working independently to assess for risk of bias, the same 2 assessors rated each item as *yes*, *no*, *unclear*, or *not applicable*. Any disagreements on the results of the bias assessment were then reviewed and discussed by the research team until a consensus was reached.

The aim of this assessment of the risk of bias was to determine the quality of each study. However, the risk of bias was not used as a criterion for the inclusion of a study in this review. A trial was judged to be at a low risk of bias overall when all items were rated as *yes*. Conversely, a study was judged to be at a high risk of bias when it reported on a procedure that could be judged as being a *no* or *unclear* in any item. Owing to the nature of the intervention, it was impossible to blind the participants;

thus, we did not include the *blinding procedures for participants* item when determining a study's overall risk of bias [21].

## Synthesis Methods

### Criteria for Grouping Studies

Following the review questions, studies for synthesis were grouped according to the study outcomes (ie, learning outcomes for research question 1 and learning experience for research question 2). In addition, studies with the same teaching domains (ie, procedural skills vs theoretical knowledge) were grouped for the analysis of the 2 outcomes.

### Standardized and Synthesis Metrics and Method

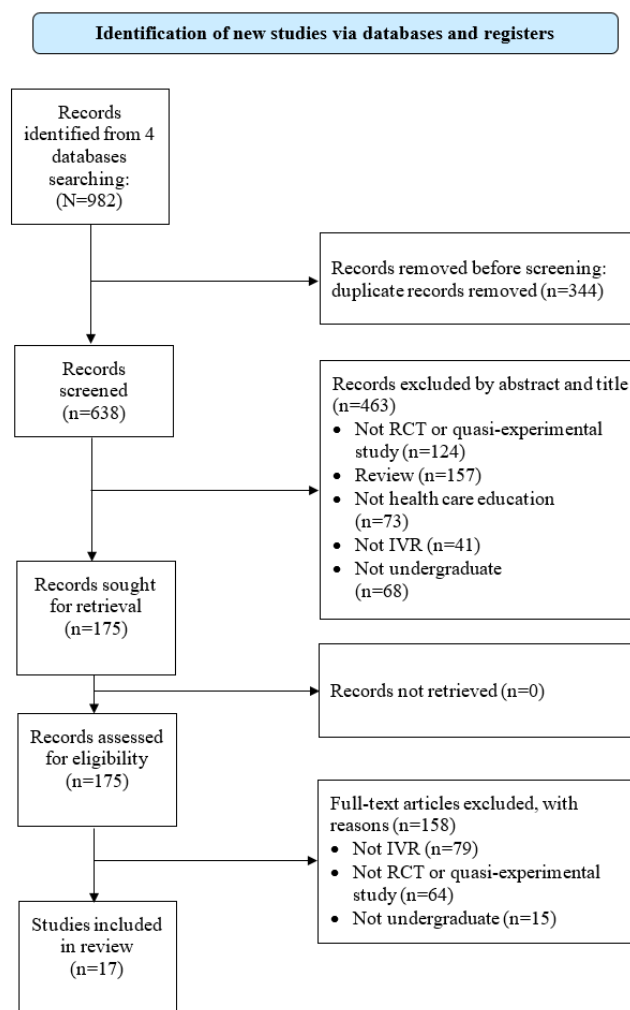
The direction of the effects (learning outcomes and learning experiences) was used as the standardized metric as there was a lack of precision, which was specific to the effects of the intervention (IVR teaching) and control on the results presented by different studies. This did not allow for the calculation of summary statistics [22]. In addition, the clinical and methodological characteristics (such as populations, intervention components, and the choice of outcome measurements and study designs) of each study were used to evaluate the heterogeneity based on the Cochrane Handbook for Systematic Reviews of Interventions [23]. The included studies were highly heterogeneous, with diverse teaching aims and outcome assessments to fit the needs of students from different health care professions. Different IVR features were adopted, with different frequencies and durations. In addition, most of the included studies were rated as having a high risk of bias. As a result, we felt that it was not appropriate to conduct a meta-analysis. Therefore, vote counting was the best match for synthesizing the results. A binomial test using SPSS (version 28; IBM Corp) was used to indicate whether there was evidence of an effect [23]. The quality of the evidence generated by different studies was assessed using the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) tool [24].

## Results

### Search Results

A total of 982 articles were found by searching the databases; after the removal of duplicates ( $n=344$ , 35%), 638 (65%) were left. After screening the titles and abstracts, 72.6% (463/638) of the articles were excluded, leaving 175 to be retrieved for a full-text screening. Following the full-text screening, a further 90.9% (159/175) of the articles were excluded, leaving 17 articles [25-41]. Of the 17 articles, 2 (12%) [31,32] were from the same study. Therefore, a total of 16 studies were included in this systematic review. Details of the selection process and the reasons for the exclusion of articles are presented in the PRISMA flowchart (Figure 1).

**Figure 1.** PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart of the literature screening process. IVR: immersive virtual reality; RCT: randomized controlled trial.



## Study Characteristics

The characteristics of the included studies are shown in [Table 1](#). The 16 studies were conducted between 2007 and 2021. An RCT design was adopted in 75% (12/16) of the studies [25,26,28-30,32-34,36-38,40], and a quasi-experimental design was adopted in 25% (4/16) of the studies [27,35,39,41]. All 12 RCTs were 2-armed, among which were 8% (1/12) [26] that had a crossover design. Of the 4 quasi-experimental trials, 1 (25%) [39] was a 4-armed study, 1 (25%) [27] was a 3-armed study, and the other 2 (50%) were 2-armed studies [35,41].

A total of 1787 participants were involved in the included studies, with the number ranging from 25 to 289 in the RCTs and from 29 to 197 in the quasi-experimental studies. In total, 44% (7/16) of the studies involved sample sizes of <60. Female students accounted for 71.73%, with 25% (4/16) of the studies [25,32,35,39] not disclosing the sex distribution. The participants

ranged in age from 19 to 35 years, whereas the age distribution was not discussed in 25% (4/16) of the studies [25,29,32,39]. The types of participants included undergraduate students majoring in medicine, nursing, rehabilitation, pharmacy, biomedicine, radiography, audiology, or stomatology.

IVR teaching was compared with other forms of VR such as desktop-based VR (ie, the active control) in 25% (4/16) of the studies [25,26,28,32]. In 50% (8/16) of the studies [29,30,34-38,40], IVR teaching was compared with conventional teaching methods such as verbal didactic instruction and hard-copy teaching materials (ie, the passive control). In total, 12% (2/16) of the studies [27,39] used passive and active controls. A total of 6% (1/16) of the studies [33] adopted role-play, and another study (1/16, 6%) [41] used a clinical practicum as their comparison group. These teaching approaches did not involve any use of VR, and we considered the role-play and clinical practicum groups as passive controls.

**Table 1.** Characteristics of the included studies (N=16).

Author, year, and country	Design	Sample size (percentage of female participants), N	Age (years)	Participants	Intervention (IVR <sup>a</sup> )	Comparator	Side effects
Gutiérrez-Maldonado et al [25], 2015, Spain	2-armed RCT <sup>b</sup>	Total: 52 (N/A <sup>c</sup> ); EG <sup>d</sup> : 26; CG <sup>e</sup> : 26	N/A	Undergraduate students (without a specific discipline)	<ul style="list-style-type: none"> <li>• Training in simulated diagnostic interview techniques for eating disorders with an IVR application</li> <li>• Number of IVR sessions: 1</li> <li>• Duration: allow students to complete the task</li> <li>• Other teaching: a basic explanation by a professor before IVR</li> <li>• Teaching domain: procedural skills</li> </ul>	<ul style="list-style-type: none"> <li>• Training in simulated diagnostic interview techniques for eating disorders with a desktop-based VR<sup>f</sup></li> <li>• Active control</li> </ul>	N/A
Harrington et al [26,42], 2018, Ireland	2-armed crossover RCT	Total: 40 (50); EG: 20; CG: 20	23	Preclinical undergraduate students (medicine, physical therapy, pharmacy, and biomedical sciences)	<ul style="list-style-type: none"> <li>• Watching a 10-minute laparoscopic cholecystectomy operation with a 360° video using the IVR</li> <li>• Number of IVR sessions: 1</li> <li>• Duration: 10 minutes</li> <li>• Other teaching: none</li> <li>• Teaching domain: procedural skills</li> </ul>	<ul style="list-style-type: none"> <li>• Watching a 10-minute laparoscopic cholecystectomy operation video in 2D format on a 75-inch LED<sup>g</sup> television</li> <li>• Active control</li> </ul>	Self-reported levels of nausea were low (0.7 out of 10) but varied substantially from 0 to 8 among participants.
Smith et al [27], 2018, United States	Quasi-experimental trial	Total: 197 (88.2); EG: 59; active CG: 58; passive CG: 55	18-25 (73.3%); 26-34 (15.5%); 35-50 (11.2%)	Senior baccalaureate-degree nursing students	<ul style="list-style-type: none"> <li>• Training in the skill of decontamination with IVR simulation using an HMD<sup>h</sup></li> <li>• Number of IVR sessions: 1</li> <li>• Duration: allow students to complete the task</li> <li>• Other teaching: 30-minute web-based video module to explain the skills</li> <li>• Teaching domain: procedural skills</li> </ul>	<ul style="list-style-type: none"> <li>• Training in the same skills with a case scenario in an emergency department room and featuring a patient on a stretcher, with available personal protective equipment and tools</li> <li>• Active CG: PC and keyboard or mouse</li> <li>• Passive CG: written instructions</li> </ul>	None
Zackoff et al [28], 2020, United States	2-armed RCT	Total: 168 (48.8); EG: 78; CG: 90	20-24 (34.5%); 25-29 (60.1%); 30-34 (4.2%); 35-39 (0.6%)	Year-3 medical students	<ul style="list-style-type: none"> <li>• In addition to receiving the same 60-minute training session as in the CG, participants received a 30-minute IVR curriculum with 3 simulations of an infant from no distress to respiratory failure</li> <li>• Number of IVR sessions: 1</li> <li>• Duration: 30 minutes</li> <li>• Other teaching: facilitation and debriefing during and after the IVR tutorial</li> <li>• Teaching domain: procedural skills</li> </ul>	<ul style="list-style-type: none"> <li>• Receiving a 60-minute session consisting of verbal didactic instructions with a subsequent high-fidelity mannequin simulation to teach participants how to identify infants with respiratory distress and failure</li> <li>• Active control</li> </ul>	None

Author, year, and country	Design	Sample size (percentage of female participants), N	Age (years)	Participants	Intervention (IVR <sup>a</sup> )	Comparator	Side effects
Francis et al [29], 2020, United States	2-armed RCT	Total: 52 (89); EG: 26; CG: 26	≤23 (40.4%); ≥24 (59.6%)	Year-2 preclinical physician assistant students	<ul style="list-style-type: none"> <li>Attending a 5-minute IVR experience in the simulated operating room with several distinct scenarios. The IVR operating room scenario was viewed with a stand-alone untethered all-in-one VR headset.</li> <li>Number of IVR sessions: 1</li> <li>Duration: 5 minutes</li> <li>Other teaching: none</li> <li>Teaching domain: orientation</li> </ul>	<ul style="list-style-type: none"> <li>Attending a traditional lecture on orientation to surgical operating room settings</li> <li>Passive control</li> </ul>	N/A
Kurul et al [30], 2020, Turkey	2-armed RCT	Total: 72 (72.2); EG: 36; CG: 36	19	Year-1 physical therapy students	<ul style="list-style-type: none"> <li>Learning head and neck region anatomy for 30 minutes with IVR</li> <li>Number of IVR sessions: 1</li> <li>Duration: 30 minutes</li> <li>Other teaching: none</li> <li>Teaching domain: theoretical knowledge</li> </ul>	<ul style="list-style-type: none"> <li>Attending a 30-minute presentation of images of the head and neck region, use of computer-based VR</li> <li>Passive control</li> </ul>	Vision discomfort (n=7), eye-strain (n=6), and general discomfort (n=6)
Gutiérrez et al [32], 2007, and Pierce et al [31], 2008, United States	2-armed RCT	Total: 25 (N/A); EG: 13; CG: 12	N/A	Year-1 medical students	<ul style="list-style-type: none"> <li>Practicing how to conduct a physical examination on a patient with traumatic head injury in an IVR environment, which was displayed with an HMD</li> <li>Number of IVR sessions: 1</li> <li>Duration: 30 minutes</li> <li>Other teaching: instructional video on using VR equipment and head injury reference materials</li> <li>Teaching domain: procedural skills</li> </ul>	<ul style="list-style-type: none"> <li>Practicing how to conduct a physical examination on a patient with traumatic head injury by using a computer screen and mouse to rotate the viewpoint, then using the joystick to perform a physical examination of the head injury</li> <li>Active control</li> </ul>	None
Sapkaroski et al [33], 2019, Australia	2-armed RCT	Total: 76 (75); EG: 38; CG: 38	21	Radiography students	<ul style="list-style-type: none"> <li>Practicing how to position a virtual avatar patient for PA<sup>i</sup> imaging of the left hand by using CESTOL<sup>j</sup> VR Clinic</li> <li>Number of IVR sessions: 1</li> <li>Duration: allow students to complete the task</li> <li>Other teaching: hand imaging lesson</li> <li>Teaching domain: procedural skills</li> </ul>	<ul style="list-style-type: none"> <li>Practicing how to position the patient for PA imaging of the left hand using conventional clinical role-play in the x-ray laboratory</li> <li>Passive control</li> </ul>	None
Stepan et al [34], 2017, United States	2-armed RCT	Total: 66 (50); EG: 33; CG: 33	21-25 (89.4%); 26-30 (9.1%); 31-35 (1.5%)	Year-1 and year-2 medical students	<ul style="list-style-type: none"> <li>Learning neuroanatomy with a 10-minute IVR model using an HMD</li> <li>Number of IVR sessions: 1</li> <li>Duration: 10 minutes</li> <li>Other teaching: 10-minute internet-based introductory lecture</li> <li>Teaching domain: theoretical knowledge</li> </ul>	<ul style="list-style-type: none"> <li>Independently studying neuroanatomy using web-based textbooks containing texts and 2D images for 20 minutes</li> <li>Passive control</li> </ul>	None

Author, year, and country	Design	Sample size (percentage of female participants), N	Age (years)	Participants	Intervention (IVR <sup>a</sup> )	Comparator	Side effects
Bakhos et al [35], 2020, France	Crossover quasi-experimental trial	Total: 29 (N/A); EG: 15; CG: 14	EG: 22; CG: 20	Year-1 audiology students (without beginning an internship)	<ul style="list-style-type: none"> <li>Receiving a 3-hour IVR audiometry training session on 3 clinical cases (ie, presbycusis, vestibular schwannoma, and sudden idiopathic deafness). The audiometric diagnosis and management were evaluated for each case, and a report was generated that summarized the errors during the evaluation.</li> <li>Number of IVR sessions: 3 cases</li> <li>Duration: 30 minutes per case</li> <li>Other teaching: 20 minutes introducing the system and debriefing</li> <li>Teaching domain: procedural skills</li> </ul>	<ul style="list-style-type: none"> <li>Receiving a 3-hour audiometry training session supervised by a teacher on basic audiometry principles and practicing audiometry techniques for different clinical cases</li> <li>Passive control</li> </ul>	None
Chao et al [36], 2021, Taiwan	2-armed RCT	Total: 45 (86.7); EG: 22; CG: 23	24	Nursing students (aged ≥20 years) who had never acquired the skills of NG <sup>k</sup> tube feeding	<ul style="list-style-type: none"> <li>Learning NG tube feeding skills through a 20-minute IVR video program</li> <li>Number of IVR sessions: 1</li> <li>Duration: 20 minutes</li> <li>Other teaching: none</li> <li>Teaching domain: procedural skills</li> </ul>	<ul style="list-style-type: none"> <li>Learning the same skills by watching the NG tube feeding demonstration DVD video</li> <li>Passive control</li> </ul>	A total of 5 students (23%) reported feeling slightly dizzy, but this did not affect their ability to watch the video.
Berg and Steinsbekk [37], 2020, Norway	2-armed RCT	Total: 289 (78.5); EG: 149; CG: 140	<20 (23.8%); 20-24 (64.6%); >25 (11.6%)	Year-1 medical and nursing students (who had started their studies no later than 2 months before this study)	<ul style="list-style-type: none"> <li>Self-practicing the ABCDE<sup>l</sup> approach in an IVR environment for assessing and managing patients who were critically ill or injured. All practice attempts were carried out on a virtual patient using virtual equipment in the IVR environment.</li> <li>Number of IVR sessions: for self-practicing, did not mention how many times the students could practice</li> <li>Duration: allow students to complete the task</li> <li>Other teaching: same as the comparator group except that they used traditional equipment to practice the skills</li> <li>Teaching domain: procedural skills</li> </ul>	<ul style="list-style-type: none"> <li>Self-practicing the ABCDE approach with traditional equipment after receiving a 1-hour teaching session, which included a 15-minute introduction, 20 minutes of individual practice, and 15 minutes of testing. All participants watched a 7-minute introduction video about the ABCDE approach. They received a printed sheet with pictures of the equipment along with simple instructions on its technical use.</li> <li>Passive control</li> </ul>	N/A
Berg and Steinsbekk [38], 2021, Norway	2-armed RCT	Total: 289 (84.6); EG: 146; CG: 143	<20 (29.6%); 20-24 (59.3%); >25 (11.1%)				N/A



Author, year, and country	Design	Sample size (percentage of female participants), N	Age (years)	Participants	Intervention (IVR <sup>a</sup> )	Comparator	Side effects
				Year-1 medical and nursing students (who had started their studies no later than 2 months before this study)	<ul style="list-style-type: none"> <li>Group practicing in the IVR platform in a virtual patient room using the ABCDE approach to immediately assess and treat patients who were critically ill or injured</li> <li>Number of IVR sessions: 1</li> <li>Duration: allow students to complete the task</li> <li>Other teaching: 6-minute lecture and 7-minute skill-demonstration video</li> <li>Teaching domain: procedural skills</li> </ul>	<ul style="list-style-type: none"> <li>Group practicing with physical equipment using the ABCDE approach, receiving a printed sheet with pictures of the equipment along with simple instructions on its technical use</li> <li>Passive control</li> </ul>	
Collaço et al [39], 2021, Brazil	Quasi-experimental trial	Total: 163 (N/A); full: 42; NP <sup>m</sup> : 40; NT <sup>n</sup> : 40; NH <sup>o</sup> : 41	N/A	Clinical dental students	<ul style="list-style-type: none"> <li>Receiving dental anesthesia skill training through an IVR application using an HMD. The training was divided into 2 phases: preceptorship and training.</li> <li>Number of IVR sessions: 2</li> <li>Duration: did not mention</li> <li>Other teaching: none</li> <li>Teaching domain: procedural skills</li> </ul>	<ul style="list-style-type: none"> <li>NP: received non-immersive VR in the preceptorship phase but IVR in the training phase.</li> <li>NT: received IVR in the preceptorship phase but nonimmersive VR in the training phase.</li> <li>NH: received non-immersive VR in both phases.</li> <li>When under non-immersive conditions, the participants visualized the preceptorship or performed the training by watching a television screen.</li> <li>Both passive and active controls</li> </ul>	Most reported experiencing no side effects (87%) and “slight” (9%), “moderate” (3%), and “severe” (1%) symptoms.
Ros et al [40], 2020, France	2-armed RCT	Total: 173 (52); EG: 85; CG: 88	N/A	Year-4 medical students (2 years before residency)	<ul style="list-style-type: none"> <li>Learning a medical procedure (ie, external ventricular drainage) through a 7-minute IVR displayed on an HMD</li> <li>Number of IVR sessions: 1</li> <li>Duration: 7 minutes</li> <li>Other teaching: reading the technical note about this procedure</li> <li>Teaching domain: procedural skills</li> </ul>	<ul style="list-style-type: none"> <li>Reading only the technical note for 7 minutes, which described the procedure for external ventricular drainage</li> <li>Passive control</li> </ul>	N/A
Yu et al [41], 2021, Korea	Quasi-experimental trial	Total: 51 (92); EG: 26; CG: 25	22.4	Senior nursing students		<ul style="list-style-type: none"> <li>Having routine clinical practice in a neonatal intensive care unit</li> <li>Passive control</li> </ul>	N/A

Author, year, and country	Design	Sample size (percentage of female participants), N	Age (years)	Participants	Intervention (IVR <sup>a</sup> )	Comparator	Side effects
					<ul style="list-style-type: none"> <li>Learning high-risk neonatal infection control skills with 3 scenarios using an IVR simulation program plus clinical routine practice as the CG</li> <li>Number of IVR sessions: 1</li> <li>Duration: allow students to complete the task</li> <li>Other teaching: prebriefing and debriefing before and after the IVR simulation</li> <li>Teaching domain: procedural skills</li> </ul>		

<sup>a</sup>IVR: immersive virtual reality.

<sup>b</sup>RCT: randomized controlled trial.

<sup>c</sup>N/A: not applicable.

<sup>d</sup>EG: experimental group.

<sup>e</sup>CG: control group.

<sup>f</sup>VR: virtual reality.

<sup>g</sup>LED: light-emitting diode.

<sup>h</sup>HMD: head-mounted display.

<sup>i</sup>PA: posterior-anterior.

<sup>j</sup>CESTOL: Clinical Education Training Solution.

<sup>k</sup>NG: nasogastric.

<sup>l</sup>ABCDE: airways, breathing, circulation, disability, exposure.

<sup>m</sup>NP: nonpreceptorship.

<sup>n</sup>NT: nontraining.

<sup>o</sup>NH: nonhaptic feedback.

### IVR Teaching Characteristics

A total of 81% (13/16) of the studies used IVR to train students in skills, including in techniques for diagnosing eating disorders [25] and respiratory distress in infants [28], laparoscopic cholecystectomy [26], decontamination skills [27], physical examinations for patients with traumatic head injuries [32], correct positioning for x-ray imaging [33], audiometry techniques [35], nasogastric tube feeding [36], immediate assessment and treatment of patients who are critically ill [37,38], dental anesthesia skills [39], external ventricular drainage [40], and neonatal infection control [41]. Only in 12% (2/16) of the studies was IVR used to teach students anatomical knowledge [30,43], whereas in 6% (1/16) of the studies, IVR was used to orient the students to the setting of the surgical operating room [29].

Most studies (12/16, 75%) featured only a single IVR experience for the students. The exceptions were the study by Bakhos et al [35], which had 3 different IVR cases, and the study by Collaço et al [39], who provided 2 IVR training episodes. Berg and Steinsbekk [37,38] allowed students to self-practice the skills but did not mention the number of sessions. The duration for students to experience IVR learning was typically short, ranging from 5 to 30 minutes in 56% (9/16) of the studies. In total, 38% (6/16) of the studies [25,27,33,37,38,41] allowed

students to be exposed to IVR environments as long as they needed to complete the specific tasks or procedures. A total of 6% (1/16) of the studies [39] did not state the exact duration of the students' IVR learning session.

The IVR products used in the included studies were Oculus VR, Samsung Gear VR, Clinical Education Training Solution VR Clinic, HTC VIVE, and a university-created platform.

Side effects were reported in 25% (4/16) of the studies. Nearly half of the students in the study by Kurul et al [30] reported different forms of slight discomfort (ie, vision discomfort, eyestrain, and general discomfort). In the study by Chao et al [36], 23% of the students reported feeling slightly dizzy, but this did not affect their viewing activities. In contrast, most students in another 12% (2/16) of the studies [26,39] did not experience any side effects.

### Theoretical Frameworks

Learning theory provides a framework to guide the development of teaching activities to help students imbibe, process, and retain the knowledge and skills that they have learned [44,45]. When applied to educational IVR, a learning theory should provide a pedagogical framework and foundation for designing IVR-related teaching and learning strategies. However, most of the included studies (15/16, 94%) did not mention any

theoretical approaches underpinning the development of IVR teaching. The exception was the study by Smith et al [27], who used the National League for Nurses Jeffries Simulation Theory [46] as their theoretical basis to guide the design of their teaching innovation. Most studies (15/16, 94%) supplemented the IVR lessons and tutorials by providing additional pedagogical practices or materials to encourage learning. In total, 25% (4/16) of the studies [27,32,34,37] included web-based modules and reading materials in addition to the IVR experience. The provision of either introduction or prebriefing or debriefing sessions before and after IVR learning was mentioned in 44% (7/16) of the studies [25,28,33,35,38,40,41]. A total of 31% (5/16) of the studies [26,29,30,36,39] used IVR as the sole method of learning.

### Risk-of-Bias Assessments

Among the included RCTs, a risk of bias was identified in most domains, with the exception of the domains Q7, Q8, and Q10 to Q13 (Table 2). Only 25% (3/12) of the studies [26,28,37,38] gave clear details on the randomization procedure, whereas other studies simply briefly stated that the design was randomized without providing further information. Allocation

concealment was not implemented or was unclear in 75% (9/12) of the studies [25,26,28-30,32,34,36,40]. In total, 33% (4/12) of the studies [25,26,33,40] did not report the baseline comparison between the groups, and in 8% (1/12) of the studies [38], there were differences in age and practical experience between the groups. The blinding of participants, intervention providers, and outcome assessors was another major concern with the RCTs. In none of the studies were the participants blinded, and in only 17% (2/12) of the studies [37,38] were both the assessors and intervention providers blinded. Among all 16 evaluated items in the JBI checklist, there were 25% (3/12) of the studies in which 1 to 3 items were viewed as having a low risk of bias and 75% (9/12) of the studies in which 4 to 6 items were viewed as having a moderate risk of bias (Figure 2).

For the quasi-experimental studies, the overall risk of bias was low (Table 3). However, in 25% (1/4) of the studies [27], 13% of the participants failed to complete all the tests, and no detailed explanations were given about which groups were involved and how this issue was handled in the study. In addition, in 50% (2/4) of the studies [35,39], the baseline difference between the groups was not reported.

**Table 2.** Critical appraisal of the included randomized controlled trials (RCTs; Joanna Briggs Institute critical appraisal checklist for RCTs).

Study, year	Q1 <sup>a</sup>	Q2 <sup>b</sup>	Q3 <sup>c</sup>	Q4 <sup>d</sup>	Q5 <sup>e</sup>	Q6 <sup>f</sup>	Q7 <sup>g</sup>	Q8 <sup>h</sup>	Q9 <sup>i</sup>	Q10 <sup>j</sup>	Q11 <sup>k</sup>	Q12 <sup>l</sup>	Q13 <sup>m</sup>
Gutiérrez-Maldonado et al [25], 2015	U <sup>n</sup>	N <sup>o</sup>	U	N	U	U	Y <sup>p</sup>	Y	Y	Y	Y	Y	Y
Harrington et al [26,42], 2018	Y	U	U	N	U	U	Y	Y	Y	Y	Y	Y	Y
Zackoff et al [28], 2020	U	U	Y	N	N	U	Y	Y	Y	Y	Y	Y	Y
Francis et al [29], 2020	U	U	Y	N	U	U	Y	Y	Y	Y	Y	Y	Y
Kurul et al [30], 2020	U	U	Y	N	— <sup>q</sup>	Y	Y	Y	Y	Y	Y	Y	Y
Gutiérrez et al [32], 2007	U	N	Y	N	—	U	Y	Y	Y	Y	Y	Y	Y
Pierce et al [31], 2008	U	N	Y	N	—	U	Y	Y	Y	Y	Y	Y	Y
Sapkaroski et al [33], 2019	U	Y	U	N	U	Y	Y	Y	Y	Y	Y	Y	Y
Stepan et al [34], 2017	U	U	Y	N	—	U	Y	Y	Y	Y	Y	Y	Y
Chao et al [36], 2021	U	N	Y	N	N	U	Y	Y	Y	Y	Y	Y	Y
Berg and Steinsbekk [37], 2020	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y
Berg and Steinsbekk [38], 2021	Y	Y	N	N	Y	Y	Y	Y	Y	Y	Y	Y	Y
Ros et al [40], 2020	U	U	U	N	U	U	Y	Y	Y	Y	Y	Y	Y

<sup>a</sup>Q1: was true randomization used for assigning participants to treatment groups?

<sup>b</sup>Q2: was the allocation to the treatment groups concealed?

<sup>c</sup>Q3: were the treatment groups similar at baseline?

<sup>d</sup>Q4: were participants blinded to the treatment assignment?

<sup>e</sup>Q5: were those delivering treatment blinded to the treatment assignment?

<sup>f</sup>Q6: were outcome assessors blinded to the treatment assignment?

<sup>g</sup>Q7: were the treatment groups treated identically other than in the intervention of interest?

<sup>h</sup>Q8: was follow-up complete, and if not, were strategies used to address incomplete follow-ups (ie, was there an analysis of patterns of those lost to follow-up)?

<sup>i</sup>Q9: were participants analyzed in the groups to which they were randomized?

<sup>j</sup>Q10: were the outcomes measured in the same way for the treatment groups?

<sup>k</sup>Q11: were the outcomes measured in a reliable way?

<sup>l</sup>Q12: was an appropriate statistical analysis used?

<sup>m</sup>Q13: was the trial design appropriate, and were any deviations from the standard RCT design (individual randomization and parallel groups) accounted for in the conduct and analysis of the trial?

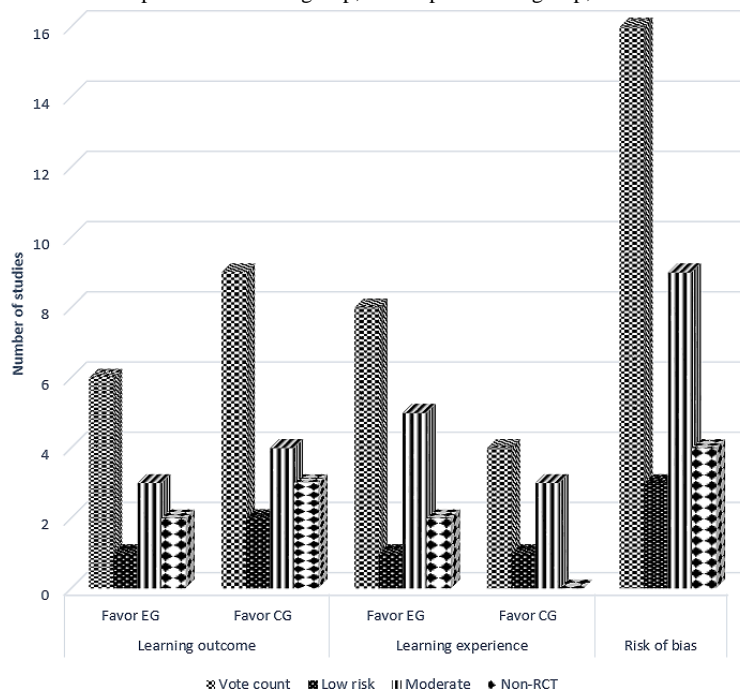
<sup>n</sup>U: unclear.

<sup>o</sup>N: no.

<sup>p</sup>Y: yes.

<sup>q</sup>Not available.

**Figure 2.** Evidence of effect: direction of effect plot. CG: control group; EG: experimental group; RCT: randomized controlled trial.



**Table 3.** Critical appraisal of the included quasi-experimental trials (Joanna Briggs Institute critical appraisal checklist for quasi-experimental trials).

Study, year	Q1 <sup>a</sup>	Q2 <sup>b</sup>	Q3 <sup>c</sup>	Q4 <sup>d</sup>	Q5 <sup>e</sup>	Q6 <sup>f</sup>	Q7 <sup>g</sup>	Q8 <sup>h</sup>	Q9 <sup>i</sup>
Smith et al [27], 2018	Y <sup>j</sup>	Y	Y	Y	Y	U <sup>k</sup>	Y	Y	Y
Bakhos et al [35], 2020	Y	U	Y	Y	Y	Y	Y	Y	Y
Collaço et al [39], 2021	Y	U	Y	Y	Y	Y	Y	Y	Y
Yu et al [41], 2021	Y	Y	Y	Y	Y	Y	Y	Y	Y

<sup>a</sup>Q1: are the *cause* and the *effect* made clear in the study (ie, there was no confusion about which variable came first)?

<sup>b</sup>Q2: were the participants who were included in any comparisons similar to each other in any way?

<sup>c</sup>Q3: were the participants who were included in any comparisons receiving similar treatment or care other than the exposure or intervention of interest?

<sup>d</sup>Q4: was there a control group?

<sup>e</sup>Q5: were there multiple measurements of the outcome both before and after the intervention or exposure?

<sup>f</sup>Q6: was the follow-up complete, and if not, were the differences between the groups in terms of their follow-up adequately described and analyzed?

<sup>g</sup>Q7: were the outcomes of the participants included in any comparisons measured in the same way?

<sup>h</sup>Q8: were the outcomes measured in a reliable way?

<sup>i</sup>Q9: was an appropriate statistical analysis used?

<sup>j</sup>Y: yes.

<sup>k</sup>U: unclear.

### Effects of IVR Teaching on Student Learning Outcomes

Of the 16 included studies (reported in 17 papers), 1 (6%) study using IVR for orientation did not include any assessment of student learning outcomes [29]. Therefore, 94% (15/16) of the studies were retained to assess the primary objective (ie, the students' learning outcomes) based on vote counting and a binomial probability test. Assuming that the true probability of favoring either IVR teaching or non-IVR teaching was equivalent to 0.05 under the null hypothesis (IVR teaching=non-IVR teaching on student learning outcomes), the results showed that 40% (6/15) of the studies [28,30,32,33,39,40] favored IVR teaching (95% CI

16.3%-67.7%;  $P=.61$ ). The remaining 60% (9/15) of the studies [25-27,34-38,41] showed similar effects between IVR and other teaching methods. These figures are below the expected binomial probability mean of 1.60 (SD 0.51) votes. Thus, we need to accept the null hypothesis.

### Procedural Skill Outcomes

In total, 87% (13/15) of the studies adopted IVR to teach procedural skills. A total of 62% (8/13) of the studies [25,27,35-38,41,42] showed that the improvement in the students' acquisition of procedural skills was similar whether they were in the IVR groups or in the groups that used other teaching methods. In other words, 38% (5/13) of the studies [28,32,33,39,40] indicated that students who received IVR

teaching performed significantly better (95% CI 13.9%-68.4%;  $P=.58$ ) than those who received other forms of teaching. These figures are also below the binomial probability mean of 1.62 (SD 0.51) votes; thus, the null hypothesis is accepted.

### Theoretical Knowledge and Other Outcomes

A total of 13% (2/15) of the studies evaluated the effects of IVR on teaching theoretical knowledge [30,34]. IVR teaching was found to have a greater effect on the acquisition of theoretical knowledge compared with the passive control in 50% (1/2) of the studies [30], but in another study [34], the effects of IVR were shown to be similar to those of the passive control. Owing to the small number of studies (only 2 studies), we did not run the binomial test. One study [29] used IVR for orienting students to the operating room, but the students' knowledge was not tested afterward. In none of the studies was IVR used to enhance the students' professional attitudes.

### Effects of IVR Teaching on the Students' Learning Experience

Of the 15 studies, 2 (13%) did not measure students' learning experience [33,40], 1 (7%) explored students' satisfaction and experiences with IVR through a focus group (only qualitative data were collected) [27], and 1 (7%) only measured the

students' learning experience in the IVR group but not in the control group [30]. These studies were excluded from the vote counting and binomial test. Of the 12 studies, 8 (67%) favored IVR teaching (95% CI 34.9%-90%;  $P=.59$ ), whereas the remaining 4 (33%) showed similar effects on the students' learning experience between IVR and other methods of teaching.

In total, 80% (12/15) of the studies [25,26,28,29,32,34-39,41] evaluated the learning experiences of students from both the IVR and control teaching groups. Overall, the students said that they had a more positive experience learning with IVR than with other teaching methods (Table 4). A total of 67% (8/12) of the studies [26,28,29,34-37,41] showed that the students favored IVR teaching, and 33% (4/12) of the studies [25,32,38,39] reported that the learning experiences of students who learned with IVR were similar to those of either the active or passive controls (95% CI 34.9%-90.1%;  $P=.39$ ). These figures are also below the binomial probability mean of 1.33 (SD 0.49) votes; thus, the null hypothesis is accepted. In 17% (2/12) of the studies [27,30], only the experimental group was evaluated, and both studies showed that students liked the approach of using IVR to learn. The effect direction plot of the different studies, together with the associated risk of bias, is shown in Figure 2.

**Table 4.** Effects of immersive virtual reality (IVR) in the included studies (N=16).

Study, year, and country or region	Type of IVR	Outcome measurements	Major findings	Learning outcomes favor group	Learning experience favor group
Gutiérrez-Maldonado et al [25], 2015, Spain	Oculus Rift DK1	Primary outcome: a diagnostic interview skill test; secondary outcome (usability): SUMI <sup>a</sup>	Primary and secondary outcome: the mean score of the EG <sup>b</sup> was higher than that of the CG <sup>c</sup> , but the difference between the 2 groups was nonsignificant ( $P=.23$ ; $P=.89$ ).	Similar	Similar
Harrington et al [26,42], 2018, Ireland	Samsung Gear VR <sup>d</sup>	Primary outcome: knowledge retention (8-point multiple choice); secondary outcomes: (1) attentiveness (engagement) and (2) appraisals	Knowledge retention: no significant variances ( $P=.14$ ). Attentiveness (engagement): a crossover analysis revealed a significantly higher level in the EG ( $P<.001$ ) and across periods ( $P<.001$ ) with no significant carryover effect ( $P=.97$ ). Appraisals: two-thirds of participants reported choosing EG.	Similar	EG
Smith et al [27], 2018, United States	Oculus Rift Developer Kit 2	Primary outcomes: (1) decontamination knowledge gain, (2) skill performance, and (3) time spent by students completing the procedure; secondary outcome: qualitative focus group	Primary outcome: decontamination knowledge gain—no significant difference among the 3 groups at all 3 time points; skill performance—no significant difference among the 3 groups at the 2 time points; performance time—no significant difference among the 3 groups at all 3 time points; secondary outcome: high levels of satisfaction in the EG	Similar	Only measured the EG
Zackoff et al [28], 2020, United States	Oculus Rift	Primary outcomes: recognition or interpretation of key examination findings, assignment of an appropriate respiratory status assessment, and recognition of the need to escalate care for patients with impending respiratory failure by using a video-based assessment; secondary outcome: self-assessed competence	Primary outcome: significantly higher in the EG than in the CG; secondary outcome: 81% of the EG demonstrated an improvement in self-assessed competency.	EG	EG
Francis et al [29], 2020, United States	Oculus VR	Self-efficacy of the participants	Self-efficacy significantly improved in the EG ( $P=.007$ ) but not for the CG ( $P=.30$ ).	N/A <sup>e</sup>	EG
Kurul et al [30], 2020, Turkey	3D Organon Anatomy (Medis Media) and 3D glasses (Oculus Rift; Oculus VR)	Primary outcome: written examination; secondary outcome: participants' perceptions, including enjoyment and learning efficiency	Primary outcome: both the EG and CG had significantly higher posttest scores, but the difference between the pretest results was found to be significantly higher in favor of the EG ( $P<.001$ ). Secondary outcome: 88.8% of students "enjoyed studying anatomy with IVR"; 83.3% of students felt that "it is easy to understand the location of structures with VR."	EG	Only measured the EG
Gutiérrez et al [32], 2007, and Pierce et al [31], 2008, United States	A platform created at the University of New Mexico	Primary outcome: written examination; secondary outcomes: the efficiency of the user interface and satisfaction with it	Primary outcome: posttest scores were significantly higher than pretest scores in both the EG and CG (within-group). There was a significant interaction between groups and time for the EG. Secondary outcome: there was no overall significant difference in efficiency and satisfaction between the groups.	EG	Similar

Study, year, and country or region	Type of IVR	Outcome measurements	Major findings	Learning outcomes favor group	Learning experience favor group
Sapkaroski et al [33], 2019, Australia	CESTOL <sup>f</sup> VR Clinic	Primary outcome: students' performance in a skill test involving taking images of hand positions with an x-ray machine	The EG performed, on average, 36% ( $P<.001$ ) better in digit separation, 11% ( $P<.001$ ) better in palm flatness, and 23% ( $P<.05$ ) better in central ray positioning onto the third metacarpal. There was no significant difference ( $P=.17$ ) in positioning between the 2 groups.	EG	N/A
Stepan et al [34], 2017, United States	Oculus Rift	Primary outcome: anatomy knowledge quizzes; secondary outcome: participants' subjective user experience via interview	Primary outcome: no significant difference was found between the EG and CG in the postintervention quiz ( $P=.87$ ) or the retention quiz ( $P=.47$ ). Secondary outcome: subjective learner experience survey—EG were more engaged ( $P<.01$ ), felt more enjoyment ( $P<.01$ ), and thought it was more useful for learning ( $P<.01$ ); IMMS <sup>g</sup> survey—the EG scored higher in the total IMMS ( $P<.01$ ) and the subscales for attention ( $P<.01$ ), confidence ( $P<.01$ ), and satisfaction ( $P<.01$ ).	Similar	EG
Bakhos et al [35], 2020, France	Oculus Rift 1	Primary outcome: examination of 20 questions; secondary outcome: posttraining satisfaction and self-confidence	Primary outcome: mean posttraining test scores showed greater improvement in the EG, but the difference between the CG and EG was not significant. Secondary outcome: satisfaction and self-confidence ratings were significantly higher for the EG than for the CG.	Similar	EG
Chao et al [36], 2021, Taiwan	HTC VIVE	Primary outcome: nasogastric tube feeding quiz; secondary outcomes: (1) confidence scale and (2) satisfaction	There were no differences in knowledge ( $P=.84$ ) and confidence ( $P=.96$ ) between the 2 groups. Within groups, the scores on knowledge and confidence improved significantly in both the EG and CG immediately and 1 month after the intervention ( $P<.01$ ). There was a significant difference in satisfaction levels between the intervention and comparison groups ( $t=2.30$ ; $P=.03$ ).	Similar	EG
Berg and Steinsbekk [37], 2020, Norway	Oculus Rift S or Oculus Quest	Primary outcome: participants' knowledge and performance of the ABCDE <sup>h</sup> approach (skill test); secondary outcome: students' experiences	Primary outcome: noninferiority of the individual IVR conducting all observations in the correct order (EG vs CG: 24.8% vs 27.1%; absolute difference: 2.3% points, one-sided 95% CI 2.3%-10.8%); secondary outcomes were similar between the groups, but more students in the EG reported liking the way they practiced and stated that it was a good way to learn. The EG also scored high on the System Usability Scale.	Similar	EG
Berg and Steinsbekk [38], 2021, Norway	Oculus Rift S or Oculus Quest	Primary outcome: participants' knowledge and performance of the ABCDE approach (skill test); secondary outcome: participants' experience	Primary outcome: 29 (20%) participants in the EG and 30 (21%) participants in the CG answered everything correctly. Knowledge and performance of the ABCDE approach were similar in the 2 groups of students, except that the EG performed better in the report on respiratory frequency and in the usability test. Secondary outcome: the EG were more displeased about the learning experience, but the difference was not significant.	Similar	Similar



Study, year, and country or region	Type of IVR	Outcome measurements	Major findings	Learning outcomes favor group	Learning experience favor group
Collaço et al [39], 2021, Brazil	Samsung	Primary outcome: skill test; secondary outcome: the participants' perceptions	Primary outcome: for the execution time—NP <sup>i</sup> and NT <sup>j</sup> took significantly longer than the full (EG) and NH <sup>k</sup> ( $P<.001$ ) groups. The full (EG) and NH groups were more accurate in needle insertion than the NP and NT ( $P<.001$ ). There were no significant differences among the groups in needle angle ( $P=.44$ ) or needle depth ( $P=.24$ ). Secondary outcome: no significant differences were found among the groups for factor 1 and factor 3. For factor 2, the NT group reported significantly ( $P<.001$ ) more difficulty than the other groups (factor 1: tactile realism; factor 2: syringe control; factor 3: ease of performance).	EG	Similar
Ros et al [40], 2020, France	Samsung Gear VR	Primary outcome: written test about indications, patient management, and preparation until the incision	The EG had significantly better results ( $P=.01$ ) in answering the questionnaire compared with the CG. The results were similar at 6 months (the scores in the EG were higher than in the CG but were nonsignificant).	EG	N/A
Yu et al [41], 2021, Korea	HTC VIVE	Student knowledge (HirNICCS <sup>l</sup> ), self-efficacy (10-point scale), and satisfaction (5-point scale)	There was no significant difference between the EG and CG ( $P=.21$ ) in knowledge. The EG showed a greater increase in self-efficacy than the CG ( $P=.02$ ). The EG had a higher satisfaction score than the CG ( $P<.001$ ).	Similar	EG

<sup>a</sup>SUMI: Software Usability Measurement Inventory.

<sup>b</sup>EG: experimental group.

<sup>c</sup>CG: control group.

<sup>d</sup>VR: virtual reality.

<sup>e</sup>N/A: not applicable.

<sup>f</sup>CESTOL: Clinical Education Training Solution.

<sup>g</sup>IMMS: Instructional Materials Motivation Survey.

<sup>h</sup>ABCDE: airways, breathing, circulation, disability, exposure.

<sup>i</sup>NP: nonpreceptorship.

<sup>j</sup>NT: nontraining.

<sup>k</sup>NH: nonhaptic feedback.





<sup>l</sup>HirNICCS: High-Risk Neonatal Infection Control Competency Scale Knowledge.

## Certainty of Evidence

The generated GRADE evidence profile was used to present a synthesis of the findings regarding objective 1 (ie, students' learning outcomes) and objective 2 (ie, students' learning

experiences) in [Table 5](#). As there were serious concerns about most of the studies with regard to the study design, inconsistent results, and a strong suspected publication bias, all the evidence was considered to have a very low level of certainty.

**Table 5.** Grading of Recommendations Assessment, Development, and Evaluation evidence profile: certainty of evidence for the learning outcomes and learning experiences (N=16).

Outcome	Studies, n (%)	Study design	Risk of bias	Inconsistency	Indirectness	Other considerations	Summary of findings			Certainty
							EG <sup>a</sup>	CG <sup>b</sup>	Direction	
Students' learning outcomes (assessed via examination and skill tests)	15 (94)	Randomized trials	Serious <sup>c</sup>	Serious <sup>d</sup>	Not serious	Publication bias strongly suspected <sup>e</sup>	6	9	EG<CG	 Very low
Students' learning on procedural skill outcomes (assessed via examination and skill tests)	13 (81)	Randomized trials	Serious <sup>f</sup>	Serious <sup>g</sup>	Not serious	Publication bias strongly suspected <sup>e</sup>	5	8	EG<CG	 Very low
Students' learning on theoretical knowledge and other outcomes (assessed via examination and skill tests)	2 (12)	Randomized trials	Serious <sup>h</sup>	Serious <sup>i</sup>	Not serious	Publication bias strongly suspected <sup>e</sup>	1	1	EG=CG	 Very low
Students' learning experience (assessed via questionnaires)	12 (75)	Randomized trials	Serious <sup>j</sup>	Serious <sup>k</sup>	Not serious	Publication bias strongly suspected <sup>e</sup>	8	4	EG>CG	 Very low

<sup>a</sup>EG: experimental group.

<sup>b</sup>CG: control group.

<sup>c</sup>A total of 27% (4/15) of the studies that evaluated the students' learning outcomes were quasi-experimental trials. According to the Joanna Briggs Institute critical appraisal checklist for randomized controlled trials, the remaining 11 randomized controlled trials had 1 to 6 items that were rated at *no* or *unclear*, which indicates that there were issues in the study design leading to a serious risk of bias.

<sup>d</sup>A total of 40% (6/15) of the studies favored immersive virtual reality teaching ( $P=.10$  in a binomial probability test showing that the null hypothesis is accepted).

<sup>e</sup>None of the papers registered their study protocol; therefore, we could not check if there was any publication bias.

<sup>f</sup>A total of 31% (4/13) of the studies were quasi-experimental trials. According to the Joanna Briggs Institute critical appraisal checklist for randomized controlled trials, the remaining 9 randomized controlled trials had 1 to 6 items that were rated at *no* or *unclear*, which indicates that there were issues in the study design leading to a serious risk of bias.

<sup>g</sup>Only 38% (5/13) of the studies favored immersive virtual reality teaching ( $P=.58$  in a binomial probability test showing that the null hypothesis is accepted).

<sup>h</sup>According to the Joanna Briggs Institute critical appraisal checklist for randomized controlled trials, these 3 randomized controlled trials have 3 to 5 items rated at *no* or *unclear*, which indicates issues in the study design.

<sup>i</sup>A total of 50% (1/2) of the studies favored immersive virtual reality teaching. One study used IVR to orient the students to the operating room.

<sup>j</sup>A total of 17% (2/12) of the studies were quasi-experimental trials. According to the Joanna Briggs Institute critical appraisal checklist for randomized controlled trials, the remaining 10 randomized controlled trials had 1 to 5 items rated at *no* or *unclear*, which indicates issues in the study design leading to a serious risk of bias.

<sup>k</sup>A total of 62% (8/13) of the studies favored immersive virtual reality teaching ( $P=.39$  in a binomial probability test showing that the null hypothesis is accepted).

## Discussion

### Principal Findings

The findings of this systematic review demonstrate that the use of IVR teaching in undergraduate health care education is effective in enhancing the procedural skills and knowledge acquisition of students. However, the effects on these learning outcomes were similar to those of other teaching approaches such as desktop-based VR or conventional classroom teaching. A vote count showed that 60% (9/15; 95% CI 16.3%-67.7%;

$P=.61$ ) of the studies identified similar learning outcomes between IVR teaching and other teaching approaches regardless of the teaching domain.

In general, in the 15 studies, the students indicated that they had positive learning experiences with IVR teaching, including increased satisfaction, self-confidence, self-assessed competency, self-efficacy, and enjoyment with IVR teaching. When compared with those in the control group who received other methods of teaching, the vote count showed that, in 62% (8/13) of the studies, the participants were in favor of using IVR as a teaching medium. However, the results of the binomial test

(95% CI 34.9%-90%;  $P=.59$ ) did not indicate a statistically significant difference. Therefore, the results indicated similarly positive learning experiences between students who received IVR and those who received other teaching approaches. Considering the low-level evidence identified based on the GRADE tool, it is inconclusive whether IVR teaching is superior to other forms of VR and to conventional teaching methods with regard to students' learning outcomes and learning experiences.

### Comparison With Prior Work

Although the application of IVR in health care teaching has become more prevalent in recent years, only a limited number of systematic reviews have focused solely on evaluating its effects. Many reviews have included all forms of VR teaching (such as 2D computer games). Therefore, the effects of IVR on teaching are yet to be confirmed. This review is one of the few to provide additional evidence on the effects of IVR in health care teaching.

The inconclusive findings on the students' learning outcomes identified in this review are similar to those of the systematic review of 29 articles by Hamilton et al [47]. Their review also reported that the effects of IVR teaching on learning outcomes and attainment levels were inconsistent compared with those of conventional desktop-based VR or the original physical training scenario [47]. These findings contradict those of the systematic review of 17 articles by Mao et al [11]. Their review reported that medical students who received IVR training performed significantly faster in the time required to complete surgical procedures and had higher scores on procedural checklists than those who received other forms of training [11]. A possible reason for the inconsistent findings may be the different target populations and objectives of the studies. Both our review and the review by Hamilton et al [47] included studies targeting the teaching of theoretical knowledge and training in procedural skills, but the review by Mao et al [11] focused mainly on training in surgical procedural skills. It is possible that IVR may be more effective for teaching procedural skills as it provides real-world simulations to give the students an immersive training experience and allows students to repeatedly practice the same procedures. However, when used in theoretical teaching, the instillation of knowledge appears to rely more on personal memorization and understanding [48].

The lack of teaching theories and pedagogies to guide the integration of IVR into health care education may be another reason for the inconsistent teaching outcomes shown in this review. This finding is also similar to that of the review by Hamilton et al [47], which identified only 1 study with a pedagogical framework to guide the development of IVR teaching [47]. Among all the included studies in this review, only the study by Smith et al [27] adopted the National League for Nurses Jeffries Simulation Theory to guide the design of simulation teaching, which is that students learn information as part of a simulated experience [46]. In general, the included studies indicated that their students' exposure to IVR learning experiences was short. A total of 75% (12/16) of the studies arranged a single IVR learning experience ranging from 5 to 30 minutes for their students. To maximize the impacts of IVR

teaching, it has been suggested that educators incorporate IVR teaching into their courses guided by a pedagogy to enrich the teaching contexts with actual experience, insightful reflections, realistic practices, and real-world connections [49].

Although the learning outcomes from IVR teaching are comparable with those from other teaching approaches, IVR still has many advantages in undergraduate health care education. In particular, it can improve the students' positive learning experiences. In total, 62% (10/16) of the studies in this review found that students favored learning with IVR, which can motivate them to learn actively rather than passively, although the results of the binomial test did not indicate a statistically significant difference. Similarly, the systematic review by Mao et al [11] reported that IVR could improve the self-confidence of medical trainees in performing surgical procedures compared with other training methods. In addition, IVR teaching can incorporate different scenarios of patients in critically ill and emergency situations, giving students valuable hands-on opportunities in a safe and controlled environment. In this review, 31% (5/16) of the studies [36-39,41] simulated real-world scenarios to train students in clinical skills such as tube feeding and conducting physical assessments of infants with respiratory distress. This provided the students with a safe environment in which to avoid unnecessary adverse events, which health care students often experience.

Although there is evidence showing that IVR can be used to enhance the professional attitudes of students, such as in the areas of empathy, decision-making, or collaborative teamwork, none of the studies in this review focused on this domain of teaching. Most (13/16, 81%) used IVR to train students in procedural skills, 12% (2/16) used it to teach anatomy, and 6% (1/16) used it for orientation. A review of 178 medical studies showed that using IVR in teaching could effectively improve medical students' understanding of the impacts of gerontological diseases on the daily life of older adults. Moreover, after learning through IVR, the students expressed more feelings of empathy toward older adults with sensory impairments and dementia [50]. Another study showed that IVR teaching can create scenarios that closely resemble those in real health care settings, allowing students to immerse themselves in practicing clinical reasoning and learn how to deal with emergencies, which can boost their ability to make decisions and determine priorities [43]. Another study tested the awareness and decision-making abilities of novice surgical residents by using IVR in surgical training. The results showed that the residents' scores improved significantly in comparison with the scores of those who received conventional PowerPoint teaching [51]. In addition, IVR teaching can help students build their communication skills. A study found that IVR teaching could significantly promote the communication skills of pediatricians to persuade parents of the merits of the influenza vaccine injection [52]. Good communication skills are a core attribute of clinical competence, and IVR teaching can allow students to practice this skill in VR clinical settings and enable them to learn how to deal with different situations [53].

This review has implications for both research and educational practice. First, the unique capabilities of IVR teaching are yet to be fully exploited. In addition to training in procedural skills

and inputting theoretical knowledge, IVR teaching has great potential to be used for other teaching domains such as health care professional attitudes. Second, health educators need to align pedagogy with IVR teaching for successful integration into undergraduate health care education [54]. Using pedagogy to guide the integration of IVR teaching into a course can maximize the impact on the students' learning outcomes by enriching teaching contexts with robust instructions, realistic practices, and real-world connections [49]. Third, it is extremely important in undergraduate health care training for skills and knowledge to be retained and applied in real clinical settings. Therefore, in addition to immediate skill tests or written examinations, other forms of assessment such as essays or group discussions can be considered for an in-depth evaluation of the students' understanding of the course [47]. It may be difficult to measure the differences between IVR and traditional teaching in terms of the transferability from knowledge to actual clinical practice. It is hard to tell whether students have transferred what they have learned from IVR into their clinical practice. Current evaluation methods focus on superficial and short-term effects. Clinically based observational studies conducted over a long period can be considered in the future. Finally, the quality of the current studies, especially the RCTs, is a concern. Standardized guidelines such as CONSORT (Consolidated Standards of Reporting Trials) [55] should be referred to when designing a study.

### Strengths and Limitations

Our review provides the most up-to-date evidence on the positive effects of IVR in undergraduate health professional education. We conducted a comprehensive search across different databases and followed the Cochrane gold-standard methodology together with a meta-synthesis while conducting this systematic review. This systematic review has some limitations. First, only articles published in English were included. It is possible that articles in other languages were overlooked. However, we adopted relatively broad eligibility criteria for the study types (RCTs and quasi-experimental trials), which led to sufficient findings. Second, the heterogeneity of the included studies prevented the data from being pooled for a meta-analysis, meaning that the effect size of the IVR on teaching outcomes could not be determined. Third, we did not analyze the learning experiences of students from data collected using qualitative methods. Although a qualitative method was only mentioned in 6% (1/16) of the included studies, it is possible that some findings related to the subjective learning experiences of students with IVR were missed. Fourth, we originally hypothesized that the types of equipment with

different levels of immersive experiences used for IVR teaching would affect students' learning outcomes and experiences. However, we found no studies that measured the level of immersive experience according to the type of VR equipment. Therefore, we could not perform any subgroup analyses to investigate these possible moderating effects on students' learning outcomes and experiences. However, in our review, we found no evidence to support this hypothesis. The use of equipment played a small role in influencing teaching and learning outcomes. Finally, this review only focused on undergraduate health care students; therefore, the findings may not be transferable to students in other majors.

### Future Directions

Despite the positive effects of IVR when adopted in undergraduate health care education in the digital era, robust assessments through high-quality, large-scale studies with long follow-up periods are still lacking. In addition, the true efficacy of IVR is best assessed through long-term integration into a real-world training guide with a pedagogical framework to maximize the effects of IVR on education. In addition to procedural training and theoretical knowledge, IVR has the potential to be used to train students in other desirable attributes of health care professionals (such as communication skills, decision-making skills, and feelings of empathy toward patients). Given its positive impacts on students' learning outcomes and experiences, we recommend further investigation with rigorous studies focusing on important outcomes for students following long-term incorporation into different health care curricula in different teaching domains.

### Conclusions

This systematic review demonstrates that the use of IVR teaching in undergraduate health care education is effective in enhancing the procedural skills and knowledge acquisition of students, although the effects on these learning outcomes were similar to those of other teaching approaches. IVR also has an advantage in enhancing the positive learning experiences of students. In total, 50% (8/16) of the studies in this review indicated that the students favored IVR teaching over other teaching methods. IVR teaching also has great potential to be used in other teaching domains such as enhancing the professional attitudes of students. Unfortunately, none of the included studies focused on those teaching domains, which are worth exploring in future studies. Finally, health educators should align pedagogy with IVR teaching for successful integration into undergraduate health care education to maximize the impact on students' learning outcomes.

---

### Acknowledgments

The research team would like to express their gratitude to the librarians at The Hong Kong Polytechnic University for recommending search strategies and to research associate Amy Cheung for formatting the manuscript.

---

### Authors' Contributions

JYWL conceived the idea for this review. JYWL and YHY searched for, reviewed, and analyzed the included studies as well as drafting the manuscript. PPKK and DSKC provided advice on methodology. Other authors critically commented on the selection of papers, the findings, and the manuscript.

---

## Conflicts of Interest

None declared.

## Multimedia Appendix 1

PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist.

[\[DOCX File , 28 KB-Multimedia Appendix 1\]](#)

## Multimedia Appendix 2

Search strategy for different databases.

[\[DOCX File , 19 KB-Multimedia Appendix 2\]](#)

## References

1. Rourke S. How does virtual reality simulation compare to simulated practice in the acquisition of clinical psychomotor skills for pre-registration student nurses? A systematic review. *Int J Nurs Stud* 2020 Feb;102:103466. [doi: [10.1016/j.ijnurstu.2019.103466](https://doi.org/10.1016/j.ijnurstu.2019.103466)] [Medline: [31783192](https://pubmed.ncbi.nlm.nih.gov/31783192/)]
2. Fealy S, Jones D, Hutton A, Graham K, McNeill L, Sweet L, et al. The integration of immersive virtual reality in tertiary nursing and midwifery education: a scoping review. *Nurse Educ Today* 2019 Aug;79:14-19. [doi: [10.1016/j.nedt.2019.05.002](https://doi.org/10.1016/j.nedt.2019.05.002)] [Medline: [31078869](https://pubmed.ncbi.nlm.nih.gov/31078869/)]
3. Chinelatto LA, Costa TR, Medeiros VM, Boog GH, Hojaij FC, Tempski PZ, et al. What you gain and what you lose in COVID-19: perception of medical students on their education. *Clinics (Sao Paulo)* 2020;75:e2133 [FREE Full text] [doi: [10.6061/clinics/2020/e2133](https://doi.org/10.6061/clinics/2020/e2133)] [Medline: [32667495](https://pubmed.ncbi.nlm.nih.gov/32667495/)]
4. Cook DA, Brydges R, Hamstra SJ, Zendejas B, Szostek JH, Wang AT, et al. Comparative effectiveness of technology-enhanced simulation versus other instructional methods: a systematic review and meta-analysis. *Simul Healthc* 2012 Oct;7(5):308-320. [doi: [10.1097/SIH.0b013e3182614f95](https://doi.org/10.1097/SIH.0b013e3182614f95)] [Medline: [23032751](https://pubmed.ncbi.nlm.nih.gov/23032751/)]
5. Hope A, Garside J, Prescott S. Rethinking theory and practice: pre-registration student nurses experiences of simulation teaching and learning in the acquisition of clinical skills in preparation for practice. *Nurse Educ Today* 2011 Oct;31(7):711-715. [doi: [10.1016/j.nedt.2010.12.011](https://doi.org/10.1016/j.nedt.2010.12.011)] [Medline: [21237536](https://pubmed.ncbi.nlm.nih.gov/21237536/)]
6. Al-Ghareeb AZ, Cooper SJ. Barriers and enablers to the use of high-fidelity patient simulation manikins in nurse education: an integrative review. *Nurse Educ Today* 2016 Jan;36:281-286. [doi: [10.1016/j.nedt.2015.08.005](https://doi.org/10.1016/j.nedt.2015.08.005)] [Medline: [26323885](https://pubmed.ncbi.nlm.nih.gov/26323885/)]
7. Chernikova O, Heitzmann N, Stadler M, Holzberger D, Seidel T, Fischer F. Simulation-based learning in higher education: a meta-analysis. *Rev Educ Res* 2020 Jun 15;90(4):499-541. [doi: [10.3102/0034654320933544](https://doi.org/10.3102/0034654320933544)]
8. Sawyer T, White M, Zaveri P, Chang T, Ades A, French H, et al. Learn, see, practice, prove, do, maintain: an evidence-based pedagogical framework for procedural skill training in medicine. *Acad Med* 2015 Aug;90(8):1025-1033. [doi: [10.1097/ACM.0000000000000734](https://doi.org/10.1097/ACM.0000000000000734)] [Medline: [25881645](https://pubmed.ncbi.nlm.nih.gov/25881645/)]
9. Cook DA. How much evidence does it take? A cumulative meta-analysis of outcomes of simulation-based education. *Med Educ* 2014 Aug;48(8):750-760. [doi: [10.1111/medu.12473](https://doi.org/10.1111/medu.12473)] [Medline: [25039731](https://pubmed.ncbi.nlm.nih.gov/25039731/)]
10. Hegland PA, Aarlie H, Strømme H, Jamtvedt G. Simulation-based training for nurses: systematic review and meta-analysis. *Nurse Educ Today* 2017 Jul;54:6-20. [doi: [10.1016/j.nedt.2017.04.004](https://doi.org/10.1016/j.nedt.2017.04.004)] [Medline: [28456053](https://pubmed.ncbi.nlm.nih.gov/28456053/)]
11. Mao RQ, Lan L, Kay J, Lohre R, Ayeni OR, Goel DP, et al. Immersive virtual reality for surgical training: a systematic review. *J Surg Res* 2021 Dec;268:40-58. [doi: [10.1016/j.jss.2021.06.045](https://doi.org/10.1016/j.jss.2021.06.045)] [Medline: [34284320](https://pubmed.ncbi.nlm.nih.gov/34284320/)]
12. Simpson R, LaViola J, Laidlaw D, Forsberg A, van Dam A. Immersive VR for scientific visualization: a progress report. *IEEE Comput Grap Appl* 2000;20(6):26-52. [doi: [10.1109/38.888006](https://doi.org/10.1109/38.888006)]
13. Brooks F. What's real about virtual reality? *IEEE Comput Grap Appl* 1999;19(6):16-27. [doi: [10.1109/38.799723](https://doi.org/10.1109/38.799723)]
14. Freina L, Ott M. A literature review on immersive virtual reality in education: state of the art and perspectives. In: *Proceedings of the eLearning and Software for Education (eLSE)*. 2015 Presented at: eLearning and Software for Education (eLSE); Apr 23-24, 2015; Bucharest, Romania.
15. Jeffries PR, Bushardt RL, DuBose-Morris R, Hood C, Kardong-Edgren S, Pintz C, et al. The role of technology in health professions education during the COVID-19 pandemic. *Acad Med* 2022 Mar 01;97(3S):S104-S109 [FREE Full text] [doi: [10.1097/ACM.0000000000004523](https://doi.org/10.1097/ACM.0000000000004523)] [Medline: [34789662](https://pubmed.ncbi.nlm.nih.gov/34789662/)]
16. McKechnie T, Levin M, Zhou K, Freedman B, Palter VN, Grantcharov TP. Virtual surgical training during COVID-19: operating room simulation platforms accessible from home. *Ann Surg* 2020 Aug;272(2):e153-e154 [FREE Full text] [doi: [10.1097/SLA.0000000000003999](https://doi.org/10.1097/SLA.0000000000003999)] [Medline: [32675522](https://pubmed.ncbi.nlm.nih.gov/32675522/)]
17. Bracq M, Michinov E, Jannin P. Virtual reality simulation in nontechnical skills training for healthcare professionals: a systematic review. *Simul Healthc* 2019 Jun;14(3):188-194 [FREE Full text] [doi: [10.1097/SIH.0000000000000347](https://doi.org/10.1097/SIH.0000000000000347)] [Medline: [30601464](https://pubmed.ncbi.nlm.nih.gov/30601464/)]

18. Rooney MK, Zhu F, Gillespie EF, Gunther JR, McKillip RP, Lineberry M, et al. Simulation as more than a treatment-planning tool: a systematic review of the literature on radiation oncology simulation-based medical education. *Int J Radiat Oncol Biol Phys* 2018 Oct 01;102(2):257-283 [FREE Full text] [doi: [10.1016/j.ijrobp.2018.05.064](https://doi.org/10.1016/j.ijrobp.2018.05.064)] [Medline: [30191859](https://pubmed.ncbi.nlm.nih.gov/30191859/)]
19. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 2009 Jul 21;6(7):e1000097 [FREE Full text] [doi: [10.1371/journal.pmed.1000097](https://doi.org/10.1371/journal.pmed.1000097)] [Medline: [19621072](https://pubmed.ncbi.nlm.nih.gov/19621072/)]
20. Campbell M, McKenzie JE, Sowden A, Katikireddi SV, Brennan SE, Ellis S, et al. Synthesis without meta-analysis (SWiM) in systematic reviews: reporting guideline. *BMJ* 2020 Jan 16;368:l6890 [FREE Full text] [doi: [10.1136/bmj.l6890](https://doi.org/10.1136/bmj.l6890)] [Medline: [31948937](https://pubmed.ncbi.nlm.nih.gov/31948937/)]
21. Shrestha N, Kukkonen-Harjula KT, Verbeek JH, Ijaz S, Hermans V, Pedisic Z. Workplace interventions for reducing sitting at work. *Cochrane Database Syst Rev* 2018 Jun 20;6(6):CD010912 [FREE Full text] [doi: [10.1002/14651858.CD010912.pub4](https://doi.org/10.1002/14651858.CD010912.pub4)] [Medline: [29926475](https://pubmed.ncbi.nlm.nih.gov/29926475/)]
22. Higgins JL, Deeks J. Chapter 6 Choosing effect measures and computing estimates of effect. In: *Cochrane Handbook for Systematic Reviews of Interventions*. Hoboken, New Jersey: Wiley; Aug 24, 2011.
23. McKenzie J, Brennan S. Chapter 12 Synthesizing and presenting findings using other methods. In: *Cochrane Handbook for Systematic Reviews of Interventions, Second Edition*. Hoboken, New Jersey: Wiley; 2019.
24. Schunemann HH, Higgins J, Vist G, Glasziou P, Skoetz N, Akl E, et al. Chapter 14 Completing 'Summary of findings' tables and grading the certainty of the evidence. In: *Cochrane Handbook for Systematic Reviews of Interventions, Second Edition*. Hoboken, New Jersey: Wiley; 2019.
25. Gutiérrez-Maldonado J, Ferrer-García M, Plasanjungano J, Andrés-Pueyo A, Talarn-Caparrós A. Virtual reality to train diagnostic skills in eating disorders. Comparison of two low cost systems. *Stud Health Technol Inform* 2015;219:75-81. [Medline: [26799883](https://pubmed.ncbi.nlm.nih.gov/26799883/)]
26. Harrington CM, Kavanagh DO, Wright Ballester G, Wright Ballester A, Dicker P, Traynor O, et al. 360° operative videos: a randomised cross-over study evaluating attentiveness and information retention. *J Surg Educ* 2018;75(4):993-1000. [doi: [10.1016/j.jsurg.2017.10.010](https://doi.org/10.1016/j.jsurg.2017.10.010)] [Medline: [29122571](https://pubmed.ncbi.nlm.nih.gov/29122571/)]
27. Smith SJ, Farra SL, Ulrich DL, Hodgson E, Nicely S, Mickle A. Effectiveness of two varying levels of virtual reality simulation. *Nurs Educ Perspect* 2018;39(6):E10-E15. [doi: [10.1097/01.NEP.0000000000000369](https://doi.org/10.1097/01.NEP.0000000000000369)] [Medline: [30335708](https://pubmed.ncbi.nlm.nih.gov/30335708/)]
28. Zackoff MW, Real FJ, Sahay RD, Fei L, Guiot A, Lehmann C, et al. Impact of an immersive virtual reality curriculum on medical students' clinical assessment of infants with respiratory distress. *Pediatr Crit Care Med* 2020 May;21(5):477-485. [doi: [10.1097/PCC.0000000000002249](https://doi.org/10.1097/PCC.0000000000002249)] [Medline: [32106189](https://pubmed.ncbi.nlm.nih.gov/32106189/)]
29. Francis ER, Bernard S, Nowak ML, Daniel S, Bernard JA. Operating room virtual reality immersion improves self-efficacy amongst preclinical physician assistant students. *J Surg Educ* 2020;77(4):947-952. [doi: [10.1016/j.jsurg.2020.02.013](https://doi.org/10.1016/j.jsurg.2020.02.013)] [Medline: [32199861](https://pubmed.ncbi.nlm.nih.gov/32199861/)]
30. Kurul R, Ögün MN, Neriman Narin A, Avci, Yazgan B. An alternative method for anatomy training: immersive virtual reality. *Anat Sci Educ* 2020 Sep;13(5):648-656. [doi: [10.1002/ase.1959](https://doi.org/10.1002/ase.1959)] [Medline: [32163659](https://pubmed.ncbi.nlm.nih.gov/32163659/)]
31. Pierce J, Gutiérrez F, Vergara VM, Alverson DC, Qualls C, Saland L, et al. Comparative usability studies of full vs. partial immersive virtual reality simulation for medical education and training. *Stud Health Technol Inform* 2008;132:372-377. [Medline: [18391324](https://pubmed.ncbi.nlm.nih.gov/18391324/)]
32. Gutiérrez F, Pierce J, Vergara VM, Coulter R, Saland L, Caudell TP, et al. The effect of degree of immersion upon learning performance in virtual reality simulations for medical education. *Stud Health Technol Inform* 2007;125:155-160. [Medline: [17377256](https://pubmed.ncbi.nlm.nih.gov/17377256/)]
33. Sapkaroski D, Baird M, Mundy M, Dimmock MR. Quantification of student radiographic patient positioning using an immersive virtual reality simulation. *Simul Healthc* 2019 Aug;14(4):258-263. [doi: [10.1097/SIH.0000000000000380](https://doi.org/10.1097/SIH.0000000000000380)] [Medline: [31274828](https://pubmed.ncbi.nlm.nih.gov/31274828/)]
34. Stepan K, Zeiger J, Hanchuk S, Del Signore A, Shrivastava R, Govindaraj S, et al. Immersive virtual reality as a teaching tool for neuroanatomy. *Int Forum Allergy Rhinol* 2017 Oct;7(10):1006-1013. [doi: [10.1002/alr.21986](https://doi.org/10.1002/alr.21986)] [Medline: [28719062](https://pubmed.ncbi.nlm.nih.gov/28719062/)]
35. Bakhos D, Galvin J, Aoustin J, Robier M, Kerneis S, Bechet G, et al. Training outcomes for audiology students using virtual reality or traditional training methods. *PLoS One* 2020;15(12):e0243380 [FREE Full text] [doi: [10.1371/journal.pone.0243380](https://doi.org/10.1371/journal.pone.0243380)] [Medline: [33270806](https://pubmed.ncbi.nlm.nih.gov/33270806/)]
36. Chao Y, Hu SH, Chiu H, Huang P, Tsai H, Chuang Y. The effects of an immersive 3d interactive video program on improving student nurses' nursing skill competence: a randomized controlled trial study. *Nurse Educ Today* 2021 Aug;103:104979. [doi: [10.1016/j.nedt.2021.104979](https://doi.org/10.1016/j.nedt.2021.104979)] [Medline: [34049120](https://pubmed.ncbi.nlm.nih.gov/34049120/)]
37. Berg H, Steinsbekk A. Is individual practice in an immersive and interactive virtual reality application non-inferior to practicing with traditional equipment in learning systematic clinical observation? A randomized controlled trial. *BMC Med Educ* 2020 Apr 22;20(1):123 [FREE Full text] [doi: [10.1186/s12909-020-02030-7](https://doi.org/10.1186/s12909-020-02030-7)] [Medline: [32326948](https://pubmed.ncbi.nlm.nih.gov/32326948/)]
38. Berg H, Steinsbekk A. The effect of self-practicing systematic clinical observations in a multiplayer, immersive, interactive virtual reality application versus physical equipment: a randomized controlled trial. *Adv Health Sci Educ Theory Pract* 2021 May;26(2):667-682 [FREE Full text] [doi: [10.1007/s10459-020-10019-6](https://doi.org/10.1007/s10459-020-10019-6)] [Medline: [33511505](https://pubmed.ncbi.nlm.nih.gov/33511505/)]

39. Collaço E, Kira E, Sallaberry LH, Queiroz AC, Machado MA, Crivello O, et al. Immersion and haptic feedback impacts on dental anesthesia technical skills virtual reality training. *J Dent Educ* 2021 Apr;85(4):589-598. [doi: [10.1002/jdd.12503](https://doi.org/10.1002/jdd.12503)] [Medline: [33274441](https://pubmed.ncbi.nlm.nih.gov/33274441/)]
40. Ros M, Debien B, Cyteval C, Molinari N, Gatto F, Lonjon N. Applying an immersive tutorial in virtual reality to learning a new technique. *Neurochirurgie* 2020 Aug;66(4):212-218. [doi: [10.1016/j.neuchi.2020.05.006](https://doi.org/10.1016/j.neuchi.2020.05.006)] [Medline: [32623059](https://pubmed.ncbi.nlm.nih.gov/32623059/)]
41. Yu M, Yang M, Ku B, Mann JS. Effects of virtual reality simulation program regarding high-risk neonatal infection control on nursing students. *Asian Nurs Res (Korean Soc Nurs Sci)* 2021 Aug;15(3):189-196 [FREE Full text] [doi: [10.1016/j.anr.2021.03.002](https://doi.org/10.1016/j.anr.2021.03.002)] [Medline: [33894407](https://pubmed.ncbi.nlm.nih.gov/33894407/)]
42. Harrington CM, Kavanagh DO, Quinlan JF, Ryan D, Dicker P, O'Keeffe D, et al. Development and evaluation of a trauma decision-making simulator in Oculus virtual reality. *Am J Surg* 2018 Jan;215(1):42-47. [doi: [10.1016/j.amjsurg.2017.02.011](https://doi.org/10.1016/j.amjsurg.2017.02.011)] [Medline: [28262203](https://pubmed.ncbi.nlm.nih.gov/28262203/)]
43. Adhikari R, Kydonaki C, Lawrie J, O'Reilly M, Ballantyne B, Whitehorn J, et al. A mixed-methods feasibility study to assess the acceptability and applicability of immersive virtual reality sepsis game as an adjunct to nursing education. *Nurse Educ Today* 2021 Aug;103:104944. [doi: [10.1016/j.nedt.2021.104944](https://doi.org/10.1016/j.nedt.2021.104944)] [Medline: [34015677](https://pubmed.ncbi.nlm.nih.gov/34015677/)]
44. Pritchard A. *Ways of Learning Learning Theories for the Classroom*. Milton Park, Abingdon-on-Thames, Oxfordshire, England, UK: Routledge; 2017.
45. Schunk D. *Learning Theories An Educational Perspective*. London, United Kingdom: Pearson; 2014.
46. Jeffries PR, Rodgers B, Adamson K. NLN Jeffries simulation theory: brief narrative description. *Nurs Educ Perspect* 2015;36(5):292-293. [doi: [10.5480/1536-5026-36.5.292](https://doi.org/10.5480/1536-5026-36.5.292)] [Medline: [26521496](https://pubmed.ncbi.nlm.nih.gov/26521496/)]
47. Hamilton D, McKechnie J, Edgerton E, Wilson C. Immersive virtual reality as a pedagogical tool in education: a systematic literature review of quantitative learning outcomes and experimental design. *J Comput Educ* 2020 Jul 11;8(1):1-32. [doi: [10.1007/s40692-020-00169-2](https://doi.org/10.1007/s40692-020-00169-2)]
48. Polanyi M. *Personal Knowledge Towards a Post-critical Philosophy*. Chicago: University of Chicago Press; 1974.
49. Bonasio A. Immersive experiences in education new places and spaces for learning. Microsoft. URL: <https://edudownloads.azureedge.net/msdownloads/MicrosoftEducation/ImmersiveExperiencesEducation2019.pdf> [accessed 2022-11-08]
50. Dyer E, Swartzlander BJ, Gugliucci MR. Using virtual reality in medical education to teach empathy. *J Med Libr Assoc* 2018 Oct;106(4):498-500 [FREE Full text] [doi: [10.5195/jmla.2018.518](https://doi.org/10.5195/jmla.2018.518)] [Medline: [30271295](https://pubmed.ncbi.nlm.nih.gov/30271295/)]
51. Pulijala Y, Ma M, Pears M, Peebles D, Ayoub A. Effectiveness of Immersive Virtual Reality in Surgical Training-A Randomized Control Trial. *J Oral Maxillofac Surg* 2018 May;76(5):1065-1072. [doi: [10.1016/j.joms.2017.10.002](https://doi.org/10.1016/j.joms.2017.10.002)] [Medline: [29104028](https://pubmed.ncbi.nlm.nih.gov/29104028/)]
52. Real FJ, DeBlasio D, Beck AF, Ollberding NJ, Davis D, Cruse B, et al. A virtual reality curriculum for pediatric residents decreases rates of influenza vaccine refusal. *Acad Pediatr* 2017;17(4):431-435. [doi: [10.1016/j.acap.2017.01.010](https://doi.org/10.1016/j.acap.2017.01.010)] [Medline: [28126612](https://pubmed.ncbi.nlm.nih.gov/28126612/)]
53. Fertleman C, Aubugeau-Williams P, Sher C, Lim A, Lumley S, Delacroix S, et al. A discussion of virtual reality as a new tool for training healthcare professionals. *Front Public Health* 2018;6:44 [FREE Full text] [doi: [10.3389/fpubh.2018.00044](https://doi.org/10.3389/fpubh.2018.00044)] [Medline: [29535997](https://pubmed.ncbi.nlm.nih.gov/29535997/)]
54. Domingo JR, Bradley EG. Education student perceptions of virtual reality as a learning tool. *J Educational Technol Sys* 2017 Oct 24;46(3):329-342. [doi: [10.1177/0047239517736873](https://doi.org/10.1177/0047239517736873)]
55. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010 Mar 23;340:c332 [FREE Full text] [doi: [10.1136/bmj.c332](https://doi.org/10.1136/bmj.c332)] [Medline: [20332509](https://pubmed.ncbi.nlm.nih.gov/20332509/)]

## Abbreviations

**CONSORT:** Consolidated Standards of Reporting Trials

**GRADE:** Grading of Recommendations Assessment, Development, and Evaluation

**IVR:** immersive virtual reality

**JBI:** Joanna Briggs Institute

**PICOS:** Population, Intervention, Comparison, Outcome, and Study Design

**PRISMA:** Preferred Reporting Items for Systematic Reviews and Meta-Analyses

**RCT:** randomized controlled trial

**SWiM:** Synthesis Without Meta-analysis

*Edited by T Leung; submitted 30.05.22; peer-reviewed by PH Liao, Y Xu, M Mbwogge; comments to author 06.09.22; revised version received 09.11.22; accepted 24.01.23; published 06.03.23*

*Please cite as:*

*Liu JYW, Yin YH, Kor PPK, Cheung DSK, Zhao IY, Wang S, Su JJ, Christensen M, Tyrovolas S, Leung AYM*

*The Effects of Immersive Virtual Reality Applications on Enhancing the Learning Outcomes of Undergraduate Health Care Students: Systematic Review With Meta-synthesis*

*J Med Internet Res 2023;25:e39989*

*URL: <https://www.jmir.org/2023/1/e39989>*

*doi: [10.2196/39989](https://doi.org/10.2196/39989)*

*PMID:*

©Justina Yat Wa Liu, Yue-Heng Yin, Patrick Pui Kin Kor, Daphne Sze Ki Cheung, Ivy Yan Zhao, Shanshan Wang, Jing Jing Su, Martin Christensen, Stefanos Tyrovolas, Angela Y M Leung. Originally published in the Journal of Medical Internet Research (<https://www.jmir.org>), 06.03.2023. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://www.jmir.org/>, as well as this copyright and license information must be included.