Outcomes, measurement instruments, and their validity evidence in randomized controlled trials on virtual, augmented, and mixed reality in undergraduate medical education: systematic mapping review

Car, L.T.; Kyaw, B.M.; Teo, A.; Fox, T.E.; Vimalesvaran, S.; Apfelbacher, C.; ... ; Chavannes, N.

Review

Outcomes, Measurement Instruments, and Their Validity Evidence in Randomized Controlled Trials on Virtual, Augmented, and Mixed Reality in Undergraduate Medical Education: Systematic Mapping Review

Lorainne Tudor Car[1,2], MD, MSc, PhD; Bhone Myint Kyaw[3], MBBS, MSc, PhD; Andrew Teo[1], PhD; Tatiana Erlikh Fox[3,4], MD, MSc; Sunitha Vimalesvaran[1], MD; Christian Apfelbacher[5,6], PhD; Sandra Kemp[7], PhD; Niels Chavannes[8], PhD

[1]Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Singapore

[2]Department of Primary Care and Public Health, School of Public Health, Imperial College London, London, United Kingdom

[3]Centre for Population Health Sciences, Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Singapore

[4]Department of Internal Medicine, Onze Lieve Vrouwen Gasthuis, Amsterdam, Netherlands

[5]Institute of Social Medicine and Health Systems Research, Otto von Guericke University Magdeburg, Magdegurg, Germany

[6]Family Medicine and Primary Care, Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Singapore

[7]Faculty of Health Sciences, Curtin Medical School, Curtin University, Bentley, Australia

[8]Department of Public Health and Primary Care, Leiden University Medical Centre, Leiden, Netherlands

Corresponding Author:
Lorainne Tudor Car, MD, MSc, PhD
Lee Kong Chian School of Medicine
Nanyang Technological University
Clinical Sciences Building
11 Mandalay Road
Singapore, 308232
Singapore
Phone: 65 69041258
Email: lorainne.tudor.car@ntu.edu.sg

Abstract

Background: Extended reality, which encompasses virtual reality (VR), augmented reality (AR), and mixed reality (MR), is increasingly used in medical education. Studies assessing the effectiveness of these new educational modalities should measure relevant outcomes using outcome measurement tools with validity evidence.

Objective: Our aim is to determine the choice of outcomes, measurement instruments, and the use of measurement instruments with validity evidence in randomized controlled trials (RCTs) on the effectiveness of VR, AR, and MR in medical student education.

Methods: We conducted a systematic mapping review. We searched 7 major bibliographic databases from January 1990 to April 2020, and 2 reviewers screened the citations and extracted data independently from the included studies. We report our findings in line with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines.

Results: Of the 126 retrieved RCTs, 115 (91.3%) were on VR and 11 (8.7%) were on AR. No RCT on MR in medical student education was found. Of the 115 studies on VR, 64 (55.6%) were on VR simulators, 30 (26.1%) on screen-based VR, 9 (7.8%) on VR patient simulations, and 12 (10.4%) on VR serious games. Most studies reported only a single outcome and immediate postintervention assessment data. Skills outcome was the most common outcome reported in studies on VR simulators (97%), VR patient simulations (100%), and AR (73%). Knowledge was the most common outcome reported in studies on screen-based VR (80%) and VR serious games (58%). Less common outcomes included participants' attitudes, satisfaction, cognitive or mental load, learning efficacy, engagement or self-efficacy beliefs, emotional state, competency developed, and patient outcomes. At least one form of validity evidence was found in approximately half of the studies on VR simulators (55%), VR patient simulations (56%), VR serious games (58%), and AR (55%) and in a quarter of the studies on screen-based VR (27%). Most studies used

assessment methods that were implemented in a nondigital format, such as paper-based written exercises or in-person assessments where examiners observed performance (72%).

Conclusions: RCTs on VR and AR in medical education report a restricted range of outcomes, mostly skills and knowledge. The studies largely report immediate postintervention outcome data and use assessment methods that are in a nondigital format. Future RCTs should include a broader set of outcomes, report on the validity evidence of the measurement instruments used, and explore the use of assessments that are implemented digitally.

Introduction

Background

Extended reality (ER) encompasses immersive technologies within the reality-virtuality continuum, such as virtual reality (VR), augmented reality (AR), and mixed reality (MR). The use of ER technologies is becoming more common in medical education. These technologies offer a wide range of educational opportunities within different medical specialties. VR is a technology that renders a fully computer-generated 3D multimedia environment in real time. It supports a first-person active-learning experience through immersion, that is, a perception of the digital world as real. VR can be integrated with other educational approaches such as virtual patients or serious games. VR patient simulations are interactive computer simulations of real-life clinical scenarios for the purpose of medical education. VR serious games incorporate gaming concepts such as different levels of difficulties, rewards, or feedback within the computer-generated 3D environment.

AR is a technology in which the real-world environment is enhanced by computer-generated virtual imagery information. In AR, virtual objects are projected over the real-world environment. MR is a hybrid technology that merges the features of VR and AR. In MR, virtual objects become a part of the real word. ER technologies can be displayed through desktop computers, mobile devices, and large screens or projected on the walls. They can be purely screen based or also involve the use of joysticks, probes, gloves, simulators, and other forms of haptic devices.

Effectiveness of VR

Our systematic review on the effectiveness of VR for health professions education showed that VR may improve postintervention knowledge and skills outcomes compared with traditional education (ie, nondigital education) or other types of digital education such as online or offline digital education [1]. Data for other outcomes were limited. Systematic reviews of randomized controlled trials (RCTs) remain the gold standard for evidence on the effectiveness of interventions. However, the heterogeneity of participants, interventions, comparison interventions, and outcomes reported in the individual studies can limit the trustworthiness of the systematic review findings and preclude a meta-analysis. Similarly, differences in measurement instruments and types of validity evidence can lead to unreliable conclusions [2]. The choice of digital education outcomes can be influenced by different factors, including types of digital education, the curriculum, and the field of study [3,4]. The process of measuring digital education outcomes can be achieved with a wide variety of measurement instruments, including multiple-choice questions, structured essays, and structured direct observations with checklists for ratings [5]. Measurement instruments used in research need to have validity evidence. Validity is defined as "the degree to which evidence and theory support the interpretations of test scores entailed by the proposed uses of tests" [6]. Validity evidence for measurement instruments is important to ensure that the instruments reliably measure what they purport to measure and to support the interpretation of assessment data. However, reporting of validity evidence of measurement instruments in health professions education literature is still suboptimal, ranging from 34.6% in studies on continuing medical education to 64% in studies on technology-enhanced health professions simulation training [7,8].

The use of measurement instruments without validity evidence severely undermines the credibility of the research results [9]. ER is increasingly used in medical education, and studies in this field should evaluate diverse outcomes using outcome measurement instruments with validity evidence. Our aim is to support this by mapping the current choice of outcomes, measurement instruments, and the prevalence of measurement instruments with validity evidence in RCTs on the use of ER in undergraduate and preregistration medical education.

Methods

Methodology, Definitions, and Eligibility Criteria

We performed this systematic review in line with the Cochrane gold standard systematic review methodology and report it according to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) standards of quality for reporting systematic reviews [10,11]. In this review, we aim to answer the following research questions:

1. Which outcomes (eg, knowledge, skills, attitudes, and behavior) are assessed and reported in RCTs on the effectiveness of VR, AR, and MR in undergraduate and preregistration medical education?

2. What type of measurement instruments were used in RCTs on the use of VR, AR, and MR in undergraduate and preregistration medical education?

XSL·FO

RenderX

3. What proportion of RCTs on the use of VR, AR, and MR in undergraduate medical education report validity evidence for the measurement instruments used, and how was the evidence reported?

We included studies meeting the following eligibility criteria:

1. RCTs
2. Studies on students participating in preregistration or undergraduate medical education in any geographical or educational setting
3. Studies evaluating any type of blended (ie, a combination of extended and nondigital, traditional education) or full ER technology, including VR, AR, and MR

4. Studies comparing VR with control interventions such as classroom-based learning, no intervention, and other types of digital and blended education

We defined different ER technologies as per Textbox 1. Preregistration or undergraduate medical education was defined in line with the World Health Organization (WHO) definition as "any type of initial study leading to a qualification that (i) is recognized by the relevant governmental or professional bodies of the country where the study was conducted and (ii) enables its holder primary entry into the healthcare workforce" [12]. Studies were excluded if they focused on traditional and complementary medicine as defined by WHO (as such education is not included in most medical schools) and used study designs other than an RCT [13].

Textbox 1. Descriptions and classification of different types of virtual reality (VR), augmented reality (AR), and mixed reality (MR).

Types of extended reality modalities in medical education

- VR is a technology that allows the user to explore and manipulate computer-generated 2D or 3D, multimedia sensory environments in real time [14]. The VR environment is the computer-generated representation of a real or artificial environment that can be interacted with by external involvement, allowing for a first-person active-learning experience through immersion [15].

- Screen-based VR interventions are computer-based 3D software applications delivered either through computer screens or head-mounted displays (ie, VR headsets). This type of VR in medical education mostly includes 3D models of organs and VR worlds.

- VR simulators or psychomotor skills trainers encompass use of VR technology and physical probes or objects that help the learners to connect with the objects from the VR environment and convey feedback or tactile sensation to the learners.

- VR patient simulation refers to the interactive computer simulations of real-life clinical scenarios in VR for the purpose of medical training, education, or assessment [16]. They include virtual patients represented by computer-generated 2D or 3D characters or avatars.

- VR serious gaming or gamification intervention involves gaming concepts such as different levels of difficulties, rewards, feedback, and so on, within the computer-generated VR environment for learning purposes.

- AR is a technology that allows a live real-time direct or indirect real-world environment to be augmented or enhanced by computer-generated virtual imagery information (eg, smart, virtually enhanced glasses). Computer-generated information is overlaid on the real-world environment. AR is distinct from VR in which only a computer-generated image is supplied to the user [17].

- MR is a hybrid technology that merges the features of VR and AR [18]. In MR, physical and virtual or digital objects are displayed together and the features of *virtuality* and reality are merged for the learners [19].

Electronic Searches

We developed a comprehensive search strategy for MEDLINE (Ovid), Embase (Elsevier), Cochrane Central Register of Controlled Trials (Wiley), PsycINFO (Ovid), Education Resources Information Center (Ovid), CINAHL (EBSCO), and Web of Science Core Collection (Thomson Reuters). Databases were searched from January 1990 until April 2020 without language restrictions.

We used 1990 as the starting year for our search because before 1990, the use of computers was uncommon for educational use. We used the MEDLINE strategy presented in Multimedia Appendix 1. This was adapted to search the other databases with the help of a librarian (Ms Yasmin Munro). To identify unpublished studies, we searched the International Clinical Trials Registry Platform Search Portal and metaRegister of Controlled Trials. We also checked reference lists of relevant systematic reviews and potentially eligible studies against the inclusion criteria.

Search results across different databases were compiled using EndNote X8 software (Clarivate), and duplicate records were removed. In all, two pairs of two reviewers (BMK, AT, TEF, and SV) independently screened the studies, extracted the data, and carried out data analysis. Any disagreements were resolved by a discussion between the 2 reviewers, with a third reviewer acting as an arbiter if needed. The PRISMA flow diagram was used to report the selection and inclusion of studies [10].

Data Extraction

The data for each of the included studies were independently extracted and managed by 2 reviewers using a structured data recording form, which included information about the study characteristics such as reference of the study, country of the study, the WHO region of the study, name of measurement instrument, description of measurement instrument, types of outcomes reported, assessment category of measurement instrument [5], assessment method of measurement instrument, types of participants, sample size, raters of the instrument, procedure of identifying the raters, and training of the raters for the instruments [20]. We recorded all information relating to validity evidence sources and measurement properties that were reported directly in the articles [5,6]. We also recorded any validity evidence recorded indirectly; for example, through a reference to a validation study focusing on a particular measurement instrument. If the studies presented more than one

outcome measure, relevant details of the second outcome measure were also recorded. The data extraction form was piloted and amended according to feedback received. We contacted the study authors for further data in case of missing information.

Data Analysis and Synthesis

We analyzed and synthesized the data as follows: (1) we ascertained the types of primary and secondary outcome measurement instruments; (2) we classified and mapped the data according to types of outcomes (eg, knowledge, skills, attitudes, satisfaction, or competencies); intervention (eg, VR vs classroom-based learning and VR vs serious gaming); year of medical studies (ie, first year, second year, or final year), types of measurement instruments (eg, written exercises [surveys with only multiple-choice questions and surveys with other types of questions and essays] vs in-person assessment where an examiner observed performance [eg, global ratings, structured direct observation, and objective structured clinical examinations]); assessment delivery mode (ie, digital vs classroom-based assessment); and discipline (eg, laparoscopic surgery, anatomy, and internal medicine); and (3) we determined the proportion of RCTs on the use of VR, AR, and MR in undergraduate medical education using measurement instruments with sufficient validity evidence in relation to the goal of the measurements (*validity evidence*). The aim of this study is to comprehensively document outcomes and measurement instruments rather than to synthesize data about the effect of the interventions [6]. Therefore, we did not undertake a risk-of-bias assessment of the studies because it was not relevant to the objectives of this review.

We assessed the validity evidence of the measurement instruments as reported in the cited validation studies using the Consensus-Based Standards for the Selection of Health Measurement Instruments (COSMIN) taxonomy of measurement properties [21]. The COSMIN taxonomy outlines three measurement properties or validity evidence domains: reliability, validity, and responsiveness. The reliability domain encompasses measurement properties such as internal consistency, reliability, and measurement error. The domain validity contains the measurement properties such as content validity (including face validity), construct validity (including structural validity, hypotheses testing, and cross-cultural validity and measurement invariance), and criterion validity [21].

Digital assessments were defined as assessments that were delivered exclusively using digital technology (ie, PCs, laptops, mobile phones, and tablets) and included online surveys,

questionnaires, computer scoring, or the use of software metrics such as time to completion, number of errors, path length, and so on. Assessments in which digital tools (eg, video recordings or Microsoft PowerPoint presentations) were used to facilitate classroom-based assessment, such as written exercises or in-person observation by the examiners, were not categorized as digital assessments.

Ethics Approval

This systematic mapping review is an analysis of published studies and as such, did not require an ethics approval.

Results

Study Characteristics

The searches identified 59,483 records through electronic databases, of which we included 126 (0.21%) RCTs. Of the 126 RCTS, 115 (91.3%) assessed different forms of VR, whereas 11 (8.7%) focused on AR simulations (Figure 1). We did not find any study evaluating the use of MR in medical student education.

Of the 115 included articles focusing on VR-based training for medical student education, 64 (55.7%) focused on VR-based psychomotor skills training [22-85], 30 (26.1%) on screen-based VR [86-115], 9 (7.8%) on VR patient simulations [116-124], and 12 (10.4%) on VR serious gaming and gamification [125-136]. Only 8.7% (11/126) of the included studies focused on AR simulations [137-147] and none focused on MR training in medical student education. The included studies were published between 1997 and 2020. Most of the studies were from high-income countries, except for 8.7% (11/126) of the studies, which were conducted in low- and middle-income countries [35,36,72,75,105,114,126,127,132,134,139]. Of the 126 studies, 31 (24.6%) cited validation studies for the measurement instruments used [23,25,27,30-32,34-36,47,48, 52,58,60,63-65,70,72,78,79,82,84,92,101,118-120,126,128,133] (Multimedia Appendices 2 and 3).

Participants included medical students from the first to sixth year of medical schools (N=9010). The studies compared the use of VR and AR training (either stand-alone intervention or blended with traditional, nondigital learning) with traditional, nondigital learning or a different form of VR and AR training or other forms of digital education such as online digital education or offline digital education. Of the 64 studies focused on the effects of VR simulators for medical student education, 61 (95%) were delivered in a university setting, whereas 3 (5%) were conducted in a hospital setting [37,72,74].

Figure 1. Study flow diagram. VR: virtual reality.
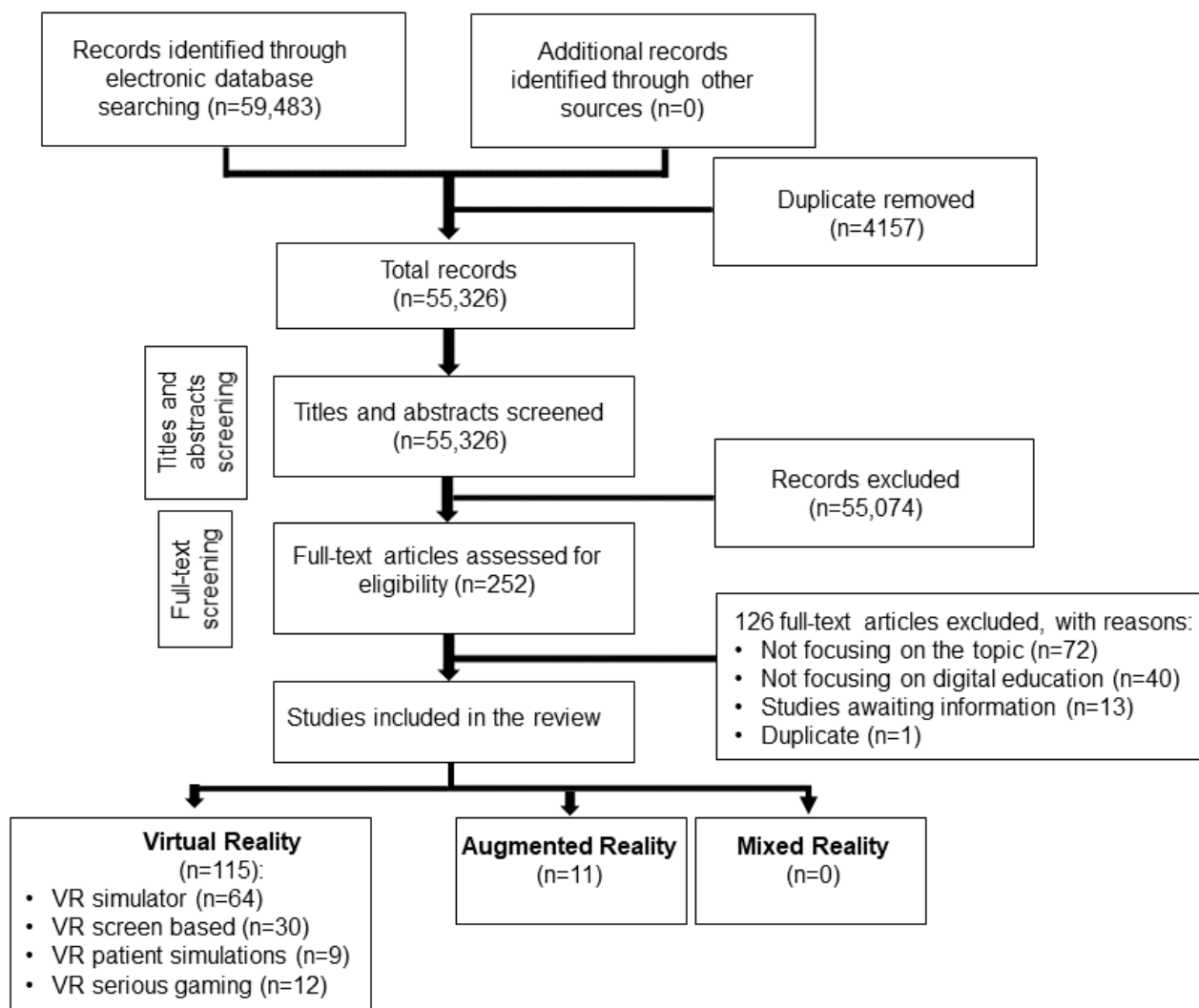
VR Simulators

Of the 115 VR studies, 64 (55.6%) with 3132 medical students evaluated the effects of VR simulators in medical student education [22-85]. The studies included first year to sixth year medical students and were published between 2001 and 2020. In terms of the topic or area of study, 53% (34/64) of the studies focused on laparoscopic surgery [22,24,27,31,35-38,40, 41,45-48,50-52,54,56-66,69,78,81-83]; 16% (10/64) on surgery [25,28,53,55,67,68,71,74,76,77]; 8% (5/64) on orthopedic surgery [39,42,73,79,84]; 8% (5/64) on ureteroscopy [30,33,34,80,85]; 5% (3/64) each on ophthalmology [26,70,75] and intravenous cannulation [29,32,72], and 2% (1/64) each on endoscopy [49], colonoscopy [23], shoulder-joint clinical anatomy [44], and empathic communication skills [43].

For the outcomes, 97% (62/64) of the studies reported on participants' postintervention skills [22-43,45-53,55-85], 8% (5/64) on knowledge [28,37,44,54,65], 14% (9/64) on attitudes toward the intervention [31,32,44,48,54,65,66,71,75], 3% (2/64) on satisfaction [68,71] and 6% (4/64) on cognitive load [25,27,39,63](Figure 2). Of the 62 studies that reported on participants' postintervention skills, 11 (18%) reported change score from baseline for the skills outcome [25,50,56,58,68,73,76-78,80,85] and 1 (2%) reported change score from baseline for the satisfaction outcome [68]. Regarding retention, 7.8% (5/64) of the studies assessed skills retention at 2-4 weeks after the intervention [25,31,33,40,83]. The remaining studies did not report retention outcomes.

Figure 2. Types of reported outcomes in virtual reality (VR)– and augmented reality (AR)–based training.
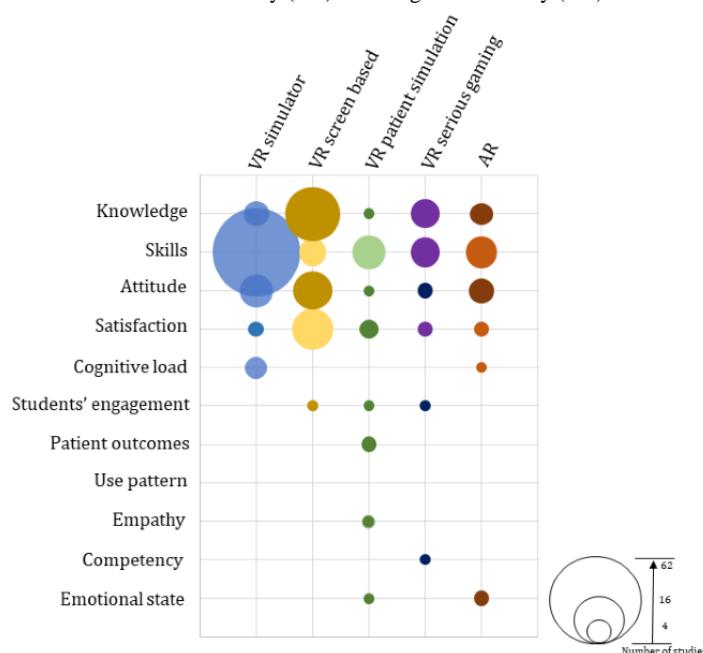
For modes of assessment, 46.8% (30/64) of the studies used paper-based written assessments or in-person assessments (ie, nondigital) using checklists by the examiners [24,25,31-37,39, 46,47,51-55,58,60,65,68,70-72,75,79-82,84]; 31% (20/64) used digital assessments such as software-based metrics (eg, time spent on training, number of errors, total path length, motion analysis, or checklists) [22,23,26,29,40-42,45,49,50,56,57,59, 61,62,64,69,73,78,85]; 11% (7/64) used a combination of digital assessments using software-based metrics, paper-based written assessments, or in-person assessments by supervising examiners [27,38,43,44,48,63,66] and 2% (1/64) used both paper-based written assessments and in-person assessments using checklists [37]. In 10.1% (7/64) of the studies, the mode of assessment was unspecified [28,30,67,74,76,77,83].

For validity evidence, 54.6% (35/64) of the studies reported a single form of validity evidence (mostly either internal consistency or reliability) for the measurement instruments largely used for assessment of skills [22,23,25,27,30-37,39,40, 47,48,51-55,58,60,63-66,68,70,72,78-80,82,84] (Multimedia Appendices 2 and 3). The remaining studies did not provide any information on the validity of assessment tools used for measuring the outcomes. Of the 64 studies, 23 (36%) referenced pertinent measurement instrument validation studies, largely used for assessment of skills (mostly either internal consistency or reliability) for the measurement instruments largely used for assessment of skills [23,25,27,30-32,34-36,47,48,52,58, 60,63-65,70,72,78,79,82,84]. Of the measurement properties, these studies mostly reported internal consistency and reliability, followed by structural validity and hypotheses testing.

Screen-Based VR

Of the 115 VR studies, 30 (26.1%) studies with 2409 medical students evaluated the effect of screen-based or nontechnical training for medical students [86-115]. The studies included

first year to sixth year medical students and were published between 1997 and 2020. In terms of the topic or area of study, 37% (11/30) of the studies focused on anatomy [87,91, 95-98,100,102,104,106,114]; 17% (5/30) on ophthalmology [93,109,112,113,115]; 17% (5/30) on surgery [88,90, 92,101,105]; 6% (2/30) each on patient examination [99,108] and one study each (3%,1/30) on operating room introduction [107], biomechanics of the spine [89], histology [111], trauma [94], traumatic head injury [86], radiology [103], and genetics [110].

For the outcomes, 80% (24/30) of the studies reported on participants' postintervention knowledge [89,91,93-107, 109-115], 17% (5/30) on skills [88,92,99,101,107], 40% (12/30) on attitudes toward topics and interventions [86,87,90,91, 95,97,102-104,107,108,115], 47% (14/30) on satisfaction [87,89,91-93,97,98,100, 102,105,109,112-114] and 3% (1/30) on students' learning engagement [89] (Figure 2). Of the 24 studies assessing knowledge, 5 (2%) also reported change score from baseline [101,104,105,113,114]. Similarly, 20% (1/5) of the studies assessing skills [101], 17% (2/12) of the studies assessing attitude toward the intervention [90,104], and 21% (3/14) of the studies assessing satisfaction [105,113,114] also reported change score from baseline. Regarding retention, only a single study assessed retention at 12 months after the intervention [112]. The remaining studies did not report outcomes at the follow-up stages.

Most of the studies (21/30, 70%) used paper-based written assessments [86,87,89-91,93,95,97,98,100,102-104,108-115]. Other forms of assessment included in-person assessments by an examiner [88], digital assessment in the form of questionnaires and ratings [94,105,106], combined paper-based written and in-person assessments [92,99,101,107], and a paper-based written assessment with questions delivered in the form of a PowerPoint presentation [96].

Of the 30 studies, 8 (27%) reported at least one form of validity evidence (mostly reliability) for the measurement instruments that were largely used to assess skills [88,91,92, 98,99,101,107,108]. Of these 8 studies, 2 (25%) referenced measurement instrument validation studies, both focusing on skills assessment and reporting on their reliability [92,101].

VR Patient Simulations

Of the 115 VR studies, 9 (7.8%) with 782 medical students evaluated the effect of VR-based patient simulations in medical student education simulations [116-124]. Of these 9 studies, 4 (44%) focused on communication skills [117-119,124]; 2 (22%) on pediatric life support [121,122]; and 1 (11%) each on clinical reasoning [123], internal medicine [116], and suicide risk assessment [120] (Figure 2).

For the outcomes, 11% (1/9) of the studies reported on participants' postintervention knowledge [122], 100% (9/9) on skills [116-124], 33% (3/9) on students' satisfaction [119,120,123], 22% (2/9) on patient-related outcomes (eg, patients' satisfaction) [119,120], and 11% (1/9) each on attitudes toward the intervention [124], engagement [123], mood changes or emotional state [124], and empathetic behavior [117]. None of the studies reported change score from baseline or retention data.

For mode of assessment, most of the studies used in-person assessments by an examiner [116-120,123,124] or paper-based written assessments [119,120,122,123]. Of the 9 studies, 2 (22%) used both paper-based written and in-person assessments by an examiner [119,120]; 1 (11%) used both digital assessments consisting of virtual patients and scoring and in-person assessment by an examiner [116]; and, finally, 1 (11%) used a combined assessment of digital assessment in the form of a survey, in-person assessment by an examiner, and paper-based written assessment for different outcomes [123].

Of the 9 studies, 5 (56%) reported at least one form of validity evidence (mostly internal consistency and reliability) for the measurement instruments used to assess skills [116-120] (Multimedia Appendices 2 and 3). Of these 5 studies, 3 (60%) referenced measurement instrument validation studies: 67% (2/3) focused on assessment of patient satisfaction [119,120] and 33% (1/3) on skills [118]. The measurement properties mentioned in the referenced validation studies were internal consistency and reliability, followed by internal validity.

VR Serious Gaming and Gamification

Of the 115 studies, 12 (10.4%) with 743 medical students evaluated the effects of VR serious gaming and gamification in medical student education [125-136]. The studies included participants from the first to fifth year of studies and were published between 2008 and 2020. Regarding the topic or area of study, 25% (3/12) of the studies focused on surgery [126,129,136] and 8% (1/12) each on acute medicine [131], advanced life support [132], basic life support [127], engagement and self-efficacy beliefs [128], geriatric medicine [130], laparoscopy [135], pediatrics [133], primary care [134], and urology [125].

For the outcomes, 58% (7/12) of the studies reported on participants' postintervention knowledge [125,127,129, 130,132-134], 58% (7/12) on skills [126,127,129, 131,132,135,136], 17% (2/12) on attitudes toward the intervention and toward the outcomes [125,132], 17% (2/12) on satisfaction [133,134], 8% (1/12) on competencies [130] and 8% (1/12) on engagement and self-efficacy belief [128](Figure 2). Of the 7 studies assessing participants' skills, 1 (14%) also reported change score from baseline [126]. Overall, 25% (3/12) of the studies assessed retention [126,133,134]. Of these 3 studies, 2 (67%) assessing the knowledge outcome also assessed retention from 4 to 6 weeks after the intervention [133,134] and 1 (33%) assessing the skills outcome also assessed retention at 3 weeks after the intervention [126].

For the assessment methods, most of the included studies used paper-based written assessments [125,130], in-person assessments by supervising clinicians [126,131,135,136], or both assessment methods [127,129,132]. Of the 12 studies, 1 (8%) used digital assessments in the form of a questionnaire in addition to paper-based written assessment [134], 1 (8%) used only digital assessments in the form of a questionnaire [133], and the mode of assessment in 1 (8%) was not mentioned [128].

Of the 12 studies, 7 (58%) reported at least one form of validity evidence (mostly internal consistency and reliability) for the measurement instruments that were mainly used to assess knowledge [125,126,128-130,133,134] (Multimedia Appendices 2 and 3). Of these 7 measurement instruments, 4 (57%) were focused on knowledge, 2 (29%) on skills, 2 (29%) on satisfaction, and 1 (14%) each on cognitive load and self-efficacy beliefs. Of the 7 studies, 3 (43%) referenced a measurement instrument validation study [126,128,133]. The reported measurement properties included internal consistency (for the skills, engagement, and satisfaction measurement instrument), reliability (for the skills and engagement measurement instrument), structural validity (for the skills and satisfaction measurement instrument), and hypothesis (for the skills measurement instrument).

AR Interventions

Of the 126 studies, 11 (8.7%) with 448 medical students used an AR intervention to assess the outcomes [137-147]. The studies included first year to fourth year medical students and were published between 2013 and 2020. The studies covered different topics, including arthroplasty [142], facet joint injection [143], needle insertion [147], general medicine [144], forensic medicine [137], ophthalmology [140], surgery [141,145], laparoscopy [146], and anatomy [138,139].

The reported outcomes included participants' postintervention knowledge [137-139,144], skills [138,140-143,145-147], attitudes toward learning experience or intervention [137,140-142,144], satisfaction [138,146], emotional state , [137,144] and cognitive load [139] (Figure 2). Most studies used paper-based written assessments [137-139,144] or in-person assessments by examiners [143,147] or both approaches [140,142,146]. Of the 11 studies, 1 (9%) used both digital and paper-based written assessments [141] and 1 (9%) used digital assessment in the form of software-based metrics [145]. Of the 8 studies assessing a skills outcome, 2 (25%) also

reported change score from baseline [138,145]. Similarly, of the 6 studies assessing knowledge and satisfaction, 1 (17%) also reported change score from baseline [138]. In terms of retention, only 25% (1/4) of the studies assessing knowledge also reported retention 2 weeks after the intervention [144].

Of the 11 studies, 6 (55%) reported at least one form of validity evidence (mostly internal consistency) for a variety of measurement instruments used [137-140,144,145]. These measurement instruments were used to assess knowledge in 18% (2/11) of the studies, attitudes in 18% (2/11), and emotional state in 18% (2/11), whereas in 9% (1/11) of the studies each, skills, cognitive load, and visuospatial assessment were assessed. None of the studies provided references for validation of the instruments used to measure the outcomes.

MR Interventions

None of the included studies assessed the effectiveness of MR interventions in medical student education.

Discussion

Principal Findings

In this review, we assessed and mapped the choice of outcomes, measurement instruments, and the prevalence of measurement instruments with validity evidence in RCTs on the use of ER technologies in undergraduate medical education. Among the 126 included studies, we found 115 (91.3%) RCTs on different forms of VR, 11 (8.7%) articles on AR simulations, and no RCTs on MR in medical student education. The included studies often reported only a single outcome and immediate postintervention assessments. The types of reported outcomes varied across different types of VR and AR simulations. Participants' skills were the most common outcomes measured in studies on VR simulators, VR patient simulations, and AR. Participants' knowledge was the most common outcome measured in studies on screen-based VR and VR serious games. Other more commonly reported primary outcomes were participants' attitudes toward the intervention or topic and satisfaction with the intervention. More than half of the studies on VR simulators, VR patient simulations, VR serious gaming, and AR as well as only a quarter of the studies on screen-based VR reported at least one form of validity evidence. The most common validity evidence for the measurement instruments used were internal consistency and reliability. Most of the studies used nondigital assessment methods such as paper-based written or in-person assessments by an examiner.

Comparison With Existing Literature

There is a lack of standardization regarding the choice of outcomes and assessments in RCTs focusing on ER for medical student education. The findings are in line with published reviews focusing on the effectiveness of digital education for pre- and postregistration health professionals [1,131,148].

Our review shows a diversity of outcomes and measurement instruments used in trials on ER in medical education. Reporting of a limited set of outcomes, immediate postintervention data, and the use of measurement instruments lacking validity evidence is common in RCTs on different digital health

professions education modalities. However, the choice of appropriate outcomes as well as robust measurement instruments to assess these outcomes is essential when designing trials. It is also important that the chosen outcomes are relevant to key stakeholders who will be able to influence policy and practice. This can be achieved through the development and use of an agreed standardized collection of outcomes and measurement instruments [21].

Strengths and Limitations

In our review, we used a comprehensive search strategy for 7 major bibliographic databases and gray literature sources without language limitations to identify relevant studies. We covered the search period starting from 1990 onward to include all available RCTs on VR-, AR-, and MR-based trainings in medical student education. We performed the screening and data extraction in parallel and independently to ensure reliability of our findings.

There are also some limitations to our study. We performed a descriptive analysis and mapping of outcomes and validity evidence for the measurement instruments used. A more in-depth analysis of the types of validity evidence used was not feasible because of limited information in the included studies. We aimed to complement this by searching for, and including, additional information on validity evidence from validation studies referenced in the included studies. However, information provided in these referenced validation studies was also often limited. We acknowledge that some of the mentioned measurement instruments may have validity evidence not reported in the included RCT papers or for which no validity study was referenced. Furthermore, the reporting of validity evidence in the included RCTs and validation studies may be incomplete and not reflect all validity evidence for a particular measurement instrument. Finally, to determine the validity evidence for the measurement instruments used in the included trials, we used COSMIN, an established taxonomy of measurement properties. Although COSMIN was originally developed for health outcome measurement instruments, it is also applicable to other types of outcomes. However, there are other validity frameworks that were developed primarily for education and may be more appropriate for future analysis of medical education outcomes [9,149].

Future Recommendations

Future studies should aim to include a broader set of outcomes, report change score from baseline, and assess learning retention. They should also aim to use measurement instruments with validity evidence. We list those used in the included trials in Multimedia Appendix 3. Most of the measurement instruments with validity evidence were used to assess participants' skills. There is a need for greater use or adaptation of existing measurement instruments with validity evidence and potentially also development of new ones assessing other relevant outcomes such as attitudes and satisfaction. In addition, digital technology offers diverse and potentially more efficient approaches to assessment and should be more extensively explored and applied in this area. This is particularly relevant given the pervasive and sudden shift to remote teaching because of the COVID-19 pandemic.

Conclusions

Studies on the use of VR and AR in undergraduate medical education often report a limited set of outcomes, mostly knowledge and skills, and usually immediate postintervention assessment data. The use of measurement instruments with validity evidence for outcomes other than skills is limited, as is the use of digital forms of assessment. Future studies should report a broader set of outcomes, change score from baseline, and retention data, as well as use measurement instruments with validity evidence.

Authors' Contributions

LTC conceived the idea for the review. BMK, AT, and TEF screened the studies. BMK, AT, TEF, and SV extracted and analyzed the data from the eligible studies. BMK and LTC wrote the review, and LTC provided methodological guidance. SK, CA, and NC critically revised the paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1

MEDLINE (Ovid) search strategy.
[DOCX File , 21 KB-Multimedia Appendix 1]

Multimedia Appendix 2

Characteristics of the included studies.
[DOCX File , 47 KB-Multimedia Appendix 2]

Multimedia Appendix 3

Types and number of reported outcomes and measurement instruments with validity in the included studies.
[DOCX File , 27 KB-Multimedia Appendix 3]

References

1. Kyaw BM, Saxena N, Posadzki P, Vseteckova J, Nikolaou CK, George PP, et al. Virtual reality for health professions education: systematic review and meta-analysis by the digital health education collaboration. J Med Internet Res 2019 Jan 22;21(1):e12959 [FREE Full text] [doi: 10.2196/12959] [Medline: 30668519]
2. George PP, Papachristou N, Belisario JM, Wang W, Wark PA, Cotic Z, et al. Online eLearning for undergraduates in health professions: a systematic review of the impact on knowledge, skills, attitudes and satisfaction. J Glob Health 2014 Jun;4(1):010406 [FREE Full text] [doi: 10.7189/jogh.04.010406] [Medline: 24976965]
3. McGaghie WC, Issenberg SB, Petrusa ER, Scalese RJ. A critical review of simulation-based medical education research: 2003-2009. Med Educ 2010 Jan;44(1):50-63. [doi: 10.1111/j.1365-2923.2009.03547.x] [Medline: 20078756]
4. Harden RM. Outcome-Based Education: the future is today. Med Teach 2007 Sep;29(7):625-629. [doi: 10.1080/01421590701729930] [Medline: 18236247]
5. Epstein RM. Assessment in medical education. N Engl J Med 2007 Jan 25;356(4):387-396. [doi: 10.1056/NEJMra054784] [Medline: 17251535]
6. American Educational Research Association, American Psychological Association, National Council on Measurement in Education. Standards for Educational and Psychological Testing. Washington, DC, US: American Educational Research Association; 2014.
7. Ratanawongsa N, Thomas PA, Marinopoulos SS, Dorman T, Wilson LM, Ashar BH, et al. The reported validity and reliability of methods for evaluating continuing medical education: a systematic review. Acad Med 2008 Mar;83(3):274-283. [doi: 10.1097/ACM.0b013e3181637925] [Medline: 18316877]
8. Cook DA, Brydges R, Zendejas B, Hamstra SJ, Hatala R. Technology-enhanced simulation to assess health professionals: a systematic review of validity evidence, research methods, and reporting quality. Acad Med 2013 Jun;88(6):872-883. [doi: 10.1097/ACM.0b013e31828ffdcf] [Medline: 23619073]
9. Kane MT. Validating the Interpretations and Uses of Test Scores. Journal of Educational Measurement 2013 Mar 14;50(1):1-73. [doi: 10.1111/jedm.12000]

XSL•FO
RenderX

10.   Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ 2009 Jul 21;339:b2700 [FREE Full text] [doi: 10.1136/bmj.b2700] [Medline: 19622552]

11.   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] [Medline: 19621072]

12.   eLearning for undergraduate health professional education: A systematic review informing a radical transformation of health workforce development. World Health Organization. 2015. URL: https://www.who.int/hrh/documents/ 14126-eLearningReport.pdf [accessed 2022-04-03]

13.   International Standard Classification of Education: Fields of education and training 2013 (ISCED-F 2013). Unesco Institute for Statistics. 2014. URL: http://uis.unesco.org/sites/default/files/documents/ international-standard-classification-of-education-fields-of-education-and-training-2013-detailed-field-descriptions-2015-en. pdf [accessed 2022-04-06]

14.   Strangman N, Hall T, Meyer A. Virtual Reality/Computer Simulations and the Implications for UDL Implementation. National Center on Accessing the General Curriculum. 2003. URL: https://ccie.ucf.edu/wp-content/uploads/sites/8/2018/ 02/VirtualRealityUDL.pdf [accessed 2022-04-06]

15.   Mantovani F, Castelnuovo G, Gaggioli A, Riva G. Virtual reality training for health-care professionals. Cyberpsychol Behav 2003 Aug;6(4):389-395. [doi: 10.1089/109493103322278772] [Medline: 14511451]

16.   Ellaway R, Candler C, Greene P, Smothers V. An Architectural Model for MedBiquitous Virtual Patients. MedBiquitous. 2006. URL: http://groups.medbiq.org/medbiq/display/VPWG/MedBiquitous+Virtual+Patient+Architecture [accessed 2022-04-06]

17.   Zhu E, Hadadgar A, Masiello I, Zary N. Augmented reality in healthcare education: an integrative review. PeerJ 2014;2:e469 [FREE Full text] [doi: 10.7717/peerj.469] [Medline: 25071992]

18.   Tepper OM, Rudy HL, Lefkowitz A, Weimer KA, Marks SM, Stern CS, et al. Mixed Reality with HoloLens: Where Virtual Reality Meets Augmented Reality in the Operating Room. Plast Reconstr Surg 2017 Nov;140(5):1066-1070. [doi: 10.1097/PRS.0000000000003802] [Medline: 29068946]

19.   Flavián C, Ibáñez-Sánchez S, Orús C. The impact of virtual, augmented and mixed reality technologies on the customer experience. Journal of Business Research 2019 Jul;100(4):547-560. [doi: 10.1016/j.jbusres.2018.10.050] [Medline: 23725941]

20.   Law GC, Apfelbacher C, Posadzki PP, Kemp S, Tudor Car L. Choice of outcomes and measurement instruments in randomised trials on eLearning in medical education: a systematic mapping review protocol. Syst Rev 2018 May 17;7(1):75 [FREE Full text] [doi: 10.1186/s13643-018-0739-0] [Medline: 29776434]

21.   Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol 2010 Jul;63(7):737-745. [doi: 10.1016/j.jclinepi.2010.02.006] [Medline: 20494804]

22.   Aggarwal R, Grantcharov T, Moorthy K, Hance J, Darzi A. A competency-based virtual reality training curriculum for the acquisition of laparoscopic psychomotor skill. Am J Surg 2006 Jan;191(1):128-133. [doi: 10.1016/j.amjsurg.2005.10.014] [Medline: 16399123]

23.   Ahad S, Boehler M, Schwind CJ, Hassan I. The effect of model fidelity on colonoscopic skills acquisition. A randomized controlled study. J Surg Educ 2013 May;70(4):522-527. [doi: 10.1016/j.jsurg.2013.02.010] [Medline: 23725941]

24.   Ahlberg G, Heikkinen T, Iselius L, Leijonmarck C, Rutqvist J, Arvidsson D. Does training in a virtual reality simulator improve surgical performance? Surg Endosc 2002 Jan;16(1):126-129. [doi: 10.1007/s00464-001-9025-6] [Medline: 11961622]

25.   Andersen SAW, Konge L, Cayé-Thomasen P, Sørensen MS. Learning Curves of Virtual Mastoidectomy in Distributed and Massed Practice. JAMA Otolaryngol Head Neck Surg 2015 Oct;141(10):913-918. [doi: 10.1001/jamaoto.2015.1563] [Medline: 26334610]

26.   Bergqvist J, Person A, Vestergaard A, Grauslund J. Establishment of a validated training programme on the Eyesi cataract simulator. A prospective randomized study. Acta Ophthalmol 2014 Nov 13;92(7):629-634 [FREE Full text] [doi: 10.1111/aos.12383] [Medline: 24612448]

27.   Bjerrum F, Sorensen JL, Konge L, Rosthøj S, Lindschou J, Ottesen B, et al. Randomized trial to examine procedure-to-procedure transfer in laparoscopic simulator training. Br J Surg 2016 Jan;103(1):44-50. [doi: 10.1002/bjs.9966] [Medline: 26511775]

28.   Bowyer CMW, Liu AV, Bonar JP. Validation of SimPL -- a simulator for diagnostic peritoneal lavage training. Stud Health Technol Inform 2005;111:64-67. [Medline: 15718700]

29.   Bowyer MW, Pimentel EA, Fellows JB, Scofield RL, Ackerman VL, Horne PE, et al. Teaching intravenous cannulation to medical students: comparative analysis of two simulators and two traditional educational approaches. Stud Health Technol Inform 2005;111:57-63. [Medline: 15718699]

30.   Brunckhorst O, Shahid S, Aydin A, McIlhenny C, Khan S, Raza S, et al. MP22-09 DEVELOPMENT AND VALIDATION OF AN INTEGRATED SKILLS CURRICULUM WITHIN URETEROSCOPY– A RANDOMISED CONTROLLED TRIAL. Journal of Urology 2015 Apr;193(4S):1018-1025. [doi: 10.1016/j.juro.2015.02.1021] [Medline: 19616797]

31.   Van Bruwaene S, Schijven MP, Napolitano D, De Win G, Miserez M. Porcine cadaver organ or virtual-reality simulation training for laparoscopic cholecystectomy: a randomized, controlled trial. J Surg Educ 2015 Oct;72(3):483-490. [doi: 10.1016/j.jsurg.2014.11.015] [Medline: 25555673]

32.   Brydges R, Carnahan H, Rose D, Rose L, Dubrowski A. Coordinating Progressive Levels of Simulation Fidelity to Maximize Educational Benefit. Academic Medicine 2010 Dec 13;85(5):806-812. [doi: 10.1097/acm.0b013e3181d7aabd] [Medline: 23493565]

33.   Bube S, Dagnaes-Hansen J, Mahmood O, Rohrsted M, Bjerrum F, Salling L, et al. Simulation-based training for flexible cystoscopy - A randomized trial comparing two approaches. Heliyon 2020 Jan;6(1):e03086 [FREE Full text] [doi: 10.1016/j.heliyon.2019.e03086] [Medline: 31922043]

34.   Chou DS, Abdelshehid C, Clayman RV, McDougall EM. Comparison of results of virtual-reality simulator and training model for basic ureteroscopy training. J Endourol 2006 Apr;20(4):266-271 [FREE Full text] [doi: 10.1089/end.2006.20.266] [Medline: 16646655]

35.   da Cruz JAS, Dos Reis ST, Cunha Frati RM, Duarte RJ, Nguyen H, Srougi M, et al. Does Warm-Up Training in a Virtual Reality Simulator Improve Surgical Performance? A Prospective Randomized Analysis. J Surg Educ 2016 Oct;73(6):974-978. [doi: 10.1016/j.jsurg.2016.04.020] [Medline: 27233673]

36.   da Cruz JAS, Sandy NS, Passerotti CC, Nguyen H, Antunes AA, Dos Reis ST, et al. Does training laparoscopic skills in a virtual reality simulator improve surgical performance? J Endourol 2010 Nov;24(11):1845-1849. [doi: 10.1089/end.2010.0328] [Medline: 20958203]

37.   De La Garza JR, Schmidt MW, Kowalewski K, Benner K, Müller PC, Kenngott HG, et al. Does rating with a checklist improve the effect of E-learning for cognitive and practical skills in bariatric surgery? A rater-blinded, randomized-controlled trial. Surg Endosc 2019 May;33(5):1532-1543. [doi: 10.1007/s00464-018-6441-4] [Medline: 30209607]

38.   Eldred-Evans D, Grange P, Cheang A, Yamamoto H, Ayis S, Mulla M, et al. Using the mind as a simulator: a randomized controlled trial of mental training. J Surg Educ 2013 Apr 23;70(4):544-551. [doi: 10.1016/j.jsurg.2013.04.003] [Medline: 23725944]

39.   Frithioff A, Frendø M, Mikkelsen PT, Sørensen MS, Andersen SAW. Ultra-high-fidelity virtual reality mastoidectomy simulation training: a randomized, controlled trial. Eur Arch Otorhinolaryngol 2020 May;277(5):1335-1341. [doi: 10.1007/s00405-020-05858-3] [Medline: 32067096]

40.   Fu Y, Cavuoto L, Qi D, Panneerselvam K, Arikatla VS, Enquobahrie A, et al. Characterizing the learning curve of a virtual intracorporeal suturing simulator VBLaST-SS©. Surg Endosc 2020 Jul;34(7):3135-3144 [FREE Full text] [doi: 10.1007/s00464-019-07081-6] [Medline: 31482354]

41.   Ganai S, Donroe JA, St Louis MR, Lewis GM, Seymour NE. Virtual-reality training improves angled telescope skills in novice laparoscopists. Am J Surg 2007 Feb;193(2):260-265. [doi: 10.1016/j.amjsurg.2005.11.019] [Medline: 17236859]

42.   Gasco J, Patel A, Ortega-Barnett J, Branch D, Desai S, Kuo Y, et al. Virtual reality spine surgery simulation: an empirical study of its usefulness. Neurological Research 2014 May 20;36(11):968-973. [doi: 10.1179/1743132814y.0000000388] [Medline: 26841278]

43.   Guetterman TC, Sakakibara R, Baireddy S, Kron FW, Scerbo MW, Cleary JF, et al. Medical Students' Experiences and Outcomes Using a Virtual Human Simulation to Improve Communication Skills: Mixed Methods Study. J Med Internet Res 2019 Nov 27;21(11):e15459-e15459 [FREE Full text] [doi: 10.2196/15459] [Medline: 31774400]

44.   Hariri S, Rawn C, Srivastava S, Youngblood P, Ladd A. Evaluation of a surgical simulator for learning clinical anatomy. Med Educ 2004 Aug;38(8):896-902. [doi: 10.1111/j.1365-2929.2004.01897.x] [Medline: 15271051]

45.   Hiemstra E, Terveer EM, Chmarra MK, Dankelman J, Jansen FW. Virtual reality in laparoscopic skills training: is haptic feedback replaceable? Minim Invasive Ther Allied Technol 2011 May;20(3):179-184. [doi: 10.3109/13645706.2010.532502] [Medline: 21438717]

46.   Hyltander A, Liljegren E, Rhodin PH, Lönroth H. The transfer of basic skills learned in a laparoscopic simulator to the operating room. Surg Endosc 2002 Sep;16(9):1324-1328 [FREE Full text] [doi: 10.1007/s00464-001-9184-5] [Medline: 11988802]

47.   Johnston TJ, Tang B, Alijani A, Tait I, Steele RJ, Ker J, Surgical Simulation Group at the University of Dundee. Laparoscopic surgical skills are significantly improved by the use of a portable laparoscopic simulator: results of a randomized controlled trial. World J Surg 2013 May;37(5):957-964. [doi: 10.1007/s00268-013-1945-5] [Medline: 23404485]

48.   Kanumuri P, Ganai S, Wohaibi EM, Bush RW, Grow DR, Seymour NE. Virtual reality and computer-enhanced training devices equally improve laparoscopic surgical skill in novices. JSLS 2008;12(3):219-226 [FREE Full text] [Medline: 18765042]

49.   Karabanov AN, Irmen F, Madsen KH, Haagensen BN, Schulze S, Bisgaard T, et al. Getting to grips with endoscopy - Learning endoscopic surgical skills induces bi-hemispheric plasticity of the grasping network. Neuroimage 2019 Apr 01;189:32-44. [doi: 10.1016/j.neuroimage.2018.12.030] [Medline: 30583066]

50.   Kothari SN, Kaplan BJ, DeMaria EJ, Broderick TJ, Merrell RC. Training in laparoscopic suturing skills using a new computer-based virtual reality simulator (MIST-VR) provides results comparable to those with an established pelvic trainer system. J Laparoendosc Adv Surg Tech A 2002 Jun 11;12(3):167-173 [FREE Full text] [doi: 10.1089/10926420260188056] [Medline: 12184901]

XSL·FO

RenderX

51. Kowalewski K, Minassian A, Hendrie JD, Benner L, Preukschas AA, Kenngott HG, et al. One or two trainees per workplace for laparoscopic surgery training courses: results from a randomized controlled trial. Surg Endosc 2019 May;33(5):1523-1531. [doi: 10.1007/s00464-018-6440-5] [Medline: 30194644]

52. Krogh C, Konge L, Bjurström J, Ringsted C. Training on a new, portable, simple simulator transfers to performance of complex bronchoscopy procedures. The Clinical Respiratory Journal 2012 Aug 20;7(3):237-244. [doi: 10.1111/j.1752-699x.2012.00311.x] [Medline: 15718699]

53. Lee CR, Rho SY, Han SH, Moon Y, Hwang SY, Kim YJ, et al. Comparison of Training Efficacy Between Custom-Made Skills Simulator (CMSS) and da Vinci Skills Simulators: A Randomized Control Study. World J Surg 2019 Nov;43(11):2699-2709. [doi: 10.1007/s00268-019-05108-6] [Medline: 31399794]

54. Lesch H, Johnson E, Peters J, Cendán JC. VR Simulation Leads to Enhanced Procedural Confidence for Surgical Trainees. J Surg Educ 2020 Feb;77(1):213-218 [FREE Full text] [doi: 10.1016/j.jsurg.2019.08.008] [Medline: 31466895]

55. Lindquist NR, Leach M, Simpson MC, Antisdel JL. Evaluating Simulator-Based Teaching Methods for Endoscopic Sinus Surgery. Ear Nose Throat J 2019 Sep 20;98(8):490-495 [FREE Full text] [doi: 10.1177/0145561319844742] [Medline: 31018690]

56. Loukas C, Lahanas V, Kanakis M, Georgiou E. The Effect of Mixed-Task Basic Training in the Acquisition of Advanced Laparoscopic Skills. Surg Innov 2015 Aug;22(4):418-425. [doi: 10.1177/1553350614556365] [Medline: 25377214]

57. Loukas C, Nikiteas N, Schizas D, Lahanas V, Georgiou E. A head-to-head comparison between virtual reality and physical reality simulation training for basic skills acquisition. Surg Endosc 2012 Sep 27;26(9):2550-2558. [doi: 10.1007/s00464-012-2230-7] [Medline: 22476832]

58. Lucas S, Tuncel A, Bensalah K, Zeltser I, Jenkins A, Pearle M, et al. Virtual reality training improves simulated laparoscopic surgery performance in laparoscopy naïve medical students. J Endourol 2008 May;22(5):1047-1051. [doi: 10.1089/end.2007.0366] [Medline: 18643722]

59. Madan AK, Frantzides CT. Prospective randomized controlled trial of laparoscopic trainers for basic laparoscopic skills acquisition. Surg Endosc 2007 Feb;21(2):209-213. [doi: 10.1007/s00464-006-0149-6] [Medline: 17122975]

60. McDougall EM, Kolla SB, Santos RT, Gan JM, Box GN, Louie MK, et al. Preliminary study of virtual reality and model simulation for learning laparoscopic suturing skills. J Urol 2009 Sep 05;182(3):1018-1025. [doi: 10.1016/j.juro.2009.05.016] [Medline: 19616797]

61. Mulla M, Sharma D, Moghul M, Kailani O, Dockery J, Ayis S, et al. Learning basic laparoscopic skills: a randomized controlled study comparing box trainer, virtual reality simulator, and mental training. J Surg Educ 2012 Sep 5;69(2):190-195. [doi: 10.1016/j.jsurg.2011.07.011] [Medline: 22365864]

62. Munz Y, Kumar BD, Moorthy K, Bann S, Darzi A. Laparoscopic virtual reality and box trainers: is one superior to the other? Surg Endosc 2004 Mar 1;18(3):485-494. [doi: 10.1007/s00464-003-9043-7] [Medline: 14752633]

63. Muresan C, Lee TH, Seagull J, Park AE. Transfer of training in the development of intracorporeal suturing skill in medical student novices: a prospective randomized trial. Am J Surg 2010 Oct;200(4):537-541. [doi: 10.1016/j.amjsurg.2009.12.018] [Medline: 20451174]

64. Nehme J, Sodergren MH, Sugden C, Aggarwal R, Gillen S, Feussner H, et al. A randomized controlled trial evaluating endoscopic and laparoscopic training in skills transfer for novices performing a simulated NOTES task. Surg Innov 2013 Dec 1;20(6):631-638. [doi: 10.1177/1553350613480854] [Medline: 23493565]

65. Nickel F, Brzoska JA, Gondan M, Rangnick HM, Chu J, Kenngott HG, et al. Virtual reality training versus blended learning of laparoscopic cholecystectomy: a randomized controlled trial with laparoscopic novices. Medicine (Baltimore) 2015 May;94(20):e764-e769 [FREE Full text] [doi: 10.1097/MD.0000000000000764] [Medline: 25997044]

66. Oussi N, Enochsson L, Henningsohn L, Castegren M, Georgiou E, Kjellin A. Trainee Performance After Laparoscopic Simulator Training Using a Blackbox versus LapMentor. J Surg Res 2020 Jun;250(3):1-11. [doi: 10.1016/j.jss.2019.12.039] [Medline: 32014696]

67. Patel A, Koshy N, Ortega-Barnett J, Chan HC, Kuo Y, Luciano C, et al. Neurosurgical tactile discrimination training with haptic-based virtual reality simulation. Neurol Res 2014 Dec;36(12):1035-1039. [doi: 10.1179/1743132814Y.0000000405] [Medline: 24984771]

68. Plana NM, Rifkin WJ, Kantar RS, David JA, Maliha SG, Farber SJ, et al. A Prospective, Randomized, Blinded Trial Comparing Digital Simulation to Textbook for Cleft Surgery Education. Plast Reconstr Surg 2019 Jan;143(1):202-209 [FREE Full text] [doi: 10.1097/PRS.0000000000005093] [Medline: 30325894]

69. Schlosser K, Alkhawaga M, Maschuw K, Zielke A, Mauner E, Hassan I. Training of laparoscopic skills with virtual reality simulator: a critical reappraisal of the learning curve. Eur Surg 2007 Jun 19;39(3):180-184. [doi: 10.1007/s10353-006-0292-2] [Medline: 17710487]

70. Selvander M, Åsman P. Virtual reality cataract surgery training: learning curves and concurrent validity. Acta Ophthalmol 2012 Aug;90(5):412-417 [FREE Full text] [doi: 10.1111/j.1755-3768.2010.02028.x] [Medline: 21054818]

71. Solyar A, Cuellar H, Sadoughi B, Olson TR, Fried MP. Endoscopic Sinus Surgery Simulator as a teaching tool for anatomy education. Am J Surg 2008 Jul;196(1):120-124. [doi: 10.1016/j.amjsurg.2007.06.026] [Medline: 18374890]

[XSL•FO]
[RenderX]

72. Sotto JAR, Ayuste EC, Bowyer MW, Almonte JR, Dofitas RB, Lapitan MCM, et al. Exporting simulation technology to the Philippines: a comparative study of traditional versus simulation methods for teaching intravenous cannulation. Stud Health Technol Inform 2009;142:346-351. [Medline: 19377182]

73. Sugand K, Akhtar K, Khatri C, Cobb J, Gupte C. Training effect of a virtual reality haptics-enabled dynamic hip screw simulator. Acta Orthop 2015 Oct 10;86(6):695-701 [FREE Full text] [doi: 10.3109/17453674.2015.1071111] [Medline: 26168925]

74. Suh IH, Mukherjee M, Park S, Oleynikov D, Siu K. Enhancing robot-assisted fundamental surgical proficiency using portable virtual simulator. Surgical Endoscopy and Other Interventional Techniques 2010;24(1):S686-S687.

75. Sun W, Konjg J, Li XY, Zhang JS. Application of operational simulation training system in the training of ophthalmic students. Int Eye Sci 2014;14(9):1567-1569.

76. Tanoue K, Ieiri S, Konishi K, Yasunaga T, Okazaki K, Yamaguchi S, et al. Effectiveness of endoscopic surgery training for medical students using a virtual reality simulator versus a box trainer: a randomized controlled trial. Surg Endosc 2008 Apr;22(4):985-990. [doi: 10.1007/s00464-007-9554-8] [Medline: 17710487]

77. Tanoue K, Yasunaga T, Konishi K, Okazaki K, Ieiri S, Kawabe Y, et al. Effectiveness of training for endoscopic surgery using a simulator with virtual reality: Randomized study. International Congress Series 2005 May 27;1281(11):515-520 [FREE Full text] [doi: 10.1016/j.ics.2005.03.367] [Medline: 31774400]

78. Torkington J, Smith SG, Rees BI, Darzi A. Skill transfer from virtual reality to a real laparoscopic task. Surg Endosc 2001 Oct 01;15(10):1076-1079. [doi: 10.1007/s004640000233] [Medline: 11727073]

79. Unger B, Tordon B, Pisa J, Hochman JB. Importance of Stereoscopy in Haptic Training of Novice Temporal Bone Surgery. Stud Health Technol Inform 2016;220:439-445. [Medline: 27046619]

80. Wilhelm DM, Ogan K, Roehrborn CG, Cadeddu JA, Pearle MS. Assessment of basic endoscopic performance using a virtual reality simulator. Journal of the American College of Surgeons 2002 Nov;195(5):675-681. [doi: 10.1016/s1072-7515(02)01346-7] [Medline: 31399794]

81. Youngblood PL, Srivastava S, Curet M, Heinrichs WL, Dev P, Wren SM. Comparison of training on two laparoscopic simulators and assessment of skills transfer to surgical performance. J Am Coll Surg 2005 Apr;200(4):546-551. [doi: 10.1016/j.jamcollsurg.2004.11.011] [Medline: 15804468]

82. Zeltser IS, Bensalah K, Tuncel A, Lucas SM, Jenkins A, Pearle MS. Training on the virtual reality laparoscopic simulator improves performance of an unfamiliar live surgical laparoscopic procedure: a randomized, controlled trial. J Endourol 2007;21(Suppl 1):A137.

83. Zhang L, Sankaranarayanan G, Arikatla VS, Ahn W, Grosdemouge C, Rideout JM, et al. Characterizing the learning curve of the VBLaST-PT(©) (Virtual Basic Laparoscopic Skill Trainer). Surg Endosc 2013 Oct;27(10):3603-3615 [FREE Full text] [doi: 10.1007/s00464-013-2932-5] [Medline: 23572217]

84. Zhao YC, Kennedy G, Yukawa K, Pyman B, O'Leary S. Can virtual reality simulator be used as a training aid to improve cadaver temporal bone dissection? Results of a randomized blinded control trial. Laryngoscope 2011 Apr;121(4):831-837. [doi: 10.1002/lary.21287] [Medline: 21433021]

85. Neumann E, Mayer J, Russo GI, Amend B, Rausch S, Deininger S, et al. Transurethral Resection of Bladder Tumors: Next-generation Virtual Reality Training for Surgeons. Eur Urol Focus 2019 Sep;5(5):906-911. [doi: 10.1016/j.euf.2018.04.011] [Medline: 29802051]

86. Alverson D, Saiki SJ, Kalishman S, Lindberg M, Mennin S, Mines J, et al. Medical students learn over distance using virtual reality simulation. Simul Healthc 2008;3(1):10-15. [doi: 10.1097/SIH.0b013e31815f0d51] [Medline: 19088637]

87. Battulga B, Konishi T, Tamura Y, Moriguchi H. The effectiveness of an interactive 3-dimensional computer graphics model for medical education. Interact J Med Res 2012 Jul 09;1(2):e2 [FREE Full text] [doi: 10.2196/ijmr.2172] [Medline: 23611759]

88. Blumstein G, Zukotynski B, Cevallos N, Ishmael C, Zoller S, Burke Z, et al. Randomized Trial of a Virtual Reality Tool to Teach Surgical Technique for Tibial Shaft Fracture Intramedullary Nailing. J Surg Educ 2020;77(4):969-977 [FREE Full text] [doi: 10.1016/j.jsurg.2020.01.002] [Medline: 32035854]

89. Courteille O, Ho J, Fahlstedt M, Fors U, Felländer-Tsai L, Hedman L, et al. Face validity of VIS-Ed: a visualization program for teaching medical students and residents the biomechanics of cervical spine trauma. Stud Health Technol Inform 2013;184:96-102. [Medline: 23400137]

90. Deladisma AM, Gupta M, Kotranza A, Bittner JG, Imam T, Swinson D, et al. A pilot study to integrate an immersive virtual patient with a breast complaint and breast examination simulator into a surgery clerkship. Am J Surg 2009 Jan;197(1):102-106. [doi: 10.1016/j.amjsurg.2008.08.012] [Medline: 19101251]

91. Drapkin ZA, Lindgren KA, Lopez MJ, Stabio ME. Development and assessment of a new 3D neuroanatomy teaching tool for MRI training. Anat Sci Educ 2015;8(6):502-509. [doi: 10.1002/ase.1509] [Medline: 25573020]

92. Flores R, DeMoss P, Klene C, Havlik RJ, Tholpady S. Digital Animation versus Textbook in Teaching Plastic Surgery Techniques to Novice Learners. Plastic and Reconstructive Surgery 2013;132(1):101e-109e. [doi: 10.1097/prs.0b013e3182910aa9]

93. Glittenberg CGO. Methods and advantages of the use of computer-assisted 3-D-design and multimedia teaching programs in the demonstration and education of the neuroophthalmological basics of the oculomotor system. Spektrum der Augenheilkunde 2003;17(6):242-246.

94. 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]

95. Hampton BS, Sung VW. Improving medical student knowledge of female pelvic floor dysfunction and anatomy: a randomized trial. Am J Obstet Gynecol 2010 Jun;202(6):601.e1-601.e8. [doi: 10.1016/j.ajog.2009.08.038] [Medline: 20430356]

96. Hisley KC, Anderson LD, Smith SE, Kavic SM, Tracy JK. Coupled physical and digital cadaver dissection followed by a visual test protocol provides insights into the nature of anatomical knowledge and its evaluation. Anat Sci Educ 2008 Jan;1(1):27-40. [doi: 10.1002/ase.4] [Medline: 19177376]

97. Hopkins R, Regehr G, Wilson TD. Exploring the Changing Learning Environment of the Gross Anatomy Lab. Academic Medicine 2011 Aug;86(7):883-888. [doi: 10.1097/acm.0b013e31821de30f] [Medline: 27307270]

98. Hu A, Shewokis PA, Ting K, Fung K. Motivation in computer-assisted instruction. Laryngoscope 2016 Aug;126 Suppl 6:S5-S13. [doi: 10.1002/lary.26040] [Medline: 27307270]

99. Kalet A, Song H, Sarpel U, Schwartz R, Brenner J, Ark T, et al. Just enough, but not too much interactivity leads to better clinical skills performance after a computer assisted learning module. Medical Teacher 2012 Aug 23;34(10):833-839. [doi: 10.3109/0142159x.2012.706727] [Medline: 21412990]

100. Keedy AW, Durack JC, Sandhu P, Chen EM, O'Sullivan PS, Breiman RS. Comparison of traditional methods with 3D computer models in the instruction of hepatobiliary anatomy. Anat Sci Educ 2011 Sep;4(2):84-91. [doi: 10.1002/ase.212] [Medline: 21412990]

101. Khatib M, Hald N, Brenton H, Barakat MF, Sarker SK, Standfield N, et al. Validation of open inguinal hernia repair simulation model: a randomized controlled educational trial. Am J Surg 2014 Aug;208(2):295-301. [doi: 10.1016/j.amjsurg.2013.12.007] [Medline: 24581993]

102. Kockro RA, Amaxopoulou C, Killeen T, Wagner W, Reisch R, Schwandt E, et al. Stereoscopic neuroanatomy lectures using a three-dimensional virtual reality environment. Ann Anat 2015 Sep;201:91-98. [doi: 10.1016/j.aanat.2015.05.006] [Medline: 26245861]

103. Lorenzo-Alvarez R, Rudolphi-Solero T, Ruiz-Gomez MJ, Sendra-Portero F. Medical Student Education for Abdominal Radiographs in a 3D Virtual Classroom Versus Traditional Classroom: A Randomized Controlled Trial. AJR Am J Roentgenol 2019 Sep;213(3):644-650. [doi: 10.2214/AJR.19.21131] [Medline: 31287725]

104. Maresky HS, Oikonomou A, Ali I, Ditkofsky N, Pakkal M, Ballyk B. Virtual reality and cardiac anatomy: Exploring immersive three-dimensional cardiac imaging, a pilot study in undergraduate medical anatomy education. Clin Anat 2019 Mar;32(2):238-243. [doi: 10.1002/ca.23292] [Medline: 30295333]

105. Motsumi MJ, Bedada AG, Ayane G. The role of Moodle-based surgical skills illustrations using 3D animation in undergraduate training. AJHPE 2019 Dec 12;11(4):149. [doi: 10.7196/ajhpe.2019.v11i4.1189]

106. Nicholson DT, Chalk C, Funnell WRJ, Daniel SJ. Can virtual reality improve anatomy education? A randomised controlled study of a computer-generated three-dimensional anatomical ear model. Med Educ 2006 Nov;40(11):1081-1087. [doi: 10.1111/j.1365-2929.2006.02611.x] [Medline: 17054617]

107. Patel V, Aggarwal R, Osinibi E, Taylor D, Arora S, Darzi A. Operating room introduction for the novice. Am J Surg 2012 Feb;203(2):266-275. [doi: 10.1016/j.amjsurg.2011.03.003] [Medline: 21703594]

108. Persky S, Eccleston CP. Medical student bias and care recommendations for an obese versus non-obese virtual patient. Int J Obes (Lond) 2011 May;35(5):728-735 [FREE Full text] [doi: 10.1038/ijo.2010.173] [Medline: 20820169]

109. Prinz A, Bolz M, Findl O. Advantage of three dimensional animated teaching over traditional surgical videos for teaching ophthalmic surgery: a randomised study. Br J Ophthalmol 2005 Nov;89(11):1495-1499 [FREE Full text] [doi: 10.1136/bjo.2005.075077] [Medline: 16234460]

110. Schutte B, de Goeij T, de Grave W, Koehorst AM. The Effects of Visual Genetics on the Learning of Students in a Problem Based Curriculum. In: Advances in Medical Education. Dordrecht, The Netherlands: Springer; 1997:336-338.

111. Scoville SA, Buskirk TD. Traditional and virtual microscopy compared experimentally in a classroom setting. Clin Anat 2007 Jul;20(5):565-570. [doi: 10.1002/ca.20440] [Medline: 17109440]

112. Succar T, Grigg J. A new vision for teaching ophthalmology in the medical curriculum: The virtual ophthalmology clinic. ascilite. 2010. URL: https://ascilite.org/conferences/sydney10/procs/Succar-poster.pdf [accessed 2022-04-06]

113. Succar T, Zebington G, Billson F, Byth K, Barrie S, McCluskey P, et al. The impact of the Virtual Ophthalmology Clinic on medical students' learning: a randomised controlled trial. Eye (Lond) 2013 Oct;27(10):1151-1157 [FREE Full text] [doi: 10.1038/eye.2013.143] [Medline: 23867718]

114. Yi X, Ding C, Xu H, Huang T, Kang D, Wang D. Three-Dimensional Printed Models in Anatomy Education of the Ventricular System: A Randomized Controlled Study. World Neurosurg 2019 May;125:e891-e901. [doi: 10.1016/j.wneu.2019.01.204] [Medline: 30743037]

115. Glittenberg C, Binder S. Using 3D computer simulations to enhance ophthalmic training. Ophthalmic Physiol Opt 2006 Jan;26(1):40-49. [doi: 10.1111/j.1475-1313.2005.00358.x] [Medline: 16390481]

116. Botezatu M, Hult H, Tessma MK, Fors UGH. Virtual patient simulation for learning and assessment: Superior results in comparison with regular course exams. Med Teach 2010;32(10):845-850. [doi: 10.3109/01421591003695287] [Medline: 20854161]

117. Deladisma AM, Cohen M, Stevens A, Wagner P, Lok B, Bernard T, Association for Surgical Education. Do medical students respond empathetically to a virtual patient? Am J Surg 2007 Jun;193(6):756-760. [doi: 10.1016/j.amjsurg.2007.01.021] [Medline: 17512291]

118. Dickerson R, Johnsen K, Raij A, Lok B, Stevens A, Bernard T, et al. Virtual patients: assessment of synthesized versus recorded speech. Stud Health Technol Inform 2006;119:114-119. [Medline: 16404028]

119. Foster A, Chaudhary N, Kim T, Waller JL, Wong J, Borish M, et al. Using Virtual Patients to Teach Empathy: A Randomized Controlled Study to Enhance Medical Students' Empathic Communication. Simul Healthc 2016 Jun;11(3):181-189. [doi: 10.1097/SIH.0000000000000142] [Medline: 26841278]

120. Foster A, Chaudhary N, Murphy J, Lok B, Waller J, Buckley PF. The Use of Simulation to Teach Suicide Risk Assessment to Health Profession Trainees-Rationale, Methodology, and a Proof of Concept Demonstration with a Virtual Patient. Acad Psychiatry 2015 Dec;39(6):620-629. [doi: 10.1007/s40596-014-0185-9] [Medline: 25026950]

121. Lehmann R, Lutz T, Helling-Bakki A, Kummer S, Huwendiek S, Bosse HM. Animation and interactivity facilitate acquisition of pediatric life support skills: a randomized controlled trial using virtual patients versus video instruction. BMC Med Educ 2019 Jan 05;19(1):7 [FREE Full text] [doi: 10.1186/s12909-018-1442-5] [Medline: 30611273]

122. Lehmann R, Thiessen C, Frick B, Bosse HM, Nikendei C, Hoffmann GF, et al. Improving Pediatric Basic Life Support Performance Through Blended Learning With Web-Based Virtual Patients: Randomized Controlled Trial. J Med Internet Res 2015 Jul 02;17(7):e162 [FREE Full text] [doi: 10.2196/jmir.4141] [Medline: 26139388]

123. McCoy L. Virtual patient simulations for medical education: Increasing clinical reasoning skills through deliberate practice. ERIC. 2015. URL: https://eric.ed.gov/?id=ED569014 [accessed 2022-04-06]

124. O'Rourke SR, Branford KR, Brooks TL, Ives LT, Nagendran A, Compton SN. The Emotional and Behavioral Impact of Delivering Bad News to Virtual versus Real Standardized Patients: A Pilot Study. Teach Learn Med 2020;32(2):139-149. [doi: 10.1080/10401334.2019.1652180] [Medline: 31437006]

125. Boeker M, Andel P, Vach W, Frankenschmidt A. Game-based e-learning is more effective than a conventional instructional method: a randomized controlled trial with third-year medical students. PLoS One 2013;8(12):e82328 [FREE Full text] [doi: 10.1371/journal.pone.0082328] [Medline: 24349257]

126. de Araujo TB, Silveira FR, Souza DLS, Strey YTM, Flores CD, Webster RS. Impact of video game genre on surgical skills development: a feasibility study. J Surg Res 2016 Mar;201(1):235-243. [doi: 10.1016/j.jss.2015.07.035] [Medline: 26850208]

127. de Sena DP, Fabrício DD, da Silva VD, Bodanese LC, Franco AR. Comparative evaluation of video-based on-line course versus serious game for training medical students in cardiopulmonary resuscitation: A randomised trial. PLoS One 2019;14(4):e0214722 [FREE Full text] [doi: 10.1371/journal.pone.0214722] [Medline: 30958836]

128. Hedman L, Schlickum M, Felländer-Tsai L. Surgical novices randomized to train in two video games become more motivated during training in MIST-VR and GI Mentor II than students with no video game training. Stud Health Technol Inform 2013;184:189-194. [Medline: 23400154]

129. Kolga Schlickum M, Hedman L, Enochsson L, Kjellin A, Felländer-Tsai L. Transfer of systematic computer game training in surgical novices on performance in virtual reality image guided surgical simulators. Stud Health Technol Inform 2008;132:210-215. [Medline: 18391288]

130. Lagro J, van de Pol MHJ, Laan A, Huijbregts-Verheyden FJ, Fluit LCR, Olde Rikkert MGM. A randomized controlled trial on teaching geriatric medical decision making and cost consciousness with the serious game GeriatriX. J Am Med Dir Assoc 2014 Dec;15(12):957.e1-957.e6. [doi: 10.1016/j.jamda.2014.04.011] [Medline: 24913210]

131. Middeke A, Anders S, Raupach T, Schuelper N. Transfer of Clinical Reasoning Trained With a Serious Game to Comparable Clinical Problems: A Prospective Randomized Study. Simul Healthc 2020 Apr;15(2):75-81. [doi: 10.1097/SIH.0000000000000407] [Medline: 32044851]

132. Phungoen P, Promto S, Chanthawatthanarak S, Maneepong S, Apiratwarakul K, Kotruchin P, et al. Precourse Preparation Using a Serious Smartphone Game on Advanced Life Support Knowledge and Skills: Randomized Controlled Trial. J Med Internet Res 2020 Mar 09;22(3):e16987 [FREE Full text] [doi: 10.2196/16987] [Medline: 32149711]

133. Sward KA, Richardson S, Kendrick J, Maloney C. Use of a Web-based game to teach pediatric content to medical students. Ambul Pediatr 2008;8(6):354-359. [doi: 10.1016/j.ambp.2008.07.007] [Medline: 19084784]

134. Tubelo RA, Portella FF, Gelain MA, de Oliveira MMC, de Oliveira AEF, Dahmer A, et al. Serious game is an effective learning method for primary health care education of medical students: A randomized controlled trial. Int J Med Inform 2019 Oct;130:103944. [doi: 10.1016/j.ijmedinf.2019.08.004] [Medline: 31442848]

135. Boyle E, Kennedy AM, Traynor O, Hill ADK. Training surgical skills using nonsurgical tasks--can Nintendo Wii™ improve surgical performance? J Surg Educ 2011;68(2):148-154. [doi: 10.1016/j.jsurg.2010.11.005] [Medline: 21338974]

136. Chien JH, Suh IH, Park S, Mukherjee M, Oleynikov D, Siu K. Enhancing fundamental robot-assisted surgical proficiency by using a portable virtual simulator. Surg Innov 2013 Apr;20(2):198-203. [doi: 10.1177/1553350612458545] [Medline: 22956399]

137. Albrecht U, Folta-Schoofs K, Behrends M, von Jan U. Effects of mobile augmented reality learning compared to textbook learning on medical students: randomized controlled pilot study. J Med Internet Res 2013 Aug 20;15(8):e182 [FREE Full text] [doi: 10.2196/jmir.2497] [Medline: 23963306]

138. Bogomolova K, van der Ham IJM, Dankbaar MEW, van den Broek WW, Hovius SER, van der Hage JA, et al. The Effect of Stereoscopic Augmented Reality Visualization on Learning Anatomy and the Modifying Effect of Visual-Spatial Abilities: A Double-Center Randomized Controlled Trial. Anat Sci Educ 2020 Sep;13(5):558-567. [doi: 10.1002/ase.1941] [Medline: 31887792]

139. Küçük S, Kapakin S, Göktaş Y. Learning anatomy via mobile augmented reality: Effects on achievement and cognitive load. Anat Sci Educ 2016 Oct;9(5):411-421. [doi: 10.1002/ase.1603] [Medline: 26950521]

140. Leitritz MA, Ziemssen F, Suesskind D, Partsch M, Voykov B, Bartz-Schmidt KU, et al. Critical evaluation of the usability of augmented reality ophthalmoscopy for the training of inexperienced examiners. Retina 2014 Apr;34(4):785-791. [doi: 10.1097/IAE.0b013e3182a2e75d] [Medline: 24670999]

141. Lemke M, Lia H, Gabinet-Equihua A, Sheahan G, Winthrop A, Mann S, et al. Optimizing resource utilization during proficiency-based training of suturing skills in medical students: a randomized controlled trial of faculty-led, peer tutor-led, and holography-augmented methods of teaching. Surg Endosc 2020 Apr;34(4):1678-1687. [doi: 10.1007/s00464-019-06944-2] [Medline: 31286252]

142. Logishetty K, Western L, Morgan R, Iranpour F, Cobb JP, Auvinet E. Can an Augmented Reality Headset Improve Accuracy of Acetabular Cup Orientation in Simulated THA? A Randomized Trial. Clin Orthop Relat Res 2019 May;477(5):1190-1199 [FREE Full text] [doi: 10.1097/CORR.0000000000000542] [Medline: 30507832]

143. Moult E, Ungi T, Welch M, Lu J, McGraw RC, Fichtinger G. Ultrasound-guided facet joint injection training using Perk Tutor. Int J Comput Assist Radiol Surg 2013 Sep;8(5):831-836. [doi: 10.1007/s11548-012-0811-5] [Medline: 23329279]

144. Noll C, von Jan U, Raap U, Albrecht U. Mobile Augmented Reality as a Feature for Self-Oriented, Blended Learning in Medicine: Randomized Controlled Trial. JMIR Mhealth Uhealth 2017 Sep 14;5(9):e139 [FREE Full text] [doi: 10.2196/mhealth.7943] [Medline: 28912113]

145. Sugand K, Wescott RA, Carrington R, Hart A, van Duren BH. Training and Transfer Effect of FluoroSim, an Augmented Reality Fluoroscopic Simulator for Dynamic Hip Screw Guidewire Insertion: A Single-Blinded Randomized Controlled Trial. J Bone Joint Surg Am 2019 Sep 04;101(17):e88. [doi: 10.2106/JBJS.18.00928] [Medline: 31483408]

146. Vera AM, Russo M, Mohsin A, Tsuda S. Augmented reality telementoring (ART) platform: a randomized controlled trial to assess the efficacy of a new surgical education technology. Surg Endosc 2014 Dec;28(12):3467-3472. [doi: 10.1007/s00464-014-3625-4] [Medline: 24962856]

147. Yeo CT, Ungi T, Leung R, Moult E, Sargent D, McGraw R, et al. Augmented reality assistance in training needle insertions of different levels of difficulty. 2018 Presented at: SPIE Medical Imaging: Image-Guided Procedures, Robotic Interventions, and Modeling; March 13, 2018; Houston, Texas, US.

148. Car J, Carlstedt-Duke J, Tudor Car L, Posadzki P, Whiting P, Zary N, Digital Health Education Collaboration. Digital Education in Health Professions: The Need for Overarching Evidence Synthesis. J Med Internet Res 2019 Feb 14;21(2):e12913 [FREE Full text] [doi: 10.2196/12913] [Medline: 30762583]

149. Cook DA, Hatala R. Validation of educational assessments: a primer for simulation and beyond. Adv Simul (Lond) 2016;1:31 [FREE Full text] [doi: 10.1186/s41077-016-0033-y] [Medline: 29450000]

Abbreviations

AR: augmented reality
COSMIN: Consensus-Based Standards for the Selection of Health Measurement Instruments
ER: extended reality
MR: mixed reality
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RCT: randomized controlled trial
VR: virtual reality
WHO: World Health Organization

XSL•FO
RenderX