Computers & Education: X Reality 3 (2023) 100034

Contents lists available at ScienceDirect

Computers & Education: X Reality

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

A systematic review of the experimental studies on the effectiveness of mixed reality in higher education between 2017 and 2021

Areej Banjar [a,b,*], Xuanhui Xu [b], Muhammad Zahid Iqbal [c], Abraham Campbell [b]

[a] Department of Computer at Jamoum University College, Umm Al-Qura University, Saudi Arabia
[b] School of Computer Science, University College Dublin, Ireland
[c] Department of Computer Science, Teesside University, United Kingdom

ARTICLE INFO

MSC:
00–01
99-00

Keywords:
Mixed reality
Mobile mixed reality
Higher education
PRISMA 2020

ABSTRACT

The integration of mixed reality technologies in higher education has gained momentum in recent years, offering promising opportunities for enhanced learning experiences. This systematic review aims to give an overview of the current evidence for the effectiveness of mixed reality use in higher education. By considering the PRISMA 2020 guidelines, the review has examined studies related to university students and explored all aspects of the PICOS model. A broad search of databases like IEEE Xplore, ProQuest, and Scopus was performed, selecting experimental studies published in English from 2017−−2021. PRISMA was chosen as a well-regarded systematic review approach, and the PICO model is specifically aimed at exploring the efficacy of an approach, hence its inclusion. The review includes 12 studies, half randomised control trials and half non-randomised. Quality assessment was performed using the Cochrane Collaboration ROB 2 and the ROBINS-I tools. The majority of these mixed reality studies concentrated more on 3D manipulation, visualisation, and understanding of the 3D object layers and components than procedural learning using HoloLens. Of the selected studies, 53% were in the medical and health sciences, particularly in anatomy, followed by 34% in engineering education, which suggests that these fields are more open to the use of MR for educational purposes than theoretical disciplines such as the humanities and social sciences. Of the 12 studies, nine used augmented reality via head-mounted displays, and five used mobile mixed reality. These studies show that mixed reality has the potential to enhance learning experiences in higher education. Although there are challenges to overcome, MR offers opportunities for innovation in pedagogical practises and curriculum development.

1. Introduction

The rapid advancement of digital technologies has led to a growing interest in employing their potential to enhance teaching and learning experiences in education (Tang et al., 2018). Mixed Reality (MR) is coming to the forefront of these digital technologies, which refers to merging virtual and real worlds to generate a new visualisation environment in which digital and physical objects interact and coexist in real time (McMillan et al., 2017). MR technology allows the integration of virtual data and the physical environment to allow users to interact with both virtual and physical content, thus enhancing their experience (Chen et al., 2020). Today, MR promises to be a truly innovative technology that becomes increasingly prominent in daily life (Pellas et al., 2020) and in various fields, including the humanities and Arts, Social Sciences, Business, Law, Engineering, and health professions, which can all benefit from the use of MR technology in higher education (Marcel, 2019) has gained considerable research attention in recent years (Knierim et al., 2018). However, despite this growing interest, no studies have examined instructional contexts to gather information on MR technology uses in education, according to Pellas et al. (2020). Therefore, there is a need for a comprehensive understanding of the effectiveness, challenges, and best practices associated with the integration of MR into higher education. This systematic literature review focuses on MR in higher education. It is systematically organised according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) (2020) (Page et al., 2021c, 2021a, 2021b) which identifies, evaluates, and summarises the findings of all relevant individual studies published in the field making the available evidence

* Corresponding author. Department of Computer at Jamoum University College, Umm Al-Qura University, Saudi Arabia.
E-mail addresses: arbanjar@uqu.edu.sa, areej.banjar@ucdconnect.ie (A. Banjar), Xuanhui.xu@ucdconnect.ie (X. Xu), Zahid.iqbal@tees.ac.uk (M.Z. Iqbal), abey.campbell@ucd.ie (A. Campbell).

https://doi.org/10.1016/j.cexr.2023.100034
Received 25 July 2022; Received in revised form 22 July 2023; Accepted 23 July 2023
Available online 9 August 2023

Table 1
Reviews conducted on MR and Education.

Authors	Title	XR type	Aim	Educational Domain	Educational level	Review type	Time
(Li & Wong, 2021)	"A literature review of augmented reality, virtual reality, and mixed reality in language learning."	AR, VR and MR	present the current developments in the field of utilisation of Augmented Reality (AR) and MR technologies in language education and to explore their future perspectives.	Language learning	Not specified	Literature review	2004 and 2018
(Maas & Hughes, 2020)	"Virtual, augmented and mixed reality in K-12 education: A review of the literature."	VR, AR and MR	The review explores the peer-reviewed scholarly studies conducted between 2006 and May 2017, which involved the use of Virtual Reality (VR), augmented reality (AR), or mixed reality (MR) technologies in the instruction of students in elementary, middle or high school.	K-12 education	students in elementary, middle or highschool	Literature review	2006 and May 2017
(Pellas et al., 2020)	"A systematic literature review of mixed reality environments in K-12 education."	MR	This study aims to review the collected studies systematically and to provide any foundation for knowledge accumulation that can assist any potential theories' expansions and improvements, identifies and closes research "gaps", unveiling further areas where previous research that has not been thoroughly addressed yet.	K-12 education	primary and secondary education	Systematic review (PRISMA)	between 2002 and the fourth quarter of 2018
(Gerup et al., 2020)	"Augmented reality and mixed reality for healthcare education beyond surgery: an integrative review."	AR and MR	This study aimed to review and synthesise the current research and state of aim Aim of this integrative review was to investigate the current research and state of AR and MR-based applications for healthcare education beyond surgery, providing an overview of the findings, strengths and weaknesses of the reported studies.	Healthcare Education	Companies, universities and hospitals (Pre-medical, medical, nursing, and health science students, novices, residents, fellows and established clinicians of different specialties, technicians, non-clinicians, non-specified participants and managers.)	Integrative review	January 2013 till September 2018.
(Barrie et al., 2019)	"Mixed reality in medical education: a narrative literature review."	AR, MR and blended reality.	Identify elements to consider when using immersive technologies in medical education	Medical education	Individuals and groups of medical students	Narrative literature review	2008–2018.

more accessible to decision-makers (Gopalakrishnan & Ganeshkumar, 2013). This review offers insight into effective pedagogical approaches and strategies for integrating MR into curricula and teaching practice; identifies best practices for successfully implementing MR in higher education; and finds gaps in the existing literature, suggesting areas that require further investigation or emerging trends that merit attention. In addition, it can provide insights to inform educational practices, policy-making, and future research. These insights may include the effectiveness of MR technologies in improving learning outcomes such as academic performance, engagement, motivation, and satisfaction. By synthesising these insights, a systematic review of mixed reality in higher education can contribute to a better understanding of the potential benefits and challenges of MR technologies and inform decision-making, planning, and implementation of MR in educational settings.

1.1. *Background and rationale*

The continual development and technological innovations in eXtended Reality (XR) make the subject attractive to many scholars (Radianti et al., 2020). The latest technological innovations in XR, such as Head-Mounted Display (HMD), give users the ability to experience objects in a more immersive way (Xu et al., 2021). Immersion refers to users' simultaneous interaction with virtual and physical environments to create a sense of being in the created hybrid realm (Johnson-Glenberg, 2018). The following sections describe the rationale for the review in the context of existing knowledge of: MR in education, reviews conducted that include MR in education, and using the population, intervention, comparison, outcomes and study design (PICOS) framework in the systematic reviews.

1.1.1. *Mixed reality in education*

Various studies have addressed the application of MR technology to the education field, reflecting increased scholarly attention (Birt et al., 2018; Burke et al., 2017). MR technology is gaining momentum in the education sector as a potential tool for learning and teaching (Banjar & Campbell, 2022; Pellas et al., 2020). According to Hauze and Marshall, the diffusive integration of MR technology in education in recent years has benefited learners by improving their motivation, problem-solving skills, and overall learning experience (Hauze & Marshall, 2020). MR gives students an outstanding opportunity to actively participate in the learning process through collaborative interaction and problem-solving using real-world objects such as robots, books, and maps. This additional layer of interactive applications creates a learning environment that positively influences learners' attention and provides them with a more engaging and fun learning experience than traditional methods (Pellas et al., 2020).

1.1.2. *Reviews of MR in education*

We chose to conduct a systematic review given that previous reviews showed only a few studies relevant for the current scope, which is experimental studies on mixed reality in higher education. Table 1 documents reviews conducted in MR in education up to the time of this review.

Most of the reviews conducted for MR in education are not systematic reviews, and only one review so far has systematically reviewed collected studies that are focused on K−−12 education. It aims to review the collected studies systematically and to provide a foundation for knowledge accumulation that can assist any potential theories' expansions and improvements; it identifies and closes research "gaps," revealing additional areas where prior research has not been thoroughly addressed (Pellas et al., 2020). Meanwhile, an integrative review that involved MR has examined the current research and state of Augmented Reality (AR) and MR-based applications for healthcare education beyond surgery, providing an overview of the reported studies' findings, strengths, and limitations (Gerup et al., 2020), which include higher education as well as companies and hospitals but are not specific to universities. The rest of the reviews conducted to this time on Mixed reality in education are Narrative or literature reviews (Li & Wong, 2021; Maas & Hughes, 2020; Barrie et al., 2019). Based on the previous reviews, they were conducted from 2004 until 2018. Therefore, this study is aimed at systematically reviewing experimental studies conducted from 2017 to January 2021 on mixed reality among university students with the guidelines of PRISMA, using the most commonly used framework to construct questions, and developing the search strategy PICOS.

1.1.3. The PICO framework

The PICO framework is used in evidence-based practice to formulate related questions (Stone, 2002). The PICO model is based on five factors which are population, intervention, comparison, outcomes and study design. In a systematic review, PICO is also used to construct literature search strategies to ensure comprehensive and unbiased searches (Division, 2020). The PICO framework emphasises population, intervention, comparison, and outcomes. It is a commonly used instrument for quantitative systematic reviews to identify various review components and is recognised by the most famous organisation to facilitate evidence-based choices about health interventions, called the Cochrane Collaboration (Higgins & Green, 2006). The Cochrane Handbook for Systematic Reviews of Interventions mentioned using the PICO framework as a model for developing a review question, thus ensuring that the relevant components of the question are well defined (Higgins & Green, 2006; Eriksen & Frandsen, 2018). In spite of the existence of other models---such as sample, the phenomenon of interest, design, evaluation, and research type (SPIDER) (Cooke et al., 2012) and setting, perspective intervention, comparison, and evaluation (SPICE) (Booth, 2006), PICO is by far the most popular framework for formulating clinical questions (Eriksen & Frandsen, 2018). However, the type of research question posed will be crucial in determining the most effective type of systematic review. This systematic review contains interventional research questions requiring the PICO framework (Bettany-Saltikov & McSherry, 2018; Pollock & Berge, 2018). A systematic review question usually focuses on limited parameters and fits into the PICO question format (Division, 2020). This review focuses on experimental studies of mixed reality in higher education, and the following is a description of each component:

- P -- Population: Most important characteristics of participants. Examples: sample size, educational level and educational domains.
- I -- Intervention or exposure: Main intervention used by experimental groups. Examples: MR including AR and AV and mixed realities such as AR and Virtual Reality (VR).
- C -- Comparison or control: Main alternative used by Control group. Examples: VR, traditional learning materials and conventional monitors.
- O −− Outcome: What you are trying to accomplish, improve, measure, or affect. Examples: learning outcomes, performance, motivation, engagement, and satisfaction.

PICO can be used along with variants such as PICOS (S-Study design) when answering a question on the effectiveness of an intervention (Higgins & Green, 2006) which is included in this review.

- S −− Study design: Experimental studies including RCTs and NRSs.

The PICO model was developed to aid in constructing a well-structured question and facilitate a search for relevant citations in the literature (Division, 2020). It has played a crucial role as a conceptual framework for evidence-based practise since its inception (Eriksen & Frandsen, 2018). The PICO framework will also aid in the reduction of time and retrieval of relevant documents, thereby ensuring a high-quality, bias-free systematic review and helping to determine the transparency of evidence synthesis results and conclusions (Division, 2020).

The rationale for a systematic review of MR in higher education at this time is based on three primary contributions that we feel it can make to the community. Firstly, it enables researchers to critically and comprehensively analyse the current body of knowledge about MR technology and its integration into higher education, providing a thorough outline of current trends, effectiveness, challenges, and best practices of MR in a higher education setting. This process helps identify gaps in the current research, which can then guide future inquiries. Secondly, the systematic review reduces the risk of bias in summarising research findings, as it follows the clearly defined and reproducible methodology of PRISMA 2020. Thirdly, by synthesising a large volume of research, this review can provide robust evidence to inform policy and practice in higher education. In a rapidly evolving field like higher education, with an ever-growing body of research, systematic reviews can play a vital role in providing clear, evidence-based insights to educators, policymakers, and researchers.

1.2. Research questions and objectives

A structured approach for framing questions may facilitate the process of formulating relevant and precise questions, which can be complex and time-consuming (Liberati et al., 2009). This five-component approach is commonly known by the acronym "PICOS": population, intervention, comparison, outcome, and study design (Liberati et al., 2009). The PICOS approach was adopted to analyse the literature mainly by extracting the key information for the study and the scope of the review (Liberati et al., 2009) and specifying the research characteristics, as suggested by Wendler (2012). The search yielded numerous study questions, mainly focused on systematising and structuring the study of MR in higher education:

- *Population:*
  - What are the characteristics of the participants using MR technology in higher education?
  - Which higher education domains use MR to examine MR effectiveness?
- *Intervention:*
  - What MR characteristics and technologies are applied to higher education?
- *Comparison:*
  - What are the characteristics of educational materials that were compared to MR used in higher education?
- *Outcomes:*
  - What are the learning outcomes of students who have experienced MR technology?
- *Study design:*
  - What characteristics of the study design of the MR studies are applied to higher education?

The main objectives of the systematic review are:

- To outline characteristics of MR in various aspects of higher education, such as population, educational domains, and curriculum.
- To outline the use of MR devices and their technological equipment in various studies.
- To demonstrate the benefits and limitations of MR in the learning environment compared to existing learning materials.
- To determine the effectiveness of MR on learning outcomes and in enhancing learning experiences in higher education.
- To explore the best educational practises, research methods, and study design employed in MR experimental research.

2. Method

The systematic review followed the PRISMA 2020 guidelines (Page et al., 2021c) for increasing and maintaining the accuracy of review studies, which was carried out until January 1, 2021. The review process are encompassed procedures, decisions, and considerations that guided the in-depth analysis of the consolidated list of articles.

2.1. Eligibility criteria

The eligibility of each study was evaluated based on its title, which had to focus on the "mixed reality" term. The review was restricted to English-language studies published from 2017 to 2020 because "a Google Trends search revealed an increasing interest in the topic of VR since 2016 when the immersive HTC Vive headset was released" (Radianti et al., 2020). Thus, starting the search from 2017 increased the likelihood of obtaining immersive VR-based learning articles. Due to the novelty of immersive HMDs, it was necessary to include conference papers; most innovative research and development using HMDs was documented in conference papers instead of journal articles.

The search was limited to scholarly journals and conference paper proceedings. Therefore, studies were included in the review if they

1. were used for higher education or university students,
2. used MR technologies,
3. were published in peer-reviewed scientific journal articles or conference papers between 2017 and 2020, and
4. were published in English.

Articles were excluded if they

1. were used for public education or schools,
2. studied non-student populations such as employees and trainees,
3. were not written in English,
4. were doctoral dissertations or pilot studies, or
5. did not include the full text of the study report.

Additional inclusion and exclusion criteria were added for the remaining articles in the screening phase to exclude all irrelevant articles. The additional inclusion criteria were as follows:

1. Experimental studies.
2. Studies that include comparative groups.
3. Studies that include only MR in the forms of AR HMDs and MMR.

The additional exclusion criteria are the following:

1. Non-experimental studies.
2. Survey papers.
3. Review articles.
4. Papers about VR.

2.2. Information sources

In the design of the review, we began by selecting keywords to search all the relevant digital libraries. The digital libraries used in the study include IEEE Xplore, ProQuest, and Scopus. IEEE Xplore is a rich repository that mainly covers the domains of computer science, engineering, information technology, and other software-related technology. ProQuest comprises articles in the areas of medicine, surgery and nursing sciences. Finally, the Scopus database offers a wide array of publication domains in the natural sciences, technology, social sciences, information technologies, and medicine.

2.3. Search strategy

In a systematic literature review, the PRISMA 2020 guidelines are an essential tool for ensuring thoroughness and transparency. To ensure comprehensive and bias-free searches, the PICOS framework was used to develop the keywords for each element as follows:

- Population: Students, educators, instructors and educational institutions (e.g colleges and universities).
- Intervention: Mixed reality technology including (augmented reality, augmented virtuality) and combination of augmented reality and virtual reality applied in educational settings.
- Comparison: Traditional learning methods, or other technology-enhanced learning methods (e.g. e-learning, multimedia, simulations).
- Outcome: Learning outcomes, engagement, motivation, satisfaction, cognitive load, knowledge retention and skills acquisition and performance etc.
- Study design: Randomised controlled trials, non-randomised studies, quasi-experimental studies, pre-experimental studies, quantitative studies, qualitative studies and mixed method studies etc.

Based on the PICOS model, we chose the following search terms:

- Population: "students", "educators", "universities", "colleges".
- Intervention: "mixed reality", "MR", "augmented reality", "AR", "augmented virtuality", "AV", "virtual reality".
- Comparison: "traditional learning", "face-to-face learning", "classroom instruction", "e-learning", "multimedia", "simulation", "computer-assisted learning".
- Outcome: "learning outcomes", "engagement", "motivation", "satisfaction", "cognitive load", "knowledge retention", "skills acquisition", "performance".
- Study design: "randomised controlled trial", "RCT", "non randomised studies", "NRS", "quasi-experimental", "pre-experimental", "quantitative", "qualitative" and "mixed method".

Using the search terms, the search strings were constructed by combining terms from different categories using Boolean operators (AND, OR, NOT). Each database has specific requirements and syntax for search strings. Due to the extensive research associated with education fields, each search term in education yielded a high number of results. Therefore, the search strings have been limited to the following strings that contain the words "educat*" and "mixed reality", which are more likely to have the desired research.

We defined the following search strings for the database search as shown in Table 2.

2.4. Selection process

The systematic review in the selected model consists of three phases of the article selection procedure: identification, screening, and inclusion. These phases are based on the PRISMA 2020 flow diagram for new systematic reviews (Page et al., 2021a), as illustrated in Fig. 1. We adopted a two-step filtering procedure in the identification phase, including (i) semi-automatic filters for the inclusion and exclusion strategies and (ii) manual filters of the potential articles based on the title.

Table 2
Search terminologies used in the selected databases.

Database	search terminologies in the titles and keywords
IEEE Xplore	"mixed reality" in educat*
ProQuest	"mixed reality" AND in AND educat*
Scopus	"mixed reality" AND in AND educat*

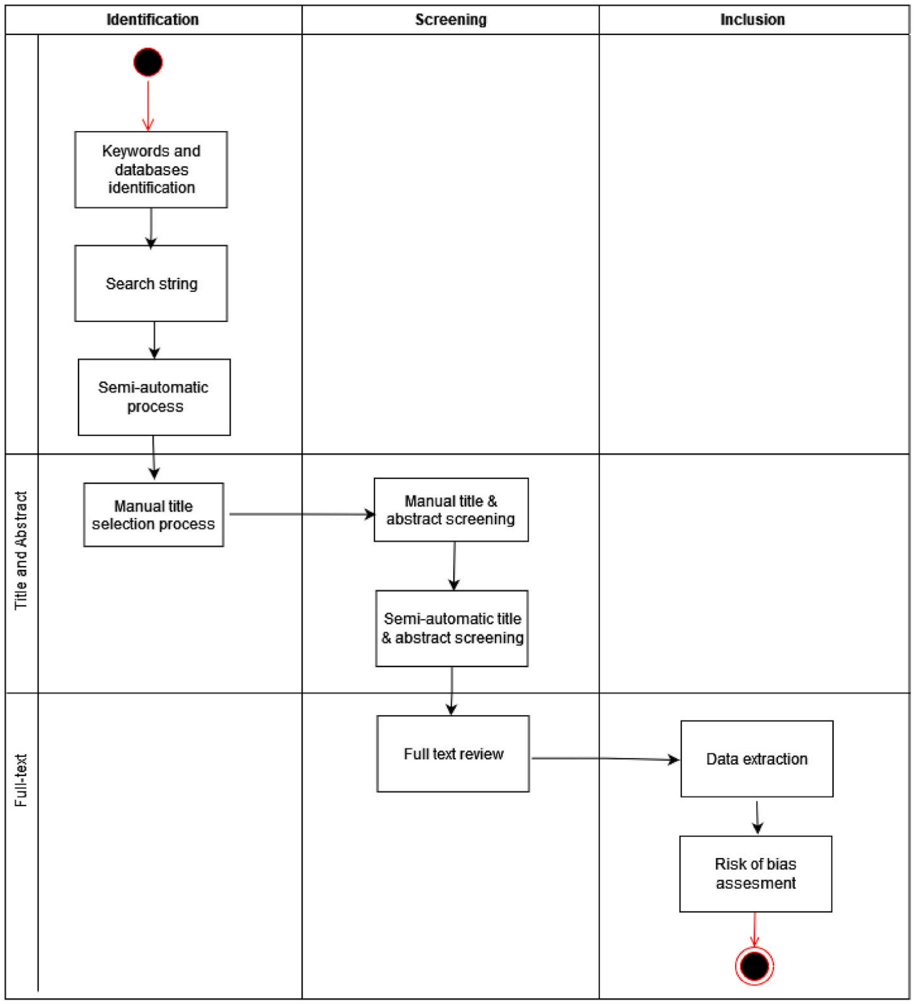

Fig. 1. Activity diagram of the systematic review strategy.

With the semi-automatic filters, articles were excluded or included by selecting representative exclusion and inclusion keywords in the digital libraries. These keywords were provided by the digital library to narrow down the list of articles for each database. After that, the manual selection process involved choosing titles that only included the "mixed reality" term. This step allowed the number of articles to be further reduced.

In the screening phase, four review authors (AB, AC, MZ, and XX) independently screened the potential articles in two steps: (i) manual and (ii) semi-automatic. In the manual step, after importing the Endnote file into Excel, the reviewers screened the title and abstract using the inclusion criteria, then marked the article as either irrelevant or relevant to the research. Each study was assessed by two reviewers, and conflicts were resolved by (AB). Each reviewer had to assess 50% of the total number, and every 25% of the total number of potential articles was assessed by different reviewers.

To increase the review's validity, a semi-automatic method was adopted for the screening phase by Covidence,[1] which manages systematic reviews to reduce selection bias. Covidence is a web-based systematic review programme created for systematic reviewers. It can import citations from reference managers like Endnote, help with abstract and full-text screening, complete risk-of-bias tables, assist with data extraction, and export to all standard formats (UTAS, 2023). The process consists of three phases: (i) screening of titles and abstracts, (ii) full-text review, and (iii) data extraction. We manually filtered in XML format an Endnote library collected from various databases, then uploaded it to the Covidence website. Then, two reviewers (AB and MZ) performed additional title and abstract selection according to the inclusion criteria. The assessment of the abstracts and titles allowed them to be designated as either "Yes", "No", or "Maybe," and conflicts were resolved by (AB). This process moved the consensus studies automatically to the next phase. Full-text reviews were performed by two reviewers (AB and AC). Any disagreements were resolved by discussion and included in the systematic review.

2.5. Data collection process

The data collection process is the process of extracting data from selected studies to facilitate the analysis of similar data from several sources. It is a critical component of any systematic or literature review process because it enables reviewers to assemble a large enough data collection to generate significant evidence (Covidence, 2023).

Data extraction is a crucial step, involving the collection of relevant information from the selected studies to answer the research questions and achieve the research objectives. With the Covidence tool, a data extraction template was created that could be used for every study processed during the title, abstract, and full-text screening phases to ensure consistency and efficiency during the systematic review. Developing a data extraction template involves creating a structured document to collect and organise the relevant information from the selected studies.For this study, the template included fields for key elements of

[1] https://www.covidence.org.

research questions and objectives, constructed based on the PICOS framework, to help standardise the process and make it more efficient, as well as any additional information relevant to the review. The primary author conducted data extraction by carefully reading every article to extract the information needed for each variable. Then, the extracted data was exported to CSV format and opened with Microsoft Excel for analysis.

2.6. Data items

Data items are variables and fields of the key elements in the data extraction template. They were carefully identified based on the research questions. These items include information that has been used to address the research questions (Zhang et al., 2020) and include two types of data: (i) data about the study as shown in Table 3, and (ii) data from the study as shown in Table 4.

2.7. Risk of bias

In terms of Quality assessment, one review author (XX) critically assessed the risk of bias for the methodological quality of Randomised Control Trial using the Cochrane collaboration risk of bias (ROB 2) tool (Sterne et al., 2019). Another review author (AB) assessed the risk of bias for Non-Randomised Study (NRS) using the Cochrane collaboration risk of bias in non-randomised studies of interventions (ROBINS-I) tool. Each potential source of bias was graded as “low”, “high”, or “some concern” for RCTs and “low”, “moderate”, “serious”, “critical”, or “no information” for NRSs. Each judgement was supported by a quote from the relevant study. The risk-of-bias plots were created using the Robvis tool.

3. Results

3.1. Study selection

Without considering any search criteria, the initial search in the three relevant libraries yielded 8302 articles. Due to the huge number of results from the initial search, we adopted two-phased filtering procedures in the identification phase, including (i) semi-automatic filters for the inclusion and exclusion strategies and (ii) manual filters for the potential articles. The semi-automatic filter led to an aggregated result of 1608 articles. The eligibility of each study was evaluated based on the title. As a result, 1144 of 1608 articles were excluded from the digital library, reducing the number of identified articles to 464. Of the articles identified for the review, 112 redundant results were removed by the “find duplicates” feature in Endnote after data was imported from the library website. Thus, the final number of articles was 352. The abstracts found in the search were assessed for the 352 remaining articles, which allowed the researchers to identify the articles with abstracts that satisfied the listed criteria. The assessment of the 352 abstracts and titles allowed them to be marked as either irrelevant or relevant to the research. This process resulted in the inclusion of 80 articles in the systematic review and the exclusion of about 272 articles. Then, further title and abstract selection were done by two reviewers (AB and MZ) using the Covidence website, and conflicts were resolved by (AB). This process resulted in 50 articles being automatically designated for full-text review. During this phase, some studies were removed from the eligible articles because of the wrong intervention (n = 7), the wrong population (n = 2), the wrong study design (n = 16), or a tool design that was not an educational study (n = 13). As a result, 12 studies were identified for inclusion in the systematic review (See Fig. 2).

Table 3
Data about the study.

Data Items	Variable description
Author	Authors’ name
Year	Study’s publishing year
Title	title of the study
Research method	Qualitative research QL, Quantitative research QN or Mixed method MM
Study design	Experimental RCT, NRS, QUASI experimental or Pre-experiment
Data collection	Surveys, forms, or questionnaires for QN, Interviews for QL and QN, Observation for QL, Documents and records and/or focus groups for QL
Data analysis	QN descriptive statistics, QN inferential statistics, QL content, Narrative, Discourse, Thematic, Grounded theory analysis and/or Interpretative phenomenological analysis IPA

Table 4
Data from the study.

Data Items	Variable description
Aims	The main goals or purpose of the research
Method	The strategies, processes, or techniques of conducting the research
Results	Scientific knowledge derived from the implementation
Population	Sample size and study domain that includes STEM (science, technology, engineering, math), medical and health science, and/or humanities and social science
Intervention	Type of MR which includes AR (smartphone/tablet), MR (AR HMDs), AV (VR HMDs), or VR + AR
Comparison	Between groups or compared to control groups
Outcome	Measures and general outcomes of the study
Limitations	Research constraints, flaws, and shortcomings
Future work	Future research suggestions

3.2. Study characteristics

The characteristics of the study for which data were extracted have been presented for each included study, guided by the PICOS framework to gauge the validity and applicability of a systematic review’s results (Liberati et al., 2009).

3.2.1. Participants

The included studies comprised 1111 participants, including 495 participants in the RCTs and 616 participants in the NRSs. There were 187 participants in the experimental groups and 120 participants in the control groups. The main inclusion criterion is university students. While not all studies identified the participants’ age and gender, most identified the students’ academic level and their major. The included studies involved a range of academic levels, with undergraduate, graduate, and postgraduate university students. Half of the included studies were in medical and health sciences (n = 6)(Al Janabi et al., 2020; Birt et al., 2017; Hauze et al., 2018; Robinson et al., 2020; Ruthberg et al., 2020; Stojanovska et al., 2019), followed by STEM (n = 3) (Frank & Kapila, 2017; Tumkor, 2018; Vasilevski & Birt, 2020). The rest were from other academic areas (n = 3) (Tang et al., 2018, 2020; Wainman et al., 2020). The number of students in each type of MR technology is shown in Fig. 3.

3.2.2. Interventions

Two types of MR interventions were applied in the studies selected for review: AR HMDs and MMR, which includes mobile AR and mobile VR. Nine studies used MR associated with AR HMDs; (Al Janabi et al., 2020; Hauze et al., 2018; Robinson et al., 2020; Ruthberg et al., 2020; Stojanovska et al., 2019; Tang et al., 2018, 2020; Tumkor, 2018; Wainman et al., 2020); four studies used mobile AR through smartphones or tablets(Birt et al., 2017; Frank & Kapila, 2017; Tumkor, 2018; Vasilevski & Birt, 2020); and one study used mobile VR HMDs (Wainman et al., 2020). Also, one study used “mixed reality” terminology for mixing realities, describing a combination of mobile AR and mobile VR technologies (Vasilevski & Birt, 2020). Similarly, some authors use the terms augmenting reality and MR interchangeably with the term AR (Akçayır & Akçayır, 2017).

Various hardware devices and software were used to evaluate the effectiveness of the interventions, including AR HMDs, such as Microsoft

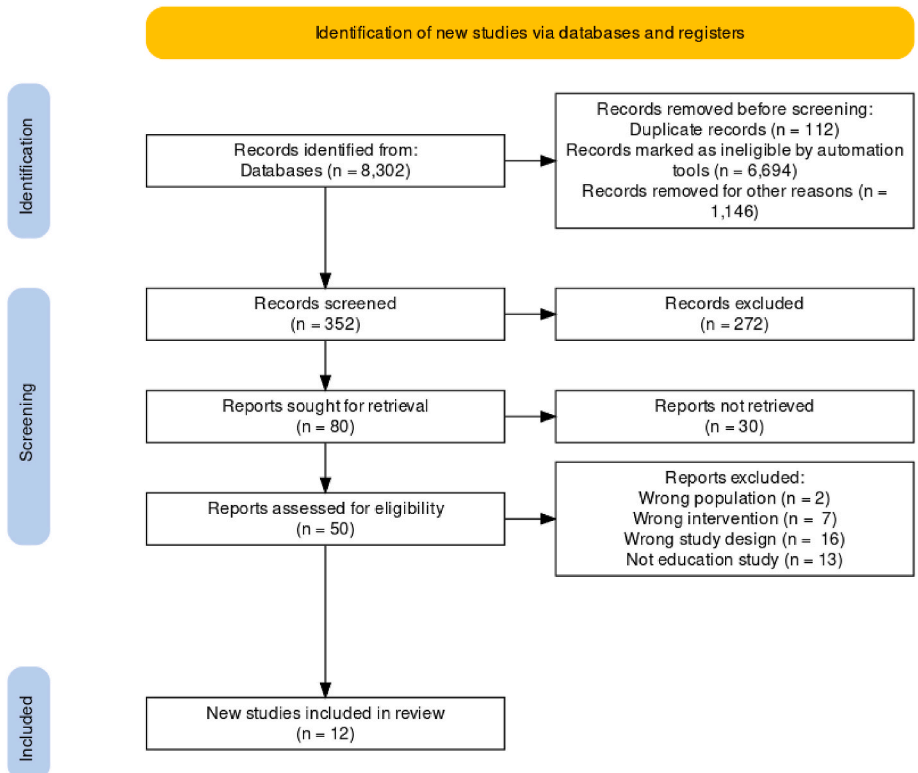

Fig. 2. PRISMA 2020 flowchart.

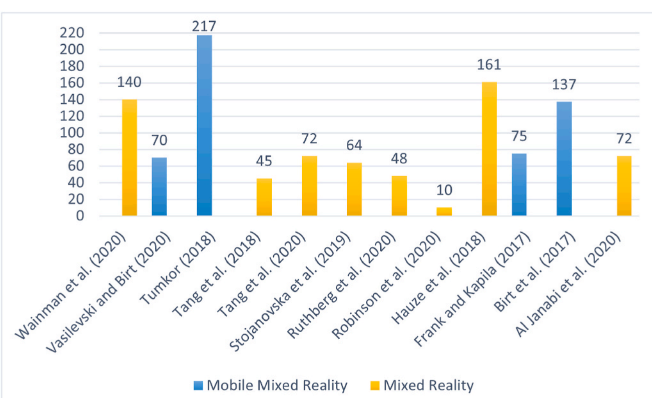

Fig. 3. Number of students in each type of MR technology.

HoloLens and Epson Moverio BT200 (Tumkor, 2018); MMR, which includes mobile VR and mobile AR: mobile VR, such as Samsung Gear VR headsets, Google Cardboard (Tumkor, 2018), and ColorCross headsets or head mount (Birt et al., 2017); mobile AR, like Samsung Galaxy S8, S7, or S6 (Vasilevski & Birt, 2020) and Apple iPad 2 (Frank & Kapila, 2017); and mobile VR HMDs such as HTC VIVE (Wainman et al., 2020).

The other hardware elements used to support the MR technology are the iBeacon IoT to support micro-location (Vasilevski & Birt, 2020) and 3D-printed equipment in the augmented environment (Frank & Kapila, 2017).

For MR development, the most common software used is Unity with the Vuforia plugin and supporting development tools such as the Gesture Manager in the HoloTool Kit (Tang et al., 2018) and the Google Cardboard Software Development Kit (SDK). In addition, 3D computer graphics software is used for modelling (Birt et al., 2017), such as Autodesk REVIT(Vasilevski & Birt, 2020) and 3D Studio MAX by Autodesk (Stojanovska et al., 2019). Some of the studies evaluated existing applications such as HoloAnatomy (Ruthberg et al., 2020). Others used applications developed by the studies' authors. In the context of educational theories, most studies rely on at least one learning theory, including simulation, constructive, game-based, multimedia, operational, experiential, contextual, and generative learning.

3.2.3. *Comparison*

In RCTs or NRSs, experimental groups were compared to control groups. Some studies compared MR with traditional learning materials such as teaching notes (Tang et al., 2020), cadaveric dissection (Ruthberg et al., 2020; Stojanovska et al., 2019), glass histology slides using light microscopes (Robinson et al., 2020), 2D video, written case study (Hauze et al., 2018), and conventional monitor(Al Janabi et al., 2020). Others compared different types of MR, either those associated with AR based on HMDs (Al Janabi et al., 2020; Hauze et al., 2018; Robinson et al., 2020; Ruthberg et al., 2020; Stojanovska et al., 2019; Tang et al., 2018, 2020; Tumkor, 2018; Wainman et al., 2020), MMR (Birt et al., 2017; Frank & Kapila, 2017; Tumkor, 2018; Vasilevski & Birt, 2020), or AVs based on VR HMDS (Wainman et al., 2020).

3.2.4. *Outcomes*

In most of the studies, the primary outcome was positive for both of the selected MR types. MR was associated with AR HMDs ($n = 9$) (Al Janabi et al., 2020; Hauze et al., 2018; Robinson et al., 2020; Ruthberg et al., 2020; Stojanovska et al., 2019; Tang et al., 2018, 2020; Tumkor, 2018; Wainman et al., 2020) and MMR ($n = 4$) that used smartphones or tablets (Birt et al., 2017; Frank & Kapila, 2017; Tumkor, 2018; Vasilevski & Birt, 2020).

Two studies were neutral, particularly for learning anatomy. These studies stated that VR and MR technologies were shown to be inferior to physical models, indicating that real stereopsis is necessary (Wainman et al., 2020). Also, regardless of the study modality, students performed similarly on the MR and the cadaver practical exams (Stojanovska et al., 2019).

3.2.5. Study design

A variety of research methods and study designs were employed in the included studies, including qualitative studies (n = 1) (Vasilevski & Birt, 2020), quantitative descriptive studies (n = 6) (Hauze et al., 2018; Ruthberg et al., 2020; Stojanovska et al., 2019; Tang et al., 2020; Tumkor, 2018; Wainman et al., 2020), mixed methods (n = 5) (Al Janabi et al., 2020; Birt et al., 2017; Frank & Kapila, 2017; Robinson et al., 2020; Tang et al., 2018), RCTs (n = 6) (Hauze et al., 2018; Robinson et al., 2020; Ruthberg et al., 2020; Stojanovska et al., 2019; Tang et al., 2020; Wainman et al., 2020), NRSs (n = 6) (Al Janabi et al., 2020; Birt et al., 2017; Frank & Kapila, 2017; Tang et al., 2018; Tumkor, 2018; Vasilevski & Birt, 2020), and, in one case, design-based research (DBR) (Birt et al., 2017).

3.3. Risk of bias

The quality information from six RCTs was collected using ROB 2.0 Tool (Sterne et al., 2019) for assessing the risk of bias. The analysis process was conducted by the author, using the Excel tool to implement ROB 2.0 .[2] All RCT studies' risk of bias is shown in Fig. 4.

One article had a low risk of bias, one article had some bias concerns, and four had a high risk of bias. All papers clarified that the outcome measurement was appropriate (D4) and that there was no selection of the reported result, as they all have ethical approval by the ethical department (D5).

D1: Some studies did not clarify whether the allocation sequence was random and did not analyse baseline differences. One had suggested a problem with the randomisation process according to the baseline, as the study group performed worse than the control in the pre-test, but their Self-Perceived Understanding of Anatomy had a higher score than the control (Robinson et al., 2020).

D2: The primary concern was bias due to deviations from the intended intervention. For instance, researchers did not mention whether the participants or the data accessors were blinded to the random assignment, and information was lacking about baseline analysis between groups (Ruthberg et al., 2020; Stojanovska et al., 2019).

D3: One had a high risk of bias in the missing outcome data domain due to the three missing data collection in the control group in the student feedback section, which may cause a perspective bias (Robinson et al., 2020). Another had a high risk of bias in the same domain because it only contained preliminary data (Hauze et al., 2018).

In the six NRS studies, the quality information was collected using ROBINS-I Tool(Sterne et al., 2016) for assessing the risk of bias. The analysis process was conducted by the author, using the Excel tool to implement ROBINS-I .[3] All NRS studies' risk of bias was shown in Fig. 5.

One article had a low risk of bias, two articles had moderate bias, and three had serious risk of bias.

D1: Some studies may have bias due to confounding as a result of students' switching between the interventions being compared. The studies evaluating only one intervention have low confounding bias. In (Al Janabi et al., 2020)'s study, novices were initially assessed with a conventional monitor, followed by an assessment with the HoloLens, which could cause confounding bias.

D2: Some studies can have bias in the selection of participants in the study (or in the analysis), which may be caused by the start of intervention not coinciding with the start of follow-up for most participants. For example, Birt et al. (2017) were not able to observe students using the simulation at a distance and usability problems they may have experienced in real-time. Also, in Vasilevski and Birt (2020) study, students submitted reflective essays after experiencing the MR activity.

D3: Bias in the classification of interventions can happen when defining an intervention group or when the status of an intervention classification has been affected by the knowledge of the outcome. For example, in Vasilevski and Birt (2020) study, the intervention group was not clearly defined, and the information used to define it was not recorded at the start of the intervention because the number of available mobile devices and headsets for the VR hands-on sessions was limited. As a result, one group participated in VR and the other in AR, while students who were not part of the hands-on groups were asked to work on their existing projects. Another example of bias is the status of intervention classification affected by knowledge of the outcome in Tumkor (2018) study, in which students were categorised as i) first-time learners, ii) learners with CAD experience, and iii) first-time learners who play video games. Likewise, Frank and Kapila (2017) classified participants based on their knowledge, and then they received the same intervention. That may have affected their intervention classification status, which could be influenced by knowledge of the outcome.

D4: Most of the studies did not mention bias due to deviations from the intended intervention. This bias can arise due to knowledge of the intervention applied. In the case of MR, blinding of the intervention to the participants or outcome assessors would not affect the outcome, but assessing mutable intervention at the same time, as happened in Vasilevski and Birt (2020) study, may cause deviation from the intended intervention.

D5: Some studies can have a bias due to missing data, which can be caused by either a lack of availability of outcome data for all or nearly all participants or by participants being excluded from the analysis. One study has a critical risk of bias (Birt et al., 2017), two have a serious risk of bias (Al Janabi et al., 2020; Vasilevski & Birt, 2020), and two have a moderate risk of bias due to missing data (Frank & Kapila, 2017; Tang et al., 2018). In all studies, outcome data was not mentioned whether it was available to the students or not. One study has a critical risk because it excluded participants for not attending the residential school for assessment after they received the MR tools (Birt et al., 2017).

D6: The main concern is about bias in the measurement of outcomes, which is usually raised when the outcome measure has been influenced by knowledge of the intervention received, outcome assessors are aware of the intervention received by study participants, the methods of outcome assessment are not comparable across intervention groups, or if there are any systematic errors in measurement of the outcome related to the intervention received. In the case of MR, blinding the intervention to participants or outcome assessors may not affect the outcome that much (Al Janabi et al., 2020; Birt et al., 2017; Frank & Kapila, 2017; Tumkor, 2018; Vasilevski & Birt, 2020).

D7: One study has a serious risk of bias in the selection of the reported result (Al Janabi et al., 2020) because the reported effect estimate is likely to be selected, on the basis of the results, from either multiple outcome measurements within the outcome or multiple analyses of the intervention-outcome relationship.

3.4. Results of individual studies

The 12 articles selected for this study were analysed critically by using the PICOS framework to explore their research methods, sample size, research design formula, and complete results, as shown in Tables 5 and 6.

4. Discussion

This systematic literature review involved an analysis of the 12 articles included in the final comprehensive review. All the included literature outlined the use of MR in different contexts and fields of higher education. The literature search results indicate numerous articles on MR, VR, and AR in higher educational settings published in recent years. This is clear evidence that the research topic attracts scholars who are interested in examining the impacts and contributions of MR on students' learning experiences, engagement, knowledge, and

[2] https://www.riskofbias.info/welcome/rob-2-0-tool/current-version-of-rob-2.

[3] https://www.riskofbias.info/welcome/home/current-version-of-robins-i.

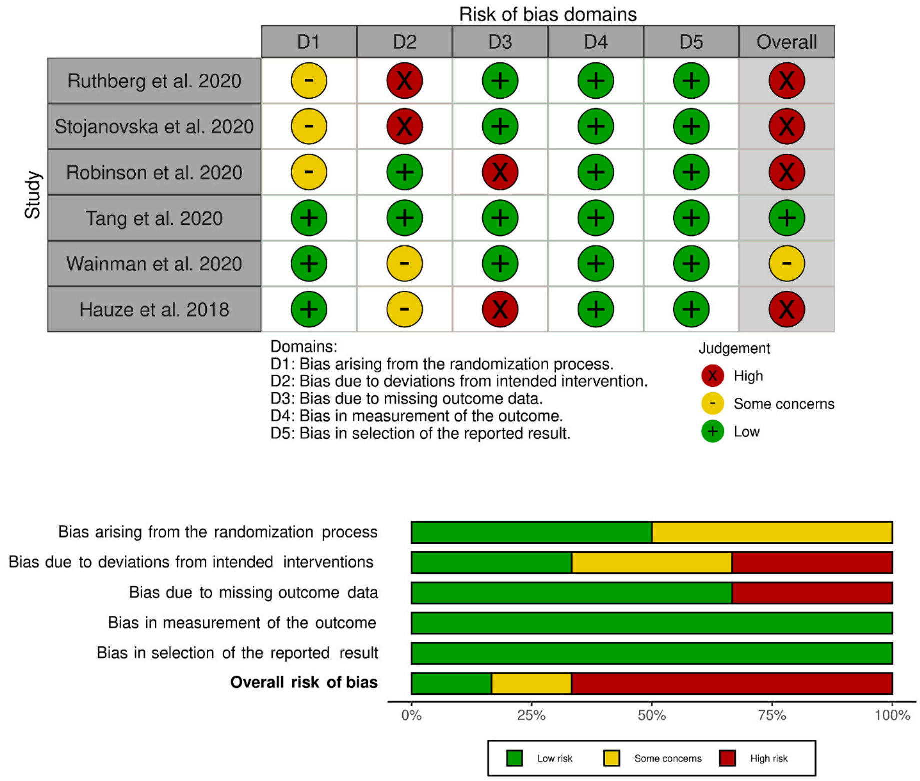

Fig. 4. Risk-of-bias analysis for RCTs.

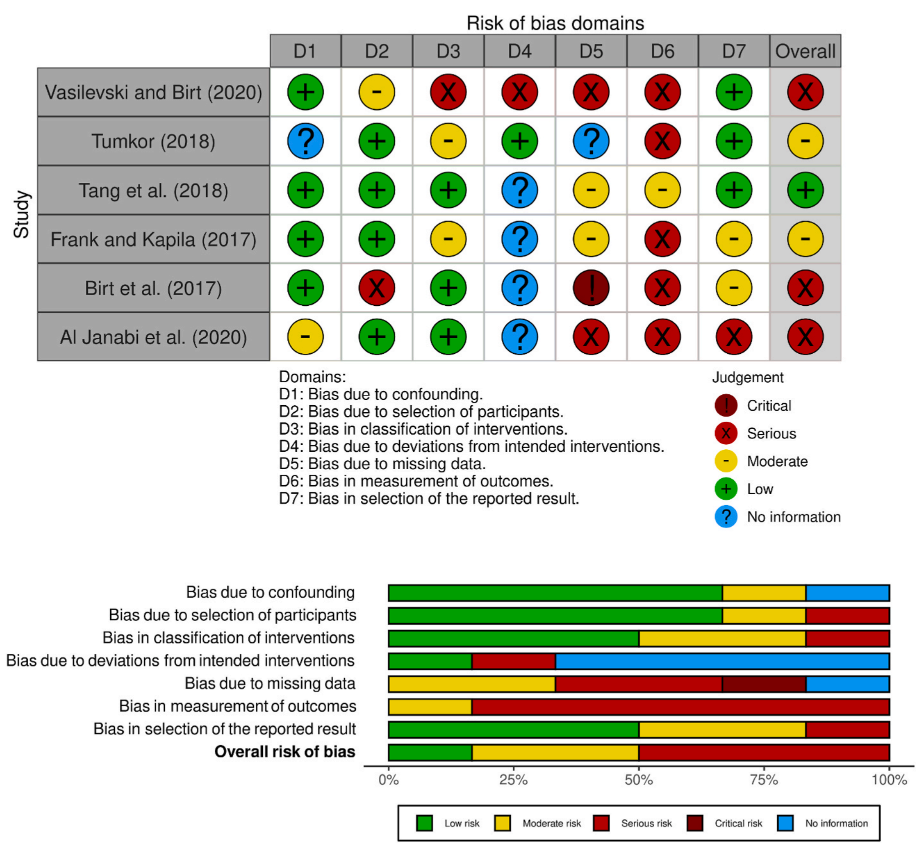

Fig. 5. Risk-of-bias analysis for NRSs.

Table 5
Characteristics of included studies.

No.	Author	Population	Intervention	Comparison	Outcomes	Study Design
1	Wainman et al. (2020)	n = 140 undergraduate engineering (n = 22), health science/science (n = 95), humanities/social science (n = 18) and other (n = 5)	MR, MR (AR HMDs), HoloLens	Physical model, AV (VR HMDs), HTC VIVE	VR and MR technologies tested are inferior to physical models, and true stereopsis is critical in learning anatomy.	RCT QN [c]
2	Vasilevski and Birt (2020)	n = 70 Postgraduate STEM Construction students	MMR (AR + VR), AR (Smartphones and tablets), Samsung Galaxy S8, S7, S6, and Samsung Gear VR headsets		MMR can result in an enhanced learning environment that provides unique learning experiences and engagement for the students throughout the learning process. Some of the core features of this delivery technique are found in the enhanced learning aspects, such as improved learning engagement and motivation, improved interaction, and increased fun and enjoyment.	NRS [b] QL [d]
3	Tumkor (2018)	n = 217 Undergraduate STEM Engineering education	MMR, AR (Smartphones and tablets), MR (AR HMDs), Google Cardboard, Epson Moverio BT200, and HoloLens		MR tools can be effectively implemented in an engineering drawing course if personalisation of large classes is needed.	NRS QN Pre-Exp [f]
4	Tang et al. (2018)	n = 45, Undergraduate and postgraduate, Design (38%), engineering (31%), business (13%), health & social sciences (9%), construction and environment (6%), and humanities (3%)	MR, MR (AR HMDs), HoloLens		Experimental results show that the proposed design positively improves students' understanding of geometric relationships and creativity.	NRS MM [e] Pre-Exp
5	Tang et al. (2020)	n = 72 (E = 44, C = 28), including, undergraduate and postgraduate, engineering (27%), design (25%)	MR, MR (AR HMDs), HoloLens	Traditional learning material	The results were positive when using MR to support their study. MR was also better than using traditional teaching notes for various measured effects.	RCT QN
6	Stojanovska et al. (2019)	n = 64 (E = 15, E = 16, C = 33), Undergraduate, 2nd-year medical students, Medical and health science, Anatomy education	MR, MR (AR HMDs), HoloLens	Cadaver	Medical students, regardless of the study modality, performed similarly on the MR and the cadaver practical exams.	RCT[a] QN

g Quasi-Exp: Quasi-Experiment.
[a] RCT: Randomised Control Trial.
[b] NRS: Non-Randomised Study.
[c] QN: Quantitative Research.
[d] QL: Qualitative Research.
[e] MM: Mixed Method.
[f] Pre-Exp: Pre-Experiment.

skills (Kounlaxay & Kim, 2020; Krach & Hanline, 2018; Wu et al., 2018).

4.1. General interpretation of the results

To interpret these results, all 12 of the included articles have been critically analysed by addressing the following research questions:

4.1.1. Population

1. **What are the characteristics of the participants using MR technology in higher education?**

The included studies involved a diverse range of participants from various higher education settings, including undergraduate and postgraduate students, but most of the studies focused on undergraduate students across multiple disciplines such as engineering, medicine, the social sciences, and other domains. This diversity highlights the potential applicability of MR across student populations and disciplines. However, students of all ages have varying technology expectations, learning needs, human interaction abilities, and attitudes, all of which change significantly as they develop (Pellas et al., 2020).

2. **Which higher education domains use MR to examine MR effectiveness?**

MR technologies have been applied to various higher education domains to examine their effectiveness. Most of the articles evaluated in the literature review are in the fields of medical and health science, especially anatomy education, followed by STEM education, especially engineering education. Therefore, these domains seem to be more suited to the use of MR for educational purposes than the humanities and social sciences, which are more theoretical. At the same time, all other academic disciplines have also investigated the use of MR, as illustrated in Fig. 6. Moreover, many studies have found that the use of MR environments can support a variety of learning subjects, as previous research has demonstrated improvements in students' subject comprehension, participation, and motivation when using MR (Pellas et al., 2020). The versatility of MR allows for its use in both theoretical and practical learning experiences, which can be adapted to the specific needs and requirements of diverse educational contexts. For instance, a large body of recent research has reported the need to use learning theories or theoretical foundations such as constructionism to inform teaching methodologies for any interactive learning experience that can be provided within MR environments (Pellas et al., 2020).

4.1.2. Intervention

3. **What MR characteristics and technologies are applied to higher education?**

The included studies employed a range of MR technologies, and the majority of the selected research on MR is related to AR HMDs such as the Microsoft HoloLens and Windows MR rather than MMR that uses mobile AR and mobile VR, as illustrated in Fig. 7.These technologies

Table 6
Characteristics of included studies.

No.	Author	Population	Intervention	Comparison	Outcomes	Study Design
7	Ruthberg et al. (2020)	n = 48, Undergraduate 2 nd year Medical and health science Anatomy education	MR, MR (AR HMDs), HoloLens	Cadaver dissection	Using HoloAnatomy may decrease the time necessary for anatomy instruction without sacrificing student understanding of the material.	RCT[a] QN[c]
8	Robinson et al. (2020)	n = 10 (E = 5, C = 5), Undergraduate 1st-year medical students Medical and health science Anatomy education	MR, MR (AR HMDs), HoloLens 1st Gen	Glass histology slides using a light microscope	MR-based education is both feasible and effective. The results of this study show that MR is capable of developing students who are knowledgeable and confident, while also enabling positive learning experiences for those involved.	RCT MM[d]
9	Hauze et al. (2018)	n = 161 (E = 54, E = 53, C = 54), Lvl1 (n = 65), Lvl2 (n = 60), Lvl3 (n = 14), and Lvl4 (n = 22) Medical and health science Baccalaureate nursing	MR, MR (AR HMDs), HoloLens	2D Video and Written Case study	Mixed reality could be used as a means of providing simulation to nursing students through devices that are significantly more affordable and accessible than traditional nursing simulation devices. The Instructional Materials Motivation Survey is a valid research instrument for measuring nursing student motivation to learn within the context of immersive simulation.	RCT QN
10	Frank and Kapila (2017)	n = 75 Undergraduate STEM Science and engineering education	MMR AR (Smartphones and tablets) Apple iPad 2	No simulation	Student participants demonstrate significant improvement in content knowledge and report having significantly beneficial experiences after using the MRLE platform compared to before using the MRLE platform and compared to student participants exposed to the content using traditional classroom and hands-on laboratory techniques.	NRS[b] MM[e] Pre-Exp[f]
11	Birt et al. (2017)	n = 137 Undergraduate 2 nd year Medical and health science Paramedic students	MMR, AR (Smartphones and tablets), ColorCross headset or head mount		A statistically significant improvement in performance for students who received the tools prior to residential school, both across the skill set and within individual skills, indicates the potential for students to develop automatic skills more rapidly, as evidenced by improved learning outcomes and student acceptance of this use of multiple forms of media.	NRS MM DBR Quasi-Exp[g]
12	Al Janabi et al. (2020)	n = 72, Medical students, urological trainees, or specialists Medical and health science Surgical education	MR, MR (AR HMDs), HoloLens	Conventional monitor	The device facilitated improved outcomes of performance in novices and was widely accepted as a surgical visual aid by all groups. The HoloLens represents a feasible alternative to the conventional setup, possibly by aligning the surgeons' visual–motor axis.	NRS MM Quasi-Exp

[a] RCT: Randomised Control Trial.
[b] NRS: Non-Randomised Study.
[c] QN: Quantitative Research.
[d] QL: Qualitative Research.
[e] MM: Mixed Method.
[f] Pre-Exp: Pre-Experiment.
[g] Quasi-Exp: Quasi-Experiment.

were used to create interactive, collaborative, and immersive learning environments that facilitated different categories of learning, such as experiential learning, game-based learning, constructive learning, generative learning, contextual learning, and operational learning (Radianti et al., 2020). Some studies that use mobile AR (smartphones and tablets) for MR in education use the term "mixed reality" to evaluate the use of AR technologies combined with physical equipment or real objects, such as medical devices or 3D printed ones, which are more commonly referred to as AR in the XR sector. Another use of MR is mixing realities, such as using both mobile AR and mobile VR and take advantage of the power of VR and AR technology to provide immersive and dynamic applications in a variety of learning settings, both in and out of classes (e.g., see Vasilevski and Birt (2020)). The use of MR technology, which combines a number of computing devices, makes this possible. Most of the applications made for MR settings that could give information about a specific item or place by using visual markers were used in museums, art shows, and chemistry field studies (Pellas et al., 2020).

4.1.3. *Comparison*

4. **What are the characteristics of educational materials that were compared to MR used in higher education?**

The included studies compared MR technologies to various traditional teaching methods, such as physical models and cadavers, glass histology, 2D videos, written case studies, and conventional monitors; other technology-enhanced learning approaches, including e-learning modules and computer simulations; and other types of XR technologies such as VR. The choice of comparison groups varied depending on the learning objectives and the specific educational context. These comparisons allowed for a better understanding of the added value of MR technologies in enhancing learning outcomes and student engagement. Thus, active learning could replace the lecture format and more traditional teaching pedagogy in such learning approaches. MR environments must provide immediate feedback on the users' actions in order to have a positive effect on students' learning performance, in addition to providing the opportunity to learn through interactive activities with

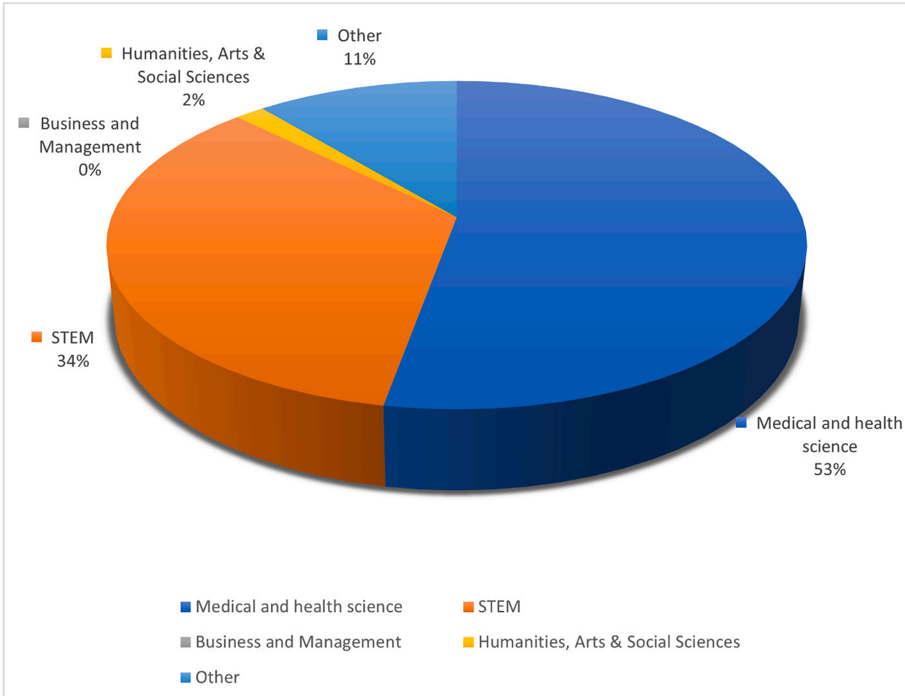

Fig. 6. Distribution of students within a university's domains using MR.

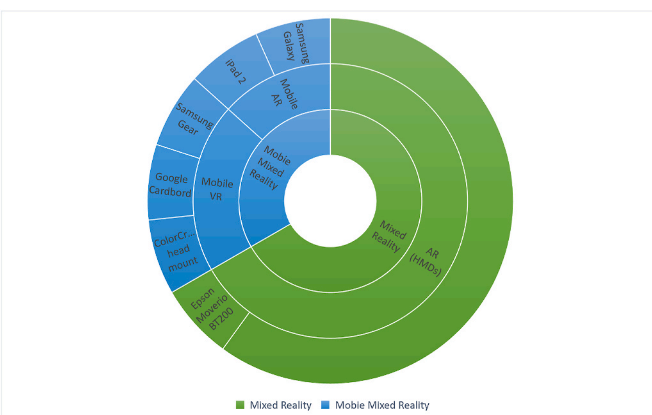

Fig. 7. Distribution of MR technologies used.

defined objectives (Pellas et al., 2020).

4.1.4. Outcomes

5. **What are the learning outcomes of students who have experienced MR technology?**

The majority of the included studies reported improvements in learning outcomes for students who used MR technologies in education in all different fields. Some research has argued that the tested VR and MR technologies are inferior to physical models, and true stereopsis is critical to learning anatomy. Also, the medical students performed similarly on the MR and the cadaver practical exam. However, other researchers have emphasised the benefits of MR in decreasing the time needed for anatomy. MR is feasible, effective, and capable of developing knowledgeable and confident students while also enabling positive learning experiences for those involved. According to Pellas et al. (2020), "MR allows these difficult conceptions to be taught so that students can try to solve complex problems by providing information related to a learning subject from the real world with virtual information". At the same time, researchers emphasise the value of traditional materials, which cannot be replaced. Using new technologies like MR is beneficial, feasible, effective, affordable, and accessible. It enhances learning environments and experiences, increases learning engagement and motivation, and improves student performance and learning outcomes.

4.1.5. Study design

6. **What characteristics of the study design of the MR studies are applied to higher education?**

The study design of the included studies is limited to experimental studies that are randomised control studies (RCTs), which provided the strongest evidence for the effectiveness of mixed reality technologies, and non-randomised controlled studies (NRSs), which offered moderate experimental or quasi-experimental designs as well. The diversity in study designs highlights the need for further research employing more robust and standardised methodologies to validate the effectiveness of MR technologies in higher education. Regarding study design, Pellas et al. (2020) claimed that few studies have followed quasi-experimental (pre- and post-tests) research method designs to compare any potential learning gains or an improvement on science concepts knowledge; therefore, it has been difficult for educators and researchers to comprehend how an MR environment and its technological equipment can be utilised to contribute to any learning subject.

4.2. Research limitations

This review has several limitations. As for any systematic review, the main limitations are the scope of the search terms and the database searched. The search was also restricted to papers published in the English language (Stretton et al., 2018). An additional limitation is the choice to focus on journals to represent the MR literature, which means that the review has excluded the knowledge of other researchers who have done their best to explore the aspects of MR, while there are other channels that may represent MR knowledge, such as books and chapters. Although conference papers were included in the eligibility criteria, none of the included studies are conference papers.

The review has excluded the meta-analysis items provided in the PRISMA 2020 checklist. Only 12 articles were analysed for the systematic review. It would not be wrong to say that this study is completely based upon these 12 articles. However, all of these articles have been critically explained in detail and depth, and as a result, they are regarded as sufficient for an investigation of perspectives on MR in education.

4.3. Implications for practice and policy

The concentration of MR studies in the medical, health sciences, and engineering disciplines suggests that practical, visualisation-heavy subjects gain the most immediate benefits from MR technologies. Thus, institutions offering these courses should prioritise the integration of MR technologies into their curriculum, particularly for modules requiring 3D visualisation and manipulation. Despite the current focus on practical disciplines, MR technologies could also be beneficial for theoretical disciplines like humanities and social sciences. Future practice should explore innovative ways to integrate MR into these disciplines, such as virtual tours of historical sites or interactive 3D models of social systems.

In terms of infrastructure, given that most studies used AR HMDs and MMR, institutions should invest in these technologies as a priority. The choice between HMDs and MMR would depend on the specific learning requirements and budget constraints.

Based on the findings of this review, the following are the implications and policy recommendations for educators and policymakers:

1. **Course Design and Delivery:** MR offers a new paradigm for designing and delivering courses. It enables interactive, immersive experiences that can enhance learners' understanding and retention of complex concepts. Policy recommendations include training faculty in MR technology and instructional design, funding MR hardware and software, and including MR in curriculum development guidelines.
2. **Accessibility and Equity:** While MR has the potential to enrich learning, it also may widen the digital divide if not implemented with consideration for equity. Policies should ensure that all students can access the necessary hardware and internet connectivity to participate fully in MR activities.
3. **Research and Development:** To stay at the forefront of educational technology, institutions should invest in research and development of MR applications tailored to their specific needs. Policies should support such research, for instance by providing funding, fostering collaborations between academics and industry, and facilitating the testing and adoption of new applications.
4. **Digital Literacy:** Proficiency in using MR technology should be recognised as an important digital literacy skill for the 21st century. Policies should integrate MR into digital literacy initiatives and ensure that students are not simply passive consumers of MR experiences but also have opportunities to create their own.
5. **Professional Development:** For MR to be effectively integrated into higher education, faculty and staff need to be comfortable and proficient with the technology. Institutions should provide ongoing professional development opportunities related to MR, and policies should encourage or require participation in such training.

The potential of MR in higher education is significant, but realising this potential requires thoughtful, proactive policy-making. Implementing these recommendations can help to ensure that MR is used in ways that are pedagogically sound, equitable, and effective.

4.4. Future research

In light of this systematic review of mixed reality in higher education, several important directions for future research have been identified. It is evident that further studies are needed to incorporate a meta-analysis approach in line with the PRISMA 2020 guidelines, as not all items in the current review have been addressed. The meta-synthesis would provide a comprehensive and overarching understanding of the impact and implications of MR technologies in higher education, thereby guiding future policy, practice, and research directions. From the intervention side, a further review study is needed that includes studies utilising VR HMDs for MR applications in higher education, such as Oculus Quest 2 and Quest Pro. Moreover, the time frame of the review must be expanded to include studies conducted from 2021 to the present day to encompass the most recent developments in the field. Alongside this, an important research endeavour would be to conduct an updated systematic review including searches of databases, registers, and other sources by following the updated guidelines for reporting systematic reviews for PRISMA 2020 (Page et al., 2021a). To reach a deeper understanding of the impact and implications of MR technologies in various domains, it is crucial to undertake a systematic review of all existing systematic reviews in the field of mixed reality for future research.

5. Conclusion

Based on our findings, medical and health sciences, as well as STEM, appear to be more adaptable to the use of MR for educational purposes than theoretical disciplines such as humanities and social sciences. This is particularly true in anatomy and engineering education. However, since most MR studies have concentrated on 3D manipulation and visualisation and understanding of the 3D object layers and components, these studies have only shown immediate short-term learning improvements. Still, there is a lack of longitudinal studies to show the impact of the use of MR in higher education. Even though MR technologies have been tested in surgical practice to enhance procedural or surgical techniques, they have not been commonly employed in education for procedural learning or integrated with instructional design except to assess the potential of MR in the field.

This focus on medicine, health sciences, and STEM may, however, be a reflection of the current state of MR technology, and as the technology develops, its use cases may expand. Since this systematic review only covered 2017 to 2021, future researchers should see whether these findings are replicated during follow-up systematic reviews in the decades that follow. As new tools and frameworks emerge that support users with less technical backgrounds in computer-related disciplines, then MR use in the humanities and social science may grow. It is our opinion that MR technology, like mobile technologies, will become more prevalent in all academic disciplines as well as daily life.

Finally, it is important to note that as seen in the papers included and rejected from this review, the most well-documented HMDs used during this period have been the Microsoft HoloLens 1 and 2. Their introduction has allowed for hundreds of MR studies worldwide by providing a common instrument for implementing and evaluating MR technology. This systematic review using the PRISMA 2020 approach so we ultimately only included a limited number of exemplar studies and the majority of these studies used the HoloLens. This is an accurate snapshot of the current field in our opinion and is an important finding to note.

The HoloLens has gained popularity due to the variety of materials available on the internet for developers. These materials, along with popular development tools such as Unity and Vuforia, have also allowed development to become increasingly accessible to a wider audience due to their rich development communities online. This success can be replicated for future MR devices for which the hardware manufacturers engage and help build development communities as Microsoft has done.

Studies conducted on MR in education find positive outcomes for students and learning environments. However, we should bear in mind the overall weight risk of bias in the selected studies. This systemic review aims to be a snapshot of the current state of the art in the field of education and help build a foundation for future research in this field.

Funding

This work was supported and funded by the Ministry of Education in Saudi Arabia and the China Scholarship Council under grant number (201908300021).

Statements on open data and ethics

The data used in this systematic review were obtained from publicly available open data sources and repositories. All included studies have been appropriately cited, and the data sources are listed in the references section. This systematic review is based on previously published studies, and no primary data collection was conducted. All included studies were assessed for compliance with relevant ethical guidelines, and only studies that reported ethical approval and informed consent were considered for inclusion.

Declaration of competing interest

The authors declare that they have no competing interests that could have influenced the research or the presented results.

Acronyms

AR	Augmented Reality
HMD	Head-Mounted Display
MR	Mixed Reality
NRS	Non-Randomised Study
PICOS	Population, Intervention, Comparison, Outcomes and Study Design
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RCT	Randomised Control Trial
VR	Virtual Reality
XR	eXtended Reality

References

Akçayır, M., & Akçayır, G. (2017). Advantages and challenges associated with augmented reality for education: A systematic review of the literature. *Educational Research Review, 20*, 1–11.

Al Janabi, H. F., Aydin, A., Palaneer, S., Macchione, N., Al-Jabir, A., Khan, M. S., Dasgupta, P., & Ahmed, K. (2020). Effectiveness of the hololens mixed-reality headset in minimally invasive surgery: A simulation-based feasibility study. *Surgical Endoscopy, 34*, 1143–1149.

Banjar, A., & Campbell, A. G. (2022). Work-in-progress—design of mixed reality educational applications using enhanced arcs+ g model for motivation. In *2022 8th international conference of the immersive learning research network (iLRN)* (pp. 1–3). IEEE.

Barrie, M., Socha, J. J., Mansour, L., & Patterson, E. S. (2019). Mixed reality in medical education: A narrative literature review. In *Proceedings of the international symposium on human factors and ergonomics in health care* (Vol. 8, pp. 28–32). Los Angeles, CA: SAGE Publications Sage CA.

Bettany-Saltikov, J., & McSherry, R. (2018). How to do a systematic review. *International Journal of Stroke, 13*, 138–156.

Birt, J., Moore, E., & Cowling, M. (2017). Improving paramedic distance education through mobile mixed reality simulation. *Australasian Journal of Educational Technology, 33*.

Birt, J., Stromberga, Z., Cowling, M., & Moro, C. (2018). Mobile mixed reality for experiential learning and simulation in medical and health sciences education. *Information, 9*, 31.

Booth, A. (2006). Clear and present questions: Formulating questions for evidence based practice. *Library Hi Tech, 24*, 355–368.

Burke, E., Felle, P., Crowley, C., Jones, J., Mangina, E., & Campbell, A. G. (2017). Augmented reality evar training in mixed reality educational space. In *2017 IEEE global engineering education conference (EDUCON)* (pp. 1571–1579). IEEE.

Chen, J. A., Tutwiler, M. S., & Jackson, J. F. (2020). Mixed-reality simulations to build capacity for advocating for diversity, equity, and inclusion in the geosciences. *Journal of Diversity in Higher Education*.

Cooke, A., Smith, D., & Booth, A. (2012). Beyond pico: The spider tool for qualitative evidence synthesis. *Qualitative Health Research, 22*, 1435–1443.

Covidence. (2023). *Better systematic review management*. URL: https://www.covidence.org/.

Division, S. R. (2020). *What are the pico elements in systematic review?*. URL: https://academy.pubrica.com/research-publication/systematic-review/what-are-the-pico-elements-in-systematic-review/.

Eriksen, M. B., & Frandsen, T. F. (2018). The impact of patient, intervention, comparison, outcome (pico) as a search strategy tool on literature search quality: A systematic review. *Journal of the Medical Library Association: JMLA, 106*, 420.

Frank, J. A., & Kapila, V. (2017). Mixed-reality learning environments: Integrating mobile interfaces with laboratory test-beds. *Computers & Education, 110*, 88–104.

Gerup, J., Soerensen, C. B., & Dieckmann, P. (2020). Augmented reality and mixed reality for healthcare education beyond surgery: An integrative review. *International Journal of Medical Education, 11*, 1.

Gopalakrishnan, S., & Ganeshkumar, P. (2013). Systematic reviews and meta-analysis: Understanding the best evidence in primary healthcare. *Journal of Family Medicine and Primary Care, 2*, 9.

Hauze, S., Hoyt, H., Marshall, J., Frazee, J., & Greiner, P. (2018). An evaluation of nursing student motivation to learn through holographic mixed reality simulation. In *2018 IEEE international conference on teaching, assessment, and learning for engineering (TALE)* (pp. 1058–1063). IEEE.

Hauze, S., & Marshall, J. (2020). Validation of the instructional materials motivation survey: Measuring student motivation to learn via mixed reality nursing education simulation. *International Journal on E-Learning, 19*, 49–64.

Higgins, J. P., & Green, S. (2006). *Cochrane handbook for systematic reviews of interventions 4.2. 6 [updated september 2006]* (Vol. 4). The cochrane library, 2006.

Johnson-Glenberg, M. C. (2018). Immersive vr and education: Embodied design principles that include gesture and hand controls. *Frontiers in Robotics and AI, 5*, 81.

Knierim, P., Kosch, T., Hoppe, M., & Schmidt, A. (2018). *Challenges and opportunities of mixed reality systems in education*.

Kounlaxay, K., & Kim, S. K. (2020). Design of learning media in mixed reality for Lao education. *CMC-COMPUTERS MATERIALS & CONTINUA, 64*, 161–180.

Krach, S. K., & Hanline, M. F. (2018). Teaching consultation skills using interdepartmental collaboration and supervision with a mixed-reality simulator. *Journal of Educational and Psychological Consultation, 28*, 190–218.

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P., Clarke, M., Devereaux, P. J., Kleijnen, J., & Moher, D. (2009). The prisma statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration. *Journal of Clinical Epidemiology, 62*, e1–e34.

Li, K. C., & Wong, B. T.-M. (2021). A literature review of augmented reality, virtual reality, and mixed reality in language learning. *International Journal of Mobile Learning and Organisation, 15*, 164–178.

Maas, M. J., & Hughes, J. M. (2020). Virtual, augmented and mixed reality in k–12 education: A review of the literature. *Technology, Pedagogy and Education, 29*, 231–249.

Marcel, F. (2019). Mobile augmented reality learning objects in higher education. *Research in Learning Technology, 27*.

McMillan, K., Flood, K., & Glaeser, R. (2017). Virtual reality, augmented reality, mixed reality, and the marine conservation movement. *Aquatic Conservation: Marine and Freshwater Ecosystems, 27*, 162–168.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021a). The prisma 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*. URL: https://www.bmj.com/content/372/bmj.n71. doi:10.1136/bmj.n71. arXiv https://www.bmj.com/content/372/bmj.n71.full.pdf.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., & Moher, D. (2021b). Updating guidance for reporting systematic reviews: Development of the prisma 2020 statement. *Journal of Clinical Epidemiology, 134*, 103–112.

Page, M. J., Moher, D., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … McKenzie, J. E. (2021c). Prisma 2020 explanation and elaboration: Updated guidance and exemplars for reporting systematic reviews. *BMJ, 372*. URL: https://www.bmj.com/content/372/bmj.n160. doi:10.1136/bmj.n160. arXiv https://www.bmj.com/content/372/bmj.n160.full.pdf.

Pellas, N., Kazanidis, I., & Palaigeorgiou, G. (2020). A systematic literature review of mixed reality environments in k-12 education. *Education and Information Technologies, 25*, 2481–2520.

Pollock, A., & Berge, E. (2018). How to do a systematic review. *International Journal of Stroke, 13*, 138–156.

Radianti, J., Majchrzak, T. A., Fromm, J., & Wohlgenannt, I. (2020). A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. *Computers & Education, 147*, Article 103778.

Robinson, B. L., Mitchell, T. R., & Brenseke, B. M. (2020). Evaluating the use of mixed reality to teach gross and microscopic respiratory anatomy. *Medical Science Educator, 30*, 1745–1748.

Ruthberg, J. S., Tingle, G., Tan, L., Ulrey, L., Simonson-Shick, S., Enterline, R., Eastman, H., Mlakar, J., Gotschall, R., Henninger, E., et al. (2020). Mixed reality as a time-efficient alternative to cadaveric dissection. *Medical Teacher, 42*, 896–901.

Sterne, J. A., Hernán, M. A., Reeves, B. C., Savović, J., Berkman, N. D., Viswanathan, M., Henry, D., Altman, D. G., Ansari, M. T., Boutron, I., et al. (2016). Robins-i: A tool for assessing risk of bias in non-randomised studies of interventions. *Bmj*, 355.

Sterne, J. A., Savović, J., Page, M. J., Elbers, R. G., Blencowe, N. S., Boutron, I., Cates, C. J., Cheng, H.-Y., Corbett, M. S., Eldridge, S. M., et al. (2019). Rob 2: A revised tool for assessing risk of bias in randomised trials. *Bmj*, 366.

Stojanovska, M., Tingle, G., Tan, L., Ulrey, L., Simonson-Shick, S., Mlakar, J., Eastman, H., Gotschall, R., Boscia, A., Enterline, R., et al. (2019). Mixed reality anatomy using microsoft hololens and cadaveric dissection: A comparative effectiveness study. *Medical Science Educator*, 1–6.

Stone, P. W. (2002). Popping the (pico) question in research and evidence-based practice. *Applied Nursing Research: ANR, 15*, 197–198.

Stretton, T., Cochrane, T., & Narayan, V. (2018). Exploring mobile mixed reality in healthcare higher education: A systematic review. *Research in Learning Technology, 26*, 2131, 2131.

Tang, Y.-M., Au, K. M., Lau, H. C., Ho, G. T., & Wu, C.-H. (2020). Evaluating the effectiveness of learning design with mixed reality (mr) in higher education. *Virtual Reality, 24*, 797–807.

Tang, Y., Au, K., & Leung, Y. (2018). Comprehending products with mixed reality: Geometric relationships and creativity. *International Journal of Engineering Business Management, 10*, Article 1847979018809599.

Tumkor, S. (2018). Personalization of engineering education with the mixed reality mobile applications. *Computer Applications in Engineering Education, 26*, 1734–1741.

UTAS. (2023). *Subject guides: Systematic reviews for health: Tools for systematic review*. URL: https://utas.libguides.com/SystematicReviews/Tools.

Vasilevski, N., & Birt, J. (2020). Analysing construction student experiences of mobile mixed reality enhanced learning in virtual and augmented reality environments. *Research in Learning Technology, 28*.

Wainman, B., Pukas, G., Wolak, L., Mohanraj, S., Lamb, J., & Norman, G. R. (2020). The critical role of stereopsis in virtual and mixed reality learning environments. *Anatomical Sciences Education, 13*, 401–412.

Wendler, R. (2012). The maturity of maturity model research: A systematic mapping study. *Information and Software Technology, 54*, 1317–1339.

Wu, W., Tesei, A., Ayer, S., London, J., Luo, Y., & Gunji, V. (2018). Closing the skills gap: Construction and engineering education using mixed reality–a case study. In *2018 IEEE frontiers in education conference (FIE)* (pp. 1–5). IEEE.

Xu, X., Mangina, E., & Campbell, A. (2021). Hmd-based virtual and augmented reality in medical education: A systematic review. Front. *Virtual Reality, 2*, 692103. https://doi.org/10.3389/frvir

Zhang, X., Tlili, A., Nascimbeni, F., Burgos, D., Huang, R., Chang, T.-W., Jemni, M., & Khribi, M. K. (2020). Accessibility within open educational resources and practices for disabled learners: A systematic literature review. *Smart Learning Environments, 7*, 1–19.