



Immersive virtual reality and augmented reality in anatomy education: A systematic review and meta-analysis

Paloma García-Robles^{1,2} | Irene Cortés-Pérez²  | Francisco Antonio Nieto-Escámez^{3,4}  |
Héctor García-López⁵  | Esteban Obrero-Gaitán²  | María Catalina Osuna-Pérez² 

¹FRATERNIDAD Muprespa, Linares, Spain

²Department of Health Sciences, University of Jaén, Jaén, Spain

³Department of Psychology, University of Almería, Almería, Spain

⁴Center for Neuropsychological Assessment and Rehabilitation (CERNEP), Almería, Spain

⁵Department of Nursing, Physiotherapy and Medicine, University of Almería, Almería, Spain

Correspondence

Esteban Obrero-Gaitán, Department of Health Sciences, University of Jaén, Campus Las Lagunillas s/n, Jaén 23071, Spain.

Email: eobrero@ujaen.es

Abstract

The purpose of this review was to (1) analyze the effectiveness of immersive virtual reality (iVR) and augmented reality (AR) as teaching/learning resources (collectively called XR-technologies) for gaining anatomy knowledge compared to traditional approaches and (2) gauge students' perceptions of the usefulness of these technologies as learning tools. This meta-analysis, previously registered in PROSPERO (CRD42023423017), followed PRISMA guidelines. A systematic bibliographical search, without time parameters, was conducted through four databases until June 2023. A meta-analytic approach investigated knowledge gains and XR's usefulness for learning. Pooled effect sizes were estimated using Cohen's standardized mean difference (SMD) and 95% confidence intervals (95% CI). A single-group proportional meta-analysis was conducted to quantify the percentage of students who considered XR devices useful for their learning. Twenty-seven experimental studies, reporting data from 2199 health sciences students, were included for analysis. XR-technologies yielded higher knowledge gains than traditional approaches (SMD=0.40; 95% CI=0.22 to 0.60), especially when used as supplemental/complementary learning resources (SMD=0.52; 95% CI=0.40 to 0.63). Specifically, knowledge performance using XR devices outperformed textbooks and atlases (SMD=0.32; 95% CI=0.10 to 0.54) and didactic lectures (SMD=1.00; 95% CI=0.57 to 1.42), especially among undergraduate students (SMD=0.41; 95% CI=0.20 to 0.62). XR devices were perceived to be more useful for learning than traditional approaches (SMD=0.54; 95% CI=0.04 to 1), and 80% of all students who used XR devices reported these devices as useful for learning anatomy. Learners using XR technologies demonstrated increased anatomy knowledge gains and considered these technologies useful for learning anatomy.

KEYWORDS

anatomy, augmented reality, health sciences courses, immersive virtual reality, meta-analysis, technology-enhanced learning

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2024 The Authors. *Anatomical Sciences Education* published by Wiley Periodicals LLC on behalf of American Association for Anatomy.

INTRODUCTION

Health sciences students must acquire a wide variety of skills and knowledge in their university studies and/or residency programs to become competent healthcare professionals. In particular, learning human anatomy is essential for health professions training^{1,2} and has traditionally been a foundational starting point for many health professions courses. Classic anatomy pedagogies are rooted primarily in donor dissections and plastic models, as well as displaying two-dimensional images in didactic lectures, textbooks, and anatomy atlases.^{3,4} With the advent of extended reality (XR) technologies and the recent plethora of investigations on their comparative efficacy, it is timely to consider how the use of XR technologies within anatomy curricula is affecting learners' acquisition of anatomical knowledge. The expanding popularity of XR technologies, partly motivated by the COVID-19 pandemic,⁵ necessitates that educators be updated on their usefulness and efficacy as tools for teaching and learning in visual disciplines.

Learning limitations and challenges of classic educational approaches

In isolation, classic educational approaches have certain limitations, such as reduced recall from passive learning (e.g., didactic lectures),^{6,7} which could foster student disinterest and a disconnection with the subject matter. For example, after receiving traditional didactic instruction, students often only reproduce 20%–30% of the learned content immediately following a session, and 2 weeks later it is estimated that learners have forgotten 90% of the information.⁸ For anatomy specifically, the three-dimensional organization of anatomical structures makes it difficult to understand their spatial relationships, which depends on students' abilities to mentally visualize, rotate, and transform flat images into three-dimensional constructs, thus, making anatomy learning a challenging cognitive process.⁹ The limitations of classic educational approaches combined with content complexity could stifle students' motivation for learning. Self-determined motivation is intrinsically linked with academic performance whereby it drives higher study effort, deeper learning, and better study strategies.^{10,11} Educational approaches that promote active participation and interactions among students and between students and educators are also thought to be more effective in knowledge acquisition.^{12,13} XR technologies, which offer hardware and software with interactive and game-like components, have been known to increase students' motivation and adherence to learning across subjects, promote knowledge acquisition, and help learners navigate difficult concepts, such as human anatomy.¹⁴ As such, strategically exploring the efficacy of XR technologies across studies and time, through meta-analysis, is necessary for summarizing the educational potential of these devices among health sciences students.

Immersive virtual reality versus augmented reality

In the last two decades, technological advancements have fostered the development of many digital tools for education, which in turn have popularized teaching through digital technologies.¹⁵ Mobility restrictions imposed by the COVID-19 pandemic, that broke out in December 2019 in Wuhan (Hubei, China),¹⁶ also accelerated the use of distance education technologies and related online resources.¹⁷ For educators, learning digital skills and capitalizing on digital solutions have become practical mechanisms for enhancing students' motivation, maintaining their attention, and increasing their knowledge base.¹⁸

Currently, immersive virtual reality (iVR) and augmented reality (AR) are widely employed to teach anatomy across health sciences courses. These interactive XR technologies (iVR and AR) permit the three-dimensional visualization of anatomical structures and their spatial relationships in virtual environments that resemble the real world.¹⁹ Virtual structures rendered through specialized software are similar to real structures; though, in the virtual environment, students interact with the anatomy using a joystick or through hand motions.²⁰ VR headsets enable the recreation of immersive experiences (including motion tracking, interaction with digital elements, a wider range of vision, and stereoscopy²¹), which allows students to visualize virtual anatomical objects as if they were real.²² Two characteristics of iVR are presence and immersion.²⁰ Presence is defined as the psychological sensation/desire to stay within the virtual environment that the subject feels is real. Immersion is the capability to interact with virtual objects.²³ Therefore, depending on the level of presence and immersion, VR can be either immersive or non-immersive.²⁴ iVR allows viewing a virtual scenario through a head-mounted display in 360 degrees whereby subjects can interact with the virtual environment and its objects using their hands as controllers.²⁵ Conversely, AR superimposes virtual elements onto the real world allowing individuals to interact with real and virtual elements, simultaneously.²⁶

Study purpose

Between 2020 and 2022, eight reviews have evaluated the effects of VR and AR interventions on performance outcomes in health sciences education.^{1,2,4,8,26-29} Six of these reviews were qualitative, non-meta-analytic, reviews. Three of these six thematic reviews assessed the effectiveness of XR interventions on anatomy education with data from 25,⁸ 20,² and 12 studies,¹ respectively. In the remaining three reviews, anatomy education was not the primary focus as very few studies provided data on the effects of XR interventions on anatomy learning outcomes (11,²⁸ 4,²⁷ and 3 studies,⁴ respectively). These six qualitative reviews had language restrictions and included studies with variable methodologies and/or studies that only evaluated a single group. To date, two meta-analyses have assessed the effectiveness of XR technologies on anatomy knowledge

gains.^{26,29} One meta-analysis reported that VR was effective in increasing knowledge gains compared to traditional learning approaches.²⁹ However, the second meta-analysis reported that XR technologies were no better than traditional methods (e.g., didactic lectures, textbooks, anatomy atlases, plastic models, and/or dissections).²⁶ This early meta-analysis included data from only 8 studies and reported low evidence and statistical power.²⁶ The present work expands upon prior studies by including newer studies not yet meta-analyzed, thus, increasing the statistical power of the analysis. For this meta-analysis exploring anatomy knowledge gains through XR exposure, five research questions were asked including: How do anatomy performance outcomes compare between (1) students who used iVR versus AR; (2) XR interventions that were complementary to educational resources versus those that replaced educational resources; (3) the use of XR versus individual educational resources (e.g., textbooks or atlases, didactic lectures); (4) undergraduate and post-graduate (i.e., medical residents and graduate students) health sciences students using XR technologies; and (5) specific anatomical regions learned using XR technologies. This work will help educators make informed decisions about the pedagogies they employ and represents the most current and robust evidence to date specific to the use of XR in anatomy education.

METHODS

Protocol review

The *Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)*,³⁰ the *Cochrane Handbook for Systematic Reviews of Interventions*,³¹ and the AMSTAR 2 (A Measurement Tool to Assess Systematic Reviews) checklist,³² were followed to undertake the present meta-analysis. The protocol of this review was previously published in PROSPERO (CRD42023423017).

Literature search and databases

The literature search was carried out by two authors who independently searched studies published in PubMed Medline, SCOPUS, Web of Science, and CINAHL Complete, since inception to June 2023. Hand searching for references in related studies, conference proceedings, the grey literature, and document excerpts was undertaken to identify references not retrieved from the electronic search. The PICOS framework was followed to establish study eligibility criteria³³: population (university students of health sciences disciplines), intervention (iVR-based and/or AR-based interventions), comparator (traditional educational approaches, such as didactic lectures, textbooks, atlases, plastic models, and/or dissections), outcomes (knowledge gains and students' perception of the usefulness of XR for learning anatomy), and study design (randomized and non-randomized controlled trials [RCTs]). The keywords used in the search strategy corresponded to PubMed

Subject Headings (MeSH) and included, but were not limited to, "virtual reality," "augmented reality," "teaching," "learning," and "anatomy." Entry terms (i.e., synonyms) related to each keyword were also used. Language and publication date filters were not employed in the search. Table S1 in Supplemental Material Appendix 1 shows the literature search strategy by database.

Study selection criteria and screening procedures

Included studies had to report data related to the PICOS model as described above and at least one of the outcomes of interest (i.e., anatomy knowledge gains in post-evaluation between intervention and control groups; and/or students' perceptions of the usefulness of XR for learning anatomy). Studies were excluded if they were editorials, commentaries, abstracts only, dissertations/theses, or were purely qualitative in design. Two authors independently screened the titles and abstracts of the retrieved references. In order to assess the agreement of inclusion judgments between the two authors Cohen's kappa coefficient (κ) was used.³⁴ The values of this coefficient are interpreted as degrees of agreement according to Landis and Koch³⁵: non-existent if $\kappa < 0$, non-significant when $0 \leq \kappa \leq 0.2$, discrete when $0.2 < \kappa \leq 0.4$, moderate if $0.4 < \kappa \leq 0.6$, substantial if $0.6 < \kappa \leq 0.8$, and excellent if $0.8 < \kappa \leq 1.0$. Each author reviewed in detail all potential inclusion records screened by the other author, with the aim of ensuring that all ineligible studies were correctly excluded. Decision conflicts regarding study inclusion were arbitrated by a third author.

Data extraction

Two authors independently collected data from each included study using a standardized locally developed data collection template. A third author resolved data extraction disagreements.

The following data points were collected: (1) Study characteristics (i.e., authorship, country, year of publication, and funding source); (2) Participant characteristics (i.e., number of participants per study, and in each group, age, gender, health science discipline, and course level); (3) Intervention group characteristics (i.e., whether students received an iVR or AR intervention, whether XR was applied as a supplemental/complementary or replacement intervention, the administration protocol, and the anatomical region studied); (4) Comparison group characteristics (i.e., type of traditional educational approaches, such as lectures, textbooks, anatomy atlases, plastic models, and/or dissections); and (5) Outcomes data (i.e., variables assessed, intervention type, assessment instruments used, sample sizes, means, and standard deviations (SD) of post-intervention outcomes of each variable in each group). All assessment instruments measured anatomical knowledge gains in a comparable manner. Students' knowledge performance scores were meta-analyzed using Cohen's standardized mean difference (SMD). When a study provided the standard error, the range, or interquartile

range from non-skewed distributions, it was transformed into a standard deviation value.^{31,36} Studies with missing data or those reporting non-meta-analyzable data were excluded. No attempt was made to retrieve missing data from corresponding authors.

Variables

The primary variable of interest was anatomy knowledge acquisition or knowledge gains. Secondly, students' perceptions of the usefulness of these technologies as learning tools were also explored.

Risk of bias and quality of evidence assessment

The risk of bias present in the included studies was analyzed using the Cochrane Risk of Bias (CRoB) tool.³⁷ This scale is comprised of seven items: sequence generation and allocation concealment for selection bias; blinding of participants and staff for performance bias, evaluators for detection bias; incomplete outcomes data for attrition bias; selective outcomes reporting for reporting bias; and others. For each item, risk is scored as low ("+"), high ("-"), and unclear ("?").

To assess the quality of study designs and their findings, we employed the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) tool³⁸; the GRADE checklist of Meader.³⁹ The quality of evidence was determined according to five items: risk of bias in each study, imprecision, inconsistency, indirect evidence, and publication bias. Inconsistency was analyzed estimating the heterogeneity of the findings. Imprecision was assessed by the number of studies included, the number of participants in each meta-analysis, and the number of participants per study (large imprecision = <100 participants or <5 studies; moderate imprecision = 100–300 participants and 5–10 studies; and low imprecision = >300 participants or >10 studies). Indirect evidence appears in studies in which the results were measured indirectly. Combining these items allows for findings to be evaluated at four different levels of evidence: high (if findings are robust), moderate (when a future study can change the current findings), low (poor level of confidence), and very low (if four or more items are not met). The quality of evidence was downgraded to one level for each item that did not meet the specified criteria. The risk of bias and quality of evidence assessments were carried out by two authors, and disagreements were resolved by a third author.

Statistical analysis

Comprehensive Meta-Analysis version 4.0 (Biostat, Englewood, NJ, USA)⁴⁰ was used to perform the meta-analyses. The pooled effect size was estimated using Cohen's SMD with a 95% confidence interval (95% CI).^{31,41,42} In the presence of heterogeneity ($I^2 > 40\%$), a random-effects model was used, as proposed by DerSimonian

and Laird,⁴³ to improve the generalization of the findings. Two types of meta-analyses were conducted in this systematic review. To investigate knowledge gains and the perceived usefulness of XR technologies for learning, a meta-analysis with two groups was performed with XR technology users serving as the intervention group and learners exposed to traditional approaches serving as the comparison group. This meta-analysis required sample sizes, means, and standard deviations from post-intervention knowledge assessments to compute the summary effect. The assessment instruments for measuring anatomy knowledge acquisition were inclusive of multiple choice questions (MCQs), matching items, and open-ended free-response questions. To compare students' perceptions of the usefulness of educational resources (i.e., learners exposed to XR technologies versus learners exposed to traditional resources), we used the sample size, mean, and standard deviation of 5-point Likert scale outcomes where higher scores represented more usefulness. Examples of questions to assess the usability for learning anatomy are "VR is useful for learning" or "VR is an effective learning tool." Although the survey instruments measuring students' perceptions were not exactly the same in all studies, they all measured the same conceptual construct (i.e., students' perceptions) and were comparable. In these meta-analyses, the summary effect sizes were expressed as SMD values and were interpreted as follows: large (SMD > 0.8), medium (SMD 0.4–0.7), small (SMD 0.1–0.3), and no effect (SMD 0).⁴⁴ Forest plots were used to graphically visualize the findings of each meta-analysis.⁴⁵

Additionally, the proportion of students who reported XR technologies as useful for learning anatomy was summarized through a proportional meta-analysis of treatment group participants who used XR technologies.⁴⁶ The data employed in this meta-analysis included the total sample size of students who were asked about the usefulness of XR and the number of students who responded satisfactorily/positively about its use.

Heterogeneity (i.e., the variation in study outcomes between studies) was calculated using the degree of inconsistency (I^2), and a p -value for the Q -test ($p < 0.05$ indicates heterogeneity).^{47,48} According to I^2 , heterogeneity can be null ($I^2 = 0$), low ($I^2 < 25\%$), medium ($I^2 25\%–50\%$), and large ($I^2 > 50\%$).^{49,50}

The risk of publication bias was assessed through the visualization of the funnel plots, with asymmetric funnel plots indicating a risk of publication bias,⁵¹ and the p -value of the Egger test (present if $p < 0.1$).⁵² Additionally, the Trim-and-Fill method was used to estimate the adjusted pooled effect size taking into account the presence of publication bias in meta-analysis with two groups.^{53,54} Variations >10% between the original and the adjusted effect size suggest a risk of publication bias and downgrade the quality of evidence one level, even if the funnel plot is symmetric.⁵⁵ For proportional meta-analysis an estimation of publication bias was not performed because the assumption that positive results are published more frequently is not necessarily true for proportional studies.⁴⁶

A sensitivity analysis was performed using the leave-one-out method.³¹ After calculating the global effects of XR devices in comparison to traditional learning resources, the following subgroup analyses

were also performed to determine the effect of knowledge gains: (1) iVR versus AR; (2) XR technologies implemented as replacements versus supplemental/complementary curricular tools; (3) XR technologies compared to individual resources (i.e., textbooks and atlases, didactic lectures); (4) XR technologies on different student populations (i.e., undergraduate versus post-undergraduate students); and (5) XR technologies according to each anatomical system or region studied.

RESULTS

Study selection

In total, 2784 references were retrieved from the primary electronic database search ($n=2781$) and other sources ($n=3$). After removing duplicates, 1931 records underwent title and abstract screening, which excluded 1847 studies leaving 90 eligible for full-text review. Sixty-three studies were further excluded for not meeting the inclusion criteria (Figure 1) leaving 27 studies for analysis.⁵⁶⁻⁸² The Kappa

value for inter-rater agreement was 0.89 indicating "excellent" agreement, according to Landis and Koch, for the title and abstract screening phase. All 27 studies were included in the meta-analysis of anatomy knowledge gains and 14 of these 27^{58,61,62,65,66,70,72,75,77-79,81,82} were included in the meta-analysis of students' perceptions of the usefulness of XR technologies for learning anatomy. Figure 1 (PRISMA flow chart) summarizes the study selection phase.

Study characteristics

Included studies were published between 2013 and 2023 and most frequently originated in the following countries: United States and Australia (22.2% each),^{57,68,70,71,75,82} Germany (18.5%),^{56,62,67,74,78} and Canada^{58,65,81} and Australia^{59,66,72} (11.1% each). The included studies yielded data from 2199 students from across various health sciences disciplines (e.g., medicine, biomedicine, nursing, and physiotherapy). Students had a mean age of 21.43 ± 2.1 years and sex representation was nearly equally distributed (48% male and 52% female). Of all

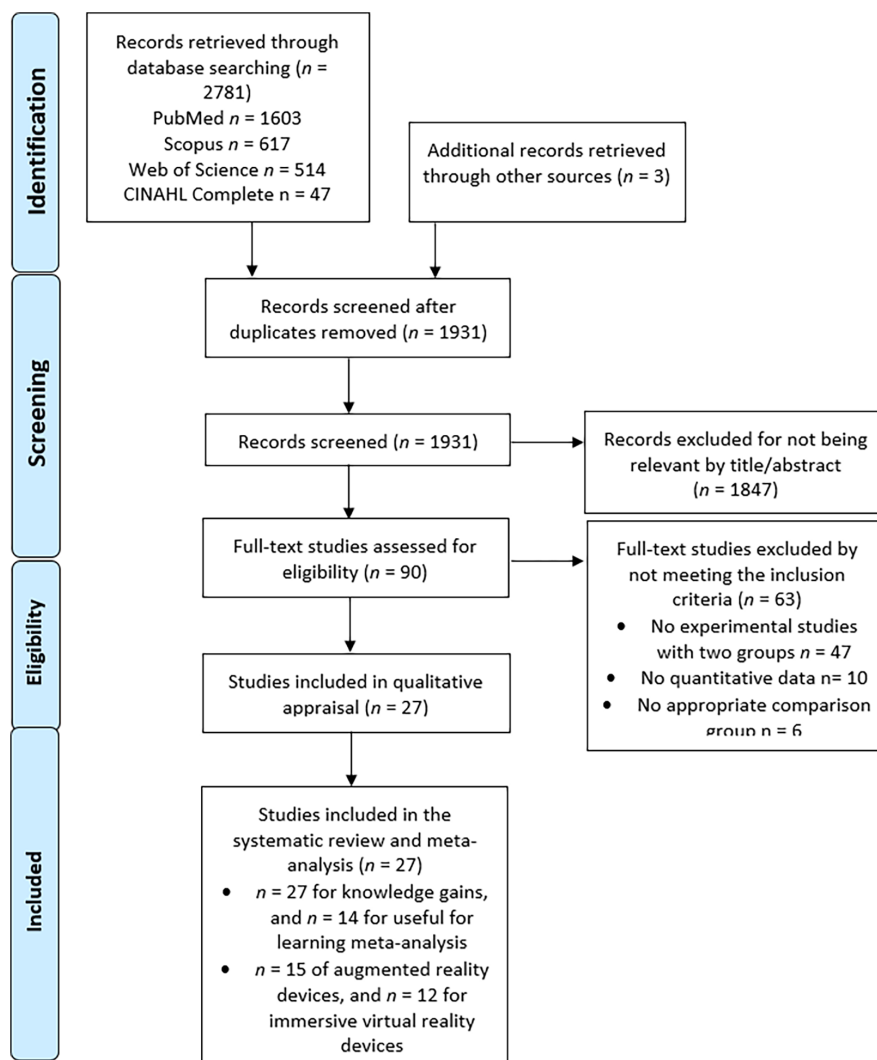


FIGURE 1 PRISMA flow diagram.

students included, 1023 learned using XR technologies (420 students iVR^{60,61,64-66,73,74,79-82} and 603 AR^{56-59,62,63,66-69,71,72,75-78}) and 1176 learned using traditional methods (i.e., textbooks, cadaveric dissections, atlases, didactic lectures, anatomical physical models, and/or images displayed in PowerPoint). The anatomical systems/regions that were studied with XR technologies included: musculoskeletal system,^{64,68,71,75-77,79} neuroanatomy,^{62,63,66,69,70,73,81} inner ear,^{58,74} cardiac system,^{65,72} skin,⁶⁷ digestive system,⁶¹ pelvic floor,⁵⁹ and breast.⁵⁷ Studies collected knowledge performance data on XR effectiveness immediately following the interventions. The length of the interventions ranged from 1 day in 25 studies (93% of all studies) to 5⁸⁰ and 6⁷¹ days in two other studies. Regarding the types of assessments used to quantify knowledge gains, 18 studies^{56,58-67,69,70,72,74,78,81} used tests comprised of multiple-choice questions (MCQ), 3 studies^{68,76,77} used open-ended free-response questions or matching labels, and the remaining studies used a combination of assessment item types (MCQ, matching, true or false questions, fill-in-the-blank, landmark questions, and open-ended free-response such as short answer or essays).^{57,70,71,73,75,79,80} For students' perceptions of the usefulness of XR technologies, the included studies provided outcomes of self-reported questionnaires that utilized 5-point Likert scales with questions such as "VR is useful for my learning," "It is an effective learning tool," or "My knowledge about anatomy improved after studying with...[XR]." Finally, 12 studies received external funding.^{59-61,65,68,70,72,74,78-81} In two studies, participants received compensation for study participation.^{76,77} Table 1 presents the characteristics of the studies included in the meta-analysis.

Risk of bias assessment

Some studies reported a probable or unclear risk of selection bias because random sequence generation and allocation concealment were not met or were unclear in 11 studies (40.7%; 11/27).^{56-58,60,61,65,67,69,70,73,78} The impossibility of blinding the participants and the staff who carried out the interventions was responsible for the risk of performance bias (100%; 27/27). However, in four studies (14.8%; 4/27),^{64,65,67,81} where the evaluators were blinded, a high risk of detection bias was present. Finally, the risk of attrition, reporting, and other biases was low in all studies. These results are detailed in Table S2.

Effects of XR technologies on learners' anatomy knowledge gains

Twenty-seven studies (k) with 33 independent comparisons (c) provided data from 2199 participants (n) to analyze the effect of XR technologies compared to traditional resources for improving knowledge after interventions.⁵⁶⁻⁸² A medium effect ($k=27$; $c=33$; $n=2199$; $SMD=0.40$; 95% $CI=0.22$ to 0.60 ; $p<0.001$) was observed in favor of XR interventions (Figure 2) with low heterogeneity ($I^2=17.1\%$; $Q=39.7$; $df=32$; $p=0.160$). The Trim-and-fill estimation

yielded an adjusted effect of 0.45 (95% $CI=0.26$ to 0.65), suggesting that the risk of publication bias was underestimating 12% of the original effect (Figure S1).

Per a subgroup analysis, learners' who utilized either iVR ($k=12$; $c=15$; $n=927$; $SMD=0.57$; 95% $CI=0.29$ to 0.85 ; $p<0.001$; $I^2=19.7\%$; $Q=16.8$; $df=14$; $p=0.260$) or AR technologies ($k=16$; $c=18$; $n=1331$; $SMD=0.27$; 95% $CI=0.01$ to 0.52 ; $p=0.042$; $I^2=11.3\%$; $Q=19.6$; $df=17$; $p=0.300$) outperformed control group participants who had access to only traditional anatomy learning resources (Figure 2).

Secondly, we compared the effectiveness of XR technologies when they were used as a unique solitary educational tool (i.e., replacements for other educational resources) versus supplementary/complementary resources. XR technologies used as a complement to existing resources demonstrated a medium effect size^{56-62,70,71,73-75,79,82} for knowledge gain ($k=15$; $c=18$; $n=1195$; $SMD=0.52$; 95% $CI=0.40$ to 0.63 ; $p<0.001$; $I^2=34.1\%$; $Q=24.3$; $df=17$; $p=0.110$). Conversely, XR technologies used as replacement tools demonstrated a small effect size^{63,64,66-69,72,76-78,80,81} ($k=12$; $c=15$; $n=1004$; $SMD=0.13$; 95% $CI=0.10$ to 0.25 ; $p=0.039$; $I^2=0.0\%$; $Q=13.2$; $df=14$; $p=0.510$; Figure S2).

Thirdly, we compared the effectiveness of XR technologies against specific traditional methods. XR technology users demonstrated significantly higher knowledge gains than learners who studied using textbooks and atlases^{56,58,63,65-68,70,72,73,76-78,80-82} ($k=16$; $c=20$; $n=1004$; $SMD=0.32$; 95% $CI=0.10$ to 0.54 ; $p=0.004$; $I^2=25.4\%$; $Q=24.5$; $df=19$; $p=0.140$) and didactic lectures^{58,59,61,62,64} ($k=5$; $c=5$; $n=363$; $SMD=1.00$; 95% $CI=0.57$ to 1.42 ; $p<0.001$; $I^2=35.1\%$; $Q=6.5$; $df=4$; $p=0.170$; Figure S3).

Upon assessing the effectiveness of XR technologies between different students populations (i.e., undergraduate versus post-undergraduate students), undergraduate students who used XR interventions demonstrated higher knowledge gains than controls^{56-72,74-78,80,81} ($k=24$, $c=28$; $n=1931$; $SMD=0.41$; 95% $CI=0.20$ to 0.62 ; $p<0.001$; $I^2=23.1\%$; $Q=37.1$; $df=27$; $p=0.100$). No effect was identified among post-undergraduate students^{73,79,82} ($k=3$; $c=5$; $n=268$; $SMD=0.38$; 95% $CI=-0.1$ to 0.86 ; $p=0.110$; $I^2=0.0\%$; $Q=2.8$; $df=4$; $p=0.600$; Figure S4).

Finally, the musculoskeletal system and neuroanatomy were the anatomical areas most commonly studied using XR. XR interventions increased learners' knowledge gains irrespective of the anatomical area studied:^{57,64,66,68,71,75-77,79,80} musculoskeletal anatomy ($k=10$; $c=13$; $n=932$; $SMD=0.20$; 95% $CI=0.06$ to 0.32 ; $p=0.003$; $I^2=0.0\%$; $Q=10.3$; $df=12$; $p=0.580$) and neuroanatomy^{62,63,69,70,73,81} ($k=6$; $c=7$; $n=596$; $SMD=0.50$; 95% $CI=0.33$ to 0.64 ; $p<0.001$; $I^2=22.1\%$; $Q=6.4$; $df=6$; $p=0.370$; Figure S5).

Students' perception of the usefulness of XR technologies as learning tool

Fourteen studies^{58,61,62,65,66,70,72,75-79,81,82} provided data from 943 students providing perceptions on whether XR technologies were

TABLE 1 Characteristics of the studies included in the review.

Study	Participants	XR intervention (n)	Control intervention (n)	Variables
Albrecht et al. 2013 (Germany)	10 third-year forensic medical students (23.7 ± 2 years old, 6M;4F)	AR using the app "Marble forensics," such as replacement to control intervention (n = 6), 1 day	Standard textbook plus anatomy images (n = 4)	Knowledge gains (test comprised of 10 multiple choice questions (MCQ))
Baratz et al. 2022 (United States)	177 first-year medical students (age-sex NR)	AR complementing the cadaveric dissection (n = 43), 1 day	Cadaveric dissection (n = 134)	Knowledge gains (quiz comprised of 7 conceptual questions)
Barmaki et al. 2019 (United States)	288 undergraduate pre-medical students (19.8 ± 2 years old, 118M;178F)	AR with "Magic Mirror," replacing traditional intervention (n = 164), 1 day	Textbooks, atlases, and virtual mirror (n = 124)	Knowledge gains (quiz comprised of 8 questions matching labels)
Bogomolova et al. 2020 (Netherlands)	38 undergraduate medical and biomedical students (18.6 ± 0.8 years old, 15M;23F)	AR through "Dynamic Anatomy" for HoloLens, replacing control (n = 20), 1 day	2D anatomical atlas images plus textbook (short description about images) (n = 18)	Knowledge gains (quiz of 30 questions (match and open-ended questions)). Learning experience was measured by a standardized self-reported questionnaire, on a 5-point Likert scale ("My knowledge about anatomy improved after studying with...")
Bogomolova et al. 2023 (Netherlands)	66 undergraduate medical and biomedical students (19.1 ± 1.6 years old, 27M;39F)	AR through "Dynamic Anatomy" for HoloLens, replacing control (n = 32), 1 day	2D anatomical atlas plus textbook (short description about images) (n = 34)	Knowledge gains are evaluated by a quiz of 30 match and open-ended questions). Learning experience was measured by a standardized self-reported questionnaire, on a 5-point Likert scale ("My knowledge about anatomy improved after studying with...")
Bork et al. 2019 (Germany)	72 first-year medical students (21.4 ± 3.4 years old, 23M;49F)	AR through "Magic Mirror" system, as replaced anatomy educative tool (n = 24), 1 day	2D anatomical atlas using Anatomage (n = 24, CG ₁) and radiological images (n = 24, CG ₂)	Knowledge gains (quiz of 20 MCQ). Learning experience (useful for learning) was assessed with standardized question (AR or traditional can be beneficial for increasing my anatomical knowledge?)
Cai et al. 2020 (China)	153 first-year medical residents (23.9 ± 1.5 years old, 153M;0F)	Immersive VR models through "Mimics 15.0" in a head-mounted display system "VR Shinecon," as supplemental intervention to preliminary lecture (n = 51), 1 day	3D-printed models created with photopolymer resin (n = 51, CG ₁); and conventional anatomic models (n = 51, CG ₂)	Knowledge gains (30 landmarks questions and clinical cases). Learning experience (usefulness for learning) was evaluated with a 5-point Likert Scale
Du et al. 2020 (Taiwan)	18 medical students (21.9 ± 0.9 years old, 14M;4F)	Immersive VR with a head-mounted display as replaced intervention (n = 12), 6 days	Textbooks plus anatomy atlases (n = 6)	Knowledge gains (questions to name the muscle highlighted in red)
Ekstrand et al. 2018 (Canada)	64 first and second-medical students (24.4 ± 2.7 years old, 36M;28F)	Immersive VR, as replaced intervention (n = 31), 1 day	Images in a textbook (n = 34)	Knowledge gains (quiz comprised of 22 MCQ)
Ellington et al. 2019 (United States)	31 medical residents (age-sex NR)	Immersive VR using, as a complement to traditional (n = 16), 1 day	Independent study with an anatomy textbook (n = 15)	Knowledge gains (quiz comprised of 28 MCQ)
Gnanasegaram et al. 2020 (Canada)	29 medical students (age-sex NR)	AR using "HoloLens," as a complement to traditional (n = 10), 1 day	Lecture (n = 9, CG ₁) and anatomy atlases and textbook (n = 9, CG ₂)	Knowledge gains (quiz comprised of 20 MCQ)

TABLE 1 (Continued)

Study	Participants	XR intervention (n)	Control intervention (n)	Variables
Gray et al. 2022 (Australia)	38 second-year nursing students (20–59 years old; 38F)	AR stereoscopic glasses in mobile-phone, as a complement to traditional (n = 20), 1 day	Lectures (n = 18)	Knowledge gains (quiz comprised of 30 MCQ)
Hu et al. 2020 (Taiwan)	101 third-year medical students (21 years old, 52M:49F)	Immersive VR through "HTC Vive Pro," as a complement to traditional training (n = 47), 1 day	Textbook plus anatomy images (n = 54)	Knowledge gains (quiz comprised of 10 MCQ)
Imai et al. 2022 (Japan)	60 fourth and fifth-year medical students (age-sex NR)	Immersive VR, as a complement to lecture (n = 30), 1 day	Lecture (n = 30)	Knowledge gains (quiz comprised by MCQ). Learning experience (useful for learning) was assessed with standardized question (<i>Do you think you now have a better understanding of the mediastinal anatomy?</i>)
Kockro et al. 2015 (Germany)	166 second-year medical students (22.5 years old, 60M:106F)	AR using the "DextroBeam system," complement to lecture (n = 86), 1 day	Lecture more visualization of 2D-images (n = 80)	Knowledge gains (quiz comprised of 10 MCQ)
Küçük et al. 2016 (Turkey)	70 second-year undergraduate medical students (18–21 years old, 34M:36F)	AR using "Axiom Neuro v1" and "Anatomy 4D" software's, replacing traditional lessons (n = 34), 1 day	Presentations of 2D-images, graphs and text (n = 36)	Knowledge gains (quiz comprised of 30 MCQ)
Kurul et al. 2020 (Turkey)	72 undergraduate physical therapy (19.2 ± 0.8 years old, 20M:52F)	Immersive VR using "3D Organon Anatomy®," replacing traditional presentations (n = 36), 1 day	Lecture with traditional presentations (n = 36)	Knowledge gains (quiz comprised of 15 MCQ)
Maresky et al. 2019 (Canada)	42 first-year medical students (18–34 years, sex data NR)	Immersive VR, as a complement to independent study (n = 28), 1 day	Traditional methods (atlas plus textbook) (n = 14)	Knowledge gains (quiz comprised of 10 MCQ). Learning experience (useful for learning) was assessed with standardized question (<i>Cardiac VR is useful for my learning</i>) on a 5-point Likert scale
Moro et al. 2017 (Australia)	59 biomedical and health sciences students (20.7 ± 5.5 years old, 28M:31F)	Immersive VR through "Oculus Rift," and others 17 in AR application, both replace interventions (n = 20), 1 day	2D-images in tablet plus text (n = 22)	Knowledge gains (quiz comprised of 20 MCQ). Learning experience (useful for learning) was assessed with standardized question (<i>It is an effective learning tool</i>) on a 5-point Likert scale
Noll et al. 2017 (Germany)	44 medical students (22.3 ± 3.3 years old, 25M:19F)	Mobile AR through "mARble-Derma" app, as replace intervention (n = 22), 1 day	Anatomical images and text in smartphone app (n = 22)	Knowledge gains (quiz comprised of 10 MCQ)
Pickering et al. 2022 (United Kingdom)	146 second-year medical students (age-sex NR)	AR "HoloLens," as replaced intervention (n = 62), 1 day	Screencasts (anatomical record information plus anatomical images) (n = 84)	Knowledge gains (quiz comprised of 10 MCQ)
Stepan et al. 2017 (United States)	66 first and second-year medical student (21–35 years old, 33M:33F)	Immersive VR using "Oculus Rift," complementing independent study (n = 33), 1 day	Independent study with textbook plus images (n = 33)	Knowledge gains (quiz comprised of 30 MCQ, labeling, and fill-in-the-blank). Learning experience with a scale from 0 (not useful for learning) to 100 (opposite)
Stojanovska et al. 2020 (United States)	31 second-year medical students (age-sex NR)	AR through "Microsoft HoloLens," as complementary intervention (n = 15), 6 days	Cadaveric dissection (n = 16)	Knowledge gains (exam comprised of 36 questions)

TABLE 1 (Continued)

Study	Participants	XR intervention (n)	Control intervention (n)	Variables
Veer et al. 2022 (Australia)	67 first-year undergraduate health science and medical students (≥ 17 years old, 22M:45F)	AR through "Microsoft HoloLens," as replaced intervention (n=33), 1 day	Textbooks (n=34)	Knowledge gains are assessed with a quiz comprised of 10 MCQs. Learning experience (useful for learning) was assessed with standardized question on a 5-point Likert scale
Vieira de Faria et al. 2016 (Brazil)	84 graduate medical students (age-sex NR)	Immersive VR through "VR Worx 2.6," as complementary intervention (n=28), 1 day	Lecture atlas (n=28, CG ₁), and traditional lessons (n=28, CG ₂)	Knowledge gains (written theory test and a lab practicum)
Von Schnakenburg et al. 2023 (Germany)	177 medical students (age-sex NR)	Immersive VR through "Oculus Rift," as complementary intervention (n=88), 1 day	Plastic models (n=89)	Knowledge gains (quiz comprised of 10 MCQ)
Weeks et al. 2021 (United States)	30 first-year medical students (24 years old, 15M:15F)	AR through "Microsoft HoloLens," as complementary intervention (n=15), 1 day	2D-lecture with radiological images using computational tomography (n=15)	Knowledge gains (quiz comprised of anatomy questions). Learning experience was measured as the percentage of students that marked AR as useful for learning

Abbreviations: AR, augmented reality; CG, control group; F, female; M, male; MCQ, multiple choice questions; N, sample size in each group; NR, not reported; p, p-value; VR, virtual reality; XR, immersive virtual reality and augmented reality interventions; XRG, XR group.

useful for learning anatomy. For this variable, two meta-analyses were performed. On the one hand, 10 studies^{61,66,70,72,75-79,81} with 14 independent comparisons compared usefulness or perceived effectiveness metrics between learners exposed to XR versus those exposed to traditional educational resources and showed a medium effect ($k=10$; $c=14$; $SMD=0.54$; 95% $CI=0.04$ to 1 ; $p=0.035$) in favor of XR (Figure 3) with medium heterogeneity ($I^2=31.2\%$; $Q=20.3$; $df=13$; $p=0.080$). The risk of publication bias was likely among these studies (Trim-and-fill estimated an adjusted effect of 1 ; 95% $CI=0.51$ to 1.48 ; Figure S6). Subgroup analysis revealed that students perceived that XR technologies were more useful for learning anatomy when they were used as complement^{61,70,75,79} ($k=4$; $c=5$; $n=309$; $SMD=1.0$; 95% $CI=0.13$ to 1.86 ; $p=0.03$; $I^2=66\%$; $Q=16.9$; $df=4$; $p=0.01$) rather than as replacement resource^{66,72,76-78,81} ($k=6$; $c=9$; $n=634$; $SMD=0.29$; 95% $CI=-0.36$ to 0.93 ; $p=0.384$; $I^2=0\%$; $Q=0.4$; $df=8$; $p=0.99$). Specifically, XR technologies were considered more useful for learning anatomy than textbooks and atlases^{66,70,76-78,81} ($k=6$; $c=9$; $SMD=0.36$; 95% $CI=0.17$ to 0.55 ; $p<0.001$; $I^2=0\%$; $Q=6.9$; $df=8$; $p=0.54$).

To complement this meta-analysis, we performed a meta-analysis of proportions using only one group (i.e., students exposed to XR interventions) with data from 233 students who used XR across 7 studies.^{58,62,65,66,75,81,82} A proportional meta-analysis was conducted to estimate the aggregated proportion of students who favored the usefulness of XR technologies across studies and comparisons. Outcomes revealed that 80% of survey participants (95% $CI=70$ to 88 ; $p<0.001$; $I^2=15.9\%$; $Q=7.1$; $df=6$; $p=0.32$) considered XR technologies useful for learning anatomy.

DISCUSSION

Human anatomy is studied across many health sciences disciplines and is often a difficult subject given its volume and conceptually challenging structural relationships in certain regions. As such, complementary tools, including interactive technologies, are being incorporated into the teaching and learning of Human Anatomy subjects.^{83,84} XR-based technologies, such as iVR and AR, have drawn more attention in recent years, especially motivated during the COVID-19 pandemic.⁸⁵ Between 2019 and 2022 several scoping and systematic reviews, without meta-analyses, were published and collectively suggested that XR-based interventions are effective for learning human anatomy.^{1,2,4,8,27,28} In 2021, however, Moro et al.'s systematic review and meta-analysis outcomes contradicted prior findings whereby they detected no statistically significant differences in anatomy knowledge gains between XR-based interventions and traditional learning resources. Compared to prior reviews, this meta-analysis included the most studies to date ($k=27$), surpassing previous meta-analyses with only fifteen²⁹ and eight studies,²⁶ and represents a comprehensive and up-to-date analysis of this topic providing additional evidence related to the efficacy of XR technologies. Additionally, this work is the first to analyze students' subjective perceptions about the

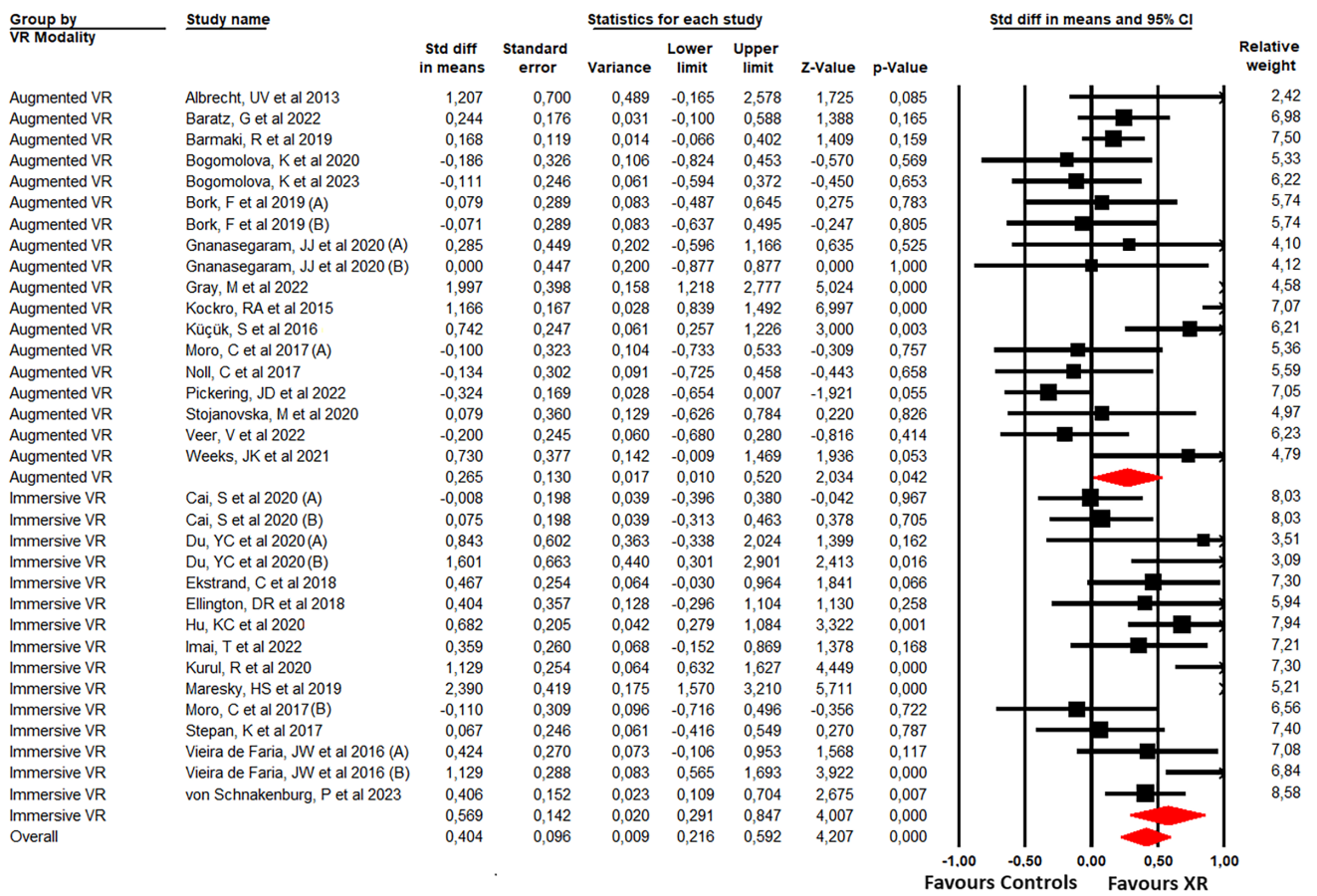


FIGURE 2 Forest plot of the meta-analysis for anatomy's knowledge gain.

usefulness of XR technologies in Human Anatomy courses through a meta-analysis.

Knowledge gains

The use of XR-based interventions compared to traditional resources improved anatomy knowledge gains with a medium effect (SMD=0.4). This finding aligns with previous systematic reviews without meta-analyses whereby iVR and AR have been considered promising tools for promoting anatomy knowledge gains among health sciences students.^{1,2,27,28} and with Zhao et al.'s meta-analysis.²⁹ The present findings contradict Moro et al.'s meta-analytic findings²⁶ which did not show statistical differences between the efficacy of XR and traditional methods. Furthermore, Moro and colleagues did not perform subgroup analyses investigating the use of XR as a complementary versus replacement tool or comparing the effects of XR against specific traditional methods.

Further subgroup analysis suggests that iVR (SMD=0.57) could be modestly more effective than AR (SMD=0.27) at increasing anatomy knowledge. iVR is proving to be an effective and versatile tool in health sciences education, especially in gross anatomy learning,

due to its customizable and "hands-on" capabilities.⁸⁶ For some, iVR is an attractive option for learning anatomy outside of donor-based dissections because it facilitates a virtual realistic rendering of spatial anatomical relationships and allows for interaction and manipulation of 3D structures to facilitate remote "hands-on" learning.⁸⁶ Additionally, in recent years the cost of developing and acquiring iVR hardware/software has been substantially reduced⁸⁷ making iVR more accessible to educational centers and universities.

An important consideration when applying these types of technologies is whether to implement them as a complement to traditional resources or to use them as a replacement for existing resources. While both modalities are effective, the greatest effect (SMD=0.52) on anatomy knowledge gains was observed when XR was used as a complementary tool alongside traditional resources. Our findings are similar to Wilson et al.'s meta-analysis which showed a similar effect for computer-aided instructional tools.⁸⁸ When implemented as a complement to traditional didactic lectures, textbooks, and/or atlases, the use of XR technologies may help to further solidify learners' spatial relationship, thus, promoting anatomy learning. In the present work, many studies compared the use of textbooks, atlases, and didactic lectures against hands-on XR technologies, which were found to be more effective than these passive learning approaches (SMD=0.32 for XR versus textbooks and atlases, and

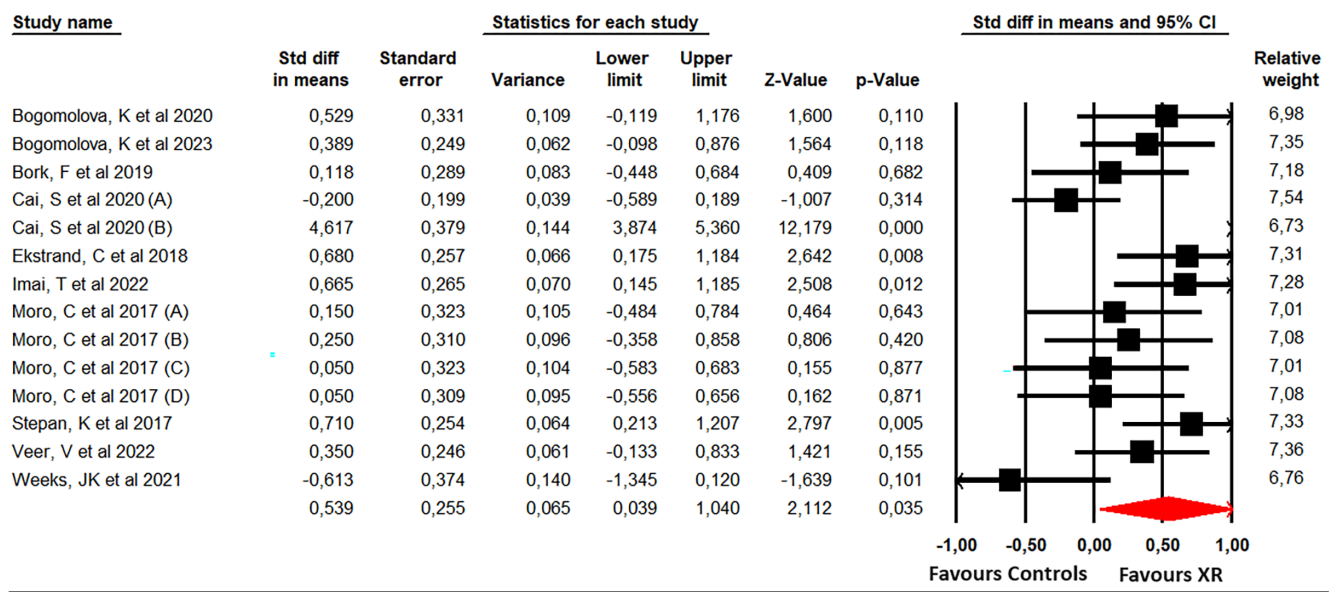


FIGURE 3 Forest plot of the meta-analysis for students' perception of the usefulness of these technologies as learning tool.

SMD=1 for XR versus didactic lectures). Didactic lectures and self-study using textbooks and atlases represent passive learning methods. Alternatively, XR technologies are a student-centered learning approach that promotes student interactions and more active participation in the teaching-learning process.^{88,89} Our findings are in line with the meta-analyses of Zhao et al. and Wilson et al. in which greater knowledge gains were identified in students who received active and interactive methodologies.^{29,88} Our findings suggest XR may be a promising interactive and student-centered learning tool for helping students to become owners of their learning.⁹⁰

Our findings also demonstrated that XR interventions are effective in undergraduate health sciences students but are not effective for post-undergraduate populations (i.e., mainly medical residents), likely because the anatomy knowledge in residents was acquired and integrated previously. These findings are consistent with the results of Zhao et al.'s meta-analysis which showed that VR was effective in promoting knowledge gains in medical students but did not show an influence on residents.²⁹

Do health sciences students perceive XR technologies as being useful for learning?

The present meta-analysis showed that students subjectively perceive XR technologies to be more useful for learning anatomy (SMD=0.54) compared to the traditional methods previously mentioned; although this effect may be confounded by publication bias. In addition, a subgroup analysis showed that, compared to traditional interventions, the perceived usefulness of XR devices is higher when it is used as a complementary intervention (SMD=1) versus a replacement intervention, especially when compared to textbooks or atlases.

A single-group proportional meta-analysis summarized the percentage of students favoring XR technologies as useful learning tools. Data from across 7 studies^{58,62,65,66,75,81,82} and 233 participants indicated that 80% of all students that used XR devices perceived XR interventions as useful for learning anatomy, as either a complement or replacement for traditional methods. Other subjective outcomes, such as whether XR interventions or devices are enjoyable, easy to use, recommendable, and/or motivating were not assessed due to an inadequate number of studies. Future studies would benefit from further summarizing these data. It is important to consider the preferences, predispositions, and perceptions of students when new didactic tools are incorporated into curricula. Overall, prior reviews suggest that university health sciences students generally report high levels of satisfaction, adherence, and motivation when using VR and AR for anatomy and physiology learning.⁹¹⁻⁹³

The virtual and hyper-realistic scenarios of human body parts that iVR and AR render likely to improve the acquisition of anatomy knowledge in health sciences education, especially when used as complementary resources.⁹⁴ Although anatomical dissection is often considered the "gold standard" for anatomy education in the health sciences, not all universities or colleges have dissection facilities. In the absence of donor tissues, XR-based interventions can augment the learning of Human Anatomy thanks to immersive 3D images.⁹⁵ However, health sciences students often show cognitive and practical difficulties transferring anatomical knowledge from the classroom to professional practice; thought to be a consequence of the challenging spatial organization of anatomical structures and their relationships with adjacent structures.⁷⁷ As such, some argue it is crucial to improve students' visual-spatial abilities by helping them to build and manipulate three-dimensional mental representations.⁹⁶ XR-based interventions, compared to traditional anatomical atlases or textbooks, allow digital anatomical structures to be

visualized and manipulated in three dimensions so that structures can be studied individually or in relationship to other surrounding structures across all spatial planes.⁹⁷ The ability of XR to improve the three-dimensional visualizations of anatomical structures explains its positive effects on knowledge acquisition.^{98,99} Finally, XR devices, which in many cases are available in low-resource settings, can promote the development of brain schemas for better long-term memory retention, internalization, and comprehension of relevant anatomical concepts.¹⁰⁰

Limitations

Although meta-analysis is a powerful method for aggregating outcomes across studies, some limitations must be acknowledged. Typical limitations of meta-analysis include publication bias, research bias, and dependence on existing, yet possibly imperfect, data.^{101,102} Publication bias was explored in the meta-analyses of independent groups but was not recommended for proportional meta-analyses.⁴⁶ Sub analyses could not be performed for certain comparison (e.g., comparing students who used XR devices vs donor dissections or XR devices versus plastic models) as a result of too few studies. Another limitation was related to the risk of biases. The impossibility of blinding participants, blinding researchers who implemented or supervised the interventions, and blinding evaluators, primarily explains the presence of performance and detection biases, which can reduce the accuracy of interventions and outcomes generalizations.¹⁰³ It was not feasible to evaluate the effectiveness of XR technologies longitudinally over time. All studies provided data about the effects of XR technologies immediately after the interventions. Too few studies provided data after long-term follow-up ($k < 6$) for proper analysis. Lastly, the presence of study heterogeneity was likely a consequence of different assessment strategies, unique learner populations, and different anatomy curricula.

CONCLUSION

The current meta-analysis showed that XR-based interventions are more effective than traditional passive learning resources as a teaching-learning resource for studying anatomy in health sciences students. Both XR technologies (iVR and AR) improved learners' anatomy knowledge gains compared to controls, especially when used alongside traditional resources as supplemental/complimentary tools. A large effect favoring XR technologies was detected when compared against didactic lectures and anatomical images in atlases and/or textbooks. Overall, 80% of students who used XR technologies subjectively reported that XR technologies are useful for learning anatomy, especially when they are used as a supplemental resource in anatomy curricula. Future studies exploring combinations of technology-based and traditional learning resources will help to elucidate the effectiveness of deliberately blended strategies on anatomy learning outcomes.

FUNDING INFORMATION

No external funding was received.

CONFLICT OF INTEREST STATEMENT

No conflicts of interest are declared by the authors.

DATA AVAILABILITY STATEMENT


Request to corresponding author.

ETHICS STATEMENT

Not applicable.

ORCID

Irene Cortés-Pérez  <https://orcid.org/0000-0002-1876-1399>

Francisco Antonio Nieto-Escámez  <https://orcid.org/0000-0001-5301-9475>

Héctor García-López  <https://orcid.org/0000-0003-3160-7692>

Esteban Obrero-Gaitán  <https://orcid.org/0000-0002-8430-5500>

María Catalina Osuna-Pérez  <https://orcid.org/0000-0003-1293-2983>

REFERENCES

- Karbasi Z, Niakan Kalhori SR. Application and evaluation of virtual technologies for anatomy education to medical students: a review. *Med J Islam Repub Iran*. 2020;34:163.
- Duarte ML, Santos LR, Guimarães Júnior JB, Peccin MS. Learning anatomy by virtual reality and augmented reality. A scope review. *Morphologie*. 2020;104(347):254–66.
- Papa V, Varotto E, Vaccarezza M, Ballestriero R, Tafuri D, Galassi FM. The teaching of anatomy throughout the centuries: from Herophilus to plastination and beyond. *Med Hist*. 2019;3: 69–77.
- Lucena-Anton D, Fernandez-Lopez JC, Pacheco-Serrano AI, Garcia-Munoz C, Moral-Munoz JA. Virtual and augmented reality versus traditional methods for teaching physiotherapy: a systematic review. *Eur J Investig Heal Psychol Educ*. 2022;12(12):1780–92.
- Sadek O, Baldwin F, Gray R, Khayyat N, Fotis T. Impact of virtual and augmented reality on quality of medical education during the COVID-19 pandemic: a systematic review. *J Grad Med Educ*. 2023;15(3):328–38.
- Quin D. Longitudinal and contextual associations between teacher–student relationships and student engagement. *Rev Educ Res*. 2017;87(2):345–87.
- Pennings HJM, van Tartwijk J, Wubbels T, Claessens LCA, van der Want AC, Brekelmans M. Real-time teacher–student interactions: a dynamic systems approach. *Teach Teach Educ*. 2014;37:183–93.
- Papa V, Varotto E, Galli M, Vaccarezza M, Galassi FM. One year of anatomy teaching and learning in the outbreak: has the COVID-19 pandemic marked the end of a century-old practice? A systematic review. *Anat Sci Educ*. 2022;15(2):261–80.
- Allen LK, Eagleson R, de Ribaupierre S. Evaluation of an online three-dimensional interactive resource for undergraduate neuroanatomy education. *Anat Sci Educ*. 2016;9(5):431–9.
- Ryan RM, Deci EL. Intrinsic and extrinsic motivation from a self-determination theory perspective: definitions, theory, practices, and future directions. *Contemp Educ Psychol*. 2020;61:101860.
- Pintrich PR, De Groot EV. Motivational and self-regulated learning components of classroom academic performance. *J Educ Psychol*. 1990;82(1):33–40.

12. Wouters A, Croiset G, Galindo-Garre F, Kusurkar RA. Motivation of medical students: selection by motivation or motivation by selection. *BMC Med Educ*. 2016;16(1):37.
13. Stegers-Jager KM, Cohen-Schotanus J, Themmen APN. Motivation, learning strategies, participation and medical school performance. *Med Educ*. 2012;46(7):678–88.
14. McCuskey RS, Carmichael SW, Kirch DG. The importance of anatomy in health professions education and the shortage of qualified educators. *Acad Med*. 2005;80(4):349–51.
15. Martinengo L, Yeo NJY, Markandran KD, Olsson M, Kyaw BM, Car LT. Digital health professions education on chronic wound management: a systematic review. *Int J Nurs Stud*. 2020;104:103512.
16. Wang C, Horby PW, Hayden FG, Gao GF. A novel coronavirus outbreak of global health concern. *Lancet*. 2020;395(10223):470–3.
17. Xiao J, Evans DJR. Anatomy education beyond the COVID-19 pandemic: a changing pedagogy. *Anat Sci Educ*. 2022;15(6):1138–44.
18. Obrero-Gaitán E, Nieto-Escamez F, Zagalaz-Anula N, Cortés-Pérez I. An innovative approach for online neuroanatomy and neuropathology teaching based on 3D virtual anatomical models using leap motion controller during COVID-19 pandemic. *Front Psychol*. 2021;12:1853.
19. Massetti T, da Silva TD, Crocetta TB, Guarnieri R, de Freitas BL, Bianchi Lopes P, et al. The clinical utility of virtual reality in neurorehabilitation: a systematic review. *J Cent Nerv Syst Dis*. 2018;10:117957351881354.
20. Weiss PL, Keshner EA, Levin MF. *Virtual reality technologies for health and clinical applications*. New York: Springer Science Business Media; 2014.
21. Nieto-Escamez F, Cortés-Pérez I, Obrero-Gaitán E, Fusco A. Virtual reality applications in neurorehabilitation: current panorama and challenges. *Brain Sci*. 2023;13(5):819.
22. Weiss PL, Kizony R, Feintuch U, Katz N. Virtual reality in neurorehabilitation. *Textb Neural Repair Rehabil*. 2006;51(8):182–97.
23. Straudi S, Basaglia N. Neuroplasticity-based technologies and interventions for restoring motor functions in multiple sclerosis. *Adv Exp Med Biol*. 2017;958:171–85.
24. Tieri G, Morone G, Paolucci S, Iosa M. Virtual reality in cognitive and motor rehabilitation: facts, fiction and fallacies. *Expert Rev Med Devices*. 2018;15(2):107–17.
25. Winter C, Kern F, Gall D, Latoschik ME, Pauli P, Käthner I. Immersive virtual reality during gait rehabilitation increases walking speed and motivation: a usability evaluation with healthy participants and patients with multiple sclerosis and stroke. *J Neuroeng Rehabil*. 2021;18(1):68.
26. Moro C, Birt J, Stromberga Z, Phelps C, Clark J, Glasziou P, et al. Virtual and augmented reality enhancements to medical and science student physiology and anatomy test performance: a systematic review and meta-analysis. *Anat Sci Educ*. 2021;14(3):368–76.
27. Barteit S, Lanfermann L, Bärnighausen T, Neuhaus F, Beiersmann C. Augmented, mixed, and virtual reality-based head-mounted devices for medical education: systematic review. *JMIR Serious Games*. 2021;9(3):e29080.
28. Ryan GV, Callaghan S, Rafferty A, Higgins MF, Mangina E, McAuliffe F. Learning outcomes of immersive technologies in health care student education: systematic review of the literature. *J Med Internet Res*. 2022;24(2):e30082.
29. Zhao J, Xu X, Jiang H, Ding Y. The effectiveness of virtual reality-based technology on anatomy teaching: a meta-analysis of randomized controlled studies. *BMC Med Educ*. 2020;20(1):127.
30. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021;372:n71.
31. Higgins J, Green S. *Cochrane handbook for systematic reviews of interventions*. London: The Cochrane Collaboration; 2011.
32. Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ*. 2017;358:j4008.
33. Eriksen MB, Frandsen TF. The impact of patient, intervention, comparison, outcome (PICO) as a search strategy tool on literature search quality: a systematic review. *J Med Libr Assoc*. 2018;106(4):420–31.
34. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas*. 1960;20(1):37–46.
35. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159.
36. Hozo SP, Djulbegovic B, Hozo I. Estimating the mean and variance from the median, range, and the size of a sample. *BMC Med Res Methodol*. 2005;5:13.
37. Higgins JPT, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ*. 2011;343:d5928.
38. Atkins D, Best D, Briss PA, Eccles M, Falck-Ytter Y, Flottorp S, et al. Grading quality of evidence and strength of recommendations. *BMJ*. 2004;328(7454):1490.
39. Meader N, King K, Llewellyn A, Norman G, Brown J, Rodgers M, et al. A checklist designed to aid consistency and reproducibility of GRADE assessments: development and pilot validation. *Syst Rev*. 2014;3:82.
40. Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. *Introduction to meta-analysis*. UK: John Wiley & Sons, Ltd; 2009. p. 1–421.
41. Cohen J. *Statistical power analysis for the behavioral sciences*. New York, NY: Academic Press; 1977.
42. Cooper H, Hedges LV, Valentine JC. *The handbook of research synthesis and meta-analysis*. 2nd ed. New York, NY: Russell Sage Foundation; 2009.
43. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials*. 1986;7(3):177–88.
44. Faraone SV. Interpreting estimates of treatment effects: implications for managed care. *P T*. 2008;33(12):700–11.
45. Rücker G, Schwarzer G. Beyond the forest plot: the drapery plot. *Res Synth Methods*. 2021;12(1):13–9.
46. Barker TH, Migliavaca CB, Stein C, Colpani V, Falavigna M, Aromataris E, et al. Conducting proportional meta-analysis in different types of systematic reviews: a guide for synthesisers of evidence. *BMC Med Res Methodol*. 2021;21(1):189.
47. Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med*. 2002;21(11):1539–58.
48. Higgins J, Thompson S, Deeks J, Altman D. Statistical heterogeneity in systematic reviews of clinical trials: a critical appraisal of guidelines and practice. *J Heal Serv Res Policy*. 2002;7(1):51–61.
49. Higgins J, Thompson S, Deeks J, Altman D. Measuring inconsistency in meta-analyses. *BMJ*. 2003;327(7414):557–60.
50. Ioannidis JPA. Interpretation of tests of heterogeneity and bias in meta-analysis. *J Eval Clin Pract*. 2008;14(5):951–7.
51. Sterne JAC, Egger M. Funnel plots for detecting bias in meta-analysis: guidelines on choice of axis. *J Clin Epidemiol*. 2001;54(10):1046–55.
52. Egger M, Smith GD, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test measures of funnel plot asymmetry. *BMJ*. 1997;315(7109):629–34.
53. Duval S, Tweedie R. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*. 2000;56(2):455–63.
54. Shi L, Lin L, Omboni S. The trim-and-fill method for publication bias: practical guidelines and recommendations based on a large database of meta-analyses. *Med (United States)*. 2019;98(23):e15987.
55. Rothman KJ, Greenland S, Lash TL. *Modern epidemiology*. Philadelphia: Lippincott Williams & Wilkins; 2008.

56. Albrecht UV, 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;15(8):e182.
57. Baratz G, Sridharan PS, Yong V, Tatsuoka C, Griswold MA, Wish-Baratz S. Comparing learning retention in medical students using mixed-reality to supplement dissection: a preliminary study. *Int J Med Educ*. 2022;13:107–14.
58. Gnanasegaram JJ, Leung R, Beyea JA. Evaluating the effectiveness of learning ear anatomy using holographic models. *J Otolaryngol Head Neck Surg*. 2020;49(1):63.
59. Gray M, Downer T, Hartz D, Andersen P, Hanson J, Gao Y. The impact of three-dimensional visualisation on midwifery student learning, compared with traditional education for teaching the third stage of labour: a pilot randomised controlled trial. *Nurse Educ Today*. 2022;108:105184.
60. Hu KC, Salcedo D, Kang YN, Lin CW, Hsu CW, Cheng CY, et al. Impact of virtual reality anatomy training on ultrasound competency development: a randomized controlled trial. *PLoS One*. 2020;15(11):e0242731.
61. Imai T, Tanaka Y, Hatanaka Y, Suetsugu T, Sato Y, Matsuhashi N, et al. Incorporation of virtual reality in the clinical training of medical students studying esophageal and mediastinal anatomy and surgery. *Surg Today*. 2022;52(8):1212–7.
62. 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 Anat Anzeiger*. 2015;201:91–8.
63. Kūçük S, Kapakin S, Göktaş Y. Learning anatomy via mobile augmented reality: effects on achievement and cognitive load. *Anat Sci Educ*. 2016;9(5):411–21.
64. Kurul R, Ögün MN, Neriman Narin A, Avci Ş, Yazgan B. An alternative method for anatomy training: immersive virtual reality. *Anat Sci Educ*. 2020;13(5):648–56.
65. 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;32(2):238–43.
66. Moro C, Štromberga Z, Raikos A, Stirling A. The effectiveness of virtual and augmented reality in health sciences and medical anatomy. *Anat Sci Educ*. 2017;10(6):549–59.
67. Noll C, von Jan U, Raap U, Albrecht UV. Mobile augmented reality as a feature for self-oriented, blended learning in medicine: randomized controlled trial. *JMIR Mhealth Uhealth*. 2017;5(9):e139.
68. Barmaki R, Yu K, Pearlman R, Shingles R, Bork F, Osgood GM, et al. Enhancement of anatomical education using augmented reality: an empirical study of body painting. *Anat Sci Educ*. 2019;12(6):599–609.
69. Pickering JD, Panagiotis A, Ntakakis G, Athanassiou A, Babatsikos E, Bamidis PD. Assessing the difference in learning gain between a mixed reality application and drawing screencasts in neuroanatomy. *Anat Sci Educ*. 2022;15(3):628–35.
70. Stepan K, Zeiger J, Hanchuk S, Del Signore A, Shrivastava R, Govindaraj S, et al. Immersive virtual reality as a teaching tool for neuroanatomy. *Int Forum Allergy Rhinol*. 2017;7(10):1006–13.
71. Stojanovska M, Tingle G, Tan L, Ulrey L, Simonson-Shick S, Mlakar J, et al. Mixed reality anatomy using Microsoft HoloLens and cadaveric dissection: a comparative effectiveness study. *Med Sci Educ*. 2020;30(1):173–8.
72. Veer V, Phelps C, Moro C. Incorporating mixed reality for knowledge retention in physiology, anatomy, pathology, and pharmacology interdisciplinary education: a randomized controlled trial. *Med Sci Educ*. 2022;32(6):1579–86.
73. de Faria JWV, Teixeira MJ, de Moura Sousa Júnior L, Otoch JP, Figueiredo EG. Virtual and stereoscopic anatomy: when virtual reality meets medical education. *J Neurosurg*. 2016;125(5):1105–11.
74. von Schnakenburg P, Heermann S, Kromeier J, Offergeld C. Einsatz von Virtual Reality in der HNO-Lehre: eine Alternative zum konventionellen Anatomiemodell. *HNO*. 2023;71(2):106–13.
75. Weeks JK, Pakpoor J, Park BJ, Robinson NJ, Rubinstein NA, Prouty SM, et al. Harnessing augmented reality and CT to teach first-year medical students head and neck anatomy. *Acad Radiol*. 2021;28(6):871–6.
76. Bogomolova K, Ham IJM, Dankbaar MEW, Broek WW, Hovius SER, 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;13(5):558–67.
77. Bogomolova K, Vorstenbosch MATM, El Messaoudi I, Holla M, Hovius SER, van der Hage JA, et al. Effect of binocular disparity on learning anatomy with stereoscopic augmented reality visualization: a double center randomized controlled trial. *Anat Sci Educ*. 2023;16(1):87–98.
78. Bork F, Stratmann L, Enssle S, Eck U, Navab N, Waschke J, et al. The benefits of an augmented reality magic Mirror system for integrated radiology teaching in gross anatomy. *Anat Sci Educ*. 2019;12(6):585–98.
79. Cai S, He Y, Cui H, Zhou X, Zhou D, Wang F, et al. Effectiveness of three-dimensional printed and virtual reality models in learning the morphology of craniovertebral junction deformities: a multicentre, randomised controlled study. *BMJ Open*. 2020;10(9):e036853.
80. Du YC, Fan SC, Yang LC. The impact of multi-person virtual reality competitive learning on anatomy education: a randomized controlled study. *BMC Med Educ*. 2020;20(1):343.
81. Ekstrand C, Jamal A, Nguyen R, Kudryk A, Mann J, Mendez I. Immersive and interactive virtual reality to improve learning and retention of neuroanatomy in medical students: a randomized controlled study. *CMAJ Open*. 2018;6(1):E103–E109.
82. Ellington DR, Shum PC, Dennis EA, Willis HL, Szychowski JM, Richter HE. Female pelvic floor immersive simulation: a randomized trial to test the effectiveness of a virtual reality anatomic model on resident knowledge of female pelvic anatomy. *J Minim Invasive Gynecol*. 2019;26(5):897–901.
83. Wainman B, Aggarwal A, Birk SK, Gill JS, Hass KS, Fenesi B. Virtual dissection: an interactive anatomy learning tool. *Anat Sci Educ*. 2021;14(6):788–98.
84. Gloy K, Weyhe P, Nerenz E, Kaluschke M, Uslar V, Zachmann G, et al. Immersive anatomy atlas: learning factual medical knowledge in a virtual reality environment. *Anat Sci Educ*. 2022;15(2):360–8.
85. Milykh I, Kafarov ES, Covantsev S, Dadashev AS, Skarlis AA, Zenin OK. A new dimension in medical education: virtual reality in anatomy during COVID-19 pandemic. *Clin Anat*. 2023;36(7):1007–15.
86. Sinha S, DeYoung V, Nehru A, Brewer-Deluce D, Wainman BC. Determinants of learning anatomy in an immersive virtual reality environment—a scoping review. *Med Sci Educ*. 2022;33(1):287–97.
87. Elmqaddem N. Augmented reality and virtual reality in education. Myth or reality? *Int J Emerg Technol Learn*. 2019;14(3):234.
88. Wilson AB, Brown KM, Misch J, Miller CH, Klein BA, Taylor MA, et al. Breaking with tradition: a scoping meta-analysis analyzing the effects of student-centered learning and computer-aided instruction on student performance in anatomy. *Anat Sci Educ*. 2019;12(1):61–73.
89. Estes CA. Promoting student-centered learning in experiential education. *J Exp Educ*. 2004;27(2):141–60.
90. Lee E, Hannafin MJ. A design framework for enhancing engagement in student-centered learning: own it, learn it, and share it. *Educ Technol Res Dev*. 2016;64(4):707–34.

91. Krause KJ, Mullins DD, Kist MN, Goldman EM. Developing 3D models using photogrammetry for virtual reality training in anatomy. *Anat Sci Educ*. 2023;16(6):1033–40.
92. Dolega-Dolegowski D, Proniewska K, Dolega-Dolegowska M, Pregarzewska A, Hajto-Bryk J, Trojak M, et al. Application of holography and augmented reality based technology to visualize the internal structure of the dental root—a proof of concept. *Head Face Med*. 2022;18(1):12.
93. da Cruz TM, Menezes JM, Belchior G, Mazzotti FP, Bittar JS, dos Santos GGR, et al. Virtual reality as a complementary learning tool in anatomy education for medical students. *Med Sci Educ*. 2023;33(2):507–16.
94. Levine AI, DeMaria S Jr, Schwartz AD, Sim AJ. *The comprehensive textbook of healthcare simulation*. New York: Springer Science & Business Media; 2013.
95. Huang Y, Hu Y, Chan U, Lai P, Sun Y, Dai J, et al. Student perceptions toward virtual reality training in dental implant education. *PeerJ*. 2023;11:e14857.
96. Gordon HW. The cognitive laterality battery: tests of specialized cognitive function. *Int J Neurosci*. 1986;29(3–4):223–44.
97. Yohannan DG, Oommen AM, Amogh BJ, Raju NK, Suresh RO, Nair SJ. “Air anatomy”—teaching complex spatial anatomy using simple hand gestures. *Anat Sci Educ*. 2022;15(3):552–65.
98. Ho S, Liu P, Palombo DJ, Handy TC, Krebs C. The role of spatial ability in mixed reality learning with the HoloLens. *Anat Sci Educ*. 2022;15(6):1074–85.
99. Brewer-Deluce D, Bak AB, Simms AJ, Sinha S, Mitchell JP, Shin D, et al. Virtual reality bell-ringer: the development and testing of a stereoscopic application for human gross anatomy. *Anat Sci Educ*. 2021;14(3):330–41.
100. Vishwanath A, Kam M, Kumar N. Examining low-cost virtual reality for learning in low-resource environments. *DIS 2017—proceedings of the 2017 ACM conference on designing interactive systems*. New York, NY: Association for Computing Machinery, Inc; 2017. p. 1277–81.
101. Wilson AB, Brooks WS, Edwards DN, Deaver J, Surd JA, Pirlo OJ, et al. Survey response rates in health sciences education research: a 10-year meta-analysis. *Anat Sci Educ*. 2024;17(1):11–23.
102. Gurevitch J, Koricheva J, Nakagawa S, Stewart G. Meta-analysis and the science of research synthesis. *Nature*. 2018;555(7695):175–82.
103. Armijo-Olivo S, Fuentes J, da Costa BR, Saltaji H, Ha C, Cummings GG. Blinding in physical therapy trials and its association with treatment effects. *Am J Phys Med Rehabil*. 2017;96(1):34–44.

AUTHOR BIOGRAPHIES

Paloma García-Robles, Physiotherapist and PhD Student at University of Jaén.

Irene Cortés-Pérez, Physiotherapist, Doctor (PhD), researcher and Professor at University of Jaén.

Francisco Antonio Nieto-Escámez, Psychologist, Doctor (PhD), researcher and Professor at University of Almería

Héctor García-López, Physiotherapist, Doctor (PhD), researcher and Professor at University of Almería

Esteban Obrero-Gaitán, Physiotherapist, Doctor (PhD), researcher and Professor at University of Jaén

María Catalina Osuna-Pérez, Physiotherapist, Doctor (PhD), researcher and Professor at University of Jaén

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: García-Robles P, Cortés-Pérez I, Nieto-Escámez FA, García-López H, Obrero-Gaitán E, Osuna-Pérez MC. Immersive virtual reality and augmented reality in anatomy education: A systematic review and meta-analysis. *Anat Sci Educ*. 2024;17:514–528. <https://doi.org/10.1002/ase.2397>