

Bond University
Research Repository



Virtual and augmented reality enhancements to medical and science student physiology and anatomy test performance: A systematic review and meta-analysis

Moro, Christian; Birt, James; Stromberga, Zane; Phelps, Charlotte; Clark, Justin; Glasziou, Paul; Scott, Anna Mae

Published in:
Anatomical Sciences Education

DOI:
[10.1002/ase.2049](https://doi.org/10.1002/ase.2049)

Licence:
Other

[Link to output in Bond University research repository.](#)

Recommended citation(APA):
Moro, C., Birt, J., Stromberga, Z., Phelps, C., Clark, J., Glasziou, P., & Scott, A. M. (2021). Virtual and augmented reality enhancements to medical and science student physiology and anatomy test performance: A systematic review and meta-analysis. *Anatomical Sciences Education*, 14(3), 368-376.
<https://doi.org/10.1002/ase.2049>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

For more information, or if you believe that this document breaches copyright, please contact the Bond University research repository coordinator.



DR. CHRISTIAN MORO (Orcid ID : 0000-0003-2190-8301)

Article type : Relevant Review

ASE-20-0254.R1

Relevant Review

**Virtual and augmented reality enhancements to medical and science student
physiology and anatomy test performance:
A systematic review and meta-analysis.**

**Christian Moro^{1,*}, James Birt², Zane Stromberga¹, Charlotte Phelps¹, Justin Clark³, Paul
Glasziou³, Anna Mae Scott³**

¹ Faculty of Health Sciences and Medicine, Bond University, Gold Coast, Queensland, Australia

² Faculty of Society and Design, Bond University, Gold Coast, Queensland, Australia

³ Institute for Evidence-Based Healthcare, Bond University, Gold Coast, Queensland, Australia

Running title: Virtual and augmented reality in medicine and science

*Correspondence to: Dr. Christian Moro. Faculty of Health Sciences and Medicine, Bond University,
Gold Coast, QLD 4229, Australia. E-mail: cmoro@bond.edu.au

This article has been accepted for publication and undergone full peer review but has not been
through the copyediting, typesetting, pagination and proofreading process, which may lead to
differences between this version and the [Version of Record](#). Please cite this article as [doi:
10.1002/ase.2049](#)

This article is protected by copyright. All rights reserved

This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Use of Self-Archived Versions.

ORCID NUMBERS:

CHRISTIAN MORO: <https://orcid.org/0000-0003-2190-8301>

JAMES BIRT: <https://orcid.org/0000-0002-0422-4867>

ZANE STROMBERGA: <https://orcid.org/0000-0003-1896-7300>

CHARLOTTE PHELPS: <https://orcid.org/0000-0002-4217-0214>

JUSTIN CLARK: <https://orcid.org/0000-0003-0133-1613>

PAUL GLASZIOU: <https://orcid.org/0000-0001-7564-073X>

ANNA MAE SCOTT: <https://orcid.org/0000-0002-0109-9001>

ABSTRACT

Virtual and augmented reality have seen increasing employment for teaching within medical and health sciences programs. For disciplines such as physiology and anatomy, these technologies may disrupt the traditional modes of teaching and content delivery. The objective of this systematic review and meta-analysis is to evaluate the impact of virtual reality or augmented reality on knowledge acquisition for students studying pre-clinical physiology and anatomy. The protocol was submitted to Prospero and literature search undertaken in PubMed, Embase, ERIC, and other databases. Citations were reviewed and articles published in full assessing learning or knowledge acquisition in pre-clinical physiology and anatomy from virtual or augmented reality were included. Of the 919 records found, 52 eligible articles were reviewed in full-text, with eight studies meeting full eligibility requirements. There was no significant difference in knowledge scores from combining the eight studies (626 participants), with the pooled difference being a non-significant increase of 2.9 percentage points (95% CI[-2.9;8.6]). For the four studies comparing virtual reality to traditional teaching, the pooled treatment effect difference was 5.8 percentage points (95% CI[-4.1;15.7]). For the five studies comparing augmented reality to traditional teaching, the pooled treatment effect difference was 0.07 (95% CI[-7.0;7.2]). Upon review of the literature, it is apparent that educators could benefit from adopting assessment processes that evaluate three-dimensional spatial understanding as a priority in physiology and anatomy. The overall evidence suggests that although test performance is not significantly enhanced with either mode, both virtual and augmented reality are viable alternatives to traditional methods of education in health sciences and medical courses.

Keywords: Systematic review, meta-analysis, health sciences education; anatomy education, physiology education, medical education; virtual reality; augmented reality; mixed reality; technology-enhanced education.

INTRODUCTION

Medical students learn anatomy and physiology most effectively through a combination of memorization, understanding, and visualization (Pandey and Zimitat, 2007). This in turn is assessed by the ability to recall three-dimensional or spatial relationships between structures (Gonzales et al., 2020). Traditionally, medical students were taught anatomy and physiology through the use of dissections or prosections alongside two-dimensional (2D) textbook materials (Estai and Bunt, 2016). As the volume of information required to learn in a modern-day medical course increases, educators are moving into technology-enhanced resources to provide engaging, interactive, and authentic learning experiences (Moro et al., 2020c). Nowadays, a range of other tools has been adopted in medical education as a way to enhance students' spatial understanding (Kuehn, 2018). This includes mobile applications (Stirling and Birt, 2014) three-dimensional (3D) printed models (Su et al., 2018), body painting (Barmaki et al., 2019), gamified learning techniques (Moro et al., 2020b) and virtual dissection tables (Darras et al., 2019; Periya and Moro, 2019).

More recently, virtual and augmented reality have emerged into disciplines such as physiology and anatomy and have shown some effectiveness for enhancing learning (Birt et al., 2018). Medical schools are increasingly incorporating aspects of virtual reality (using head-mounted displays) and augmented reality into their science curriculum to facilitate teaching and learning (Moro et al., 2017; Kuehn, 2018). The value of virtual and augmented reality has been assessed in anatomical education by implementing these tools to display complex 3D anatomical structures, replicate dissection courses (Bork et al., 2019), and even immerse medical students in surgical procedures (Cao and Cerfolio, 2019). These modes of education delivery have been found to be powerful for increasing spatial understanding of medical students, particularly in low spatial ability students (Bork et al., 2019). However, the benefits and effectiveness of these devices for student learning remain unclear (Elmqaddem, 2019).

Determining the viability of these teaching tools is of importance as universities have been consistently increasing their use of technology to supplement learning within health sciences in recent years (Zargaran et al., 2020). The present study's aim was to systematically review the evidence for

the impact of virtual reality or augmented reality, on knowledge acquisition for students studying pre-clinical physiology and anatomy.

Background

Comprehension of human anatomy involves significant intellectual effort, where the learner must identify the diverse anatomical structures, internal organization, and their relationship to other structures of the body (Pandey and Zimitat, 2007). Similarly, the study of physiology involves understanding the functions of activities of life, such as in cells, tissues, or organs. Traditional modes of learning physiology and anatomy, such as textbook illustrations or lecture slides, are two-dimensional with an increasing criticism that they are becoming outdated, passive, and lacking experiential learning affordances (Estai and Bunt, 2016). This has led to a shift away from these educational practices through digital enrichment using rotatable models or computer-based learning (Stirling and Birt, 2014; Stirling and Moro, 2020), as well as enhancing variety through gamification methods such as interactive quizzing (Moro and Stromberga, 2020) and serious games (Moro et al., 2020b).

The anatomy of the human body, physiology, pathology, and other spatial structures can be visualized in three-dimensional (3D) space using virtual and augmented reality. Using this type of technology may assist in the learner's overall understanding of the organ and concept in question, while also facilitating an understanding of its relationships with nearby structures (Langlois et al., 2020). However, there are ever-increasing concerns over the introduction of these technologies with a potential lack of academic skill base to teach effectively, as well as higher costs, and reduced scalability to cohorts of larger students (Elmqaddem, 2019; Losco et al., 2017; Moro et al., 2020c). In addition, in many cases academics may be unwilling to adopt the new learning technologies (Liu et al., 2020), preferring to teach in more traditional means, creating barriers in faculties looking to integrate new modes of instruction such as virtual and augmented reality. As institutions consider increasing the amounts of funding and time invested in the incorporation of new technologies, such as virtual and augmented reality into medical schools, there needs to be evidence to support their inclusion.

Boyles (2017) explored the use of augmented and virtual reality within the United States Military Academy across several different disciplines. It was found that the immersive technology offered significant strengths in shifting the role of the teacher from didactic deliverer to learner centered facilitator. The ability to learn experientially and proceed at their own pace were major positives in the use of the technology. Liou et al. (2017) explored the use of augmented reality and virtual reality on student learning and found that real objects presented using augmentation reduced the mental load on students because learners could anchor objects as reference and encourage students to conduct tasks. Huang et al., (2019) also found that augmented reality was effective in conveying information through the pathway of spatial presence. Sattar et al., (2020) explored the motivating factors in medical students using virtual reality and found that user experience, perceived competence, usefulness and motivation were high for medical students. Vasilevski and Birt (2020) conducted a thematic analysis of student experiences of virtual and augmented reality. It was found that using immersive technology resulted in an enhanced learning environment that facilitated unique learning experiences, engagement, and motivation however the processes leading to these learning aspects required further analysis.

There have been systematic reviews completed on health professions education (Kyaw et al., 2019), as well as surgical skill development in a range of specializations (Vitale et al., 2020). However, these reviews did not assess whether or not virtual reality or augmented reality were viable educational tools for pre-clinical physiology and anatomy education. Zhao et al. (2020) completed a systematic review on anatomy teaching using virtual reality across randomized control trials although differed from the present study as its focus included learners in clinical environments and without the inclusion of physiology (Hu et al., 2019). The present study addresses this by systematically reviewing the impact of virtual reality or augmented reality on knowledge acquisition for students studying anatomy and pre-clinical physiology.

Defining spatially immersive visualization technologies

Although anatomical structures and how they relate to health and disease states can be relatively complex, three-dimensional shapes and disease education is primarily taught using two-dimensional resources. These traditional teaching modes include lecture slides and tutorial notes utilizing 2D-based illustrations and text-focused resources. As a result of technological advancements, traditional learning modes have now been disrupted, providing the educator with a range of technology-enhanced devices to implement as supplementary tools (Moro et al., 2020a). In particular, virtual reality, augmented reality, and mixed reality applications have demonstrated the potential to enhance knowledge acquisition and the overall learning experience. The following definitions of spatially immersive visualization technologies have been adapted from Moro et al. (2017):

Virtual reality: The user is fully immersed in an artificial environment, which is experienced through sensory stimuli (sight, hearing, and motion) that mimics the properties of the real world through constantly updating, high-resolution head-mounted displays, stereo headphones and motion-tracking systems.

Augmented reality: Using a camera and screen (i.e. smartphone or tablet), digital models are superimposed onto the real-world in augmented reality (AR). The user is then able to interact with both the real and virtual elements of their surrounding environment.

Mixed reality: While augmented reality overlays digital information onto real-world elements, mixed reality (MR) allows for an additional layer of interactivity. The visual displays within mixed reality devices are presented as holographic renderings of images anchored into the real world that can be interacted with as if they were situated objects.

Extended Reality: The umbrella term used in this study to refer to virtual reality, augmented reality and mixed reality hardware collectively (XR) as a spatially immersive ecosystem to enable immersive learning.

In this study the term 'traditional teaching modes' will be employed for studies utilizing lectures, cadaveric material, workshops, textbooks, or educational technology that can only depict models and objects in 2D. Alternatively, studies were merged under XR when the technological visualization methods were capable of rendering immersive 3D representations of the human body. This allowed a direct comparison between these modes when compared to traditional 2D teaching methods for instruction. The rationale of not using the term 'mixed reality' as defined by Milgram and Kishino (1994), is the current misunderstanding in the literature. The original definition by Milgram states that MR is a continuum and not necessary as a classification or a device type. Therefore, XR will be used as the umbrella term with a focus on VR and AR as categories of devices. Studies including MR will be categorized under VR or AR where best appropriate.

MATERIALS AND METHODS

Protocol

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement was followed for the review's outline. The protocol was developed prospectively and submitted to Prospero, an International Prospective Register of Systematic Reviews (Centre for Reviews and Dissemination, University of York, York, UK) on 11 November 2019 (registration number 158040). This protocol was followed with only very minor deviations where required, and these are reported in the discussion section.

Included studies

Randomized controlled trials of any design (e.g. parallel, cluster) were included, which involved participants currently enrolled in health sciences or medical programs in universities. Studies that compared the use of virtual reality via head-mounted displays, or augmented reality, as teaching tools, to any traditionally used methods in teaching (e.g. textbooks, lecture slides, computer or tablet models, smartphone applications) were included. Studies were only included if they assessed learning, such as through examination or quiz scores from students studying pre-clinical physiology and anatomy. Studies based on simulations, simulated patients, clinical skills or clinical environments were not included, as these are not core components of a contemporary physiology and anatomy

This article is protected by copyright. All rights reserved

curriculum. The primary outcome was pre-clinical knowledge scores, and the secondary outcome was skills scores. Studies investigating devices not yet commercially available, such as the Microsoft HoloLens (Microsoft Corp., Redmond, WA) or Magic Leap (Magic Leap, Inc., Plantation, FL), were not included.

Search strategy

Databases searched were PubMed (U.S. National Library of Medicine, Bethesda MD), Embase (Elsevier, Amsterdam, The Netherlands), Cochrane Central Register of Controlled Trials (CENTRAL, Cochrane, London, UK), the Education Resources Information Center (ERIC, Institute of Education Sciences (IES), Washington, DC), and clinical trial registries (U.S. National Library of Medicine, Bethesda MD), and the WHO International Clinical Trials Registry Platform (ICTRP, World Health Organization, Geneva, Switzerland), from January 1990 to 19 November 2019 (see Supplemental Material Appendix 2). The search start date was restricted to 1990, as that year saw the first widespread commercial releases of virtual reality Head-Mounted Display consumer headsets and augmented reality devices (Moro et al., 2020c). Forward- and backward-citation searches were also conducted from the included studies (21 January 2020). Only studies published in-full were included; studies published as abstract only (e.g. conference abstract) were excluded. No language restrictions were applied.

Study selection and data extraction

Screening of the literature was conducted by two authors independently (J.B., C.M.), first by title/abstract, and subsequently in full text. Discrepancies were resolved by consensus or referring to the third author if necessary (A.M.S.). Three data extraction forms were used: Table of Characteristics form, Primary Outcomes data form, Risk of Bias form; the data extraction forms were pre-piloted on two studies. Data from included studies were extracted independently by two authors (Z.S., C.P.), with discrepancies resolved by discussion, or by the third author (A.M.S.). In five cases, the primary author was contacted and asked to provide additional information, with all responding. In four cases, authors provided raw data, which was then included in the calculations in place of the summarized data in their original publication.

This article is protected by copyright. All rights reserved

Risk of bias

The risk of bias was assessed by two authors independently (Z.S., C.P.) using the Cochrane Collaboration's Risk of Bias tool 1 (Higgins et al., 2019). Discrepancies were resolved by discussion or reference to a third author (A.M.S.).

The following domains were assessed: (1) Random sequence generation, (2) Allocation concealment, (3) Blinding of participants and personnel, (4) Blinding of outcome assessment, (5) Incomplete outcome data, (6) Selective outcome reporting, and (7) Other bias (focusing on potential for biases due to funding or conflict of interest).

Each potential source of bias was graded as low, high, or unclear. Each judgement was supported by a quote from the relevant trial.

Statistical analysis

Data were sufficient to conduct a meta-analysis; Review Manager software, version 5.4 (Cochrane Collaboration, London, UK) was used to calculate the effect size. Results were reported using mean difference (MD) and 95% confidence intervals, and a random effects model used. Heterogeneity among the included trials was measured using the I^2 statistic. A One ad-hoc sensitivity analysis was completed by including versus excluding a study with three domains rated at high risk of bias.

Where data were presented as mean and confidence interval in the original studies, the confidence interval was converted to standard deviation (Higgins et al., 2019). As fewer than 10 trials were included, a funnel plot was not created.

RESULTS

Search results

Electronic searches identified 760 references, and the forward and backward citation searches identified 438 references. After de-duplication, 919 references were screened on title/abstract, and

861 were excluded. Full texts for 58 references were obtained, excluding 50 (see Supplemental Material Appendix 3 for reasons), and 8 references describing 8 studies were included (Figure 1).

[Insert Figure 1 here]

Risk of bias

The risk of bias was low or unclear for most of the domains. High risk of bias was reported for blinding of participants and personnel, as it is not possible to blind participants to use of virtual reality equipment. Allocation concealment was generally unclearly reported, but no bias from selective reporting was identified across the included studies (see Figure 2).

[Insert Figure 2 here]

Characteristics of the included studies

Eight trials were included (Albrecht et al., 2013; Küçük et al., 2016; Moro et al., 2017; Noll et al., 2017; Stepan et al., 2017; Ekstrand et al., 2018; Barmaki et al., 2019; Maresky et al., 2019) (see Table 1, and Supplemental Material Appendix 1 for additional detail). Studies were conducted in Germany (n = 2), United States (n = 2), Canada (n = 2), Turkey (n = 1), and Australia (n = 1).

Studies were two-arm parallel randomized trials, except for Moro et al. (2017) (a three-arm parallel trial) and (Barmaki et al., 2019) (a cluster RCT). The studies included a total of 626 participants (the AR arm from one three-arm study evaluating VR vs AR vs traditional teaching was excluded from the analysis (Albrecht et al., 2013; Noll et al., 2017; Maresky et al., 2019)). Most trials included were medium-sized, with only 3 containing fewer than 50 participants (Albrecht et al., 2013; Noll et al., 2017; Maresky et al., 2019); the largest study included 288 participants.

Four studies assessed virtual reality (Moro et al., 2017; Stepan et al., 2017; Ekstrand et al., 2018; Maresky et al., 2019), using VR head-mounted displays; 3 studies evaluated Oculus Rift (Moro et al., 2017; Stepan et al., 2017; Maresky et al., 2019) and one evaluated HTC Vive (Ekstrand et al., 2018).

This article is protected by copyright. All rights reserved

The remaining studies evaluated augmented reality, using: mobile AR applications (Albrecht et al., 2013; Küçük et al., 2016; Noll et al., 2017), or mirrored AR displayed on a television screen (Barmaki et al., 2019).

Comparators included two phone applications (Albrecht et al., 2013; Noll et al., 2017), one textbook with virtual mirror (Barmaki et al., 2019), two paper-based (Küçük et al., 2016; Ekstrand et al., 2018), one tablet (Moro et al., 2017), one web-based textbook (Stepan et al., 2017) and one did not report materials (Maresky et al., 2019).

Four studies measured the knowledge scores of participants by multiple-choice pre- and post-tests (Albrecht et al., 2013; Noll et al., 2017; Ekstrand et al., 2018; Maresky et al., 2019). Other methods of measuring knowledge scores included a single multiple-choice post-test (Küçük et al., 2016; Moro et al., 2017), matching labels pre- and post-test (Barmaki et al., 2019), as well as Stepan et al. (2017) undertaking a multiple-choice and fill-in-the-blanks test for their pre-test and a multiple-choice, fill-in-the-blanks and labelling models post-test (Table 1).

[Insert Table 1]

Outcomes

Knowledge scores using XR devices (VR or AR) were reported in eight (8) studies, incorporating a total of 626 participants. The mean difference in knowledge scores between XR and traditional teaching groups was 2.0% (95% CI -2.9 to 8.6, $P = 0.33$), however the difference was not statistically significant. Heterogeneity was high ($I^2 = 72\%$) (Figure 3).

[Insert Figure 3 here]

A post-hoc sensitivity analysis was conducted to establish the source of heterogeneity. Removal of one study (Maresky et al., 2019) from the meta-analysis decreases the heterogeneity considerably (I^2

= 52%), however, the difference between groups in knowledge scores remains non-significant (MD 0.8 95% CI -3.6 to 5.3, P = 0.71, See Supplemental Material Appendix 4).

As data was sufficient, a sub-group analysis was conducted, comparing VR to traditional teaching methods and AR to traditional teaching methods. Differences between groups in knowledge scores remained non-significant (MD 2.4%, 95% CI -3.0 to 7.9, P = 0.38), and heterogeneity remained high (I^2 of 76% and 68% respectively, Figure 4).

[Insert Figure 4 here]

DISCUSSION

This analysis was conducted to examine the impacts of learning physiology and anatomy through virtual and augmented reality on student learning and achievement. As discussed in Estai and Bunt (2016), new teaching resources and strategies are being used in physiology and anatomy education to increase motivation, delivery, and cost. Traditional methods to teach anatomy such as dissection, prosections, plastic and 3D printed models, medical imaging, and living anatomy are shown to be successful in educational delivery. However, there is a need to direct future research to the evaluation of new teaching methodologies to satisfy learning styles and delivery methods. This is further reiterated by Losco et al. (2017) that supports alternative methods through computer-assisted learning to improve the quality of materials and increase education efficiency.

It is noted in Zhao et al. (2020), and replicated in this present study, that there is no significance in the test mean scores between the control and XR interventions. These findings may not be entirely unexpected as the reported assessments and tests in the included studies focus on the traditional 2D structural recall, rather than the combination of spatial connection and structures. Application of these 3D skills is not usually performed until later years when simulated patients are fully utilized. Due to the highly variable nature of testing instruments used in assessing new technology in education (Lai and Bower, 2019), there is a need for future studies to use enhanced testing methods. These could

incorporate innovations that could assess the spatial skills of learners in both recognition and recall, which would be more beneficial in understanding whether traditional or spatial XR methods had an impact on real-world visualisation of physiological and anatomical content. This testing could be done before the administration of formal spatial assessments using cadavers and manikins to familiarize learners before more expensive testing methods. In addition, increased optimization of manikin and 3D physical model usage in the classroom could benefit learners who are more spatially prepared before these more costly interventions. This presents the requirement for new forms of computer assisted learning to allow for more efficient and effective methods of assessment.

A counter argument to the non-significant result is presented by Reeves and Lin (2020). It is argued that it is difficult for any single educational paper to have significant impact on practice and the focus of educational technology research should be on larger problems affecting education not specific questions about technology (Reeves 1995, 2004). Reeves (2015) proposes the use of design-based research which focuses on solutions and interventions to problems rather than the focus being on the significance of trials. McKenney and Reeves (2019) in their book outline the educational challenges but also the theoretical understanding that can reimaging how researchers can understand their work in practice and craft effective solutions to complex educational challenges. In terms of the heterogeneity of results as noted in the Zhao et al. (2020) study, experience has an impact on the result and therefore learners in earlier years can be more easily engaged with the XR tools. As the learners' gain experience, so do the requirements upon the visualization as learners bring in existing knowledge filling in the gaps not presented with these tools. The effectiveness of virtual and augmented reality depicted within the study is promising, as it enables learning when cadaveric specimens or hands-on approaches to instruction are unavailable. Modern trends include the introduction of newer, mixed reality augmented reality devices which may allow for both heightened learning and a more engaging experience from 3D models (Moro et al., 2020a). However, methods of assessment also need to be introduced to properly ensure equity in any introduced learning tool. For example, Wainman et al. (2021) identified that VR may be detrimental to students with low visuospatial ability when compared to using physical modes for learning anatomy.

The fact that there is no difference in learning between AR, VR or using traditional approaches is highly positive and validates the possibilities of using these innovative technologies in the classroom. This finding is reiterated by Losco et al. (2017) and Estai and Bunt (2016), noting that computer-assisted learning may not improve the scores, but the capacity to scale and have efficiency returns warrants the continued research in computer-assisted learning methods. This observation is also highlighted in other studies such as Stirling and Birt (2014), which reported no significant increases in learner test scores. This also highlights the emerging design based approach proposed by Reeves (1995, 2004, 2015) as a way to counter the problems associated with non-significant results.

Limitations of the study

This systematic review focused on knowledge gained. However, there are a number of other potential intrinsic benefits for students that are using this technology for learning, such as enhanced engagement with their learning materials and the ability to problem-solve collaboratively with their fellow students. In addition, only pre-clinical physiology and anatomy curricula were evaluated in the studies included in the review. This excluded the applications of these disciplines, such as assessing simulated patients or learning within a clinical environment. As such, this study's findings are restricted to the first two years of a medical or health sciences program, before students enter their clinical training. It is also noted that none of the included studies reported any secondary outcomes such as skill scores.

As outlined in the protocol, studies based on simulations, simulated patients, clinical skills or clinical environments were not included as these are not core components of contemporary physiology and anatomy in the pre-clinical years of medical and health sciences training. There were also limitations within the included studies. In particular, the lack of blinding of participants and personnel from the knowledge of which intervention a participant received. However, it is impossible to blind the participant on which they received as the technological devices are visually different from each other. While studies could not blind the participants, one study did (Maresky et al., 2019) successfully blinded the anatomy professor and teaching assistants as to student participation and had no access to the collected data, which future studies should consider. Also notable is the considerable degree of

unexplained heterogeneity which is worth exploring in future research. Finally, there were several necessary changes from the original protocol: (1) it was more straightforward to use mean differences rather than standardized mean differences, and (2) the sensitivity analyses was not pre-specified, but the need for these based on the high statistical heterogeneity was recognized.

CONCLUSIONS

This systematic review has identified consistent test results from students learning with either traditional teaching methods, or via the use of VR and AR in physiology and anatomy courses. This review provides evidence-based support for the introduction of VR and AR as delivery modes, without any adverse impact on student achievement. There is the chance that the use of this technology may have a highly positive impact on students spatial understanding and 3D comprehension of anatomical structures. However, this has not been adequately investigated in the literature, highlighting the need for academic programs to integrate more innovative techniques to accurately assess three-dimensional learning of physiology and anatomy in the modern age. This systematic review has also identified a need for additional long-term studies on learner retention of knowledge gained after receiving a lesson on either VR or AR devices. Overall, based on this study's findings it can be concluded that educators wishing to adopt VR or AR within medical or biomedical sciences programs can be confident that the new technology will not adversely impact learning or assessment results, and is a suitable alternative to traditional teaching practices.

ACKNOWLEDGEMENTS

The authors wish to thank Sevda Küçük, Katelyn Stepan and Chelsea Ekstrand for providing additional information about their published studies. The authors also appreciate the responses from Dee Ballyk and Hillel Maresky regarding their studies. The authors have no conflict of interest to report.

NOTES ON CONTRIBUTORS

CHRISTIAN MORO, B.Sc., B.Ed., M.Bus., Ph.D. (Physiology), is an associate professor and the science and scholarship domain lead for the medical program in the Faculty of Health Sciences and Medicine at Bond University, Gold Coast, Queensland, Australia. His laboratory research investigates the physiology associated with diseases of the urinary bladder, while his medical education research focusses on the implementation of novel technological tools to enhance student learning, participation and interaction.

JAMES BIRT, B.I.T. (Hon.), Ph.D., is an associate professor of information and computer sciences in the Faculty of Society and Design at Bond University, Gold Coast, Queensland, Australia. His research spans computer science, games, and visual arts, with an emphasis on applied design and development of interactive virtual and augmented reality experiences assisting learning, skills acquisition, and knowledge discovery.

ZANE STROMBERGA, B.Biom.Sci. (Hon.), Ph.D., is a researcher in physiology and pharmacology in the Faculty of Health Sciences and Medicine at Bond University, Gold Coast, Queensland, Australia. Her research interests involve urological functional studies and medical education.

CHARLOTTE PHELPS, B.Biom.Sci (Hon.), is a doctoral student in physiology and pharmacology, and also an education researcher in the Medical and Biomedical Sciences Program, Faculty of Health Sciences and Medicine at Bond University, Gold Coast, Queensland, Australia. Her research interests involve disorders of the lower urinary tract, and using mixed reality to enhance teaching and student engagement in organ system's education.

JUSTIN CLARK, B.A., Lib.Tech., is a Cochrane information specialist for the Cochrane Acute Respiratory Infections Group and is the senior research information specialist in the Institute of Evidence-Based Healthcare at Bond University, Gold Coast, Queensland, Australia. His research interests involve improving the speed and quality of the conduct of systematic reviews.

This article is protected by copyright. All rights reserved

This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Use of Self-Archived Versions.

PAUL GLASZIOU, M.B.B.S., Ph.D., is a professor in the Institute of Evidence-Based Practice, at Bond University, Queensland, Australia. His research interests include antibiotic resistance, overdiagnosis, neglected non-pharmaceutical treatments and waste in medical research.

ANNA MAE SCOTT, B.A., M.A., M.Sc., Ph.D., is an assistant professor in the Institute of Evidence-Based Practice at Bond University, Queensland, Australia. Her research focuses on advancing the methodology for conducting systematic reviews, most recently developing and trailing the "2 week systematic review" process.

LITERATURE CITED

Albrecht UV, Folta-Schoofs K, Behrends M, von Jan U. 2013. Effects of mobile augmented reality learning compared to textbook learning on medical students: Randomized controlled pilot study. *J Med Internet Res* 15:e182.

Barmaki R, Yu K, Pearlman R, Shingles R, Bork F, Osgood GM, Navab N. 2019. Enhancement of anatomical education using augmented reality: An empirical study of body painting. *Anat Sci Educ* 12:599–609.

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.

Bork F, Stratmann L, Enssle S, Eck U, Navab N, Waschke J, Kugelmann D. 2019. The benefits of an augmented reality magic mirror system for integrated radiology teaching in gross anatomy. *Anat Sci Educ* 12:585–598.

Boyles B. 2017. Virtual reality and augmented reality in education. West Point, NY: Center For Teaching Excellence, United States Military Academy. Master Teaching Program Dissertation. 12 p.

Cao C, Cerfolio RJ. 2019. Virtual or augmented reality to enhance surgical education and surgical planning. *Thorac Surg Clin* 29:329–337.

Darras KE, Spouge R, Hatala R, Nicolaou S, Hu J, Worthington A, Krebs C, Forster BB. 2019. Integrated virtual and cadaveric dissection laboratories enhance first year medical students' anatomy experience: A pilot study. *BMC Med Educ* 19:366.

Ekstrand C, Jamal A, Nguyen R, Kudryk A, Mann J, Mendez I. 2018. Immersive and interactive virtual reality to improve learning and retention of neuroanatomy in medical students: a randomized controlled study. *CMAJ Open* 6:E103–E109.

Elmqaddem N. 2019. Augmented reality and virtual reality in education. Myth or reality? *Int J Emerging Technol Learn* 14:234–242.

Estai M, Bunt S. 2016. Best teaching practices in anatomy education: A critical review. *Ann Anat* 208:151–157.

Gonzales RA, Ferns G, Vorstenbosch M, Smith CF. 2020. Does spatial awareness training affect anatomy learning in medical students? *Anat Sci Educ* 13:707–720.

Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (Editors). 2019. *Cochrane Handbook for Systematic Reviews of Interventions*. 2nd Ed. Chichester, UK: John Wiley & Sons Ltd. 728 p.

Hu HZ, Feng XB, Shao ZW, Xie M, Xu S, Wu XH, Ye ZW. 2019. Application and prospect of mixed reality technology in medical field. *Curr Med Sci* 39:1–6.

Huang KT, Ball C, Francis J, Ratan R, Boumis J, Fordham J. 2019. Augmented versus virtual reality in education: An exploratory study examining science knowledge retention when using augmented reality/virtual reality mobile applications. *Cyberpsychol Behav Social Netw* 22:105–110.

Küçük S, Kapakin S, Göktaş Y. 2016. Learning anatomy via mobile augmented reality: Effects on achievement and cognitive load. *Anat Sci Educ* 9:411–421.

Kuehn BM. 2018. Virtual and augmented reality put a twist on medical education. *JAMA* 319:756–758.

This article is protected by copyright. All rights reserved

This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Use of Self-Archived Versions.

Kyaw BM, Saxena N, Posadzki P, Vseteckova J, Nikolaou CK, George PP, Divakar U, Masiello I, Kononowicz AA, Zary N, Tudor Car L. 2019. Virtual reality for health professions education: Systematic review and meta-analysis by the digital health education collaboration. *J Med Internet Res* 21:e12959.

Lai JW, Bower M. 2019. How is the use of technology in education evaluated? A systematic review. *Comput Educ* 133:27–42.

Langlois J, Bellemare C, Toulouse J, Wells GA. 2020. Spatial abilities training in anatomy education: A systematic review. *Anat Sci Educ* 13:71–79.

Liou HH, Yang SJ, Chen SY, Tarng W. 2017. The influences of the 2D image-based augmented reality and virtual reality on student learning. *J Educ Technol Soc* 20:110–121.

Liu Q, Geertshuis S, Grainger R. 2020. Understanding academics' adoption of learning technologies: A systematic review. *Comput Educ* 151:103857.

Losco CD, Grant WD, Armson A, Meyer AJ, Walker BF. 2017. Effective methods of teaching and learning in anatomy as a basic science: A BEME systematic review: BEME guide no. 44. *Med Teach* 39:234243.

Maresky HS, Oikonomou A, Ali I, Ditkofsky N, Pakkal M, Ballyk B. 2019. Virtual reality and cardiac anatomy: Exploring immersive three-dimensional cardiac imaging, A pilot study in undergraduate medical anatomy education. *Clin Anat* 32:238–243.

McKenney S, Reeves TC. 2019. *Conducting Educational Design Research*. 2nd Ed. Abingdon, Oxon, UK: Routledge. 256 p.

Milgram P, Kishino F. 1994. A Taxonomy of mixed reality visual displays. *IEICE Trans Info Syst* 12:1321–1329.

Moher D, Liberati A, Tetzlaff J, Altman DG. 2009. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Med* 6:e1000097.

Moro C, Phelps C, Redmond P, Stromberga, Z. 2020a. , HoloLens and mobile augmented reality in medical and health science education: A randomised controlled trial. *Br J Educ Technol* (in press; <https://doi.org/10.1111/bjet.13049>).

Moro C, Phelps C, Stromberga Z. 2020b. Utilizing serious games for physiology and anatomy learning and revision. *Adv Physiol Educ* 44:505–507.

Moro C, Stromberga Z, Birt J. 2020c. Technology considerations in health professions and clinical education. In: Nestel D, Reedy G, McKenna L, Gough S (Editors). *Clinical Education for the Health Professions: Theory and Practice*. 1st Ed. Singapore, Singapore: Springer Nature Singapore Pte Ltd. Pp 1-22.

Moro C, Štromberga Z, Raikos A, Stirling A. 2017. The effectiveness of virtual and augmented reality in health sciences and medical anatomy. *Anat Sci Educ* 10:549–559.

Moro C, Stromberga Z. 2020. Enhancing variety through gamified, interactive learning experiences. *Med Educ* 54:1180–1181.

Noll C, von Jan U, Raap U, Albrecht UV. 2017. Mobile augmented reality as a feature for self-oriented, blended learning in medicine: Randomized controlled trial. *JMIR Mhealth Uhealth* 5:e139.

Pandey P, Zimitat C. 2007. Medical students' learning of anatomy: Memorisation, understanding and visualisation. *Med Educ* 41:7–14.

This article is protected by copyright. All rights reserved

This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Use of Self-Archived Versions.

Periya NS, Moro C. 2019. Applied learning of anatomy and physiology: Virtual dissection tables within medical and health sciences education. *Bangkok Med J* 15:121–127.

Reeves TC. 1995. Questioning the questions of instructional technology research. In: Simonson M, Anderson M (Editors). *Proceedings of the 1995 Annual National Convention of the Association for Educational Communications and Technology (AECT 1995)*; Anaheim, CA, 1995 November 2-7. p 459–470. Association for Educational Communications and Technology, Washington, DC.

Reeves TC. 2004. No significant differences revisited: A historical perspective on the research informing contemporary online learning. In: Kearsley G (Editor). *Online Learning: Personal Reflections on the Transformation of Education*. 1st Ed. Englewood Cliffs, NJ: Educational Technology Publications. p 299–308.

Reeves TC. 2015. Educational design research: Signs of progress. *Australas J Educ Technol* 31:613–620.

Reeves TC, Lin L. 2020. The research we have is not the research we need. *Educ Technol Res Dev* 68:1991–2001.

Sattar M, Palaniappan S, Lokman A, Shah N, Khalid U, Hasan R. 2020. Motivating medical students using virtual reality based education. *Int J Emerging Technol Learn* 15:160–174.

Stepan K, Zeiger J, Hanchuk S, Del Signore A, Shrivastava R, Govindaraj S, Iloreta A. 2017. Immersive virtual reality as a teaching tool for neuroanatomy. *Int Forum Allergy Rhinol* 7:1006–1013.

Stirling A, Birt J. 2014. An enriched multimedia eBook application to facilitate learning of anatomy. *Anat Sci Educ* 7:19–27.

Stirling AC, Moro C. 2020. The use of virtual, augmented and mixed reality in anatomy education. In: Chan LK, Pawlina W (Editors). *Teaching Anatomy: A Practical Guide*. 2nd Ed. Cham, Switzerland: Springer Nature Switzerland AG, p 359–366.

Su W, Xiao Y, He S, Huang P, Deng X. 2018. Three-dimensional printing models in congenital heart disease education for medical students: A controlled comparative study. *BMC Med Educ* 18:178.

Vasilevski N, Birt JR. 2020. Analysing construction student experiences of mobile mixed reality enhanced learning in virtual and augmented reality environments. *Res Learn Technol* 28:2329.

Vitale SG, Caruso S, Vitagliano A, Vilos G, Di Gregorio LM, Zizolfi B, Tesarik J, Cianci A. 2020. The value of virtual reality simulators in hysteroscopy and training capacity: A systematic review. *Minim Invasive Ther Allied Technol* 29:185–193.

Wainman B, Aggarwal A, Birk SK, Gill JS, Hass KS, Fenesi B. 2021. Virtual dissection: An interactive anatomy learning tool. *Anat Sci Educ* (in press; doi: 10.1002/ase.2035).

Zargaran A, Turki MA, Bhaskar J, Spiers HV, Zargaran D. 2020. The role of technology in anatomy teaching: Striking the right balance. *Adv Med Educ Pract* 11:259–266.

Zhao J, Xu X, Jiang H, Ding Y. 2020. The effectiveness of virtual reality-based technology on anatomy teaching: a meta-analysis of randomized controlled studies. *BMC Med Educ* 20:127.

FIGURE LEGENDS

Figure 1: PRISMA flow diagram. Adapted from Moher et al. (2009).

Figure 2: Review authors' judgements about each risk of bias item presented as percentages across all included studies.

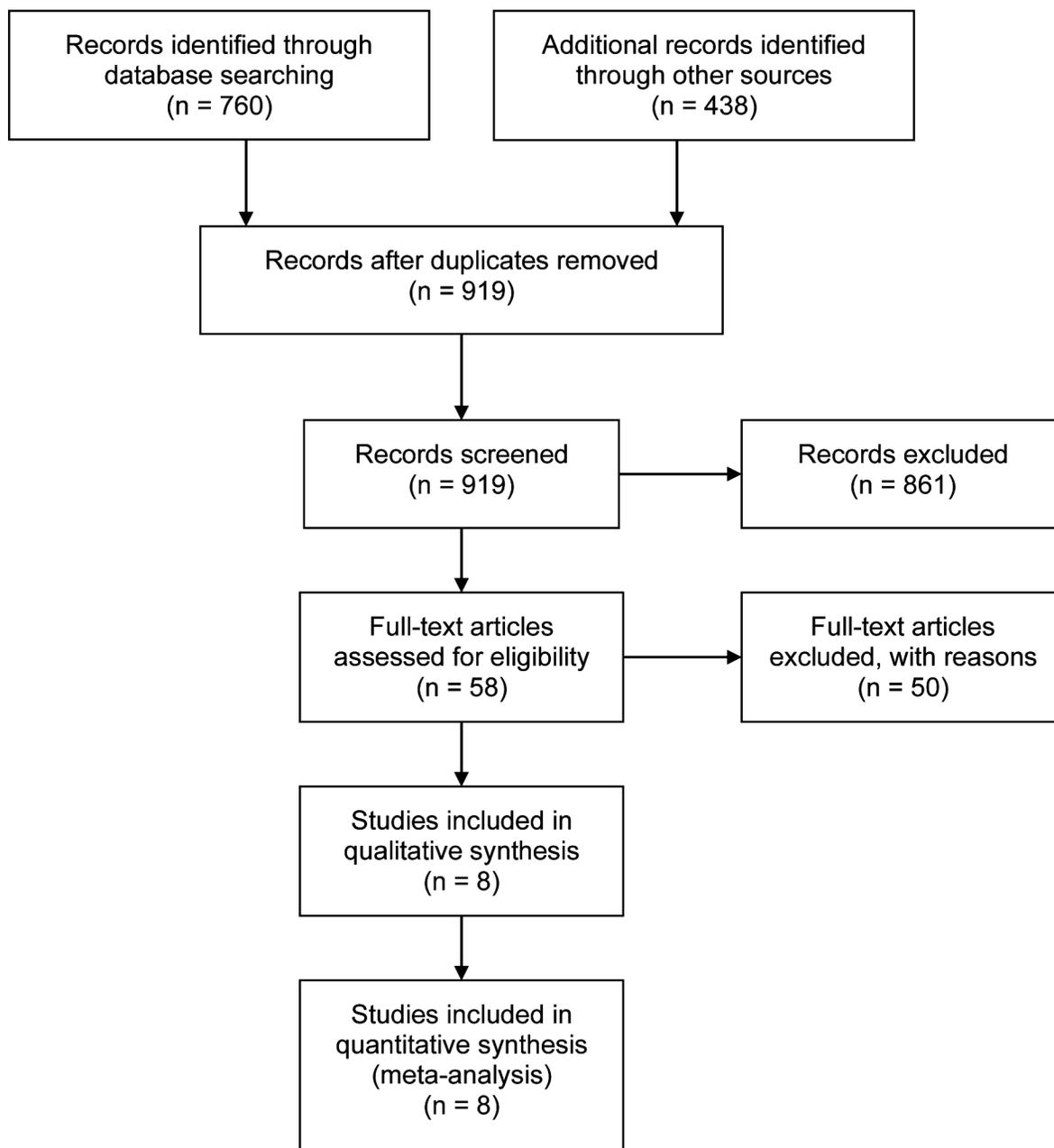
Figure 3: Mean difference between groups in knowledge scores (using percentages). “Control” indicates traditional teaching methods approaches; “Experimental” indicates augmented or virtual reality approaches. Risk of bias legend: A, Random sequence generation (selection bias); B, Allocation concealment (selection bias); C, Blinding of participants and personnel (performance bias); D, Blinding of outcome assessment (detection bias); E, Incomplete outcome data (attrition bias); F, Selective reporting (reporting bias); G, Other bias. Green color indicates low risk of bias; yellow indicated unclear risk of bias; and red color indicates high risk of bias.

Figure 4: Mean difference between groups in knowledge scores (using percentages); studies of virtual reality subgrouped separately from studies of augmented reality. More than one intervention was found within the Moro et al. (2017) study and the controls split to calculate intervention. “Control” indicates traditional teaching methods approaches; “Experimental” indicates augmented or virtual reality approaches. VR, virtual reality; AR augmented reality; Risk of bias legend: A, Random sequence generation (selection bias); B, Allocation concealment (selection bias); C, Blinding of participants and personnel (performance bias); D, Blinding of outcome assessment (detection bias); E, Incomplete outcome data (attrition bias); F, Selective reporting (reporting bias); G, Other bias. Green color indicates low risk of bias; yellow indicated unclear risk of bias; and red color indicates high risk of bias.

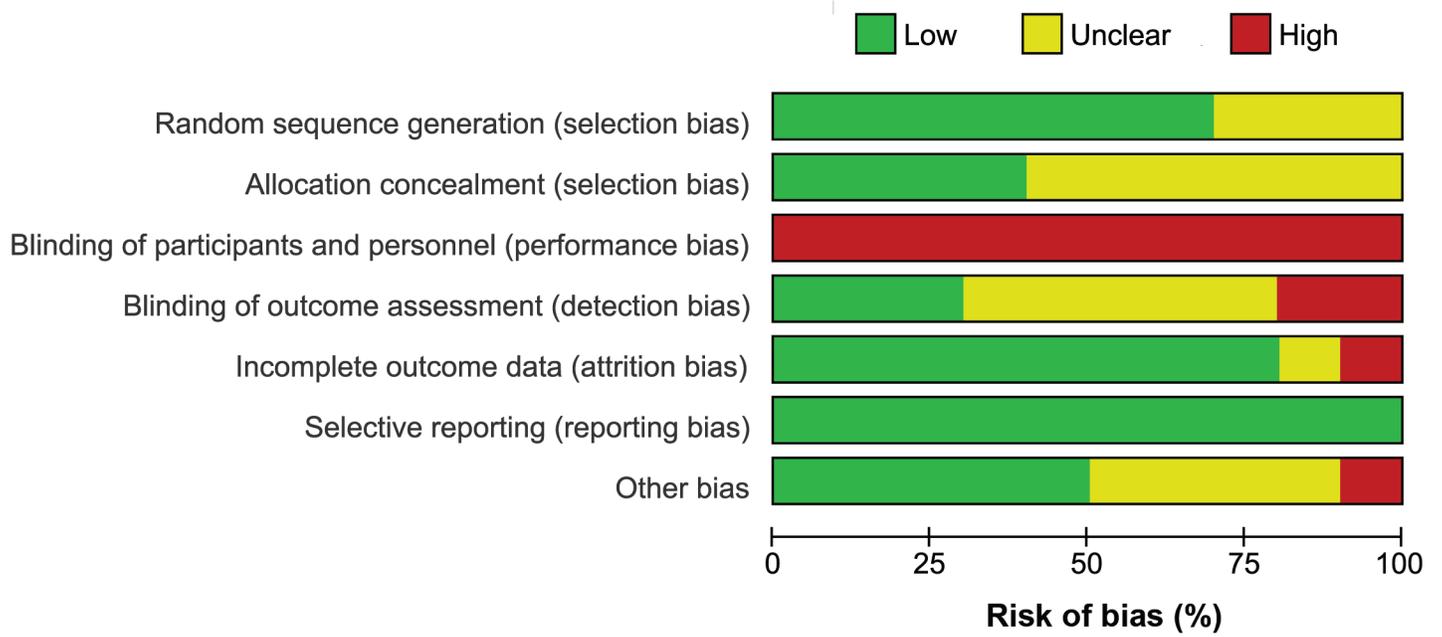
Table 1. Major Characteristics of Eight Studies Evaluating the Effectiveness of Virtual or Augmented Reality-Based Enhancements to Medical and Science Student Physiology and Anatomy Test Performance.

No.	Authors (Year)	Study design	Intervention	Comparator	Number of study participants	Measurement of Knowledge Scores
1.	Albrecht et al. (2013)	RCT – 2 parallel arms	Mobile AR	Phone application	10	10-question multiple-choice pre- and post-test
2.	Barmaki et al. (2019)	Cluster RCT – 4 clusters (class-sized each)	Mirrored AR	Textbook and virtual mirror	288	8-question matching labels pre- and post-test
3.	Ekstrand et al. (2018)	RCT – 2 parallel arms	VR	Paper-based	64	22-question multiple-choice pre- and post-test
4.	Küçük et al. (2016)	RCT – 2 parallel arms	Mobile AR	Presentation material (2D pictures, graphs, text)	70	30-question multiple-choice post-test
5.	Maresky et al. (2019)	RCT – 2 parallel arms	VR	Independent study (materials NR)	42	10-question multiple-choice pre- and post-test
6.	Moro et al. (2017)	RCT – 3 parallel arms	VR and tablet AR ^a	Tablet	59	20-question multiple-choice post-test
7.	Noll et al. (2017)	RCT – 2 parallel arms	Mobile AR	Phone application	44	10-question multiple-choice pre- and post-test
8.	Stepan et al. (2017)	RCT – 2 parallel arms	VR	Web-based textbook	66	10-question multiple-choice and fill-in-the-blanks pre-test and 30-question multiple-choice, fill-in-the-blank and labelling models post-test

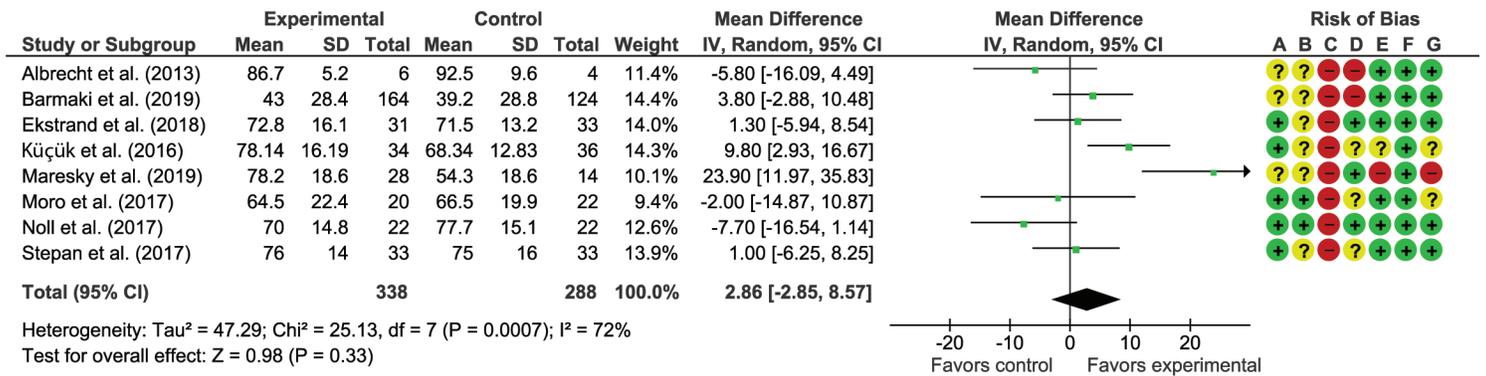
AR, Augmented Reality; VR, Virtual Reality; 2D, Two-dimensional; RCT, Randomized control trial; NR, Not reported; ^aAR arm excluded from the analysis.



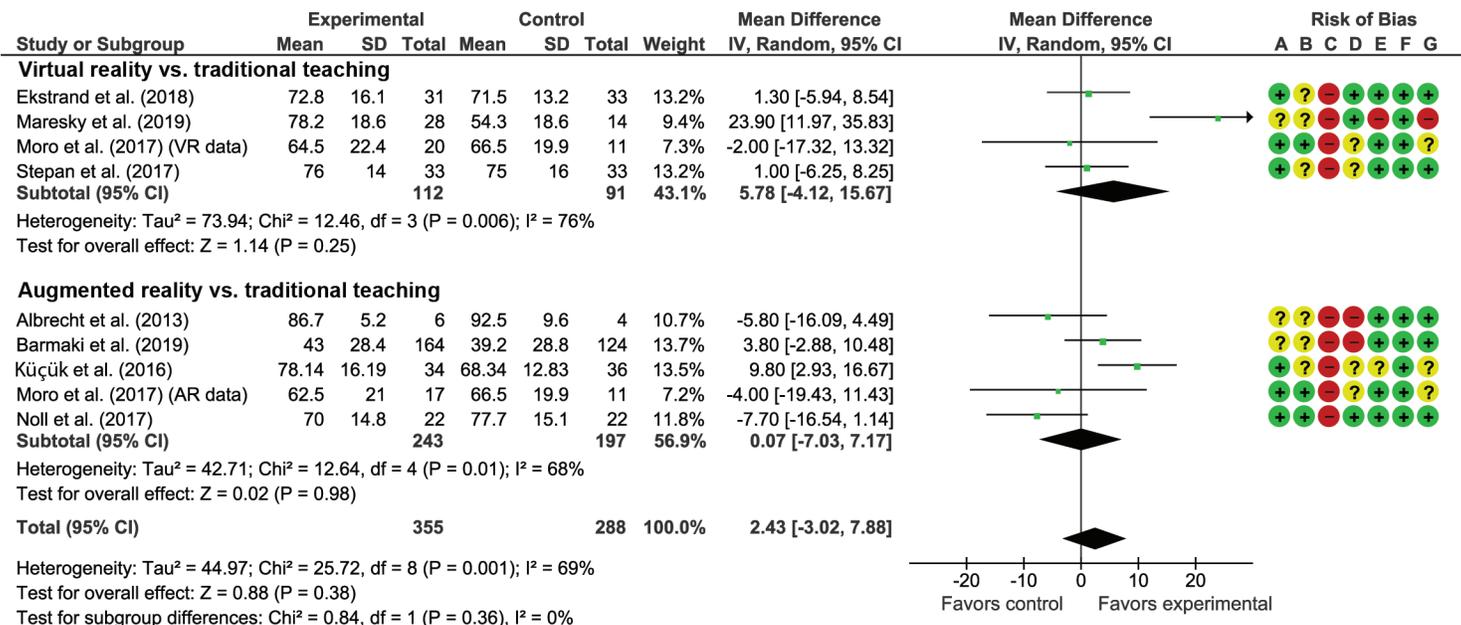
ase_2049_f1.tif



ase_2049_f2.tif



ase_2049_f3.tif



ase_2049_f4.tif