Medical Teacher

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/imte20

Technology enhanced neuroanatomy teaching techniques: A focused BEME systematic review of current evidence: BEME Guide No. 75

Hamish J. Newman, Amanda J. Meyer, Tim J. Wilkinson, Nalini Pather & Sandra. E. Carr

MEDICAL TEACHER
https://doi.org/10.1080/0142159X.2022.2039382

MEDICAL TEACHER

Taylor & Francis
Taylor & Francis Group

BEME GUIDE

Technology enhanced neuroanatomy teaching techniques: A focused BEME systematic review of current evidence: BEME Guide No. 75

Hamish J. Newman[a,b], Amanda J. Meyer[a], Tim J. Wilkinson[c], Nalini Pather[d] and Sandra. E. Carr[b]

[a]Department of Anatomy, Physiology and Human Biology, School of Human Sciences, The University of Western Australia, Perth, Australia; [b]Health Professions Education, School of Allied Health, The University of Western Australia, Perth, Australia; [c]Education Unit, University of Otago, Christchurch, New Zealand; [d]Department of Anatomy, Faculty of Medicine and Health, University of New South Wales, Sydney, Australia

ABSTRACT

Background: In response to growing curriculum pressures and reduced time dedicated to teaching anatomy, research has been conducted into developing innovative teaching techniques. This raises important questions for neuroanatomy education regarding which teaching techniques are most beneficial for knowledge acquisition and long-term retention, and how they are best implemented. This focused systematic review aims to provide a review of technology-enhanced teaching methods available to neuroanatomy educators, particularly in knowledge acquisition and long-term retention, compared to traditional didactic techniques, and proposes reasons for why they work in some contexts.

Methods: Electronic databases were searched from January 2015 to June 2020 with keywords that included combinations of 'neuroanatomy,' 'technology,' 'teaching,' and 'effectiveness' combined with Boolean phrases 'AND' and 'OR.' The contexts and outcomes for all studies were summarised while coding, and theories for why particular interventions worked were discussed.

Results: There were 4287 articles identified for screening, with 13 studies included for final analysis. There were four technologies of interest: stereoscopic views of videos, stereoscopic views of images, augmented reality (AR), and virtual reality (VR). No recommendation for a particular teaching method was made in six studies (46%) while recommendations (from weak to moderate) were made in seven studies (54%). There was weak to moderate evidence for the efficacy of stereoscopic images and AR, and no difference in the use of stereoscopic videos or VR compared to controls.

Conclusions: To date, technology-enhanced teaching is not inferior to teaching by conventional didactic methods. There are promising results for these methods in complex spatial anatomy and reducing cognitive load. Possible reasons for why interventions worked were described including students' engagement with the object, cognitive load theory, complex spatial relationships, and the technology learning curve. Future research may build on the theorised explanations proposed here and develop and test innovative technologies that build on prior research.

KEYWORDS
Neuroanatomy; cranial anatomy; pedagogy; student; teaching; mixed reality; stereopsis; augmented reality; virtual reality; mobile technology; virtual assistants; learning; technology-enhanced learning

Background

Learning neuroanatomy is known to be challenging for medical students, junior doctors, and many specialist doctors due to the complexity and interconnectedness of anatomical structures (Giles 2010). The difficulty in learning neuroanatomy was summarised by Jozefowicz who introduced the term 'neurophobia' in 1994 (Jozefowicz 1994). More recently, there has been a decrease in anatomy teaching hours within medical courses with associated claims of poorer knowledge of anatomy of medical students (McKeown et al. 2003; Prince et al. 2005; Waterston and Stewart 2005); particularly with respect to nervous system anatomy (Waterston and Stewart 2005). Practically, lower levels of neuroanatomy knowledge are associated with reduced confidence of junior doctors in managing neurological conditions (Schon et al. 2002; Zinchuk et al. 2010), or with unsafe medical practice (Waterston and Stewart 2005).

Practice points

- Augmented reality and virtual reality can be used in place of, or complementary to, traditional learning methods.
- Technology could be utilised where complex spatial relationships must be understood, and in advanced presentations where cognitive load must be reduced.
- To optimise teaching and learning, reduce extraneous cognitive load and learning curves required for the use of technologies.

Traditional teaching in anatomy laboratories involves human prosected materials and plastinated specimens, models, and medical imaging (Craig et al. 2010). In efforts to improve long-term knowledge acquisition, spatial

CONTACT Hamish J. Newman newmahj@gmail.com School of Allied Health, Health Professions Education, The University of Western Australia, 35 Stirling Hwy, PERTH WA 6009, Australia

understanding and reduce neurophobia, in the context of reduced teaching hours and evolving curricula, research has also been conducted into developing innovative (largely digital) teaching strategies and techniques (Moxham et al. 2015; Arantes et al. 2018). No consensus for use of these methods exists, despite there being a core neuroanatomy curriculum, described by Moxham et al. (2015) and Craig et al. (2010). Without clear evidence to guide academics, the curriculum content, instruction methodology, and assessment are at the discretion of individual institutions. It is, therefore, perhaps unsurprising that educational practice is highly variable among universities around the world (Craig et al. 2010; Sotgiu et al. 2020). This raises globally important questions in neuroanatomy education regarding which practices are most beneficial for knowledge acquisition and long-term retention, and how they are best implemented.

Two systematic reviews have been conducted into this area of research. Arantes et al. (2018) reviewed 29 papers that assessed the impact of using 15 methods of teaching on students' learning of neuroanatomy as a guide for curricular improvements (Arantes et al. 2018). The second, by Sotgiu et al. (2020), reviewed 16 papers covering eight teaching techniques in an attempt to identify the most effective method/s to teach neuroanatomy (Sotgiu et al. 2020). These techniques included 3D- and 2D-based computer tools, 3D physical models (PMs), tablet applications, near-peer teaching, equivalence-based instruction, face-to-face teaching, flipped classroom, inquiry-based learningand intensive modes of delivery among others. Sotgiu et al. recommended a combination of pedagogical resources when teaching neuroanatomy (Sotgiu et al. 2020). A limitation of both reviews is that the most recent types of technology-enhanced teaching methods, including augmented or virtual reality (VR), were not included. Arantes et al. noted 83% of studies included were published between 2010 and 2018 and both papers illustrated this is a fast-moving area of research. The annual HORIZON Report describes emerging technologies in higher education that are likely to have an impact on teaching and learning. The 2019 report suggested four relevant technologies will have a major impact in the next 5 years: mobile learning, mixed reality, artificial intelligence and virtual assistants (Alexander et al. 2019). These educational tools are increasingly used in neuroanatomy education and have appeared more frequently in the literature, even since publication of these recent reviews (Henn et al. 2002; Levinson et al. 2007; Kockro et al. 2015; de Faria et al. 2016; Küçük et al. 2016; Moro et al. 2017; Stepan et al. 2017; Ekstrand et al. 2018; Henssen et al. 2020).

The implementation and measurement of effect of educational interventions is complex. A systematic review that not only identifies and compares available teaching techniques, but through evidence synthesis raises possibilities for why they work in some contexts and not others, would form a useful guide for educators and researchers into modern practices and their application. Neither systematic review (Arantes et al., 2018; Sotgiu et al., 2020) commented on context and/or description of educational tools in sufficient detail. If neuroanatomy educators understood the context in which an educational tool was investigated (in terms of curricula alignment and learning outcomes being assessed) and/or understood the circumstances in which tools are the most effective, they are better placed to apply the research into practice.

New educational technologies offer highly realistic learning experiences supportive of complex learning and transfer, and their specific application in neuroanatomy teaching is interesting to educators (Kamphuis et al. 2014). A focused review that updates educators on the application of the most recent technology-enhanced teaching methods in neuroanatomy education, and offers clarification as to why they may work in some contexts over others, is required. This review may form a useful update for educators regarding these technologies and their evidence for use.

Review questions

Regarding neuroanatomy teaching methods:

- Which technology-enhanced teaching methods are available? How can they be applied?
- How do technology-enhanced teaching methods compare with traditional learning techniques? Which are associated with improved knowledge acquisition and long-term retention?
- How/why do some methods work in some contexts and not others?

Materials and methods

A focused review is a form of knowledge synthesis that embraces the core principles of systematicity, while addressing a relatively narrow research question (Gordon et al. 2019). This study is a focused systematic review of quantitative research, undertaken according to the PRISMA checklist of systematic review sections (Moher et al. 2009). It applied Cook's Framework for the descriptive outcomes (what was done), justification outcomes (did it work) and clarification outcomes (why or how did it work) for each study (Cook et al. 2008). This allowed for a discussion of the context of selected papers, and a synthesis of the proposed reasons for the relationship observed between outcomes and the contexts in which technologies were investigated.

A full protocol was published by Best Evidence Medical Education (BEME) Collaboration (Newman et al. 2020).

Search strategy

Electronic databases were searched from January 2015 to June 2020 including PubMed, Medline (EBSCO), Cinahl Plus (EBSCO), Academic Search Premier (EBSCO), ProQuest Central (ERIC), EBook Central (ERIC), ERIC, Scopus, Web of Science, and SAGE. Keywords included combinations of 'neuroanatomy,' 'technology,' 'teaching,' and 'effectiveness' combined with Boolean phrases 'AND' and 'OR' (Table 1).

In addition, authors hand searched the references of all included studies and any relevant reviews. A hand search of key journals in medical education was conducted using combinations of key search terms including: *Medical Teacher, Medical Education, Academic Medicine, Anatomical Sciences Education*. Grey literature was excluded.

A population, intervention, comparison, and outcome-based strategy informed our inclusion and exclusion criteria.

Table 1. Search terms used in PubMed, Cinahl Plus, EBSCO, Scopus, Web of Science, and Sage Databases.

Neuroanatomy	AND technology	AND teaching	AND effectiveness
OR cranial anatomy	OR technology	OR learning	OR instructional effectiveness
OR skull base anatomy	OR mixed reality	OR education	OR knowledge
OR brain anatomy	OR virtual reality		OR retention
OR head anatomy	OR augmented reality		OR memory
OR central nervous system anatomy	OR mobile technology		OR understanding
OR peripheral nervous system anatomy	OR virtual assistants		OR application
OR cranial nerves	OR artificial intelligence		OR enhance
OR peripheral nerves	OR 3 D		
OR spinal cord	OR computer-assisted instruction/methods		
OR deep brain structures	OR computer simulation		

Inclusion and exclusion criteria

Population

Preference was given to those learning in health sciences contexts (e.g. medical and allied health) but studies were not excluded on this basis (e.g. veterinary sciences were included).

Intervention

Advanced learning methods identified in the HORIZON report including augmented reality (AR), VR, mobile learning, mixed reality, artificial intelligence, and virtual assistants (Alexander et al. 2019). Three-dimensional computer-based technology was excluded unless augmented by modern devices (e.g. stereoscopic 3D learning technologies).

Comparison

Any comparison of two different teaching methods. An ideal control cohort was those learning under traditional circumstances, but we did not exclude comparisons of two technology-enhanced learning methods (e.g. stereotactic learning compared to 3D computer-based learning). We define traditional learning as didactic lectures supplemented by cadaveric- or model-based laboratories.

Outcome

Any quantitative assessment of knowledge acquisition and retention in neuroanatomy. Secondary outcomes included a description of teaching methods or analysis of other learning factors, such as spatial ability.

Screening, data collection, analysis, and synthesis

A standard proforma for screening and data collection was followed. HN and AM screened all references and abstracts using Covidence software (Veritas Health Innovation, Melbourne, Australia). Disagreements were resolved by discussion (SC). A coding sheet was designed and discussed as a group, with all reviewers coding the same paper independently to ensure consistency. Each author was allocated three included papers for coding, with HN coding all papers. All conflicts were able to be resolved by discussion.

Papers were graded for their quality of evidence, determined by incorporating risk of bias and then critiquing the reported study design and methodology. Papers were appraised according to PRISMA guidelines, with risk of bias assessed with the Cochrane Risk of Bias tool (Higgins et al. 2011). Methodological quality was reported according to the grades outlined by Colthart et al.'s (2008) method: Grade 1 (no clear conclusions can be drawn; not significant), Grade 2 (results ambiguous, but there appears to be a trend), Grade 3 (conclusions can probably be based on the results), Grade 4 (results are clear and very likely to be true), and Grade 5 (results are equivocal) (Colthart et al. 2008). Where a recommendation was made, the strength of recommendation was made by HN informed by the strength of the quantitative findings/result (e.g. weak $r = 0.10–0.39$, moderate $r = 0.4–0.69$, strong $r = 0.70–0.89$) (Schober et al. 2018), quality of evidence, and risk of bias.

The context and outcomes for all studies were summarised while coding. HN theorised proposed reasons for why interventions worked by tabulating key results of similar papers against their methodology, identifying recurrent patterns between similar studies, proposing various reasons, and discussing this synthesis with the review team. Where the authors of individual papers suggest explanations/hypotheses, explicit mention is made to differentiate their input from ours. We checked this process for consistency. Key terms are defined in a glossary shown in Table 2.

Findings

The screening process is summarised in Figure 1, with 4287 articles identified for screening. HN and AM demonstrated 'moderate' inter-rater reliability, with a Cohen's Kappa of 0.503 (Landis and Koch 1977). The main features of included articles are summarised in Table 3. The 13 included studies are summarised in Table 4.

There were five papers published in the United States of America (USA), and one each in Australia, Canada, New Zealand, Brazil, the Netherlands, Turkey, Germany, and France. There were a total of 1208 students, with a mean of 93 students for each study. Medical students were the most common type of participant ($n = 8$ articles), with biomedical students ($n = 1$ article), and high school students ($n = 1$ article) included. There were three studies that included students from non-specific university backgrounds. Pre-intervention knowledge tests were completed in nine studies (69%). Post-intervention knowledge tests were done immediately after the intervention in nine studies (69%), while four studies assessed retention by testing students between one- and eight-week after the intervention (31%). No recommendation for a particular teaching method was made in six studies (46%) while recommendations (from weak to moderate) were made in seven studies (54%). Strength of evidence and risk of bias according to Colthart et al. (2008) and the Cochrane Risk of Bias Assessment Tool are summarised in Table 3.

Stereoscopic 3D video tools

Assessed in three studies (Kockro et al. 2015; Goodarzi et al. 2017; Bernard et al. 2020), stereoscopic 3D video tools made no difference to overall knowledge acquisition

Table 2. Glossary of key terms.

Term	Definition
Artificial intelligence	'the theory and development of computer systems able to perform tasks normally requiring human cognition. In education, this relates to technologies that personalise learning experiences and reduce workloads'
Augmented reality	'adding digital elements to a live view'
Cognitive load theory	'to acquire biologically secondary knowledge, a learner must obtain novel information, with very limited amounts of novel information able to be processed at a given time'
Curriculum	'a prescriptive term that defines subjects comprising a course of study'
Effectiveness	'the ability of an educational tool to enhance student's neuroanatomy knowledge in an appropriate, cost-effective, timely manner. Often used relative to other education tools'
Engagement	'Zimmerman's definition (1989) of self-regulated learning is used for this paper; that is, the extent to which students are 'metacognitively, motivationally and behaviourally active participants in their own learning process'
Extraneous cognitive overload	'when cognitive processing (the ability of a learner to understand the essential material) and extraneous cognitive processing (ability of a learner to interpret or overcome confusing layout of presented material) exceed the learner's cognitive capacity'
Interaction	'Moore's definition (1989) of learner-content interaction is used for this paper, as opposed to other types of interaction such as learner-learner interaction or learner-teacher interaction'
Mixed reality	'intersection between digital technology and the real world. Current available technologies are augmented reality or virtual reality. Holographic devices are emerging in this space'
Mobile learning	'education or training conducted by means of portable computing devices such as smartphones or tablet computers'
Neurophobia	'a fear of the neural sciences and clinical neurology that is due to the students' inability to apply their knowledge of basic sciences to clinical situations' (Jozefowicz 1994)
Performance	'in the context of neuroanatomy education, this will refer to student's knowledge acquisition, ability to understanding the content and the long-term retention of information'
Stereopsis	'the ability of both eyes to create the perception of depth by seeing an object as one image. Stereoscopy is a technique for creating or enhancing the illusion of depth in an image'
Technology enhanced teaching techniques	'novel teaching techniques researched with regards to their application to neuroanatomy, particularly in the last 5 years. Examples include those mentioned in the HORIZON report such as AR, VR, mobile technology, artificial intelligence and virtual assistants'
Temporal split-attention affect	'where multiple elements of interest are displayed simultaneously, a learner is often required to distribute their attention across multiple areas'
Virtual assistant	'the use of spoken commands, voice recognition and a natural user interface to connect students to the virtual environment'
Virtual reality	'implies a complete immersion experience and shuts out the physical world'

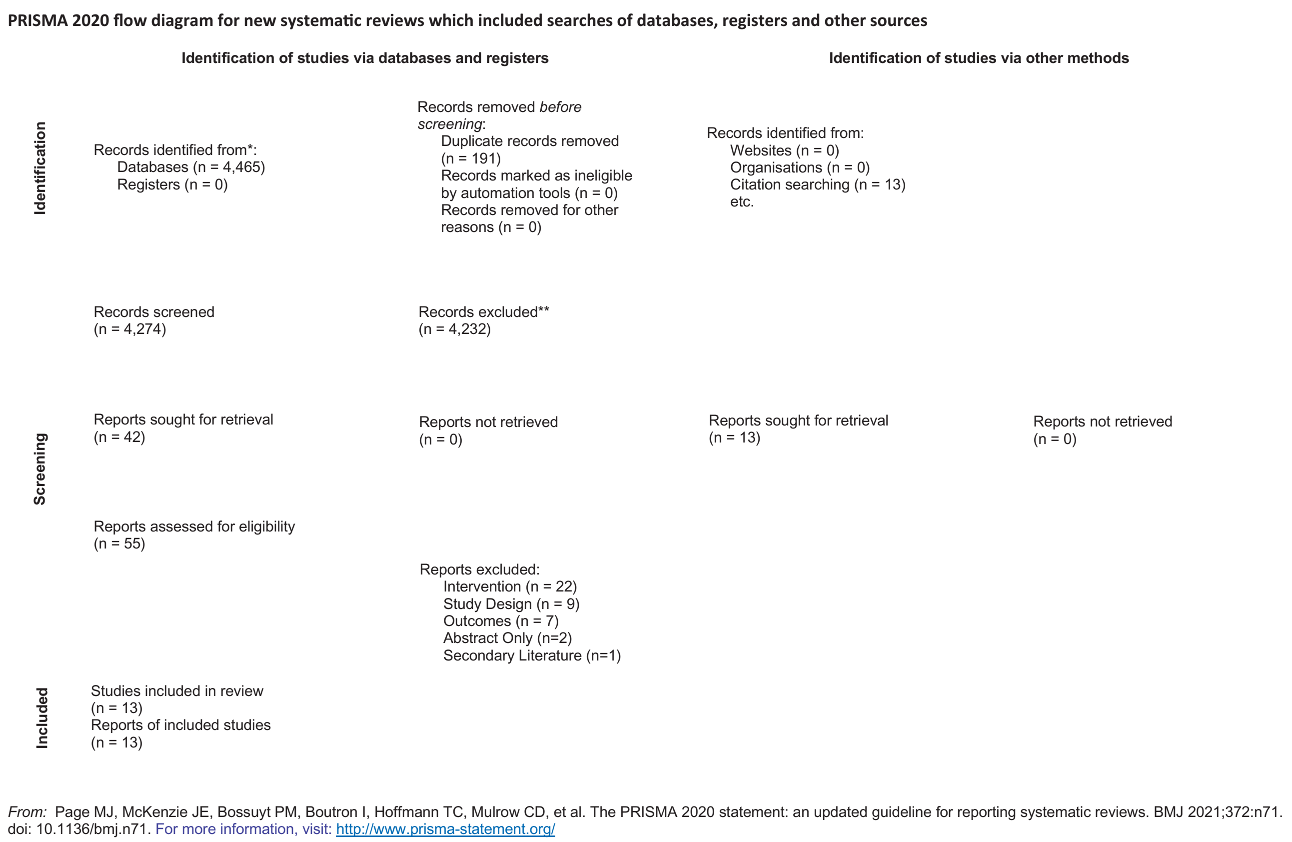

Figure 1. PRISMA Flow Chart of studies included for analysis.

compared to 2D videos of the same content. Stereoscopic resources were popularised for their use in neuroanatomy teaching in 2011 by Dr Albert Rhoton Jr, who designed an entire collection of images of the brain for use by neurosurgeons (Sorenson et al. 2016). Stereoscopy is a technique used to give the illusion of depth, or 3-dimensions, by using stereopsis for binocular vision. Traditionally, this is done by presenting two offset images to the left and right eye independently (practically, students wear glasses that are often red- and blue-coloured, anaglyph glasses) which

Table 3. Main features of manuscripts included for analysis ($n = 13$).

Features	Number (%)	Studies (#)
Year		
2015	2 (15.38)	1, 4
2016	2 (15.38)	5, 7
2017	4 (30.77)	2, 6, 9–10
2018	2 (15.38)	11,12
2019	0 (0)	
2020	3 (23.08)	3, 8, 13
Country		
USA	5 (38.46)	2, 4, 6, 10, 12
France	1 (7.69)	3
Brazil	1 (7.69)	5
Netherlands	1 (7.69)	8
Canada	1 (7.69)	11
Germany	1 (7.69)	1
Turkey	1 (7.69)	7
Australia	1 (7.69)	9
New Zealand	1 (7.69)	13
Number of participants		
0–50	2 (15.38)	6, 8
51–100	8 (53.85)	4–5, 7, 9–13
101–150	0 (0)	
151–200	2 (15.38)	1, 3
200+	1 (7.69)	2
Type of participants		
Medical students	8 (61.54)	1, 3, 5–7, 10–11, 13
Biomedical students	1 (7.69)	8
High school students	1 (7.69)	4
Non-specified university- students	3 (23.08)	2, 9, 12
Type of teaching tool		
Stereoscopic 3D video	3 (23.08)	1–3
Stereoscopic D images	3 (23.08)	4–6
Augmented reality	2 (15.38)	7–8
Augmented and Virtual reality	1 (7.69)	9
Virtual reality	4 (30.77)	10–13

are then combined by our visual cortex to give the perception of 3D. In stereoscopic 3D video tools, students watch audiovisual material while wearing a stereoscopic device (glasses or otherwise) allowing them to appreciate the video in 3D.

Kockro et al. (2015) randomly allocated 169 medical students to receive a pre-recorded audio lecture accompanied by either a 2D Microsoft PowerPoint presentation or a 3D animated tour of the third ventricle with DextroBeam (Bracco Advanced Medical Technologies, Princeton), a stereoscopic viewing device. On a 10-question test administered immediately after the lecture, there was no significant difference between the 2D and 3D groups (5.19 *vs.* 5.45, $p > .05$).

Similarly, Goodarzi et al. (2017) sought to investigate the difference that visualising a video with stereopsis had on the short-term acquisition of skull-base anatomy knowledge compared to 2D. The authors compared the performance of School of Education students ($n = 249$) in tests conducted immediately before and after watching a video of the anatomy of the skull base from the Rhoton Collection of the American Association of Neurologic Surgeons. Students were assigned to watch the video in 2D or 3D (stereoscopic). The 3D group performed better in both pre- and post-intervention tests. However, the magnitude of improvement between 2D and 3D groups was no different (48.7% *vs.* 53.5%, $p = .855$).

Bernard et al. (2020) also conducted a similar experiment, comparing 175 students' results on a 30-item test done immediately before, and 1 month after watching a 5-min instructional video on the cerebrovascular system of the brain. Students were randomised into two groups, a 3D group (stereoscopic) and a 2D group (non-stereoscopic). Scores were similar in the fundamental knowledge test 1 month after the intervention (mean 73.2% and 74.4%, $p = .37$).

Participants in all studies were observers, with limited interaction with the content. This was likely done to isolate stereopsis as the variable of interest. While a legitimate teaching tool, the ability to learn from watching audio-visual material is limited by the extent to which a learner can engage with the video (Ertelt 2007). A problem with any video animation is the difficulty people have with the real-time perception of animated visuals and 'extraction' of the message (Proffitt et al. 1990). We hypothesised to be due to a so called 'temporal split-attention effect' (Lowe 2003). One element of this is where multiple elements of interest are displayed simultaneously, requiring the learner to distribute their attention across multiple areas. The extraneous cognitive load negatively influences learning (Lowe 2003). This effect would be independent of what modality the video is viewed through and may be an explanation for why Kockro et al. (2015), Goodarzi et al. (2017), and Bernard et al. (2020) found no difference in knowledge acquisition between stereoscopic views of videos and normal 2D.

There may be other factors for why no significant differences were observed in these studies. For example, post-intervention tests may not have had the number or type of questions suitable to find a difference, increasing the risk of type II error. Power analysis of question type prior to test administration would be an important addition for future studies. Further, by selecting students with minimal or no prior neuroanatomy knowledge, the findings from both papers may not be generalised to all fields of education. Further evidence may be gathered in different populations (for example, neurosurgical residents) to validate the study findings.

One finding of interest was in Bernard et al. (2020)'s paper, the 3D group performed better in questions requiring an understanding of anatomical relationships (86.4% *vs.* 63.5%, $p = .004$). It is possible that where students were able to perceive and extract the content of the videos, the stereoscopic views displayed anatomical relationships in a way that was easier to understand and facilitated students' learning. This is discussed further in reference to other technologies.

Stereoscopic 3D image tools

Three papers assessed the value of learning 2D images with stereoscopic 3D image visualisation (Ferdig et al. 2015; de Faria et al. 2016; Cui et al. 2017). As opposed to the previous section on stereoscopic videos, students would view sequences of still images while wearing a stereoscopic viewing device, creating the illusion of 3D.

Ferdig et al. (2015) compared the results of 89 high-school students on an eight-item test immediately after a laboratory from which they were allocated to learn brain structure and function from either 2D or 3D stereoscopic images. The students in the 3D stereoscopic group scored significantly better than the control group (80% *vs.* 73.6%, *p* value unknown).

de Faria et al. (2016) described the development of a library of stereoscopic images, evaluating the pedagogy by randomly allocating 84 medical students into a 2D images, 2D interactive and non-stereoscopic, and 3D interactive

Table 4. Summary of included studies.

#	References	Title	Country	Participants	Study design; Technology (control *vs.* experimental)	Method	Outcome	Conclusion	Strength of evidence (Risk of bias)	Limitations
1	Kockro et al. (2015)	Stereoscopic neuroanatomy lectures using a three-dimensional virtual reality environment	Germany	169 (second year medical students, University of Mainz)	RCT; 2D lecture *vs.* Stereoscopic 3D video	10-multiple choice question anatomy test immediately after 20-min laboratory	No difference was reported in knowledge between 2D and 3D stereoscopic groups (0.519 ± 0.212 *vs.* 0.545± −0.216, $p = .215$)	No recommendation for either	3 (medium)	Test of short-term memory Too few test items
2	Goodarzi et al. (2017)	Effect of Stereoscopic Anaglyphic 3-Dimensional Video Didactics on Learning Neuroanatomy	USA	249 (school of education students, University of California)	Case–control study; Online 2D video *vs.* Stereoscopic 3D video	Anatomy tests (details unknown) immediately before and immediately after one-hour laboratory	No difference was reported in knowledge gained between 2D and 3D stereoscopic groups (0.487 ± 0.213 *vs.* 0.53 ± −0.227, $p = .0855$)	No recommendation for either	2 (low)	Test of short-term memory
3	Bernard et al. (2020)	Does 3D stereoscopy support anatomical education?	France	190 (second year medical students, University of Angers)	RCT; Online 2D video vs Stereoscopic 3D video	30-mixed question type anatomy tests immediately before and 1 month after 5-min laboratory	No difference was reported in knowledge gained between 2D and 3D stereoscopic groups (0.732 *vs.* 0.744, $p = .37$). 3D stereoscopic group performed better in questions relating to anatomical relationship (0.864 vs. 0.635, $p = .004$).	Weak recommendation for stereoscopic 3D in anatomical relationship knowledge acquisition and retention	4 (medium)	
4	Ferdig et al. (2015)	Using stereoscopy to teach complex biological concepts	USA	89 (high school students)	Mixed, randomised; 2D images *vs.* Stereoscopic 3D	8-multiple choice question anatomy test immediately after laboratory	Stereoscopic 3D group scored significantly better than the classmates learning traditionally (0.800 *vs.* 0.736, $p < .05$)	Moderate recommendation for stereoscopic 3D in knowledge acquisition	2 (low)	Test of short-term memory Too few test items
5	de Faria et al. (2016)	Virtual and stereoscopic anatomy: when virtual reality meets medical education	Brazil	84 (graduate medical students, University of São Paulo)	RCT; Lecture *vs.* 3D Computer-Based *vs.* Stereoscopic 3D	1- short answer (10-part) anatomy test immediately before and repeated immediately after one-hour laboratory	3D-Computer Based and Stereoscopic 3D groups both performed equally, and better than lecture only (0.472 ± 0.120 *vs.* 0.597 ± 0.128 *vs.* 0.603 ± 0.120, $p = .05$)	Weak recommendation for stereoscopic 3D and 3D computer-based in knowledge acquisition over lecture only	3 (medium)	Test of short-term memory Too few test items
6	Cui et al. (2017)	Evaluation of the Effectiveness of 3D Vascular Stereoscopic Models in Anatomy Instruction for First Year Medical Students	USA	39 (first year medical students, University of Mississippi Medical Center)	RCT; 2D images *vs.* Stereoscopic 3D	15-question anatomy tests immediately before and after 20-min laboratory	Stereoscopic 3D group scored significantly better than the 2D learning group (0.762 ± 0.186 *vs.* 0.583 ± 0.188, $p = .003$)	Strong recommendation for stereoscopic 3D in knowledge acquisition, particularly for students with low spatial ability	4 (medium)	
7	Küçük et al. (2016)	Learning anatomy via mobile augmented reality: Effects on achievement and cognitive load	Turkey	70 (second year medical students, Ataturk University)	Mixed, randomised; Textbook *vs.* Augmented Reality	30-multiple choice question anatomy test immediately after a three-hour laboratory	Mixed augmented reality group had higher anatomy test scores (0.781 ± 0.162 *vs.* 0.683 ± 0.128, $p < .05$) and lower cognitive scores compared to the control	Moderate recommendation for AR in knowledge acquisition and cognitive load	4 (medium)	Test of short-term memory

(*continued*)

Table 4. Continued.

#	References	Title	Country	Participants	Study design; Technology (control *vs.* experimental)	Method	Outcome	Conclusion	Strength of evidence (Risk of bias)	Limitations
8	Henssen et al. (2020)	Neuroanatomy Learning: Augmented Reality vs. Cross-Sections	Netherlands	31 (first year biomedicine students, Radboud University)	RCT; 2D images *vs.* Augmented Reality	25-mixed question type anatomy tests immediately before and after two-hour laboratory	Students in the control group (2D learning only) performed significantly better in the post-laboratory test than the AR group (0.543 ± 0.140 *vs.* 0.373 ± 0.084, $p < .01$)	Moderate recommendation for traditional learning in knowledge acquisition	3 (medium)	Test of short-term memory
9	Moro et al. (2017)	The Effectiveness of Virtual and Augmented Reality in Health Sciences and Medical Anatomy	Australia	59 (medical, biomedical and health sciences students, Bond University)	RCT; Tablet *vs.* augmented reality *vs.* virtual reality	20-, multiple choice question anatomy test immediately after 10-min laboratory	No difference was reported between tablet-based, augmented reality or virtual reality groups (0.665, 0.625, 0.645, $p = .874$)	No recommendation for any particular method	3 (medium)	Test of short-term memory Too few test items Limited laboratory time
10	Stepan et al. (2017)	Immersive virtual reality as a teaching tool for neuroanatomy	USA	66 (first and second year medical students, Icahn School of Medicine)	RCT; Online textbook *vs.* virtual reality	10-, 25-, 15-mixed question type anatomy tests immediately before, immediately after and 8-weeks after a 30-min laboratory	No difference was reported between traditional learning or virtual reality groups (0.75 ± 0.16 *vs.* 0.76 ± 0.14, $p = .95$)	No recommendation for either	3 (medium)	
11	Ekstrand et al. (2018)	Immersive and interactive virtual reality to improve learning and retention of neuroanatomy in medical students: a randomised controlled study	Canada	64 (first and second year medical students, University of Saskatchewan)	RCT; 2D images *vs.* virtual reality	22-multiple choice question anatomy tests immediately before, after and seven days after a 12-min laboratory	No difference was reported in knowledge gained between 2D and virtual reality groups	No recommendation for either	4 (medium)	Limited laboratory time
12	Wismer et al. (2018)	A Workload Comparison During Anatomical Training with a Physical or Virtual Model	USA	61 (university students, University of Central Florida)	Case–control study; Physical models *vs.* virtual reality	25-mixed question type anatomy tests immediately before and after 10-min laboratory	No difference was reported in knowledge gained between 2D and virtual reality groups	No recommendation for either	4 (medium)	Test of short-term memory
13	Wang et al. (2020)	A Randomised Control Trial and Comparative Analysis of Multi-Dimensional Learning Tools in Anatomy	New Zealand	52 (second year medical students, University of Otago)	RCT; Textbook *vs.* 3D Computer-Based *vs.* virtual reality	18-short answer question anatomy tests before and 1 month after laboratory	Mixed reality performs worse in acquiring nominal based information, with no difference in mixed/spatial questions compared to 3DM or text-only. However, MR group retained information better in nominal and spatial type questions after a month	Weak recommendation for MR in long-term retention of neuroanatomy	4 (medium)	

and stereoscopic groups. Students studied the components of the limbic system for 50–60 min, and then were asked to recall how many they could remember immediately after the laboratory. The interactive groups scored significantly higher than the 2D image group (6.03 ± 1.20 and 5.97 ± 1.28 *vs.* 4.72 ± 1.20, $p < .05$), and the addition of a stereoscopic view of the images did not result in a significant difference in knowledge acquisition compared to 2D viewing (6.03 ± 1.20 and 5.97 ± 1.28, $p > .05$).

Cui et al. (2017) randomly allocated 39 first-year medical students to learn the cerebral vasculature for 20-min by either 3D stereoscopic models or 2D images captured as snapshots of these models and the radiographic images from which they were built. There was a significant difference in the post-intervention test results between the 2D and 3D groups (mean = 58.3% *vs.* 76.19%, $p = .003$). Further, students in the 3D learning session with a low-spatial ability measured using the Mental Rotations Test (MRT) (Shepard and Metzler 1971), improved in their post-laboratory test to a level comparable to that demonstrated by students with high-spatial ability. This was not the same for the 2D learning group, where students with a higher spatial ability improved relatively more.

Ferdig et al. (2015) and de Faria et al. (2016) both incorporated elements of learner-content interaction in their studies to facilitate learning, while in Cui et al. (2017), participants were passive observers (Moore 1989). Specifically, Ferdig et al. (2015) allowed students to rotate through images at their own pace, while de Faria et al. (2016) had students choose the viewpoint from which they observed anatomical specimens, with or without stereopsis. Interaction is naturally a key element of learning, as without engaging with the content, learners cannot change their understanding, or perspective, or any other cognitive structure (Moore 1989). Holmberg described the 'internal didactic conversation' where learners 'talk to themselves' about the information they encounter (Holmberg 1986). By facilitating this internal dialogue with increased interaction, Ferdig et al. (2015) and de Faria et al. (2016) demonstrated improved learning.

There was conflicting data on how useful the addition of stereoscopy was in learning from 2D images. While Ferdig et al. (2015) and Cui et al. (2017) demonstrated improved performance in knowledge tests with stereopsis, it did not result in a significance difference in de Faria et al. (2016)'s study. It is difficult to say why this is the case. Theoretically, and as the authors of both studies propose, stereopsis would allow learners to appreciate anatomical relationships between structures better by displaying specimens in 3D (Hilbelink 2009; Held and Hui 2011). It is possible that virtual manipulation of the object, whether in 2D or 3D was enough for learners to appreciate these anatomical relationships sufficiently (de Faria et al. 2016). An interesting study to investigate this theory would be a comparison of an interactive 3D software that is displayed in 2D and a stereoscopic view of a 2D image. As discussed, there may be other reasons for why significant differences were not found, including the types of knowledge tests utilised by the authors and insufficient power of the studies.

Augmented reality

Coming in many forms, AR allows the user to 'augment' their view of reality with additional material, typically with projection onto a real-time visual display of the viewed subject matter. The technology requires a processor, display, sensors, and input devices. Smart phones or tablets have these elements including a camera, and sensors such as accelerometer, solid-state compass and global positioning systems that make them suitable AR platforms. AR was assessed specifically in two studies, with mixed results in terms of knowledge acquisition (Küçük et al. 2016; Henssen et al. 2020). A consistent finding was that AR reduced the cognitive load of students in learning neuroanatomy.

Henssen et al. (2020) investigated the performance of GreyMapp-AR (GreyMapp-AR, Radboud University, the Netherlands) compared to traditional, cross-sectional learning of general brain anatomy and subcortical structures. Medical and biomedical students ($n = 31$) were randomly allocated into the two conditions and completed a 2-h laboratory followed by a test immediately after. Students learning by 2D images performed significantly better on test scores than students who worked with AR (21.1 *vs.* 23.9, $p < .05$). However, the difference could be reversed by excluding cross-sectional anatomy questions. Further, the AR group experienced lower germane and extraneous cognitive load than the 2D group, although the sample size limited the reliability of these conclusions.

Küçük et al. (2016) investigated the use of mobile AR (mAR) compared to traditional 2D presentation materials when studying the anatomy of the spinal cord tracts. Mobile AR is a sub-type of AR while the smart phone device is the platform of choice. Undergraduate medical students ($n = 70$) had five hours of teaching on the subject. The authors randomly allocated students to study the material either with mAR, or with the 2D learning resources, before independently completing a 30-item knowledge test. There was a significant difference in test results of students who studied the anatomy with mAR compared to traditional learning resources (78.1% *vs.* 68.3%, $p < .05$). Students studying with mAR reported significantly lower cognitive load than their counterparts in the control group.

The mixed results between these studies may have been due to study design and participant number. Küçük et al. (2016) showed improved performance when using mAR. In interpreting this study, and the broader literature, we suggest this may have been due to the application's ability to improve visualisation of abstract structures and clarify complex topics with overlayed information (Yuen et al. 2011; Bujak et al. 2013; Wu et al. 2013). Henssen et al. (2020)'s experiment had less than half the number of participants; introduced to the GreyMapp software through a short introductory training session using a similar, but not the same, application. Compared to the relatively intuitive mAR utilised by Küçük et al. (2016), participants using GreyMapp-AR may have required a longer introduction to familiarise themselves with the technology. Across one laboratory, this may have been a significant enough hindrance in learning to affect the result.

Virtual reality

VR was assessed in five studies (Moro et al. 2017; Stepan et al. 2017; Ekstrand et al. 2018; Wismer et al. 2018; Wang et al. 2020). Whereas AR involves digital overlays onto perceptions of the real world, VR replaces the real world with a simulated one. Various devices enable the user to immerse themselves in this environment, commonly with a head-mounted device and eye-tracking technology.

Moro et al. (2017) investigated the utility of both augmented and VR compared to tablet-based learning. Investigators randomly allocated 59 medical, biomedical, and health science students to one of the three learning modes before completing a 10-min lesson in skull anatomy and sitting a 20-item anatomical knowledge test. There was no significant difference found in mean assessment scores between the groups. However, students in the VR group experienced side effects such as headaches (25%), dizziness (40%), or blurred vision (35%).

Stepan et al. (2017) assessed the utility of a head-mounted display for learning the ventricular system and cerebral vasculature compared to online textbooks. There were 66 students randomly allocated to each learning mode, before completing a 30-min laboratory. There was no significant difference in anatomy knowledge assessed in pre-intervention, post-intervention, and 8-week retention quizzes between learning methods.

Ekstrand et al. (2018) conducted a randomised control study investigating the performance of 66 university students on a 22-item test sat before, immediately after, and 1-week after studying the spatial relationships between nine neural structures for 12 min by either VR or paper-based means. There was no significant difference between experimental and control groups for any of the testing intervals.

Wismer et al. (2018) conducted an experiment where university students studied gross brain anatomy for 10 min using either a plastic PM ($n = 29$) or models presented in VR ($n = 32$). There was no difference reported in knowledge gained between 2D and VR groups in a post-laboratory knowledge test conducted immediately after the lesson. However, by using surveys of students, Wismer et al. (2018) describe learners with VR experienced less spatial workload, mental demand, and frustration.

Wang et al. (2020) randomly assigned 52 medical students to three learning tools: text-only, three-dimension visualisation on a 2D screen or mixed reality. Students learned about the anatomy of the visual system for 20 min, and then completed a knowledge test 1 month after. The mixed reality group performed worse in acquiring nominal-based information, with no difference in mixed/spatial questions compared to 3D visualisation or text-only groups. However, the mixed-reality group retained information better in nominal and spatial type questions after a month.

Across the five studies investigating the utility of VR, there was no advantage in using VR as an educational tool. The methodologies and results were similar enough between papers to be synthesised together. The 'redundancy principle' of the cognitive theory of multimedia learning may apply to VR in this setting. The principle states that people learn better from graphics and narration than from graphics, narration, and printed text (Mayer and Johnson 2008). Extraneous overload occurs when cognitive processing (the ability of a learner to understand the essential material) and extraneous cognitive processing (ability of a learner to interpret or overcome confusing layout of presented material) exceed the learner's cognitive capacity (Mayer and Fiorella 2014). This may be applicable to the results of studies investigating VR, where the extra information provided to the learner does not necessarily aid in understanding the content, or in its translation to memory (Mayer and Fiorella 2014).

Further explanation for lack of significant results were the required training of participants in learning to navigate a new technology, termed the 'technology learning curve' (Meyer et al. 2016). Familiarising themselves with VR may have taken critical time away from learning the proposed content. As VR becomes more accessible and commonplace, the extraneous cognitive processing and time required to learn how to use and understand VR may offer the learner a valuable perspective on neuroanatomical material. The adverse physical side-effects of using the VR devices such as dizziness must also be taken into consideration, especially amongst a VR-naïve group. Benefits of VR discussed by the authors were a reduction in workload, mental demand and frustration and authors found this technology more engaging, specifically due to greater immersion, clarity of the learning tool and novelty of the instrument.

Discussion

Based on the current evidence, technology-enhanced teaching is similar to learning by traditional techniques in terms of knowledge acquisition, with only weak-to-moderate evidence for the use of stereoscopic 3D images and AR, and only in particular settings. Where there was a difference, technology aided most when students were required to learn anatomy with complex spatial arrangements or that were difficult to visualise in a laboratory setting.

The included studies demonstrated significant heterogeneity in terms of the teaching tool they were investigating and their methodology, with many being statistically underpowered. Studies often assessed immediate recall rather than long-term knowledge retention. Further, outcome measures of studies, being mostly knowledge tests, were also heterogeneous and difficult to evaluate for quality, with few details provided that would allow for re-producible trials.

We have synthesised four explanations as to why some technology-enhanced methods were more effective than others in promoting neuroanatomy knowledge acquisition and long-term retention. When assessing educational pedagogy, clarifying reasons why some interventions work in some contexts over others can be difficult However, by using existing theories of learning or inferring from data available to reviewers in the papers, one may be able to theorise potential explanations. Below we discuss in greater detail, the four areas: engagement with the object, cognitive load theory, complex spatial relationships, and the technology learning curve.

Engagement with the object

A problem with any audio-visual material is the difficulty people have with the real-time perception of visuals and 'extraction' of the message (Proffitt et al. 1990). A 'temporal split-attention effect' describes the phenomenon where users must split their attention between the materials, for example, an image and text, to understand the information being conveyed. This is more likely to occur where multiple elements of interest are displayed simultaneously, requiring the learner to distribute their attention. The 'redundancy principle' of the cognitive theory of multimedia learning may apply here as well, where cognitive processing (the ability of a learner to understand the essential material) and extraneous cognitive processing (ability of a learner to interpret or overcome confusing layout of presented material) exceed the learner's cognitive capacity (Mayer and Fiorella 2014). This extraneous cognitive overload may limit learning and is discussed further below. Learning through stereoscopic views of videos and VR are examples of where students may be limited by the extent to which they can cognitively comprehend and engage with material (Ertelt 2007). This may be an important consideration in the design of future technologies, where the intention should be to stagger the material exposure in a spatial and temporal manner.

Engagement may also be considered as the extent to which there is learner-content interaction. Between studies, this interaction varied from allowing users to select the viewpoint from which they viewed virtual anatomical specimens or choosing the order in which they viewed images of objects of interest. de Faria et al. (2016) demonstrated that learner-content interaction, independent of the learning modality (traditional *vs.* stereoscopic visualisation), was associated with increased neuroanatomy knowledge acquisition. A student's ability to interact with an object encourages Holmberg's 'internal didactic conversation,' where learners 'talk to themselves' about the information they encounter (Holmberg 1986). This encourages a change from observation to understanding and perception of knowledge, thus enhances learning. The level of learner-content interaction, particularly how much a learner can choose their views of an object, is a therefore a key variable to be considered in future studies of technology-enhanced neuroanatomical resources.

Cognitive load theory

Cognitive load theory also explains why some modalities may work in some contexts over others; in particular extraneous overload, where a learner's cognitive capacity is exceeded in certain situations. The cognitive load theory states that to acquire biologically secondary knowledge, a learner must obtain novel information in small amounts compatible with the working memory's ability to process it (Sweller 2011). Stereoscopic 3D video and VR are good examples of where learners may experience the split-attention affect when presented with too much information simultaneously, and the extraneous cognitive load applied may have a negative influence on learning. These modalities may be better suited where less content is presented, and the novelty of the learning tools may be beneficial for engagement of the students.

AR and some select instances of VR (where structure labelling was presented in a more user-friendly manner) decreased cognitive load and facilitated student learning. This difference may have been due to presentation formats. For example, in Wismer et al. (2018), the VR model had labelled structures that could be toggled on and off, compared to the physical model that had numbered structures labelled on a piece of paper. The extra mental step in connecting numbered labels to structure identifiers may have explained the difference in spatial workload, mental demand, and frustration experienced by students in the physical model group. Similarly, AR allows information to be virtually overlayed onto the object, decreasing the time, and mental workload involved in matching information presented in a laboratory textbook and the specimen. AR, in terms of ease of access to smart phone technology and ease of user interface, may serve as the ideal intermediate.

Complex spatial relationships

Three-dimensional stereoscopic videos were no more advantageous in neuroanatomy knowledge acquisition than viewing in 2D. However, there seemed to be marginally more evidence for the efficacy of learning by 3D stereoscopic images than 2D. A unifying theme is that stereoscopy is most effective compared to learning by 2D images when teaching students complex spatial anatomy. That is, a 3D representation of images is most useful when anatomical relationships are difficult to conceptualise, such as deep-brain structures or C-shaped features of the brain including the cingulum, corpus callosum, fornix, lateral ventricles, and caudate nucleus. There may be additional benefit to those students with a low-spatial ability, and may also be used in identifying small structures, or those with complex relationships where the learner may have difficulty viewing during routine anatomy laboratories, such as the middle and inner ear anatomy of the temporal bone (Cui et al. 2017). A possible explanation for the limited role of stereoscopic *vs.* 2D views of videos is that in a video, students have a chance to appreciate complex anatomical relationships while the subject matter moves through various view-points. A stereoscopic view may not offer the student any additional information compared to 2D. A useful study may be the comparison of 2D video of a 3D anatomical feature, and a 3D stereoscopic image of that feature. Similarly, comparison of an interactive 3D software that is displayed in 2D compared to stereoscopic view of a 2D image would rely on similar theoretical principles, with the added variable of interaction. Equivocal results may be supportive of this hypothesis.

The benefit of technology-enhanced teaching in understanding complex spatial anatomy was also true in selective studies that investigated AR or VR, particularly where spatial question types were investigated separately. AR facilitated the learning of spinal cord pathways in one paper (Küçük et al. 2016), but resulted in poorer knowledge acquisition, especially in cross-sectional anatomy, in the other (Henssen et al. 2020). The study methodology, in particular the training needed by participants in how to use the technology, may have resulted in differences in the findings as discussed. In one study, a group learning by mixed reality retained information better in nominal- and spatial-type questions after a month (Wang et al. 2020). Similar to the findings of stereoscopic and AR technologies, there may be an advantage in VR allowing the user to understand complex spatial relationships more easily than 2D presentations. Other studies have investigated alternative pedagogies such as play-doh or printed models that improve understanding of spatial relationships effectively, and may be a cheaper alternative (Estevez et al. 2010; McMenamin et al. 2014).

It is possible that novelty of the technology encouraged students to study more, and the measurable differences in knowledge acquisition can be attributed to it requiring 'work' on the part of the student, rather than the technology itself. However, it is difficult to quantify this relative affect. While not inferior, the expense of VR technologies may be prohibitive for many institutions to roll-out this teaching method broadly. Finally, an important consideration was that VR made some students physically ill, a side-effect that will have to be further investigated and minimised in future iterations for duty of care.

Technology learning curve

A limitation across many mixed reality studies was the time required to teach students how to use the teaching method. This may have resulted in student performance being spuriously low in experimental groups due to differences in familiarity with the learning tool. While a mixture of pedagogical techniques is theorised to be the best method of teaching neuroanatomy (Sotgiu et al. 2020), incorporating too many novel techniques may overwhelm students, who spend as much time learning how to use the tools as they do learning content. This may change as novel techniques become more mainstream, and technology literacy grows.

Limitations

While some papers specify the number of assessment items and question-types, most studies did not provide the test itself making evaluation of outcomes difficult to evaluate for quality. Further, the methodology behind the test administration was often flawed, as the post-intervention tests were done immediately following the intervention. The results of these studies, therefore, give a measure of short-term memory, as opposed to knowledge acquisition and retention. Tests conducted 1 week after the experiment or longer give a more robust assessment of learning. One feature that papers do not comment on is the social aspect of learning, and the relative collaborative/cooperative characteristics of novel technologies. Future studies may comment on these elements and their impact on learning.

As a focused review of the most recent technology-enhanced learning methods, this review included analysis of only 13 studies. Two authors' screening search items demonstrated moderate inter-rater reliability, raising the possibility of missed papers, and may be due to a 10-year difference in experience levels. New studies into technology-enhanced teaching continue to be added so it may be relevant to establish a running commentary in a public forum. However, the intention is to update this review in three years for a systematic appraisal of new evidence, as per the original protocol.

Conclusion

To date, technology-enhanced teaching methods are not inferior to conventional didactic methods, although their efficacy in learning neuroanatomy may not be as groundbreaking as one would originally have hoped. So far, there is only weak-to-moderate evidence for the use of stereoscopic 3D images and AR, and only in particular settings. Limited engagement with content due to extraneous cognitive overload, and technology learning curves associated with new technologies are amongst the possible reasons for why these technologies are not performing as expected. However, there are promising results for technology-enhanced teaching in complex spatial anatomy and reducing cognitive load in some instances. Future research may validate the theorised reasons proposed in this review and develop and test innovative technologies that build on prior research.

Statement of compliance

The authors confirm that we have read the Journal's position on issues involved in ethical publication and affirm that this report is consistent with those guidelines.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The author(s) reported there is no funding associated with the work featured in this article.

ORCID

Hamish J. Newman http://orcid.org/0000-0001-7213-7290
Amanda J. Meyer http://orcid.org/0000-0001-7953-6040
Tim J. Wilkinson http://orcid.org/0000-0002-4080-4164
Nalini Pather http://orcid.org/0000-0001-5288-7713
Sandra. E. Carr http://orcid.org/0000-0003-0043-4362

References

Alexander B, Ashford-Rowe K, Barajas-Murph N, Dobbin G, Knott J, McCormack M, Pomerantz J, Seilhamer R, Weber N. 2019. EDUCAUSE horizon report 2019 higher education edition. EDU19.

Arantes M, Arantes J, Ferreira M. 2018. Tools and resources for neuroanatomy education: a systematic review. BMC Med Educ. 18(1):94.

Bernard F, Richard P, Kahn A, Fournier H-D. 2020. Does 3D stereoscopy support anatomical education? Surg Radiol Anat. 42(7):843–852.

Bujak KR, Radu I, Catrambone R, MacIntyre B, Zheng R, Golubski G. 2013. A psychological perspective on augmented reality in the mathematics classroom. Comput Educ. 68:536–544.

Colthart I, Bagnall G, Evans A, Allbutt H, Haig A, Illing J, McKinstry B. 2008. The effectiveness of self-assessment on the identification of learner needs, learner activity, and impact on clinical practice: BEME Guide no. 10. Med Teach. 30(2):124–145.

Cook DA, Bordage G, Schmidt HG. 2008. Description, justification and clarification: a framework for classifying the purposes of research in medical education. Med Educ. 42(2):128–133.

Craig S, Tait N, Boers D, McAndrew D. 2010. Review of anatomy education in Australian and New Zealand medical schools. ANZ J Surg. 80(4):212–216.

Cui D, Wilson TD, Rockhold RW, Lehman MN, Lynch JC. 2017. Evaluation of the effectiveness of 3D vascular stereoscopic models in anatomy instruction for first year medical students. Anat Sci Educ. 10(1):34–45.

de Faria JWV, Teixeira MJ, Júnior LMS, Otoch JP, Figueiredo EG. 2016. Virtual and stereoscopic anatomy: when virtual reality meets medical education. J Neurosurg. 125(5):1105–1111.

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(1):E103–E109.

Ertelt A. 2007. On-screen videos as an effective learning tool. The effect of instructional design variants and practice on learning achievements, retention, transfer and motivation dissertationsschrift. Baden-Württemberg, Germany: Albert-Ludwigs-Universität Freiburg.

Estevez ME, Lindgren KA, Bergethon PR. 2010. A novel three-dimensional tool for teaching human neuroanatomy. Anat Sci Ed. 3(6):309–317.

Ferdig R, Blank J, Kratcoski A, Clements R. 2015. Using stereoscopy to teach complex biological concepts. Adv Physiol Educ. 39(3): 205–208.

Giles J. 2010. Clinical neuroscience attachments: a student's view of 'neurophobia'. Clin Teach. 7(1):9–13.

Goodarzi A, Monti S, Lee D, Girgis F. 2017. Effect of stereoscopic anaglyphic 3-dimensional video didactics on learning neuroanatomy. World Neurosurg. 107:35–39.

Gordon M, Grafton-Clarke C, Hill E, Gurbutt D, Patricio M, Daniel M. 2019. Twelve tips for undertaking a focused systematic review in medical education. Med Teach. 41(11):1232–1238.

Held RT, Hui TT. 2011. A guide to stereoscopic 3D displays in medicine. Acad Radiol. 18(8):1035–1048.

Henn JS, Lemole GM, Jr Ferreira MA, Gonzalez LF, Schornak M, Preul MC, Spetzler RF. 2002. Interactive stereoscopic virtual reality: a new tool for neurosurgical education. Technical note. J Neurosurg. 96(1): 144–149.

Henssen DJHA, Heuvel L, Jong GD, Vorstenbosch MATM, Walsum A-M, Hurk MMVd, Kooloos JGM, Bartels RHMA. 2020. Neuroanatomy learning: augmented reality vs. cross-sections. Anat Sci Educ. 13(3): 353–365.

Higgins JP, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, Savović J, Schulz KF, Weeks L, Sterne JA, et al., 2011. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ. 343(2):d5928–d5928.

Hilbelink AJ. 2009. A measure of the effectiveness of incorporating 3D human anatomy into an online undergraduate laboratory. Br J Educ Technol. 40(4):664–672.

Holmberg B. 1986. Growth and structure of distance education.

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(10):1006–1013.

Jozefowicz RF. 1994. Neurophobia: the fear of neurology among medical students. Arch Neurol. 51(4):328–329.

Kamphuis C, Barsom E, Schijven M, Christoph N. 2014. Augmented reality in medical education? Perspect Med Educ. 3(4):300–311.

Kockro RA, Amaxopoulou C, Killeen T, Wagner W, Reisch R, Schwandt E, Gutenberg A, Giese A, Stofft E, Stadie AT. 2015. Stereoscopic neuroanatomy lectures using a three-dimensional virtual reality environment. Ann Anat-Anatomisch Anzeiger. 201:91–98.

Küçük S, Kapakin S, Göktaş Y. 2016. Learning anatomy via mobile augmented reality: effects on achievement and cognitive load. Anat Sci Educ. 9(5):411–421.

Landis JR, Koch GG. 1977. The measurement of observer agreement for categorical data. Biometrics. 33(1):159–174.

Levinson AJ, Weaver B, Garside S, McGinn H, Norman GR. 2007. Virtual reality and brain anatomy: a randomised trial of e-learning instructional designs. Med Educ. 41(5):495–501.

Lowe RK. 2003. Animation and learning: selective processing of information in dynamic graphics. Learn Inst. 13(2):157–176.

Mayer RE, Fiorella L. 2014. 12 principles for reducing extraneous processing in multimedia learning: coherence, signaling, redundancy, spatial contiguity, and temporal contiguity principles. The Cambridge handbook of multimedia learning. Cambridge: Cambridge University Press.

Mayer RE, Johnson CI. 2008. Revising the redundancy principle in multimedia learning. J Educ Psychol. 100(2):380–386.

McKeown P, Heylings D, Stevenson M, McKelvey K, Nixon JR, McCluskey D. 2003. The impact of curricular change on medical students' knowledge of anatomy. Med Educ. 37(11):954–961.

McMenamin PG, Quayle MR, McHenry CR, Adams JW. 2014. The production of anatomical teaching resources using three-dimensional (3D) printing technology. Anat Sci Educ. 7(6):479–486.

Meyer AJ, Stomski NJ, Losco CD, Armson AJ. 2016. The influence of anatomy app use on chiropractic students' learning outcomes: a randomised controlled trial. Chiropr Man Therap. 24(1):1–5.

Moore MG. 1989. Three types of interaction. New York (NY): Taylor & Francis.

Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med. 6(7):e1000097.

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(6):549–559.

Moxham B, McHanwell S, Plaisant O, Pais D. 2015. A core syllabus for the teaching of neuroanatomy to medical students. Clin Anat. 28(6): 706–716.

Schon F, Hart P, Fernandez C. 2002. Is clinical neurology really so difficult. J Neurol Neurosurg Psychiatry 72: 557–559.

Newman H, Meyer AJ, Carr S, Pather N, Wilkinson TJ. 2020. Comparative analysis of modern neuroanatomy teaching techniques: a BEME systematic review Protocol. Best Evidence Medical Education.

Prince KJ, Scherpbier AJ, Van Mameren H, Drukker J, Van Der Vleuten CP. 2005. Do students have sufficient knowledge of clinical anatomy? Med Educ. 39(3):326–332.

Proffitt DR, Kaiser MK, Whelan SM. 1990. Understanding wheel dynamics. Cogn Psychol. 22(3):342–373.

Schober P, Boer C, Schwarte LA. 2018. Correlation coefficients: appropriate use and interpretation. Anesth Analg. 126(5):1763–1768.

Shepard RN, Metzler J. 1971. Mental rotation of three-dimensional objects. Science. 171(3972):701–703.

Sorenson J, Khan N, Couldwell W, Robertson J. 2016. The rhoton collection. World Neurosurg. 92:649–652.

Sotgiu MA, Mazzarello V, Bandiera P, Madeddu R, Montella A, Moxham B. 2020. Neuroanatomy, the Achille's heel of medical students. A systematic analysis of educational strategies for the teaching of neuroanatomy. Anat Sci Educ. 13(1):107–116.

Sweller J. 2011. Cognitive load theory. Psychology of learning and motivation. Amsterdam, Netherlands: Elsevier; p. 37–76.

Wang C, Daniel BK, Asil M, Khwaounjoo P, Cakmak YO. 2020. A randomised control trial and comparative analysis of multi-dimensional learning tools in anatomy. Sci Rep. 10(1):1–10.

Waterston SW, Stewart IJ. 2005. Survey of clinicians' attitudes to the anatomical teaching and knowledge of medical students. Clin Anat. 18(5):380–384.

Wismer A, Reinerman-Jones L, Teo G, Willis S, McCracken K, Hackett M. 2018. Assessing performance and usability of 3D visualization technologies for anatomical training. MODSIM World. 17:1–11.

Wu HK, Lee SWY, Chang HY, Liang JC. 2013. Current status, opportunities and challenges of augmented reality in education. Comput Educ. 62:41–49.

Yuen SCY, Yaoyuneyong G, Johnson E. 2011. Augmented reality: An overview and five directions for AR in education. J Educ Technol Dev Exchange. 4(1):11.

Zinchuk AV, Flanagan EP, Tubridy NJ, Miller WA, McCullough LD. 2010. Attitudes of US medical trainees towards neurology education: "Neurophobia" - a global issue. BMC Med Educ. 10(1):49.