
Article:
Clunie, L, Morris, N orcid.org/0000-0003-4448-9381, Joynes, VCT et al. (1 more author) (2018) How Comprehensive are Research Studies Investigating the Efficacy of Technology-Enhanced Learning Resources in Anatomy Education? A Systematic Review. Anatomical Sciences Education, 11 (3). pp. 303-319. ISSN 1935-9772

https://doi.org/10.1002/ase.1762

ASE-17-0149.R2

Relevant Review

How Comprehensive are Research Studies Investigating the Efficacy of Technology-Enhanced Learning Resources in Anatomy Education? A Systematic Review

Lauren Clunie[1], Neil P. Morris[2], Viktoria C. T. Joynes[3], James D. Pickering[1,2]*

[1]Division of Anatomy, Leeds Institute of Medical Education, School of Medicine, University of Leeds, Leeds, United Kingdom

[2]Research Centre in Digital Learning, School of Education, University of Leeds, Leeds, United Kingdom

[3]Institute of Clinical Sciences, School of Medicine, University of Liverpool, Liverpool, United Kingdom

Running title: Anatomy TEL Evaluation

*Correspondence to: Dr. James Pickering, Division of Anatomy, Leeds Institute of Medical Education, School of Medicine, 9.06 Worsley Building, Clarendon Way, University of Leeds, Leeds, LS2 9NL, UK. E-mail: j.d.pickering@leeds.ac.uk

ABSTRACT

Anatomy education is at the forefront of integrating innovative technologies into its curricula. However, despite this rise in technology numerous authors have commented on the shortfall in efficacy studies to assess the impact such technology-enhanced learning (TEL) resources have on learning. To assess the range of evaluation approaches to TEL across anatomy education, a systematic review was conducted using MEDLINE, the Educational Resources Information Centre (ERIC), Scopus and Google Scholar, with a total of 3,345 articles retrieved. Following the PRISMA method for reporting items, 153 articles were identified and reviewed against a published framework – the technology-enhanced learning evaluation model (TELEM). The model allowed published reports to be categorized according to evaluations at the level of (1) learner satisfaction, (2) learning gain, (3) learner impact and (4) institutional impact. The results of this systematic review reveal that most evaluation studies into TEL within anatomy curricula were based on learner satisfaction, followed by module or course learning outcomes. Randomized controlled studies assessing learning gain with a specific TEL resource were in a minority, with no studies reporting a comprehensive assessment on the overall impact of introducing a specific TEL resource (e.g., return on investment). This systematic review has provided clear evidence that anatomy education is engaged in evaluating the impact of TEL resources on student education, although its remains at a level that fails to provide comprehensive causative evidence.

Key Words: Technology-enhanced learning, evaluation, anatomy education, curriculum design

INTRODUCTION

Anatomy education is at the forefront of utilizing the latest technological advancements to develop increasingly blended learning environments. As Trelease (2016) recently described in a comprehensive review, this approach has significantly shifted the learning and teaching of anatomy from a relatively unenhanced position to the cutting edge. Such changes to anatomy curricula are becoming increasingly documented for all aspects of teaching and learning, including face-to-face sessions supported by faculty members, and periods of self-directed learning, where students consolidate and revise course material.

This changing approach to anatomy education delivery is underpinned by a number of multi-factorial drivers, including: the availability and logistics of cadaveric resources (McLachlan et al., 2004; McLachlan and Patten, 2006), the relevance of anatomy in a modern and expanding medical curriculum (Cottam, 1999; McKeown et al., 2003; Turney, 2007; Louw et al., 2009), increasing student numbers, decreasing available curriculum time to teach the required anatomy, and pedagogical approaches (Heylings, 2002; Drake et al., 2009; Bergman et al., 2014; Drake et al., 2014; Freeman et al., 2014; Chen et al., 2017). This change in anatomy education approach has been long-standing and can be tracked back to the introduction of the personal computer (PC) almost 30 years ago (Trelease, 2008), with great strides being made since, including: two-dimensional (2D) and three-dimensional (3D) applications (Evans, 2011; Lewis et al., 2014; Pickering, 2015a, 2016a), eBooks (Mayfield et al., 2013; Stirling and Birt, 2014; Pickering, 2015b; Stewart and Choudhury, 2015), social media (Jaffar, 2014; Raikos and Waidyasekara, 2014; Hennessy et al., 2016; Pickering and Bickerdike, 2016) lecture webcasts (Vaccani et al., 2016), 3D printing of replica specimens (McMenamin et al., 2014; Reilly et al., 2016), discussion fora (Choudhury and Gouldsborough, 2012; Green et al., 2014) massive open

online courses (MOOCs; Reinders and de Jong, 2016; Swinnerton et al., 2017), and virtual and augmented reality (Moro et al., 2017), all becoming established mediums through which anatomy content can be delivered. This diffusion of innovation into higher education can be observed alongside changing approaches to curriculum design with the increasing use of active learning techniques (Freeman et al., 2014) and flipped classrooms (Chen et al., 2017) enabled by such TEL resources.

However, given the well documented change in approach to anatomy education, it is important that upon the introduction of TEL resources a robust evaluation of efficacy is conducted. This desire has been longstanding with McLachlan and Patten commenting over a decade ago that the field of evaluation was *'the single most desirable improvement in anatomy teaching'* (McLachlan and Patten, 2006), and more recently Trelease (2016) commenting that *'e-learning innovations in anatomical sciences education currently suffer from a scarcity of statistically reliable learning efficacy evidence'*. Furthermore, and despite this desire, there remains only an emerging level of evaluation into both the short- and long-term impact individual TEL resources have on student education (Tworek et al., 2013; Colliver and Cianciolo, 2014; Cook and Ellaway, 2015; Pickering and Joynes, 2016; Trelease, 2016; Pickering, 2017a). Recently, some comprehensive studies have attempted to address this issue with a series of meta-analyses detailing the impact optical and virtual microscopy, 3D visualization technologies, physical models, and laboratory pedagogies, have on anatomy education (Wilson et al., 2016; Yammine and Violato, 2015, 2016; Wilson et al., 2017). This in-depth understanding of the impact such TEL resources have on student learning is of paramount importance if faculty wish to make informed decisions into the best options available when developing, reviewing or wanting to introduce and a new learning tool.

Across the medical education discipline, this desire to evaluate has been supported by a growing number of evaluation frameworks that endeavor to understand the impact teaching interventions have on student learning (Frye and Hemmer, 2012). Of these, Kirkpatrick's model of evaluation is the most widely cited and influential (Kirkpatrick, 1994, 2017), consisting of four levels that are based on learning outcomes as a measure of program impact and behavioral change. However, due to the multi-faceted nature of anatomy curricula currently being developed, and as these frameworks typically attempt to assess the impact at the level of the program or course, utilizing such an approach can fail to draw out the specific impact individual TEL resources have on student outcomes. This approach to evaluation has been criticized due to its reductionist approach, in that the changes observed at the program level are solely attributed to the new intervention (Holton, 1996; Yardley and Dornan, 2012). Moreover, this assumes a certain linearity of the program, with a clear cause and effect that is often difficult - if not impossible - to achieve in educational settings. In an attempt to remedy this shortfall, two evaluation frameworks have been proposed that are specifically focused on the role of TEL in medical education (Cook and Ellaway, 2015; Pickering and Joynes, 2016b). The framework put forward by Cook and Ellaway (2015) suggests that a thorough evaluation of TEL resources must encompass seven broad areas that are unique to TEL (e.g., usability, student experience and cost-analysis), with the intention of providing meaningful comparison between institutions. Similar to the desired outcomes of Kirkpatrick (1994, 2017), this particular framework provides a protocol that is heavily based on evaluation at the program or course level. Most recently, an additional TEL evaluation model (TELEM) has been developed that builds on existing frameworks and focuses on understanding the impact individual resources have within a resource heavy curriculum (Pickering and Joynes, 2016b). Building on the work of Kirkpatrick (1994, 2017), the

TELEM encompasses four-levels of evaluation that aims to examine: learner satisfaction, learning gain, learner impact and institutional impact, through a diverse and extensive range of both qualitative and quantitative methodologies to achieve a more holistic overview of the TEL resources' impact.

BASIS FOR SYSTEMATIC REVIEW

With an increasingly diverse range of TEL resources being introduced into anatomy education, and with the increasing levels of acceptance for their use within medical education, it is inconceivable that any institution would consider withdrawing its use from their curriculum (Fuller and Joynes, 2015; Lumsden et al., 2015). Although, the levels of student satisfaction and the types of devices available are generally well understood (Koehler, 2012; Wallace et al., 2012; Chen and Denoyelles, 2013), further empirical research is needed to fully explore the efficacy of such TEL-based resources to ensure their effective integration into anatomy curricula. It is therefore, within the context of increasing reliance on, and integration of, technology in anatomy education, that this systematic review has examined the scope of evaluation within research studies evaluating the impact of TEL resources using the TELEM as a benchmarking tool. Only when faculty are fully aware of the efficacy of such tools, can meaningful decisions be made on their introduction into learning environments.

Summary of the Technology-Enhanced Learning Evaluation Model

The TELEM consists of four levels (Fig. 1), with each level summarized below (for more information on the model please refer to Walsh et al., (2013) and Pickering and Joynes (2016b)):

- **Level 0** is a preliminary evaluation of need that assesses the requirements for introducing a TEL resource. The intention of this stage is to ensure that technology is the most appropriate solution to either a curriculum problem that needs remedying or an alternative approach to meeting the course's learning objectives. Once the need has been established a development phase begins that leads to either the in-house creation or commercial procurement of the relevant TEL resource.
- **Level 1** is divided into two parts: 1a – learner satisfaction and 1b – learning gain.
  Level 1a of the TELEM model examines the levels of satisfaction with the newly introduced resource by way of well–developed Likert-style questionnaires and qualitative approaches, such as focus–groups. Although the primary goal of the evaluation model is to examine the efficacy of a specific resource, for students to engage with the resources it must be user-friendly and enjoyable (Van Nuland et al., 2016; Kirschner, 2016).
  Level 1b assesses the specific impact the resource has on learning gain in a controlled environment via a randomized-controlled trial format. Given the ethical and educational restrictions on this approach, this level would utilize volunteers and a within-subject/repeated-measures design to determine whether the resource is effective and efficient in enhancing learning gain compared to an alternative resource. Recruiting volunteers and deploying a well-established experimental protocol, such as a pre- and post-test design, causative data can be obtained to assess TEL resource efficacy, with confounding variables limited.
- **Level 2** takes a holistic and correlational approach to assess how the TEL resource impacts the student in regard to summative assessment outcomes within an active curriculum. This level utilizes a combination of quantitative (e.g., learning analytics on

usage) and qualitative (e.g., questionnaire and focus groups) approaches in an attempt to link the level of usage and assessment outcomes by comprehensively investigating student access and utilization. By combining the evaluation data from Levels 1 and 2 a holistic view of the impact a specific TEL resource has on student learning can be achieved that goes beyond the reductionist approach of other frameworks.

- **Level 3** is concerned with assessing the specific TEL resource's cost-feasibility, that is, given the information obtained on the efficacy of the TEL resource, is its continued deployment viable in regard to changes in learning gain, the impact on the individual learner and institution? This level draws on the work by Walsh et al. (2013) and is the most complex level, requiring input from a broad range of students and faculty to create the necessary institutional benchmarks that the TEL resource will be judged against. For a full cost-feasibility analysis to be achieved each level is associated with a specific cost-analysis approach (Fig. 1.). Level 1a (learner satisfaction) is associated with cost-utility. This uses a subjective assessment to assign a monetary cost per student for providing the TEL resource in relation to the levels of satisfaction received. A judgement is then formed by comparing the monetary cost and satisfaction level of the TEL resource. Level 1b (learning gain) is associated with cost-effectiveness and compares the monetary cost of developing the TEL resource, in relation to its impact on learning gain. Level 2 (learner impact) is linked to a cost-benefit analysis, where the monetary cost of introducing the TEL resource into the curriculum is linked to the learning outcomes of the target student cohort. A study that attempts to undertake this multi-level evaluation would constitute a full cost-feasibility analysis and reach Level 3 of the TELEM.

Aim and Research Questions

The overall aim of this systematic review is to assess the scope of evaluation within research studies that evaluate the use of TEL resources in anatomy education. Given the extensive range of methodologies detailed within the TELEM, it was deemed an appropriate benchmarking tool to achieve this aim. In order to achieve this aim, the following research questions were developed: (1) How comprehensive are TEL resources across anatomy education being evaluated? and (2) What types of evaluation are currently being reported?

MATERIALS AND METHODS

Search Strategy

An electronic search of the following databases was conducted: MEDLINE (U.S. National Library of Medicine, Bethesda, MD), the Educational Resources Information Centre (ERIC) (United States Department of Education, Washington, DC), Scopus (Elsevier, Amsterdam, The Netherlands), and Google Scholar (Google Inc., Mountain View, CA) from the beginning of the research period until November 2, 2016. The search terms under the three categories, including: type of education, educational delivery method and technology-enhanced learning are detailed in Table 1. No date restriction was implemented since the use of technology in anatomy education is a relatively new phenomenon and is therefore self-limited to the last two decades. Additional articles were identified by manually searching reference lists of other reviews, related review articles and authors' files. The Preferred Reporting Items for Systematic Reviews and Meta Analyses (PRISMA) was used to report findings (Moher et al., 2009).

Inclusion and Exclusion Criteria

All citation titles and abstracts were initially screened, with the full text retrieved for all potentially eligible abstracts with insufficient information. Inclusion of citations was only considered after review of the full text. Citations were included if they were found to be specific to basic science anatomy education (including neuroanatomy, histology and embryology) and contained an evaluation protocol that detailed student learning. Citations were excluded if they focused on clinical training, including interpretation of radiological images in clinical settings, and if the citation was based on the evaluation of the technology itself (e.g., interactivity and usability). There were no geographical restrictions and only texts that were available in English were included. The full text of those citations retained were reviewed by two of the authors (L.C., J.D.P.), with conflicts discussed until a consensus was reached.

Data Extraction

The following data was extracted from the citations selected for inclusion in the review: sample size and subject area (e.g., medicine, dentistry, nursing, allied healthcare and biomedical science), length and type of study, year of publication, learning and teaching setting (e.g., classroom, self-directed and anatomy laboratory), evaluation methodology (e.g., pre/post testing, comparison of summative assessment scores and learner satisfaction surveys) and instructional modality (e.g., computer assisted learning tool, mobile devices, online learning, 3D printing, virtual reality, augmented reality). Once the data had been extracted, each citation was categorized by the level of evaluation reported in accordance with those documented in the TELEM. To assess inter-rater reliability the percent agreement was calculated, along with the Kappa coefficient to take into consideration the possibility of chance agreement.

Data Analysis

Descriptive analysis was performed on the citations included and reported the number that corresponded to at least one TELEM level and its year of publication. The number of citations that included no evaluation, but highlighted an innovative new approach to anatomy educational using technology, were also included to inform historical perspectives (referred to as 'show and tell'). Data handling was performed using Microsoft Excel 2015, version 15.14 (Microsoft Corp., Redmond, WA), with figures exported to Illustrator, Adobe CS6, version 16.0.4 (Adobe Systems Software, Ireland Ltd., Dublin, Ireland) for editing.

RESULTS

Descriptive Analysis of Included Studies

Following the approach outlined in Figure 2, the electronic databases yielded 3,259 citations, with a further 86 identified from additional sources, resulting in a total of 3,345. Once duplicates were removed, 1,371 citations were identified with 553 of these removed as they were deemed ineligible. The remaining 818 citations were considered for full review, with 665 of these then excluded for the following reasons: not human anatomy education (25.6%; 170 of 665), show and tell (18.6%; 124 of 665), clinical training (13.7%; 91 of 665), viewpoint or review article (12.6%; 84 of 665), technicality (12.2%; 81 of 665), curriculum assessment based (11.7%; 78 of 665), TEL evaluation out of context (3.0%; 20 of 665), TEL development instruction (1.7%; 11 of 665), and duplicates (0.9%; 6 of 665). Citations that described the development and/or implementation of a TEL resource, but did not mention if any evaluation was performed were classified as show and tell. Overall, 153 citations were deemed to be eligible, retained for the systematic review, and assigned to one or more of the four TELEM levels by L.C. Of these,

25% were independently assessed by author (J.D.P.) for confirmation. The reviewers disagreed on two studies, yielding a percent agreement of 94.7% and a Kappa coefficient of 0.92. The two papers were discussed and a consensus reached.

Scrutinizing the approach to evaluation detailed in each of the eligible studies against the TELEM revealed that a small majority (52.3%; 80 of 153) carried out a multi-level approach, with 47.7% (73 of 153) only reporting a single level. Figure 3 details the proportion of studies that conducted an evaluation at either a single (1a, 1b or 2 only) or multiple (1a and 1b or 1a and 2) level. The most common combination of evaluation levels was at 1a and 2 (39.2%; 60 of 153), followed by level 1a only (30.7%; 47 of 153). These two approaches were identified to be the most popular with level 1a and 1b (13.1%; 20 of 153), level 1b only (8.5%; 13 of 153) and level 2 only (8.5%; 13 of 153) all being similarly lower in proportion. No studies reported an evaluation at level 3. By pooling the data by individual level of evaluation (i.e., the total number of studies that included each evaluation level), the most popular method of evaluation was level 1a (83.0%; 127 of 153), then level 2 (47.7%; 73 of 153), followed by level 1b (21.6%; 33 of 153).

Assessing the citations by year of publication revealed that the earliest study to report an evaluation in anatomy education was in 1987 and detailed student satisfaction with anatomy videotapes (Ogunranti, 1987). Since then the number of published studies that include an evaluation has increased annually over the last two decades, with 64.7% (99 of 153) of these published since 2010. The number of evaluation studies published in five year periods is displayed in Figure 4 and includes the frequency for each level of evaluation achieved starting from 1996, respectively. Only three show and tell articles were published prior to 1996, in 1980, 1992 and 1994 (Bellardini et al., 1980; Conley et al., 1992; Packer, 1994), respectively, and one

at level 1a in l987 (Ogunranti, 1987). In addition, Figure 4, also tracks the number of 'show and tell' papers that have been published in the same time period and although a steady increase in such papers is revealed, they have subsequently been superseded by evaluation based studies since 2006.

Types of Technology-Enhanced Learning Resources Evaluated

Reviewing the published literature on TEL in anatomy education yielded a wide variety of resources that are currently embedded within curricula. These included: (1) instructor-developed resources, which accounted for the highest proportion (39.2%; 60 of 153), and includes resources such as videos, podcasts and computer assisted instructional tools; (2) virtual 3D computer models (33.4%; 51 of 153), which includes any virtual or augmented reality representation of anatomical or histological structures; (3) online repository resources (12.4%; 19 of 153), including any resource stored on virtual learning environment (VLEs)/learning management systems (LMSs), such as discussion fora, online lectures and massive open online courses (MOOCs); (4) mobile devices (9.8%; 15 of 153), including mobile applications (apps) and eBooks; (5) purpose-built resources (2.6%; 4 of 153) such as "Virtual Dissection" tables, holograms and 3D printed anatomical specimens and; (6) social media (2.6%; 4 of 153) such as Facebook (Facebook Inc., Menlo Park, CA) and Twitter (Twitter Inc., San Francisco, CA).

Types of Evaluation within Eligible Studies

Table 2 details the eligible studies by resource type, along with the assigned level of evaluation in accordance with the TELEM.

Level 1a (Learner satisfaction). The favored method for achieving this level of evaluation was via student surveys, with intervention-specific surveys (71.7%; 91 of 127; e.g., Brewer et al., 2012; Stirling and Birt, 2014; Ferrer-Torregrosa et al., 2015), adaptations to existing module evaluations (18.1%; 23 of 127; e.g., Choudhury et al., 2010; Barbeau et al., 2013; Wilkinson and Barter, 2016), or surveys to compare traditional resources with TEL resources (10.2%; 13 of 127; e.g., Corton et al., 2006; Adamczyk et al., 2009; Hopkins et al., 2011), all reported. The most popular approach was via Likert scale questions (e.g., McNulty et al., 2009; Wright and Hendricson, 2010; O'Reilly et al., 2016), with 81.8% (104 of 127) of the total number of papers reporting level 1a utilizing this approach. Of these, 35.6% (37 of 104) used either statements presented as standalone questions (e.g., McNulty et al., 2000; Hu et al., 2010), 55.8% (58 of 104) incorporated additional open-ended questions (e.g., Beale et al., 2014; Traser et al., 2015), and 8.6% (9 of 104) incorporated qualitative approaches with either focus groups or interviews (e.g., Tworek et al., 2013; Ocak and Topal, 2015; Swinnerton et al., 2017). A number of studies (13.4%; 17 of 127) mentioned student feedback (e.g., Ogunranti, 1987; Chopra et al., 2012), but did not reveal the details of the methods used to collect this information.

Level 1b (Learning gain). From the studies that reported a 1b level of evaluation, the majority (57.6%; 19 of 33) deployed a pre- and post-test methodology using controlled conditions to limit the influence of any confounding variables on test scores (e.g., Tan et al., 2012; Stirling and Birt, 2014; Pickering 2016a). Other methods included the use of post-test data alone (36.4%; 12 of 33; e.g., Bogacki et al., 2004; Chan et al., 2015), the individual's existing GPA as a measure of baseline knowledge (3.0%; 1 of 33; Hallgren et al, 2002), or other subjective measures, such as a drawing test (3.0%; 1 of 33; Das and Michell, 2013). Variations in the approach to conducting

this level of evaluation was observed, with randomized control approaches using a control group and either one (e.g., Levinson et al., 2007; Pickering, 2016a) or two (Hopkins et al., 2011) experimental groups. Furthermore, crossover study designs were also reported with a pre- and post-test deployed either side of a teaching intervention, and students then permitted to experience the alternative intervention (e.g., Allen et al., 2016).

Level 2 (Learner impact). A large proportion (65.8%; 48 of 73) of studies that evaluated at level 2 compared a previous cohort of students ('control') with subsequent cohort(s) who had access to the new TEL resource ('experimental'; e.g., Pereira et al., 2004; Braun and Kearns, 2008; Evans, 2011). For example, Morris et al. (2016) assessed examination performance in a neuroanatomy course over three years, with the first cohort acting as the control group, and the subsequent two cohorts provided with neuroanatomy apps on tablet devices during a tutorial class, acting as the experimental group. The remaining 34.2% (25 of 73) evaluated the impact of a TEL resource within the same cohort (e.g., Chopra et al., 2012; Pickering, 2015b). The effectiveness of such approaches to TEL resource evaluation was measured by either comparing the assessment scores at the end of the respective course (79.5%; 58 of 73; e.g., O'Byrne et al., 2008; Lee et al., 2012; Traser et al., 2015), or by making correlations between assessment scores and usage metrics (20.5%; 15 of 73; e.g., Green et al., 2013; Green et al., 2014; Choi-Lundberg et al., 2015).

Level 3 (Institutional impact). From the total number of included articles, none reported carrying out a full cost –feasibility analysis at level 3. Some studies made reference to the cost of introducing a TEL resource into the curriculum (34.6%; 53 of 153). Raney (2015) reported on the costs of mobile applications, but more often the only reference made was to the financial cost

of the resource itself (e.g., Richardson-Hatcher et al. 2014; Rinaldi et al. 2016). In a small number of studies a dedicated section of the article was devoted to the topic of cost, with these varying from extended passages on the financial costs of the resources (e.g., Attardi and Rogers, 2015), to discussions on the cost-effectiveness of implementing a new resource (e.g., O'Byrne et al., 2008; Traser et al., 2015), and comparisons with the cost of a new resource compared to the traditional resources, such as printed text (e.g., Raynor and Iggulden, 2008) or cadaveric dissection (e.g., Hisley et al., 2008).

DISCUSSION

The rapid rise of technology integration into anatomy education has supported the creation of novel blended learning approaches to support student education. However, although this active integration of technology into anatomy curricula is extensive, as many authors have noted, there persists a paucity of empirical evidence on the efficacy of such interventions to meaningfully justify their inclusion as effective learning tools (McLachlan and Patten, 2006; Tworek et al., 2013; Colliver and Cianciolo, 2014; Cook and Ellaway, 2015; Pickering and Joynes, 2016b; Pickering, 2017a;). This systematic review has aimed to provide a picture of the current scope of research within anatomy education, and highlights the need for further robust evaluation that moves beyond student satisfaction. As the findings from this review suggest much work is currently underway, but this is rooted in student satisfaction and user perceptions rather than quantifiable changes in learning outcomes. Given the current changes in regard to anatomy education (e.g., reduction in available teaching hours and renewed focus on relevance), it is important that such TEL resources are evaluated to ensure they are providing equitable learning

gains irrespective of satisfaction and enjoyment. Only when educators have a clear understanding on the efficacy can meaningful decisions on deployment be made.

Using the TELEM to scrutinize the level of TEL evaluation currently underway across anatomy education, a propensity for understanding the impact on student satisfaction (Level 1a – Learner satisfaction) was revealed as the preferred approach for most studies. Studies that assessed the impact on learning and knowledge gain either in controlled settings (Level 1b – Learning gain) or as part of a wider curriculum (Level 2 – Learner impact), however, were observed much less frequently. This strong emphasis on student satisfaction as a measure of TEL resource evaluation supports a growing theme in higher education and aligns with recent reports that place students central to discussions on resource development and curricula design (Davis et al., 2014; Healey et al., 2014; Roberts et al., 2016; Border, 2017). Moreover, this desire to assess levels of student satisfaction and engagement are widespread within the literature with much attention focusing on the various forms of engagement (Krause and Coates, 2008; Dixson, 2015). However, although student satisfaction data can yield rich and valuable information on the utility of a resource, it should be noted that favorable attitudes or levels of engagement do not necessarily correlate with enhanced and sustained learning outcomes, or provide an accurate reflection of student behavior (Dixon, 1990; Holton, 1996; Kruger and Dunning, 1999; Jamieson-Noel and Winne, 2002). This latter point is not intended to necessarily discount or undermine the value of student satisfaction data entirely, but merely to try and distinguish it from any 'novelty effect' that may be present with the introduction of a novel resource. Student feedback on perceptions, satisfaction and interface design are all essential in supporting the development of a TEL resource, (Wiers-Jenssen et al., 2002; Van Nuland et al., 2016), however,

it remains the contention of the authors here that such data should not be presented as the only measure of success.

The underlying reasons for the high proportion of studies focusing on student satisfaction and not exclusively learning gain is likely to be multi-factorial, with faculty workload, curriculum design and the overall rationale for evaluating the TEL resource all contributing. This is evidenced with the very low proportion of studies that embarked on a detailed qualitative assessment of a TEL resource alongside the traditional questionnaire approach. It could be argued strongly that follow-up analysis focusing on a qualitative understanding, if conducted broadly and inclusive of a wide range of student profiles, can provide authentic insight into the underlying themes governing TEL resource utility (Stalmeijer et al., 2014; Tavakol and Sandars, 2014). This follow up analysis is particularly important if the underlying rationale for evaluating the resource is to measure levels of satisfaction and engagement with the resource, as talking with students about their experience, via interviews and focus groups, is a well-established approach that can draw out pertinent findings (Chan, 2009; Kirkwood and Price, 2014). The combination of in-depth qualitative approaches with quantitative data, can provide extensive and detailed insights into the self-perceived satisfaction levels, and clarify the role the TEL resource played in supporting learning (Dixon, 1990; Kirkwood and Price, 2014). This detailed analysis using multiple forms of data to create a clear picture of satisfaction, engagement and utility is essential, as it is widely recognized that an individual's own perceptions of their ability alone can often be over-inflated and not reflect their true knowledge base (Kruger and Dunning, 1999).

The second most reported method for evaluating the efficacy of a TEL resource was what the TELEM ascribes to level 2 (Learner impact). This level looks for changes in the overall assessment scores across the curriculum in which the TEL resource has been embedded, and

when collating the research studies a combination of level 2 with an assessment of student satisfaction (level 1a) was the most popular multi-level approach. This indicates a strong desire from faculty to explore both student satisfaction and assessment outcomes holistically. However, although providing a more diverse evaluation than using level 1a alone, attributing changes in assessment scores to TEL resource usage assumes a certain linearity in cause and effect that may be misleading due to a number of reasons. Firstly, although a TEL resource would have been deployed, it is the teacher's craft that decides how this tool is integrated into the course in conjunction with other immovable curriculum factors such as timetabling and room design. Secondly, at present usage statistics can often portray a level of usage that is inaccurate and unrepresentative of actual student use. For example, if usage data reveal that a student downloaded a resource or spent a specific period of time accessing a specific webpage, there is no easy way of knowing whether the student ever opened the resource after it had been downloaded or actually accessed an alternative webpage when the 'learning' webpage was also open. As mentioned in the future directions section, this level of data analysis will be much improved when learning analytics have become sufficiently sophisticated to counter such issues. Finally, the individual differences within students, such as spatial abilities and the daily fluctuations in motivation and cognitive load, will determine the impact such a resource has on learning. Many of these confounding variables are well documented in the literature and should be used in conjunction with the evaluation's findings on learner impact when drawing conclusions (Krause and Coates, 2008; Burgoon et al., 2012; Pizzimenti and Axelson, 2015; Abdel Meguid and Khalil, 2016; Iqbal, 2016; Pickering, 2017b).

Although these confounding variables will have varying degrees of impact on the students' ability to breach the desired learning objectives, and notwithstanding the issues

mentioned previously, courses with a large cohort of learners are likely to observe patterns that may indicate if any underlying impact exists. This may be particularly evident with studies that compare the assessment scores of the same cohort during one iteration of the course where confounding variables can be limited (Pickering and Bickerdike, 2017c), compared to those from different cohorts enrolled on different iterations of the same course (O'Byrne et al., 2008; Hoyek et al., 2014; Ahmad and Wright, 2014). Despite these drawbacks and viewed in the context of anatomy teachers having to design studies pragmatically to get around the curriculum's fixed components, the combination of level 1a and 2 was the most frequently reported type of evaluation, with the majority of studies that undertook a level 2 using usage metrics to compare assessment scores across cohorts.

Measuring any increase in usable and retained knowledge, often termed learning gain, can be achieved in a number of ways within the context of level 1b, with pre- and post-testing a reliable and popular methodology (Hake, 1998; Dimitrov and Rumrill, 2003; Issa et al., 2011; McGrath et al., 2015). The benefit of this approach is that it allows causation to be explored to a much greater extent than by comparing end of course assessments at level 2, however, it was found to be used only in a limited number of research studies. Although this approach can yield valuable causative data by controlling for confounding variables that can contribute to an individuals' learning gain, within an educational setting it is often difficult to create these conditions due to curriculum time constraints, student recruitment issues and ethical considerations (Boileau et al., 2017). Unlike the biomedicine or engineering industries, which routinely use randomized controlled trials as the gold standard, it is simply not possible to create ideal control and experimental groups with human participants. Furthermore, even with studies that were able to create such experimental conditions, a commonly understood and appreciated

limitation of pre- and post-test design is the inability to control an individual's acquisition of knowledge by factors outside the study's design (Bonate, 2000; Pickering, 2016a). However, it should be noted that the degree of influence from these confounding variables is much reduced compared to an evaluation at level 2. Together, all these factors may account for evaluations at level 1b being reported least often and it is therefore unsurprising to observe variations in how this level of evaluation was conducted.

One of the benefits from those studies that were able to develop such experimental conditions, is the ability to control for variables such as spatial ability (Hu et al., 2010; Tan et al., 2012) and cognitive load (Van Nuland and Rogers, 2015). By embarking on a study that controls for such variables, clear insights can be ascertained to support how a resource can be deployed throughout a curriculum, and importantly what supporting material is required to prepare students for using technology within their course. Although many students may enter university-level education with a seemingly intuitive understanding of technology and mobile devices, the so-called 'digital native', numerous reports have highlighted how it is important not to assume such individuals exist (Kirschner and van Merriënboer, 2013; Selwyn, 2016; Kirschner and De Bruyckere, 2017). Although difficult, a study at this level of evaluation in conjunction with an in-depth understanding of student satisfaction with a resource, can provide strong causal evidence to support the inclusion or exclusion of a TEL resource within a curriculum (Pickering and Joynes, 2016b).

The final level of the TELEM is a cost-feasibility analysis that aims to assess the overall cost of introducing a new TEL resource into the curriculum, with cost-utility, cost-effectiveness and cost-benefit all matched to specific levels (Fig. 1). Although it is recognized that understanding the underlying costs of introducing a resource into a curriculum can be extensive

and subjective, calculating the specific return on investment can be particularly difficult to achieve (Walsh et al., 2013). This may account for the low number of studies that embarked on a full cost-feasibility analysis. Some models have attempted to identify the key components required to conduct an analysis, both prior to, and after the intervention has been embedded, but it is generally accepted that an analysis of this nature will consume considerable time and effort (Laurillard, 2007; Cook and Ellaway, 2015). While some studies did make reference to certain aspects of cost (Raynor and Iggulden, 2008; Hisley et al., 2008; O'Byrne et al., 2008; Richardson-Hatcher et al., 2014; Raney, 2015; Traser et al., 2015; Rinaldi et al., 2016;), the vast majority made no reference whatsoever. However, it must be noted that the low level of cost-feasibility studies throughout the anatomy education literature does not necessarily mean they are not being conducted within institutions, just that the data is not reaching the academic community through journal articles or other outlets.

This systematic review has attempted to highlight the current level of TEL evaluation within anatomy education using the TELEM as a benchmarking tool. Despite the widely held view that sufficient evaluation into the efficacy of TEL is limited (McLachlan and Patten, 2006; Tworek et al., 2013; Cook and Ellaway, 2015; Pickering and Joynes, 2016b; Trelease, 2016; Pickering, 2017a), anatomy education does in fact appear to be embarking on considerable evaluation. This is clear from Figure 4 that indicates the cross over from show and tell citations to studies of evaluation. This observation is to be expected as new technologies and innovations diffuse into the market and appear in various curricula, before being followed up with detailed evaluations. This theme is evident in the work by Petersson et al. (2009), who described the development of videos based on 3D vascular models, before going on to assess student satisfaction and comparison with assessment scores. Similarly, work by O'Reilly et al. (2016)

reported the methods for generating a 3D printed model of the lower limb, and then followed up this work with an evaluation of student perceptions and efficacy measured using a pre- and post-test design.

The steady increase in evaluation studies over time is an enlightening outcome from the systematic review and shows the development of evidence-based curricular design. As previously discussed, educational research comes with inherent issues, however, it is clear from those studies that attempted to gather causal evidence, that the necessary evidence required to support curriculum design is achievable. The results also reveal an expected lag-time between the more mature TEL resources, such as instructor-developed resources, and the newest forms, such as social media and virtual reality, which have the least amount of evaluations due to their emerging presence. This is a predictable occurrence as the balance between innovation and evaluation coexists as part of an iterative process that informs the development and integration of TEL into modern anatomy curricula. Although the level of evaluation across anatomy education is positive and shows a commitment to understand the role of TEL in improving curricula, the lack of causative studies means the current evidence base may not be sufficient to make sweeping recommendations and proposals for substantial change. However, within the context of education and the inherent difficulties of conducting research of this type, with the appropriate support offered to faculty members to successfully pursue such endeavors, anatomy education appears to be well placed to continue understanding the role of TEL.

In light of these findings within the relevant literature obtained through the systematic review process, three key themes have emerged on the range of approaches to evaluating TEL resources in anatomy education. Firstly, there is no pedagogical 'silver-bullet' that can cut through the complexity of educational research and provide a methodology that enables anatomy

teachers to make meaningful decisions with one evaluation tool. With all methodologies having advantages and disadvantages in providing clear empirical data it will often be the individual responsible for delivering the evaluation, alongside the overarching rationale for wanting to understand its impact, which will determine the adoption of such approaches. These factors will be diverse, plentiful and exhaustive. Educational research and scholarship is not purely scientific in the same way a clinical scientist or biologist would conduct and design a 'gold-standard' science experiment. Given the multi-modal nature of higher education it is not feasible given the ethical and often 'messy' environments of educational scholarship to create experimental and control groups that would be considered the 'gold-standard' in other disciplines (Sullivan 2011). Therefore, the way in which educational research is judged needs to reflect the diverse setting in which it is positioned, with anatomy teachers having to build up the individual pixels of a much larger picture using a broad range of quantitative and qualitative methodologies.

Secondly, and notwithstanding the issues mentioned above, the results presented from this systematic review highlight that the full range of experimental designs are possible. However, there appears to be a clear inclination for evaluation studies to focus on student satisfaction, perception and engagement in regard to anatomy TEL resources. Given the changes underway across anatomy education pedagogy, and therefore for anatomy teachers to make informed decision on the deployment of a TEL resource, more information on its learning efficacy would be beneficial and improve greatly the decision-making process. Currently, few meta-analyses are available that allow anatomy teachers to make informed decisions on the use of TEL (Yammine and Violato, 2015), with the number of research studies available to conduct these often small (Wilson et al., 2016; Yammine and Violato, 2016), reflecting the low level of learning efficacy citations eligible as highlighted in this systematic review. Those that do have a

high number of eligible studies are typically broad and include TEL studies alongside more traditional pedagogical approaches (Wilson et al., 2017). Although a valuable tool to inform practice, given the low level of studies eligible for such complex analyses and the varying approaches to deploy and evaluate TEL resources, issues of generalizability between student groups, courses and TEL resource is high.

Finally, reviewing the literature has highlighted the diverse settings in which anatomy education integrates TEL resources. Without doubt this will lead to how the resource is evaluated and also the potential impact it has on the individual student. A TEL resource based on a smart phone, tablet, desktop computer or even a larger more substantial piece of hardware, does not run the curriculum, it does not lead the teaching session. The way these tools are integrated are down to the ideas and experience of the teacher who has decided to approach a specific set of learning objectives with this tool. How learning efficacy can be separated to that which was purely down to the TEL resource and that which was purely down to the teacher is yet another elusive variable that will confound researchers ad infinitum.

Future Directions

The role of technology in higher education will continue to expand as new technologies are discovered and applied to the educational setting (Sharples et al., 2016). This continual innovation, however, needs to be matched with robust evaluation strategies that can provide answers to the why, how, and when questions. For example, why should this TEL resource be integrated into the anatomy curriculum? What tangible benefits are the students gaining in learning with this tool? How best can this resource be integrated? By formulating evaluation strategies that answer some, if not all, of these questions students will receive robust curricula

that appropriately utilizes TEL. As detailed in the systematic review, the low level of controlled studies is an area of concern, and although the randomized control approach will always be viewed as the gold standard in achieving causal findings, alternatives such as learning analytics and other holistic approaches may help to produce additional insights. Recent work across all education disciplines has focused on learning analytics as a tool to monitor, track and understand the interaction students have with the learning process (Saqr et al., 2017). By having access to this data, faculty are able to monitor, predict and identify earlier the students who are on the path to poor performance. This level of data can provide valuable insights into the role TEL resources have on student achievement.

A final area worthy of future exploration, given the high emphasis on student satisfaction as a clear metric for understanding TEL impact, is to ensure that these quantitative approaches are sufficiently robust. Although used commonly throughout higher education, self-report instruments used to monitor and assess certain behaviors can be imperfect, with learners often providing inaccurate and misjudged findings (Jamieson-Noel and Winne, 2002). Therefore, further work needs to be done on developing a robust quantitative approach to accurately measure student behavior with TEL using a validated survey instrument that can provide comparable insights across resources and institutions.

Limitations of the study

As with all systematic reviews, there are a number of limitations that should be documented to provide suitable background and context. This systematic review focused on the use of TEL resources in anatomy education and the degree to which these have been evaluated. It is therefore to be expected that some resources may have been missed if the keywords selected to locate such

studies failed to adequately locate them within the selected database. Although the range of keywords was extensive, there is the possibility that some studies may have been missed. In an attempt to counter this limitation, a manual search across a number of journals known to routinely publish such studies was also conducted. Similarly, studies may have been missed if the TEL resource, although used within a teaching intervention, was not clearly identified as a central component. An additional limitation could include the subjective, inconsistent or erroneous coding of eligible studies. Although possible, such discrepancies were mitigated for by having an in-depth understanding of the evaluation framework used to code the studies and a suitable methodology employed to assess inter-rater reliability.

Given the range of keywords used to select the studies, the period of time selected to gather appropriate studies, utilisation of relevant and extensive databases, and the manual search across known journals, the failure to identify eligible studies is likely to be limited.

CONCLUSIONS

This systematic review has comprehensively assessed the current level of TEL evaluation across anatomy education. The main conclusions from the review reveal that despite an increasing amount of TEL evaluation over the last two decades, the majority is descriptive and looking to draw simple correlations between the introduction of TEL resources and improved student feedback, rather than exploring for more meaningful causative relationships between TEL resources and improvements in learning. This is clearly evidenced with the majority of evaluation approaches addressing student satisfaction and course assessment outcomes, respectively. Only a minority of studies evaluated at the level of an individual TEL resource with a causative approach. This disparity is to be expected due to the nature of educational research

and the lack of opportunities afforded to faculty to conduct 'gold standard' approaches within an active curriculum. Although the high number of evaluation studies indicates a desire to understand the underlying efficacy of such resources, the lack of causative studies prevents overly authoritative conclusions being drawn on the impact specific TEL resources have on anatomy learning.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the supportive and constructive feedback received from the reviewers and editors on this manuscript.

NOTES ON CONTRIBUTORS

LAUREN CLUNIE, B.Sc., M.Sc., is a graduate (Ph.D.) student and anatomy demonstrator in the Division of Anatomy, School of Medicine at the University of Leeds. Her research interests are in the evaluation of technology-enhanced learning resources for anatomy education.

NEIL P. MORRIS, B.Sc., Ph.D., P.G.C.L.T.H.E., F.H.E.A., is a professor of educational technology, innovation and change in the School of Education, University of Leeds, Leeds, United Kingdom. He has considerable experience in teaching neuroscience to biomedical science and medical students with the use of technology, and is currently the university's Director of Digital Learning.

VIKTORIA C. T. JOYNES, B.A., M.Sc., Ph.D., is a lecturer in medical education and the Interim Director of Medical Studies at the School of Medicine, University of Liverpool, Liverpool, United Kingdom. She teaches professionalism, identity, and sociology and psychology as applied to medicine, and researches and publishes regularly in the field of technology-enhanced learning in relation to medical education.

JAMES D. PICKERING, B.Sc., Ph.D., P.G.C.L.T.H.E., S.F.H.E.A., is an associate professor of anatomy in the Division of Anatomy, School of Medicine, University of Leeds, United Kingdom and leads the anatomy curriculum for the M.B.Ch.B. He teaches anatomy of the trunk region and neuroanatomy to medical students, and has a strong interest in technology-enhanced learning and how it can be used to support learning gain.

LITERATURE CITED

Abdel Meguid EM, Khalil MK. 2016. Measuring medical students' motivation to learning anatomy by cadaveric dissection. Anat Sci Educ 10:363–371.

Adamczyk C, Holzer M, Putz R, Fischer MR. 2009. Student learning preferences and the impact of a multimedia learning tool in the dissection course at the University of Munich. Ann Anat 191:339–348.

Ahmad N, Wright A. 2014. Three-dimensional temporal bone reconstruction from histological sections. J Laryngol Otol 128:416–420.

Alexander CJ, Crescini WM, Juskewitch JE, Lachman N, Pawlina W. 2009. Assessing the integration of audience response system technology in teaching of anatomical sciences. Anat Sci Educ 2:160–166.

Allen LK, Eagleson R, de Ribaupierre S. 2016. Evaluation of an online three-dimensional interactive resource for undergraduate neuroanatomy education. Anat Sci Educ 9:431–439.

André MH. 2016. Physical education students' perceived competence in using technology integration. Int J Kinesiol High Educ 7:23–32.

Ang ET, Yip G, Lim ECH, Sugand K. 2014. Learning undergraduate human anatomy –

Reflections on undergraduate preferences in Singapore: A pilot study. J NUS Teach Acad 4:36–52.

Attardi SM, Choi S, Barnett J, Rogers KA. 2016. Mixed methods student evaluation of an online systemic human anatomy course with laboratory. Anat Sci Educ 9:272–285.

Attardi SM, Rogers KA. 2015. Design and implementation of an online systemic human anatomy course with laboratory. Anat Sci Educ 8:53–62.

Bacro TR, Gebregziabher M, Ariail J. 2013. Lecture recording system in anatomy: possible benefit to auditory learners. Anat Sci Educ 6:376–384.

Barbeau ML, Johnson M, Gibson C, Rogers KA. 2013. The development and assessment of an online microscopic anatomy laboratory course. Anat Sci Educ 6:246–256.

Battulga B, Konishi T, Tamura Y, Moriguchi H. 2012. The effectiveness of an interactive 3-dimensional computer graphics model for medical education. Interact J Med Res 1:e2.

Beale EG, Tarwater PM, Lee VH. 2014. A retrospective look at replacing face-to-face embryology instruction with online lectures in a human anatomy course. Anat Sci Educ 7:234–241.

Bellardini H, Cashin G, Stroble L, Way J. 1980. Project SuperHeart—Year one. Health Educ

11:11–12.

Bergman EM, Verheijen IW, Scherpbier AJ, Van der Vleuten CP, De Bruin AB. 2014. Influences on anatomical knowledge: The complete arguments. Clin Anat 27:296–303.

Bogacki RE, Best A, Abbey LM. 2004. Equivalence study of a dental anatomy computer-assisted learning program. J Dent Educ 68:867–871.

Boileau E, Patenaude J, St-Onge C. 2017. Twelve tips to avoid ethical pitfalls when recruiting students as subjects in medical education research. Med Teach (in press; doi: 10.1080/0142159X.2017.1357805).

Bonate PL. 2000. Analysis of Pretest-Posttest Designs. 1st Ed. London, UK: Chapman & Hall/CRC. 224 p.

Border S. 2017. Working with students as partners in anatomy education. Anat Sci Educ 10:613–614.

Braun MW, Kearns KD. 2008. Improved learning efficiency and increased student collaboration through use of virtual microscopy in the teaching of human pathology. Anat Sci Educ 1:240–246.

Brewer DN, Wilson TD, Eagleson R, De Ribaupierre S. 2012. Evaluation of neuroanatomical

training using a 3d visual reality model. Stud Health Technol Inform 173:85–91.

Brown PM, Hamilton NM, Denison AR. 2012. A novel 3D stereoscopic anatomy tutorial. Clin Teach 9:50–53.

Brown GA, Bice MR, Shaw BS, Shaw I. 2015. Online quizzes promote inconsistent improvements on in-class test performance in introductory anatomy and physiology. Adv Physiol Educ 39:63–66.

Bryner BS, Saddawi-Konefka D, Gest TR. 2008. The impact of interactive, computerized educational modules on preclinical medical education. Anat Sci Educ 1:247–251.

Burgoon JM, Meece JL, Granger NA. 2012. Self-efficacy's influence on student academic achievement in the medical anatomy curriculum. Anat Sci Educ 5:249–255.

Carmichael SW, Pawlina W. 2000. Animated PowerPoint as a tool to teach anatomy. Anat Rec 261:83–88.

Chan D. 2009. So why ask me? Are self-report data really that bad? In: Lance C, Vandenberg R (Editors). Statistical and Methodology Myths and Urban Legends: Doctrine, Verity and Fable in Organizational and Social Sciences. 1st Ed. New York, NY: Routledge. p 309–336.

Chan I, D'Eon M, Haggag H, Burbridge B. 2015. The effectiveness of learning anatomy and medical imaging using the Anatomage table compared with prosections. In: Abstracts of Joint Congress on Medical Imaging and Radiation Sciences; Montreal, QC, Canada; 2015 May 28-30; Abstract RT015. The Canadian Association of Radiologists: Ottawa, ON, Canada.

Chen B, Denoyelles A. 2013. Exploring students' mobile learning practices in higher education. Educause Rev Online. URL: http://er.educause.edu/articles/2013/10/exploring-students-mobile-learning-practices-in-higher-education [accessed 13 August 2017].

Chen F, Lui AM, Martinelli SM. 2017. A systematic review of the effectiveness of flipped classrooms in medical education. Med Edu 51:585–597.

Choi-Lundberg DL, Low TF, Patman P, Turner P, Sinha SN. 2015. Medical student preferences for self-directed study resources in gross anatomy. Anat Sci Educ 9:150–160

Choi-Lundberg DL, Cuellar WA, Williams AM. 2016. Online dissection audio-visual resources for human anatomy: Undergraduate medical students' usage and learning outcomes. Anat Sci Educ 9:545–554.

Chopra J, Rani A, Rani A, Verma RK. 2012. Traditional versus computer assisted teaching of human osteology: A randomized control trial study. Ind J App Basic Med Res 5:370–374.

Choudhury B, Gouldsborough I. 2012. The use of electronic media to develop transferable skills

in science students studying anatomy. Anat Sci Educ 5:125–131.

Choudhury B, Gouldsborough I, Gabriel, S. 2010. Use of interactive sessions and e-learning in teaching anatomy to first-year optometry students. Anat Sci Educ 3:39–45.

Codd AM, Choudhury B. 2011. Virtual reality anatomy: Is it comparable with traditional methods in the teaching of human forearm musculoskeletal anatomy? Anat Sci Educ 4:119–125.

Colliver JA, Cianciolo AT. 2014. When is enough enough? Judging the sufficiency of evidence in medical education. Med Educ 48:740–741.

Conley DM, Kastella KG, Sundsten JW, Rauschning W, Rosse C. 1992. Computer-generated three-dimensional reconstruction of the mediastinum correlated with sectional and radiological anatomy. Clin Anat 5:185–202.

Cook DA Ellaway RH. 2015. Evaluating technology-enhanced learning: A comprehensive framework. Med Teach 37:961–970.

Corton MM, McIntire DD, Wai CY, Ling FW, Wendel GD Jr. 2006. A comparison of an interactive computer-based method with a conventional reading approach for learning pelvic anatomy. Am J Obstet Gynecol 195:1438–1443.

Cottam WW. 1999. Adequacy of medical school gross anatomy education as perceived by

certain postgraduate residency programs and anatomy course directors. Clin Anat 12:55–65.

Cui D, Lynch JC, Smith AD, Wilson TD, Lehman MN. 2015. Stereoscopic vascular models of the head and neck: A computed tomography angiography visualization. Anat Sci Educ 6:179–185.

Das S, Mitchell P. 2013. Comparison of three aids for teaching lumbar surgical anatomy. Br J Neurosurg 27:475–478.

Davis CR, Bates AS, Ellis H, Roberts AM. 2014. Human Anatomy: Let the students tell us how to teach. Anat Sci Educ 7:262–272.

de Faria JW, Teixeira MJ, de Moura Sousa Júnior L, Otoch JP, Figueiredo EG. 2016. Virtual and stereoscopic anatomy: When virtual reality meets medical education. J Neurosurg 125:1105–1111.

Devitt P, Palmer E. 1999. Computer-aided learning: An overvalued educational resource? Med Educ 33:136–139.

Dimitrov DM, Rumrill PD. 2003. Pretest-posttest designs and measurement of change. Work. 20:159–165.

Dixon NM. 1990. The relationship between trainee responses on participant reaction forms and

posttest scores. Hum Resour Dev Q 1:129–137.

Dixson MD. 2015. Measuring student engagement in the online course: The online student engagement scale (OSE). Online Learn 19:4.

Donnelly L, Patten D, White P, Finn G. 2009. Virtual human dissector as a learning tool for studying cross-sectional anatomy. Med Teach 31:553–555.

Doubleday AF, Wille SJ. 2014. We are what we do: Examining learner-generated content in the anatomy laboratory through the lens of activity theory. Anat Sci Educ 7:361–369.

Drake RL, McBride JM, Lachman N, Pawlina W. 2009. Medical education in the anatomical sciences: The winds of change continue to blow. Anat Sci Educ 2:253–259.

Drake RL, McBride JM, Pawlina W. 2014. An update on the status of anatomical sciences education in United States medical schools. Anat Sci Educ 7:321–325.

Durham JA, Brettell S, Summerside C, McHanwell S. 2009. Evaluation of a virtual anatomy course for clinical undergraduates. Eur J Dent Educ, 13:100–109.

Evans DJ. 2011. Using embryology screencasts: A useful addition to the student learning experience? Anat Sci Educ 4:57–63.

Ferrer-Torregrosa J, Torralba J, Jimenez MA, García S, Barcia JM. 2015. ARBOOK: Development and assessment of a tool based on augmented reality for anatomy. J Sci Educ Technol 24:119–124.

Elizondo-Omaña RE, Morales-Gómez JA, Guzmán SL, Hernández IL, Ibarra RP, Vilchez FC. 2004. Traditional teaching supported by computer-assisted learning for macroscopic anatomy. Anat Rec 278B:18–22.

Ernst RD, Sarai P, Nishino T, Collins T, Oto A, Hernandez A, Walser EM, Chaljub G. 2003. Transition from film to electronic media in the first-year medical school gross anatomy lab. J Digit Imag 16:337–340.

Farah CS, Maybury TS. 2009a. Implementing digital technology to enhance student learning of pathology. Eur J Dent Educ 13:172–178.

Farah CS, Maybury TS. 2009b. The e-evolution of microscopy in dental education. J Dent Educ 73:942–949.

Foreman, KB, Morton DA, Musolino GM, Albertine, KH. 2005. Design and utility of a web-based computer-assisted instructional tool for neuroanatomy self-study and review for physical and occupational therapy graduate students. Anat Rec 285B:26–31.

Freeman S, Eddy SL, McDonough M, Smith MK, Okoroafor N, Jordt H, Wenderoth MP. 2014.

Active learning increases student performance in science, engineering, and mathematics. Proc Natl Acad Sci U S A 111:8410–8415.

Frye AW, Hemmer PA. 2012. Program evaluation models and related theories: AMEE guide No. 67. Med Teach 34:288–299.

Fuller R, Joynes VK. 2015. Should mobile learning be compulsory for preparing students for learning in the workplace? Brit J Educ Technol. 46:153–158.

Gopal T, Herron SS, Mohn RS, Hartsell T, Jawor JM, Blickenstaff JC. 2010. Effect of an interactive web-based instruction in the performance of undergraduate anatomy and physiology lab students. Comput Educ 55:500–512.

Granger N, Calleson DC, Henson OW, Juliano E, Wineski L, McDaniel MD, Burgoon JM. 2006. Use of web-based materials to enhance anatomy instruction in the health sciences. Anat Rec 289B:121–127.

Granger N, Calleson DC. 2007. The impact of alternating dissection on student performance in a medical anatomy course: Are dissection videos an effective substitute for actual dissection? Clin Anat 20:315–321.

Green SM, Weaver M, Voegeli D, Fitzsimmons D, Knowles J, Harrison M, Shephard K. 2006. The development and evaluation of the use of a virtual learning environment (Blackboard 5) to

support the learning of pre-qualifying nursing students undertaking a human anatomy and physiology module. Nurse Educ Today 26:388–395.

Green RA, Hughes DL. 2013. Student outcomes associated with use of asynchronous online discussion forums in gross anatomy teaching. Anat Sci Educ 6:101–106.

Green RA, Farchione D, Hughes DL, Chan SP. 2014. Participation in asynchronous online discussion forums does improve student learning of gross anatomy. Anat Sci Educ 7:71–76.

Green RA, Whitburn LY. 2016. Impact of introduction of blended learning in gross anatomy on student outcomes. Anat Sci Educ 9:422–430.

Guerri-Guttenberg RA. 2008. Web-based method for motivating 18-year-old anatomy students. Med Educ 42:1119.

Guy R, Pisani HR, Rich P, Leahy C, Mandarano G, Molyneux T. 2015. Less is more: Development and evaluation of an interactive e-atlas to support anatomy learning. Anat Sci Educ 8:126–132.

Hake RR. 1998. Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. Am J Phys 66:64–74.

Hallgren RC, Parkhurst PE, Monson CL, Crewe NM. 2002. An interactive, web-based tool for

learning anatomic landmarks. Acad Med 77:263–265.

Harris T, Leaven T, Heidger P, Kreiter C, Duncan J, Dick F. 2001. Comparison of a virtual microscope laboratory to a regular microscope laboratory for teaching histology. Anat Rec 265:10–14.

Healey M, Flint A. Harrington K. 2014. Engagement through Partnership: Students as Partners in Learning and Teaching in Higher Education. 1st Ed. Heslington, York, UK: Higher Education Academy. 77 p. URL: https://www.heacademy.ac.uk/system/files/resources/engagement_through_partnership.pdf [accessed 4 August 2017].

Heidger PM Jr, Dee F, Consoer D, Leaven T, Duncan J, Kreiter C. 2002. Integrated approach to teaching and testing in histology with real and virtual imaging. Anat Rec 269:107–112.

Helle L, Nivala M, Kronqvist P. 2013. More technology, better learning resources, better learning? Lessons from adopting virtual microscopy in undergraduate medical education. Anat Sci Educ 6:73–80.

Helle L, Nivala M, Kronqvist P, Gegenfurtner A, Björk P, Säljö R. 2011. Traditional microscopy instruction versus process-oriented virtual microscopy instruction: A naturalistic experiment with control group. Diagn Pathol 6:S8.

Hennessy CM, Kirkpatrick E, Smith CF, Border S. 2016. Social media and anatomy education: Using twitter to enhance the student learning experience in anatomy. Anat Sci Educ 9:505–515.

Heylings DJ. 2002. Anatomy 1999-2000: The curriculum, who teaches it and how? Med Educ 36:702–710.

Higazi TB. 2011. Use of interactive live digital imaging to enhance histology learning in introductory level anatomy and physiology classes. Anat Sci Educ 4:78–83.

Hisley KC, Anderson LD, Smith SE, Kavic SM, Tracy JK. 2008. Coupled physical and digital cadaver dissection followed by a visual test protocol provides insights into the nature of anatomical knowledge and its evaluation. Anat Sci Educ 1:27–40.

Holton EF. 1996. The flawed four-level evaluation model. Hum Resour Dev Q 7:5–21.

Hopkins R, Regehr G, Wilson TD. 2011. Exploring the changing learning environment of the gross anatomy lab. Acad Med 86:883–888.

Hoyek N, Collet C, Di Rienzo F, De Almeida M, Guillot A. 2014. Effectiveness of three-dimensional digital animation in teaching human anatomy in an authentic classroom context. Anat Sci Educ 7:430–437.

Hoyt A, McNulty JA, Gruener G, Chandrasekhar A, Espiritu B, Ensminger D, Price Jr R,

Naheedy R, Price R, Naheedy R. 2010. An audience response system may influence student performance on anatomy examination questions. Anat Sci Educ 3:295–299.

Hu A, Wilson T, Ladak H, Haase P, Doyle P, Fung K, Doyle P, Fung K, Doyle P, Fung K, Doyle P, Fung K. 2010. Three-dimensional educational computer model of the larynx: Voicing a new direction. J Otolaryngol Head Neck Surg 39:315–322.

Husmann PR, O'Loughlin VD, Braun MW. 2009. Quantitative and qualitative changes in teaching histology by means of virtual microscopy in an introductory course in human anatomy. Anat Sci Educ 2:218–226.

Iqbal H. 2016. Anatomy "peer teaching" in medical school: A literature review. MedEdPublish 5:6. URL: https://doi.org/10.15694/mep.2016.000033 [accessed August 4 2017].

Inwood MJ, Ahmad J. 2005. Development of instructional, interactive, multimedia anatomy dissection software: A student-led initiative. Clin Anat 18:613–617.

Issa N, Schuller M, Santacaterina S, Shapiro M, Wang E, Mayer RE, DaRosa DA. 2011. Applying multimedia design principles enhances learning in medical education. Med Educ 45:818–826.

Jaffar AA. 2012. YouTube: An emerging tool in anatomy education. Anat Sci Educ 5:158–164.

Jaffar AA. 2014. Exploring the use of a Facebook page in anatomy education. Anat Sci Educ 7:199–208.

Jamieson-Noel D, Winne PH. 2002. Exploring students' calibration of self-reports about study tactics and achievement. Contemp Educ Psychol 27:551–572.

Javadian P, Shobeiri S. 2016. Comparison of dissection-based vs. Internet-based pelvic anatomy education for 3rd year medical students. J Educ Tech Health Sci 3:38–41.

Johnson IP, Palmer E, Burton J, Brockhouse M. 2013. Online learning resources in anatomy: What do students think? Clin Anat 26:556–563.

Keedy AW, Durack JC, Sandhu P, Chen EM, O'Sullivan PS, Breiman RS. 2011. Comparison of traditional methods with 3D computer models in the instruction of hepatobiliary anatomy. Anat Sci Educ 4:84–91.

Kelc R. 2012. Zygote body: A new interactive 3-dimensional didactical tool for teaching anatomy. WebmedCentral Anat 3:WMC002903.

Khalil MK, Mansour MM, Wilhite DR. 2010. Evaluation of cognitive loads imposed by traditional paper-based and innovative computer-based instructional strategies. J Vet Med Educ 37:353–357.

Khot Z, Quinlan K, Norman GR, Wainman B. 2013. The relative effectiveness of computer-based and traditional resources for education in anatomy. Anat Sci Educ 6:211–215.

Kirkpatrick DL. 1994. Evaluating Training Programs: The Four Levels. 1st Ed. San Francisco, CA: Berrett-Koehlar. 229 p.

Kirkpatrick DL. 2017. The New World Kirkpatrick Model. Kirkpatrick Partners, LLC., Newnan, GA. URL: http://www.kirkpatrickpartners.com/OurPhilosophy/TheNewWorldKirkpatrickModel/tabid/303/ [accessed 14 May 2017].

Kirkwood A, Price L. 2014. Technology-enhanced learning and teaching in higher education: what is 'enhanced' and how do we know? A critical literature review. Learn Media Technol 39:6–36.

Kirschner P. 2016. Guest post: An interview with an educational realist and grumpy old man. The Learning Scientists, Providence, RI. URL: http://www.learningscientists.org/blog/2016/8/16-1 [accessed 5 July 2017].

Kirschner PA, De Bruyckere P. 2017. The myths of the digital native and the multitasker. Teach Teach Educ 67:135–142.

Kirschner P, van Merriënboer JJ. 2013. Do learners really know best? Urban legends in education. Educ Psychol 48:169–183.

Koehler N. 2012. Medical students' use of and attitudes towards medical applications. J Mob Technol Med 1:16–21.

Krause KL, Coates H. 2008. Students' engagement in first-year university. Assess Eval High Educ 33:493–505.

Krippendorf BB, Lough J. 2005. Complete and rapid switch from light microscopy to virtual microscopy for teaching medical histology. Anat Rec 285B:19–25.

Kruger J, Dunning D. 1999. Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. J Pers Soc Psychol 77:1121–1134.

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.

Kumar RK, Velan GM, Korell SO, Kandara M, Dee FR, Wakefield D. 2004. Virtual microscopy for learning and assessment in pathology. J Pathol 204:613–618.

Laurillard D. 2007. Modelling benefits-oriented costs for technology enhanced learning. High Educ 54:21–39.

Lee LM, Nagel RW, Gould DJ. 2012. The educational value of online mastery quizzes in a

human anatomy course for first-year dental students. J Dent Educ 76:1195–1199.

Levine MG, Stempak J, Conyers G, Walters JA. 1999. Implementing and integrating computer-based activities into a problem-based gross anatomy curriculum. Clin Anat 12:191–198.

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

Lewis TL, Burnett B, Tunstall RG, Abrahams PH. 2014. Complementing anatomy education using three-dimensional anatomy mobile software applications on tablet computers. Clin Anat 27:313–320.

Lim KH, Loo ZY, Goldie SJ, Adams JW, McMenamin PG. 2015. Use of 3D printed models in medical education: A randomized control trial comparing 3D prints versus cadaveric materials for learning external cardiac anatomy. Anat Sci Educ 9:213–221.

Limpach AL, Bazrafshan P, Turner PD, Monaghan MS. 2008. Effectiveness of human anatomy education for pharmacy students via the internet. Am J Pharm Educ 72:145.

Lochner L, Wieser H, Waldboth S, Mischo-Kelling M. 2016. Combining traditional anatomy lectures with e-learning activities: How do students perceive their learning experience? Int J Med Educ 7:69–74.

Lombardi SA, Hicks RE, Thompson KV, Marbach-Ad G. 2014. Are all hands-on activities equally effective? Effect of using plastic models, organ dissections, and virtual dissections on student learning and perceptions. Adv Physiol Educ 38:80–86.

Louw G, Eizenberg N, Carmichael SW. 2009. The place of anatomy in medical education: AMEE Guide No 41. Med Teach 31:373–386.

Lumsden CJ, Byrne-Davis LM, Mooney JS, Sandars J. 2015. Using mobile devices for teaching and learning in clinical medicine. Arch Dis Child Educ Pract 100:244–251.

Ma M, Fallavollita P, Seelbach I, Von Der Heide AM, Euler E, Waschke J, Navab N. 2016. Personalized augmented reality for anatomy education. Clin Anat 29:446–453.

Maggio MP, Hariton-Gross K, Gluch J. 2012. The use of independent, interactive media for education in dental morphology. J Dent Educ 76:1497–1511.

Mahmud W, Hyder O, Butt J, Aftab A. 2011. Dissection videos do not improve anatomy examination scores. Anat Sci Educ 4:16–21.

Mars M, McLean M. 1996. Students' perceptions of a multimedia computer-aided instruction resource in histology. S Afr Med J 86:1098–1102.

Mathiowetz V, Yu CH, Quake-Rapp C. 2016. Comparison of a gross anatomy laboratory to

online anatomy software for teaching anatomy. Anat Sci Educ 9:52–59.

Maybury TS, Farah CS. 2010. Electronic blending in virtual microscopy. J Learn Design 4:41–51.

Mayfield CH, Ohara PT, O'Sullivan PS. 2013. Perceptions of a mobile technology on learning strategies in the anatomy laboratory. Anat Sci Educ 6:81–89.

McGrath CH, Guerin B, Harte E, Frearson M, Manville C. 2015. Learning Gain in Higher Education. 1st Ed. Santa Monica, CA: Rand Corp. 125 p.

McKeown PP, Heylings DJA, Stevenson M, McKelvey KJ, Nixon JR, McCluskey DR. 2003. Basic science the impact of curricular change on medical students' knowledge of anatomy. Med Educ 37:954–961.

McLachlan JC, Bligh J, Bradley P, Searle J. 2004. Teaching anatomy without cadavers. Med Educ 38:418–424.

McLachlan JC, Patten D. 2006. Anatomy teaching: Ghosts of the past, present and future. Med Educ 40:243–253.

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

McNulty JA, Halama J, Dauzvardis MF, Espiritu B. 2000. Evaluation of Web-based computer-aided instruction in a basic science course. Acad Med 75:59–65.

McNulty JA, Halama J, Espiritu B. 2004. Evaluation of computer-aided instruction in the medical gross anatomy curriculum. Clin Anat 17:73–78.

McNulty JA, Sonntag B, Sinacore JM. 2009. Evaluation of computer-aided instruction in a gross anatomy course: a six-year study. Anat Sci Educ 2:2–8.

Meyer AJ, Stomski NJ, Innes SI, Armson AJ. 2015. VARK learning preferences and mobile anatomy software application use in pre-clinical chiropractic students. Anat Sci Educ 9:247–254.

Miller M. 2016. Use of computer-aided holographic models improves performance in a cadaver dissection-based course in gross anatomy. Clin Anat 29:917–924.

Mione S, Valcke M, Cornelissen M. 2013. Evaluation of virtual microscopy in medical histology teaching. Anat Sci Educ 6:307–315.

Mione S, Valcke M, Cornelissen M. 2015. Remote histology learning from static versus dynamic microscopic images. Anat Sci Educ 9:222–230.

Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. 2009. Preferred reporting items

for systematic reviews and meta-analyses: The PRISMA statement. PLoS Med 6:e1000097.

Moorman SJ. 2006. Prof-in-a-box: Using internet-videoconferencing to assist students in the gross anatomy laboratory. BMC Med Educ 6:55.

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.

Morris NP, Lambe J, Ciccone J, Swinnerton B. 2016. Mobile technology: Students perceived benefits of apps for learning neuroanatomy. J Comp Assist Learn 32:430–442.

Nicholson DT, Chalk C, Funnell WRJ, Daniel SJ. 2006. Can virtual reality improve anatomy education? A randomised controlled study of a computer-generated three-dimensional anatomical ear model. Med Educ 40:1081–1087.

Nieder GL, Borges NJ. 2012. An eight-year study of online lecture use in a medical gross anatomy and embryology course. Anat Sci Educ 5:311–320.

Nieder GL, Nagy F. 2002. Analysis of medical students' use of web-based resources for a gross anatomy and embryology course. Clin Anat 15:409–418.

Nieder GL, Scott JN, Anderson MD. 2000. Using QuickTime virtual reality objects in computer-assisted instruction of gross anatomy: Yorick--The VR skull. Clin Anat 13:287–293.

O'Byrne PJ, Patry A, Carnegie JA 2008. The development of interactive online learning tools for the study of anatomy. Med Teach 30:260–271.

O'Reilly MK, Reese S, Herlihy T, Geoghegan T, Cantwell CP, Feeney RN, Jones JF. 2016. Fabrication and assessment of 3D printed anatomical models of the lower limb for anatomical teaching and femoral vessel access training in medicine. Anat Sci Educ 9:71–79.

Ocak MA, Topal AD. 2015. Blended learning in anatomy education: A study investigating medical students' perceptions. Eurasia J Math Sci Tech Educ 11:647–663.

Ogunranti JO. 1987. Video technology in integrated anatomy education. J Educ Televis 13:63–67.

Packer JE. 1994. A computer-based pictorial anatomy museum catalogue. Med Teach 16:97–100.

Patasi B, Boozary A, Hincke M, Jalali A. 2009. The utility of podcasts in Web 2.0 human anatomy. Med Educ 43:1116.

Patel SG, Rosenbaum BP, Chark DW, Lambert HW. 2006. Design and implementation of a web-based, database-driven histology atlas: Technology at work. Anat Rec 289B:176–183.

Pereira JA, Merí A, Masdeu C, Molina-Tomás MC, Martinez-Carrio A. 2004. Using videoclips to improve theoretical anatomy teaching. Eur J Anat 8:143–146.

Pereira JA, Pleguezuelos E, Merí A, Molina-Ros A, Molina-Tomás MC, Masdeu C. 2007. Effectiveness of using blended learning strategies for teaching and learning human anatomy. Med Educ 41:189–195.

Peterson DC, Mlynarczyk GS. 2016. Analysis of traditional versus three-dimensional augmented curriculum on anatomical learning outcome measures. Anat Sci Educ 9:529–536.

Petersson H, Sinkvist D, Wang C, Smedby Ö. 2009. Web-based interactive 3D visualization as a tool for improved anatomy learning. Anat Sci Educ 2:61–68.

Pickering JD. 2015a. Anatomy drawing screencasts: Enabling flexible learning for medical students. Anat Sci Educ 8:249–257.

Pickering JD. 2015b. Introduction of an anatomy eBook enhances assessment outcomes. Med Educ 49:522–523.

Pickering JD. 2016a. Measuring learning gain: Comparing anatomy drawing screencasts and paper-based resources. Anat Sci Educ 10:307–316.

Pickering JD, Joynes VK. 2016b. A holistic model for evaluating the impact of individual

technology-enhanced learning resources. Med Teach 38:1242–1247.

Pickering JD, Bickerdike SR. 2016c. Medical student use of Facebook to support preparation for anatomy assessments. Anat Sci Educ 10:205–214.

Pickering JD. 2017a. Developing the evidence-base to support the integration of technology-enhanced learning in healthcare education. Med Sci Educ (in press; doi: 10.1007/s40670-017-0424-2).

Pickering JD. 2017b. Cognitive engagement: A more reliable proxy for learning? Med Sci Educ (in press; doi:10.1007/s40670-017-0447-8).

Pizzimenti MA, Axelson RD. 2015. Assessing student engagement and self-regulated learning in a medical gross anatomy course. Anat Sci Educ 8:104–110.

Raikos A, Waidyasekara P. 2014. How useful is YouTube in learning heart anatomy? Anat Sci Educ 7:12–18.

Raney MA. 2015. Dose- and time-dependent benefits of iPad technology in an undergraduate human anatomy course. Anat Sci Educ 9:367–377.

Raynor M, Iggulden H. 2008. Online anatomy and physiology: piloting the use of an anatomy and physiology e-book-VLE hybrid in pre-registration and post-qualifying nursing programmes

at the University of Salford. Health Info Lib J 25:98–105.

Reeves RE, Aschenbrenner JE, Wordinger RJ, Roque RS, Sheedlo HJ. 2004. Improved dissection efficiency in the human gross anatomy laboratory by the integration of computers and modern technology. Clin Anat 17:337–344.

Reinders ME, de Jong PG. 2016. Innovating clinical kidney transplant education by a massive open online course. Transpl Immunol 38:1–2.

Ribeiro M, Amaral M, Ribeiro H, Machado J, Povo A, Severo M, Ferreira MA. 2009. "Virtual Quiz": A tool for active learning and assessment in clinical anatomy. REIT - International Conference Record Book, University of Porto, Porto, Portugal. URL: http://hdl.handle.net/10216/13426 [accessed 24 October 2016].

Rich P, Guy R. 2013. A "do-it-yourself" interactive bone structure module: development and evaluation of an online teaching resource. Anat Sci Educ 6:107–113.

Richardson-Hatcher A, Hazzard M, Ramirez-Yanez G. 2014. The cranial nerve skywalk: A 3D tutorial of cranial nerves in a virtual platform. Anat Sci Educ 7:469–478.

Rinaldi VD, Lorr NA, Williams K. 2016. Evaluating a technology supported interactive response system during the laboratory section of a histology course. Anat Sci Educ 10:328–338.

Rizzolo LJ, Aden M, Stewart WB. 2002. Correlation of Web usage and exam performance in a human anatomy and development course. Clin Anat 15:351–355.

Rizzolo LJ, Rando WC, O'Brien MK, Haims AH, Abrahams JJ, Stewart WB. 2010. Design, implementation, and evaluation of an innovative anatomy course. Anat Sci Educ 3:109–120.

Roberts LD, Howell JA, Seaman K, Gibson, DC. 2016. Student attitudes toward learning analytics in higher education: "The fitbit version of the learning world". Front Psychol 7:1959.

Rondon S, Sassi FC, Furquim de Andrade CR. 2013. Computer game-based and traditional learning method: A comparison regarding students' knowledge retention. BMC Med Educ 13:30.

Rosas C, Rubí R, Donoso M, Uribe S. 2012. Dental students' evaluations of an interactive histology software. J Dent Educ 76:1491–1496.

Said CS, Shamsudin K, Mailok, Johan R, Hanaif HF. 2015. The development and evaluation of a 3D visualization tool in anatomy education. J Sci Math Tech 2:48–56.

Saqr M, Fors U, Tedre M. 2017. How learning analytics can early predict under-achieving students in a blended medical education course. Med Teach 39:757–767.

Saltarelli AJ, Roseth CJ, Saltarelli WA. 2014. Human cadavers vs. multimedia simulation: A study of student learning in anatomy. Anat Sci Educ 7:331–339.

Sander B, Golas MM. 2013. HistoViewer: An interactive e-learning platform facilitating group and peer group learning. Anat Sci Educ 6:182–190.

Saxena V, Natarajan P, O'Sullivan PS, Jain S. 2008. Effect of the use of instructional anatomy videos on student performance. Anat Sci Educ 1:159–165.

Scoville SA, Buskirk TD. 2007. Traditional and virtual microscopy compared experimentally in a classroom setting. Clin Anat 20:565–570.

Selwyn N. 2016. Digital downsides: Exploring university students' negative engagements with digital technology. Teach High Educ 21:1006–10021.

Sharples M, de Roock R, Ferguson R, Gaved M, Herodotou C, Koh E, Kukulska-Hulme A, Looi CK, McAndrew P, Rienties B, Weller M, Wong LH. 2016. Innovating Pedagogy 2016: Open University Innovation Report 5. 1st Ed. Milton Keynes, UK: The Open University. 45 p.

Shoepe TC, Cavedon DK, Derian JM, Levy CS, Morales A. 2015. The ATLAS project: The effects of a constructionist digital laboratory project on undergraduate laboratory performance. Anat Sci Educ 8:12–20.

Silén C, Wirell S, Kvist J, Nylander E, Smedby O.2008. Advanced 3D visualization in student-

centred medical education. Med Teach 30:115–124.

Sivamalai S, Murthy SV, Gupta TS, Woolley T. 2011. Teaching pathology via online digital microscopy: Positive learning outcomes for rurally based medical students. Aust J Rural Health 19:45–51.

Stalmeijer RE, McNaughton N, Van Mook WN. 2014. Using focus groups in medical education research: AMEE guide No. 91. Med Teach 36:923–939.

Stewart S, Choudhury B. 2015. Mobile technology: Creation and use of an iBook to teach the anatomy of the brachial plexus. Anat Sci Educ 8:429–437.

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

Sullivan GM. 2011. Getting off the "gold standard": Randomized controlled trials and education research. J Grad Med Educ 3:285–289.

Swinnerton BJ, Morris NP, Hotchkiss S, Pickering JD. 2017. The integration of an anatomy massive open online course (MOOC) into a medical anatomy curriculum. Anat Sci Educ 10:53–67.

Tan S, Hu A, Wilson T, Ladak H, Haase P, Fung K. 2012. Role of a computer-generated three-

dimensional laryngeal model in anatomy teaching for advanced learners. J Laryngol Otol 126:395–401.

Tavakol M, Sandars J. 2014. Quantitative and qualitative methods in medical education research: AMEE guide No 90: Part II. Med Teach 36:838–848.

Tian Y, Xiao W, Li C, Liu Y, Qin M, Wu Y, Xiao L, Li H. 2014. Virtual microscopy system at Chinese medical university: An assisted teaching platform for promoting active learning and problem-solving skills. BMC Med Educ 14:74.

Thompson AR, Lowrie DJ Jr. 2017. An evaluation of outcomes following the replacement of traditional histology laboratories with self-study modules. Anat Sci Educ 10:276–285.

Topping DB. 2014. Gross anatomy videos: Student satisfaction, usage, and effect on student performance in a condensed curriculum. Anat Sci Educ 7:273–279.

Traser CJ, Hoffman LA, Seifert MF, Wilson AB. 2015. Investigating the use of quick response codes in the gross anatomy laboratory. Anat Sci Edu 8:421–428.

Trelease RB. 2008. Diffusion of innovations: Smartphones and wireless anatomy learning resources. Anat Sci Educ 1:233–239.

Trelease RB. 2016. From chalkboard, slides, and paper to e-learning: How computing

technologies have transformed anatomical sciences education. Anat Sci Educ 9:583–602.

Triola MM, Holloway WJ. 2011. Enhanced virtual microscopy for collaborative education. BMC Med Educ 11:4.

Turney B. 2007. Anatomy in a modern medical curriculum. Ann R Coll Surg Engl 89:104–107.

Tworek JK, Jamniczky H A, Jacob C, Hallgrímsson B, Wright B. 2013. The LINDSAY Virtual human project: An immersive approach to anatomy and physiology. Anat Sci Educ 6:19–28.

Van Nuland SE, Eagleson R, Rogers KA. 2017. Educational software usability: Artifact or design? Anat Sci Educ 10:190–199.

Van Nuland SE, Rogers KA. 2016. E-learning, dual-task, and cognitive load: The anatomy of a failed experiment. Anat Sci Educ 9:186–196

Vaccani JP, Javidnia H, Humphrey-Murto S. 2016. The effectiveness of webcast compared to live lectures as a teaching tool in medical school. Med Teach 38:59–63.

Venail F, Deveze A, Lallemant B, Guevara N, Mondain M. 2010. Enhancement of temporal bone anatomy learning with computer 3D rendered imaging software. Med Teach 32:e282–e288.

Veneri DA, Gannotti M. 2014. A comparison of student outcomes in a physical therapy

neurologic rehabilitation course based on delivery mode: Hybrid vs traditional. J Allied Health 43:e75–e81.

Venkatiah J. 2010. Computer-assisted modules to enhance the learning of anatomy by dissection. Med Educ 44:523–524.

Wait KR, Cloud BA, Forster LA, Jones TM, Nokleby JJ, Wolfe CR, Youdas JW. 2009. Use of an audience response system during peer teaching among physical therapy students in human gross anatomy: perceptions of peer teachers and students. Anat Sci Educ 2:286–293.

Wallace S, Clark M, White J. 2012. 'It's on my iPhone': Attitudes to the use of mobile computing devices in medical education, a mixed-methods study. BMJ Open 2:e001099.

Walsh K, Levin H, Jaye P, Gazzard J. 2013. Cost analyses approaches in medical education: There are no simple solutions. Med Educ 47:962–968.

Wiers-Jenssen J, Stensaker B, Grøgaard JB. 2002. Student satisfaction: Towards an empirical deconstruction of the concept. Qual High Educ 8:183–195.

Wilkinson K, Barter P. 2016. Do mobile learning devices enhance learning in higher education anatomy classrooms? J Pedagog Dev 6:14–23.

Wilson AB, Taylor MA, Klein BA, Sugrue MK, Whipple EC, Brokaw JJ. 2016. Meta-analysis

and review of learner performance and preference: Virtual versus optical microscopy. Med Educ 50:428–440.

Wilson AB, Miller CH, Klein BA, Taylor MA, Goodwin M, Boyle EK, Brown K, Hoppe C, Lazarus M. 2017. A meta-analysis of anatomy laboratory pedagogies. Clin Anat (in press; doi: 10.1002/ca.22934).

Wright EF, Hendricson WD. 2010. Evaluation of a 3-D interactive tooth atlas by dental students in dental anatomy and endodontics courses. J Dent Educ 74:110–122.

Wright SJ. 2012. Student perceptions of an upper-level, undergraduate human anatomy laboratory course without cadavers. Anat Sci Educ 5:146–157.

Yammine K, Violato C. 2015. A meta-analysis of the educational effectiveness of three-dimensional visualization technologies in teaching anatomy. Anat Sci Educ 8:525–538.

Yammine K, Violato C. 2016. The effectiveness of physical models in teaching anatomy: A meta-analysis of comparative studies. Adv Health Sci Educ Theory Pract 21:883–895.

Yao WC, Regone RM, Huyhn N, Butler EB, Takashima M. 2014. Three-dimensional sinus imaging as an adjunct to two-dimensional imaging to accelerate education and improve spatial orientation. Laryngoscope 124:596–601.

Yardley S, Dornan T. 2012. Kirkpatrick's levels and education 'evidence'. Med Educ 46:97–106.

TABLES

Table 1. Key words used to identify studies for the systematic review.

Search term theme		
Type of education	**Delivery method**	**Resource type**
Medical education	Technology-enhanced learning	Animation
Undergraduate Medicine	e-learning	3D models
Gross anatomy	Computer-assisted learning (/instruction)	eBooks
Regional anatomy	Web-based learning	Virtual reality
Gross anatomy (/anatomical) education	Blended learning	Augmented reality
Anatomy teaching	Flipped classroom	Three-dimensional model
Anatomy (/anatomical) education	Flexible learning	Anatomy videos
Anatomical sciences	Multimedia learning	Anatomical reconstruction
	Mobile learning	Digital anatomy
	Virtual learning	Mobile devices[a]
	Educational technology	(Mobile) Applications
		3D printing
		Virtual dissection

[a]including specific terms such as laptop, tablet, smartphone, and eBook.

Table 2. Details of the 153 eligible studies extracted from the systematic review by resource type, with the assigned level of evaluation documented in accordance with the Technology-enhanced learning evaluation model (TELEM) as a benchmarking tool.

First author (Year)	Level				First author (Year)	Level			
	1a	1b	2	3		1a	1b	2	3
Instructor–Developed Resources									
Adamczyk et al. (2009)	+	-	+	-	Mahmud et al. (2011)	+	-	+	-
Ahmad and Wright (2016)	+	-	+	-	Mars and McLean (1996)	+	+	-	-
Allen et al. (2008)	+	-	+	-	Mathiowetz (2016)	+	-	+	-
Ang et al. (2014)	+	-	-	-	McNulty et al. (2000)	-	-	+	-
Bogacki et al. (2004)	+	+	-	-	McNulty et al. (2004)	-	-	+	-
Bryner et al. (2008)	+	+	-	-	McNulty et al. (2009)	+	-	+	-
Choi-Lundberg et al. (2015)	+	-	-	-	Moorman (2006)	+	-	-	-
Choi-Lundberg et al. (2016)	+	-	+	-	Nieder et al. (2000)	+	-	-	-
Chopra et al. (2012)	+	-	+	-	O'Byrne et al. (2008)	+	-	+	-
Choudhury et al. (2010)	+	-	+	-	Ocak and Topal (2015)	+	-	-	-
Corton et al. (2006)	+	+	-	-	Ogunranti (1987)	+	-	-	-
Devitt and Palmer (1999)	-	+	-	-	Patasi et al. (2009)	+	-	-	-
Doubleday and Wille (2014)	+	-	-	-	Pereira et al. (2004)	+	-	+	-
Durham et al. (2009)	+	-	-	-	Pereira et al. (2007)	+	-	+	-

Elizondo-Omaña et al. (2004)	-	-	+	-	Petersson et al. (2009)	+	-	+	-
Ernst et al. (2003)	+	-	-	-	Pickering (2014)	+	-	+	-
Evans (2011)	+	-	+	-	Pickering (2016a)	-	+	-	-
Foreman et al. (2005)	+	-	-	-	Reeves et al. (2004)	+	-	-	-
Gopal et al. (2010)	-	+	-	-	Rich and Guy (2013)	+	-	+	-
Granger and Calleson (2007)	+	-	+	-	Richardson-Hatcher et al. (2014)	+	-	-	-
Granger et al. (2006)	+	-	-	-	Rizzolo et al. (2002)	+	-	+	-
Green and Whitburn (2016)	+	-	+	-	Rizzolo et al. (2010)	+	-	+	-
Guy et al. (2015)	+	-	-	-	Rondon et al. (2013)	-	-	+	-
Hallgren et al. (2002)	-	+	-	-	Saltarelli et al. (2014)	-	-	+	-
Inwood and Ahmad (2005)	+	-	-	-	Saxena et al. (2008)	+	-	+	-
Johnson et al. (2013)	+	-	+	-	Shoepe et al. (2015)	+	-	+	-
Khalil et al. (2010)	+	-	+	-	Topping (2014)	+	-	+	-
Levine et al. (1999)	+	-	-	-	Veneri and Gannotti (2014)	+	-	+	-
Levinson et al. (2007)	-	+	-	-	Venkatiah (2010)	+	-	+	-
Maggio et al. (2012)	+	+	-	-	Wright and Hendricson (2010)	+	-	-	-
Mobile Devices									

Alexander et al. (2009)	+	-	+	-	Raynor and Iggulden (2008)	+	-	-	-
André (2016)	+	-	+	-	Stewart and Choudhury (2015)	+	+	-	-
Hoyt et al. (2010)	+	-	+	-	Stirling and Birt (2014)	+	+	-	-
Mayfield et al. (2012)	+	-	-	-	Traser et al. (2015)	+	-	+	-
Meyer et al. (2015)	+	-	-	-	Wait et al. (2009)	+	-	-	-
Morris et al. (2016)	+	-	+	-	Rinaldi et al. (2016)	+	-	+	-
Pickering (2015b)	+	-	+	-	Wilkinson and Barter (2016)	+	-	+	-
Raney (2015)	+	-	+	-					
Online Repository Resources									
Attardi and Rogers (2015)	-	-	+	-	Guerri-Guttenberg (2008)	+	-	+	-
Attardi et al. (2016)	+	-	-	-	Lee et al. (2012)	+	-	+	-
Bacro et al. (2013)	-	-	+	-	Limpach et al. (2008)	+	-	+	-
Barbeau et al. (2013)	+	-	+	-	Lochner et al. (2016)	+	-	-	-
Beale et al. (2014)	+	-	+	-	Javadian and Shobeiri (2016)	-	-	+	-
Brown et al. (2015)	-	-	+	-	Nieder and Nagy (2002)	+	-	-	-
Carmichael and Pawlina	+	-	-	-	Nieder and Borges	-	-	+	-

(2000)					(2012)				
Green and Hughes (2013)	+	-	+	-	Ribeiro et al. (2007)	+	-	+	-
Green et al. (2014)	-	-	+	-	Swinnerton et al. (2016)	+	-	-	-
Green et al. (2006)	+	-	+	-					
Purpose-built Resources									
Chan et al. (2015)	+	+	-	-	Miller (2016)	-	-	+	-
Lim et al. (2015)	-	+	-	-	O'Reilly et al. (2016)	+	+	-	-
Social Media									
Jaffar (2012)	+	-	-	-	Pickering (2016a)	+	-	+	-
Jaffar (2014)	+	-	-	-	Hennessy et al. (2016)	+	-	+	-
Virtual 3D Computer Model									
Allen et al. (2016)	+	-	+	-	Khot et al. (2013)	-	+	-	-
Battulga et al. (2012)	+	+	-	-	Lombardi et al. (2014)	+	+	-	-
Brewer et al. (2012)	+	-	+	-	Nicholson et al. (2006)	-	+	-	-
Brown et al. (2012)	+	+	-	-	Peterson and Mlynarczyk (2016)	+	-	+	-
Codd and Choudhury (2011)	+	-	-	-	Silén et al. (2008)	+	-	-	-

Das and Mitchell (2013)	+	+	-	-	Said et al. (2015)	+	-	-	-
Donnelly et al. (2009)	-	+	-	-	Tan et al. (2012)	+	+	-	-
Hisley et al. (2008)	-	+	-	-	Tworek et al. (2013)	+	-	-	-
Hopkins et al. (2011)	+	+	-	-	Van Nuland and Rogers (2015)	-	+	-	-
Hoyek et al. (2014)	+	-	+	-	Venail (2010)	+	-	+	-
Hu et al. (2010)	+	+	-	-	de Faria et al. (2016)	-	+	-	-
Keedy et al. (2011)	+	+	-	-	Yao et al. (2014)	+	-	-	-
Kelc (2012)	+	-	-	-					
Augmented Reality									
Ferrer-Torregrosa et al. (2015)	+	-	+	-	Ma et al. (2015)	+	-	-	-
Küçük et al. (2016)	+	+	-	-					
Virtual Microscopy									
Braun and Kearns (2008)	+	-	-	-	McCready et al. (2013)	+	-	-	-
Farah and Maybury (2009a)	+	-	+	-	Mione et al. (2013)	-	+	-	-
Farah and Maybury (2009b)	+	-	-	-	Mione et al. (2015)	-	-	+	-
Harris et al. (2001)	+	-	-	-	Patel et al. (2006)	+	-	-	-
Heidger et al. (2002)	+	-	-	-	Rosas et al. (2012)	+	-	-	-
Helle et al. (2011)	+	+	-	-	Sander and Golas	+	-	-	-

					(2013)				
Helle et al. (2013)	+	-	+	-	Scoville and Buskirk (2007)	+	-	+	-
Higazi (2011)	+	-	+	-	Sivamalai et al. (2011)	+	-	-	-
Husmann et al. (2009)	+	-	+	-	Tian et al. (2014)	+	+	-	-
Krippendorf and Lough (2005)	+	-	+	-	Thompson and Lowrie (2017)	+	-	+	-
Kumar et al. (2004)	+	-	-	-	Triola and Holloway (2011)	-	-	+	-
Maybury and Farah (2010)	+	-	-	-					

FIGURE LEGENDS

Figure 1. The Technology-Enhanced Learning (TEL) Evaluation model used as a benchmarking tool for studies evaluating the effect of TEL on anatomy learning. The model groups research methodologies into either Level 0 (TEL resource development), Level 1a (Learner satisfaction), Level 1b (Learning gain), Level 2 (Learner impact) and Level 3 (Institutional impact; modified from Pickering and Joynes, 2015).

Figure 2. A summary of the selection process presented in the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) flow diagram (Moher et al., 2009). TEL, technology-enhanced learning.

Figure 3. Quantitative data displayed in a bar chart detailing the proportion of papers as a percentage (%) that were assigned to one of the levels detailed in the Technology-enhanced learning evaluation model (TELEM). 1A, Level 1a (Learner satisfaction [white]); 1B, Level 1b (Learning gain [blue]); 1A and 1B, studies that combined Level 1a (Learner satisfaction) and Level 1b (Learning gain [green]); 2, Level 2 (Learner impact [grey]); 1A and 2, studies that combined Level 1a (Learner satisfaction) and Level 2 (Learner impact [red]).

Figure 4. Quantitative data displayed in a bar chart detailing the proportion of papers per Technology-enhanced learning evaluation model (TELEM) level within each time period. The black circle indicates the number of show and tell (S&T) articles published within each time period. 1A, Level 1a (Learner satisfaction [white]); 1B, Level 1b (Learning gain [blue]); 1A and

1B, studies that combined Level 1a (Learner satisfaction) and Level 1b (Learning gain [green]); 2, Level 2 (Learner impact [grey]); 1A and 2, studies that combined Level 1a (Learner satisfaction) and Level 2 (Learner impact [red]).