Research Article

Are Multimedia Resources Effective in Life Science Education? A Meta-Analysis

Vivien E Rolfe and Douglas Gray

School of Allied Health Sciences, De Montfort University, Leicester, UK

Abstract

Multimedia learning is widely used in life science education where the use of pictures and text can bring complex structures and processes to life. However the impact on academic performance and deeper understanding is not well documented. We therefore carried out a systematic review to evaluate the effectiveness of multimedia resources in tertiary level life science education. Comprehensive literature searches were conducted; studies were selected based on stringent pre-set criteria, and data were extracted for meta-analysis. In total, 17 studies were used in the meta-analyses with a total population of 2,290 students. The results show that, when used as a substitute for laboratory practicals, multimedia improved student learning gains assessed with an end-of-year examination, (mean difference 7.06, ±4.61). Although it did not improve short-term learning gains in this scenario, multimedia improved learning gains in 10 of the 16 sub-group comparisons made across all the studies.

Overall, multimedia learning was more effective than many traditional educational methods although the numbers of studies included in the analysis were ultimately considered to be small due to many exclusions from the studies included in the analysis. Therefore, more good quality trials are required to evaluate a broader range of scenarios relevant to modern practices. Studies would benefit from being rigorous in design with good quality reporting of all aspects of methodology and study results.

Keywords: Multimedia instruction, bioscience teaching, meta-analysis, higher education

Introduction

Multimedia instruction is a well established means of instructional delivery in the life sciences and is often used to complement or blend with traditional didactic elements (Pereira *et al.,* 2007), or replace other 'traditional' teaching methods altogether (Dewhurst *et al.,* 1994; Gibbons, 2004). The use of mechanical devices as educational tools emerged in the 1950s with Skinner's "teaching machine", a machine that allowing students to respond to questions (Skinner, 1960), and later the notion of computers as educational tools became established (Suppes, 1972). The affordability and availability of desk-top computers fuelled the growth in development of electronic educational resources, so-called computer assisted learning or instruction (CAL/ CAI). Early software required some degree of programming knowledge (Dewhurst *et al.,* 1994) whereas later resources could more easily be created using commercial authoring solutions (Gibbons, 2004).

The term multimedia appeared in the 1990s and was defined by Reddi and Mishra (2003) as:

> *"an integration of multiple media elements (audio, video, graphics, text, animation etc.) into one synergetic and symbiotic whole that results in more benefits for the end user than any one of the media element can provide individually".*

Mayer (2005) extended the definition:

> *" a multimedia instructional message is a presentation consisting of words and pictures that is designed to foster meaningful learning".*

The use of multimedia components soon became an important part of e-learning strategies for teaching a wide range of subjects including physiology, although studies that have explored the educational impact of multimedia more often focused on mathematics, engineering and computer science perhaps due to the more technical inkling of staff in these areas (Mayer, 2005). The educational impact of multimedia resources for life science subjects is less well studied which is surprising considering the 3-dimensional and real-time nature of these subjects and the fact that animation can easily depict processes and concepts that students often find hard to grasp.

The purpose of this systematic review is to explain whether multimedia resources are effective in undergraduate life science education. The review covers medical and health-related education where life science subjects are taught. A systematic review investigates a precise research question by conducting exhaustive searches, being transparent in the resource selection and exclusion process, and clearly states the findings. A systematic review conducted on a medical subject will include randomised-controlled trial designs, (Evans and Benefield, 2001), however difficulties arise in applying these approaches to education research due to the less rigorous study quality. Slavin (1986) suggests that in the absence of good quality randomised controlled trials (which would be the mainstay of a medical review); other designs should be accepted if controlled in a matched or longitudinal design. The main conclusion of Evans and Benefield (2001) is that when conducting systematic reviews of education research, papers are not excluded on the basis of whether they were randomised or not, as long as there is an equivalent control or comparison group. Once a body of papers has been identified, the results of studies can be combined into a meta-analysis. This is a statistical procedure that expresses the results in a standardised way and provides a visual and numerical representation of the overall effect of any intervention (Egger *et al.,* 1997).

Methods

Review question
This review explores whether multimedia resources are effective in enhancing undergraduate life science education. The multimedia resources outlined in the review may be delivered in a variety of methods: as stand alone internet-based applications, via a CD-rom or through a virtual learning environment (VLE). Resources may be used as a substitute for educational methods or used in a blended approach.

Criteria for including studies
The participants included in the review were any undergraduate student studying life science, biomedical sciences or a health or medical-related programme e.g. pre-clinical medicine, nursing, dentistry or veterinary science. Participants studying at university, post-secondary education institutions and graduate schools were included.

The types of intervention selected for the study included any multimedia resource combining animation, video with audio and text, comprising of a range of granularities - learning objects, CAL, CAI, web-based learning, blended learning, and entire modules delivered via a VLE. All randomised controlled studies were selected, as were quasi-randomised studies that allocated groups based on pre-scheduled timetabled sessions. Non-randomised studies with a comparison or control group were included (e.g. one group or more with a pre- and post-test; two groups or more with a post-test). The primary outcome measure of interest was a knowledge or learning gain indicated by test or examination results, or percentage of students passing a test or examination.

Criteria for excluding studies

Studies were excluded if the participants were from school, high-school, continuing professional development or clinical medicine. Other exclusions were papers about other forms of education e.g. public health or patient advice. Articles that evaluated student satisfaction as a study outcome were excluded since this was not the primary outcome of interest of this systematic review.

Search strategy for identifying studies

A search strategy was developed in conjunction with a university librarian, and used an iterative process with the citations of initial searches scanned for additional terms to generate a broader list of key words. Medline was searched using MESH Subject Headings and non-categorised terms (e.g. learning object). Search terms included the intervention (CAL, multimedia, animation, learning object); participants (undergraduate) and subject matter (e.g. physiology, anatomy, biomedical).
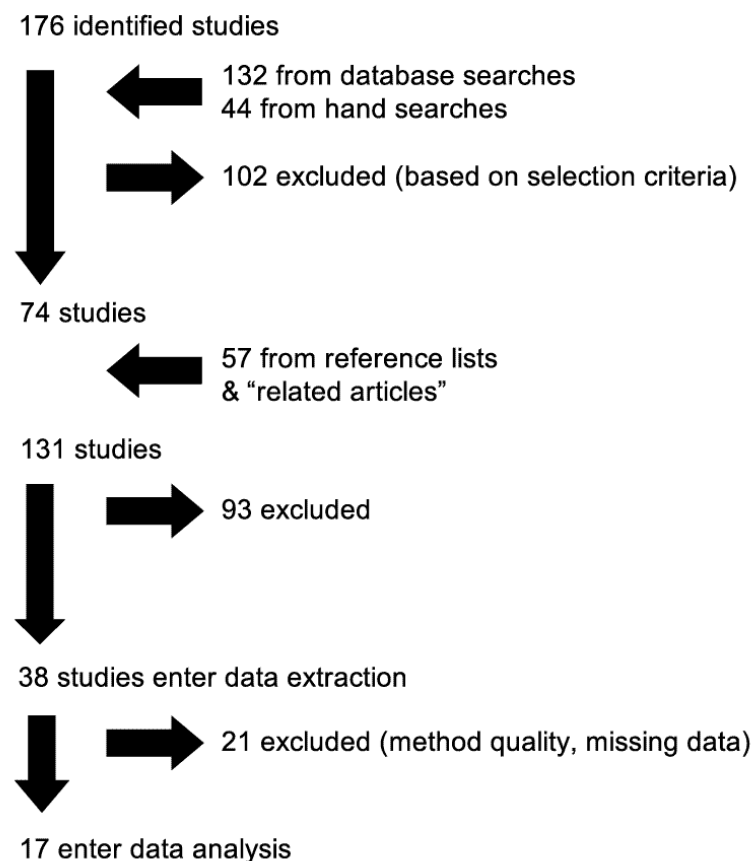
176 identified studies

132 from database searches
44 from hand searches

102 excluded (based on selection criteria)

74 studies

57 from reference lists
& "related articles"

131 studies

93 excluded

38 studies enter data extraction

21 excluded (method quality, missing data)

17 enter data analysis

Figure 1 *Details of literature searching leading to excluded and selected studies*

The following electronic databases were searched:

- Cochrane group, EPPI Centre, The Campbell Library, BEME
- ISI Web of Knowledge Database (Web of Science, BIOSIS previews, Medline, Journal Citation Reports).
- EBSCO (British Nursing Index, CINAHL, Education Research Complete, ERIC, Library Information Science and Technology, Psych Articles, Psych Info.
- Additional studies were found through the "related article" tab on many databases, and by scanning of reference lists of review articles and identified studies.

In addition, high yield journals were searched by hand e.g. Advances in Physiological Education, experts in the field were approached for data and authors were approached for missing data and also asked if they were aware of any unpublished studies.

Study selection and data extraction

Studies were selected using the criteria independently by two subject specialist reviewers (VR, DG) and any discrepancies resolved by discussions with a third reviewer a methodological expert (WC). Full papers were obtained where there was doubt as to whether a study was relevant just from the title and abstract. Study data were extracted onto a pre-designed spreadsheet and included details of the institution, sample population, course and level of study, educational intervention and technology used, outcome data and a brief comment on the author's conclusions. The outcome data were extracted independently by two reviewers (VR, DG), and disagreements resolved by the third (WC).

Assessment of risk of bias

An assessment of the risk bias was made using the criteria detailed in the Cochrane Handbook (Higgins and Green, 2011), with studies designated quality scores. A low risk of bias (allocation to groups is randomised and concealment of groups is explicit); a moderate risk of bias (randomised allocation in which concealment is not explicit), and high risk of bias (allocation to groups is neither randomised nor concealed). The risk of bias was not a reason to exclude studies in this review.

Heterogeneity

The level of similarity between the studies included in the analyses was estimated using the I-squared (i^2) statistic, which displays a percentage of variability not due to chance that results in diversity of the study population or methods. An I-squared value greater than 75% represent a high level of dissimilarity and levels less than 40% represent a good level of consistency (Higgins and Green, 2011).

Data analysis and statistics

Meta-analyses were constructed and statistical analyses performed using the Cochrane Review Manager software (RevMan 5, The Cochrane Collaboration, 2011). The measures of the effect of the intervention were generally continuous data based on the numerical grades obtained in a test or exam, or the percentage of students passing an exam, so were thus represented by the mean difference of the actual numerical result or % pass rate. Due to the variable level of heterogeneity in the analyses, a random effects model was used which includes the variability in the result as reflected in wider confidence intervals. The level of statistical significance was set at $p<0.05$.

The meta-analysis results represent the overall effect of the intervention, and when displayed as a forest plot, each individual study effect is represented by a square and the pooled results of all the studies are represented by a diamond. The centre of the square or diamond displays the overall effect, and the spread of the square or diamond represents the confidence interval. The results are also displayed numerically (Higgins and Green, 2011).

The effects estimate is expressed as the mean difference and confidence intervals. The mean difference represents predominantly the numerical grade obtained in a test or exam with the exception of one study that looked at the percentage of correct answers in an exam.

Results

The 17 publications identified carried out 9 comparisons (e.g. multimedia instruction *versus* traditional teaching), and there were 16 sub-groupings depending on how the author's evaluated learning gain (numerical result of a test or exam, or percentage of correct answers in an exam). In total, the meta-analyses pooled data from 2,290 student participants (Table 1).

Table 1 *Details of included study participants and methodologies*

Study	Participants		Methods	
	No.	**Course**	**Design**	**Comparison**
Dewhurst 1994	14	Year 2 physiology	NRCT, cross over	Practical vs multimedia
Lilienfield 1994	60	Pre-clinical medicine	NRCT	Practical vs multimedia
Kronke 2010	295	Pre-clinical medicine	Quasi RCT	Practical vs multimedia
Dewhurst 1998	39	Year 1 physiotherapy	NRCT, cross over	Lecture vs multimedia
Gibbons 2004a	30	Year 2 science	RCT, cross over	Lecture vs multimedia
Gibbons 2004a	47	Year 1 science	Quasi RCT	Practical vs multimedia
Bogacki 2004	45	Year 1 dentistry	RCT	Lecture vs multimedia
Pereira 2007	138	Year 1 biology	NRCT	Lecture vs blended
Evans 2004	32	Year 3 biological science	NRCT, cross over	Static web vs interactive
Glittenberg 2006	40	Opthalmics	RCT	Textbook vs animation
McGrath 2003	1541	Year 1 undergraduates	NRCT	Multimedia + lecture vs Tutorial + lecture
Nicholson 2006	61	Pre-clinical medicine	RCT	3D multimedia vs Multimedia
Jenkins 2008	73	Pre-clinical medicine	RCT	Lecture vs multimedia?
Marsh 2009	?	Undergraduate science	RCT	Lecture vs lecture + Animation
O'Day 2006	86	Year 3 biological science	RCT	Graphics vs animations
O'Day 2007	393	Year 3 biological science	NRCT	Graphics vs animations
Thatcher 2006	22	Year 1 osteopathy	Quasi RCT	Textbook vs multimedia CD
Starbek 2009	468	Pharmacy/ nursing	Quasi RCT	Multiple; Textbook, illustrated text, lecture, multimedia

The study outcome measures in the sub-groupings included post-test results completed immediately after using the resource, retention tests completed one to several months afterwards, examination results and percentages of correct examination answers completed at the end of the academic year.

Overall, the use of multimedia embedded into a wide variety of eLearning strategies was found to have a positive outcome on undergraduate student learning gains in 10 of the 16 sub-group comparisons (Table 1). It must be noted that the heterogeneity or variability of the studies was high and this was not surprising since studies were not excluded on the basis of quality.

Literature searches

As Figure 1 illustrates, a total of 176 studies were retrieved from the searches (132 from database searches and 44 from hand searches) with 57 articles added in from reference lists and "related article" suggestions; no additional studies were identified through contact with experts in the field. Titles and abstracts were reviewed to identify studies that met the inclusion criteria, and full articles were obtained where required. This meant that 195 studies were excluded due to a failure to meet the selection criteria; 56 were not medical related (e.g. engineering papers); 46 were school and pre-18 college age groups; 44 were review articles and not intervention studies; 11 evaluated student satisfaction as an outcome; 18 studies identified were duplicates in the searches; and there were 20 papers that evaluated only 1 study group with no control.

Excluded studies

A total of 38 studies entered the data extraction phase and, after obtaining the full articles for critical appraisal, a further 21 were excluded on the basis of inappropriate study design. Four papers had no comparison or control group and did not measure a learning gain (Adamczyk, 2009; Blake, 2003; Dantas, 2008; McAteer, 1996); three evaluated an inappropriate student population (Dunsworth, 2007; McLean, 2005; Corton, 2006); eight had incomplete and/or irretrievable data (Fawver, 1990; Garg, 2002; Stith, 2004; McFarlin, 2008; Kohlmeier, 2003; Goldberg, 2000; Petersson, 2009 and Guy, 1992). Three other papers were also excluded on the basis of being a single-cohort study with no comparison to either a control group or as part of a before-and-after design (Dewhurst, 2000; McNulty, 2000; McNulty, 2004). Three papers were an evaluation of online delivery without a comparison group (Buchowski, 2010; Rawson, 2002; and Kohlmeier, 2000).

Included studies

Details of the 17 final papers are shown in Table 2.

Table 2 Details of study interventions

Study	Subject	Technology	Multimedia
Dewhurst 1994	Intestinal absorption	Turbo C++ (Borland)	Graphics, animation, text, simulation
Lilienfield 1994	Cardiac cycle	Computer programme	Graphics, audio, animation, text
Kronke 2010	Oxygen consumption	Audiovisual animations	Simulation, animation, quiz, text
Dewhurst 1994	Cardiovascular physiology	CBL programme	Simulation, animation ,quiz, text
Gibbons 2004a	Bioinformatics	Macromedia Authorware	Graphics, text, quiz
Gibbons 2004a	Karotyping	Macromedia Authorware	Graphics, drag and drop
Bogacki 2004	Tooth morphology	CAL package	Text, photos, graphics, 3D, quiz
Pereira 2007	Human anatomy	Web pages	Images, PDF text, quiz
Evans 2004	Infertility genetics	Macromedia Authorware	Audio, photos, graphics, text, quiz
Glittenberg 2006	Eye anatomy	3D animation software	3D models, 3D animations
McGrath 2003	Neurology	NeuroLab (web + CDrom)	Simulations, graphics, audio, text, animation
Nicholson 2006	Ear anatomy	MRI and CT 3D images	3D models, 2D graphics, text
Jenkins 2008	Dermatology	CAI online atlas	Text, graphics, quiz
Marsh 2008	Embryology	2D and 3D multimedia	Animations, 2D graphics, 3D models, text
O'Day 2006	Cell signalling	CorelDraw, PPT	Animations, graphics, narrations
O'Day 2007	Cell biology	CorelDraw, PPT	Animations, graphics, narrations
Thatcher 2006	DNA replication	CDrom	Graphics, animation, text, quiz
Starbek 2009	Protein synthesis	Multimedia package	Graphics, text, animation

N.B. Gibbons presented 2 studies in one paper.

Participants included healthcare, medicine and science undergraduates. Studies evaluated a range of pedagogical approaches. Four studies evaluated whether resources containing multimedia were effective replacements for laboratory physiology and biochemistry practicals (Dewhurst, 1994; Lilienfield, 1994; Kronke, 2010; Gibbons, 2004). In Gibbons (2004), the authors also evaluated electronic resources in replacement of lecture sessions as did three other studies (Dewhurst, 1998; Bogacki, 2004; Jenkins, 2008). McGrath (2003) and Marsh (2008) compared the use of multimedia resources as an adjunct to delivering a traditional lecture, and two compared the use of conventional textbooks with online textbook delivery (Glittenberg, 2006; Thatcher, 2006).

In terms of exploring different multimedia formats, one study looked at using 3D animations and images (Nicholson, 2006); one study compared static versus interactive resources (Evans, 2004). Two studies compared static graphics with animations (O'Day, 2006; O'Day, 2007), and one compared a range of multimedia formats – text, text with illustration, online multimedia delivery versus traditional lecture (Starkbek, 2010).

Risk of bias in included studies
After applying the procedure for assessing bias as described in the Cochrane Handbook, (Higgins and Green, 2011), 7 studies were identified as randomised controlled studies (RCT) but with no details of allocation concealment, so fell into the category – "moderate risk of bias". This level of bias might have resulted in a study being excluded from a medical systematic review, but as Evans and Benefield (1991) discuss, a more lenient approach to quality should be considered in education research. The remaining studies were quasi- and non-randomised (QRCT and NRCT) which are designated "high risk of bias" category, since students were neither randomly allocated to groups nor was allocation concealed. The types of multimedia intervention are shown in Table 2. The studies used different combinations of static graphics, animations text and audio as interventions, whereas none used video.

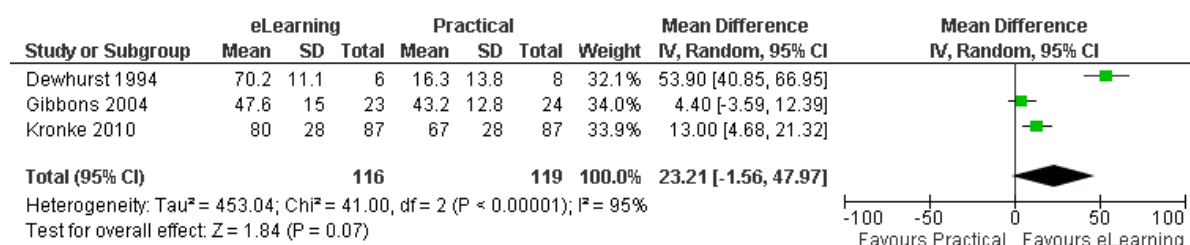*Comparison 1:* Multimedia versus traditional laboratory practical.

Study or Subgroup	eLearning Mean	SD	Total	Practical Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Dewhurst 1994	70.2	11.1	6	16.3	13.8	8	32.1%	53.90 [40.85, 66.95]
Gibbons 2004	47.6	15	23	43.2	12.8	24	34.0%	4.40 [-3.59, 12.39]
Kronke 2010	80	28	87	67	28	87	33.9%	13.00 [4.68, 21.32]
Total (95% CI)			116			119	100.0%	23.21 [-1.56, 47.97]

Heterogeneity: Tau² = 453.04; Chi² = 41.00, df = 2 (P < 0.00001); I² = 95%
Test for overall effect: Z = 1.84 (P = 0.07)

Figure 2A Effects of Multimedia Interventions — Forest plot of Comparison 1: Multimedia versus Practical. Outcome: 1.1 Post-Test results.

Study or Subgroup	eLearning Mean	SD	Total	Practical Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Lilienfield 1994	81.73	7.69	48	75.29	9.65	129	79.9%	6.44 [3.70, 9.18]
Lilienfield 1994b	82	7.93	12	72.46	12.62	63	20.1%	9.54 [4.08, 15.00]
Total (95% CI)			60			192	100.0%	7.06 [4.61, 9.51]

Heterogeneity: Tau² = 0.00; Chi² = 0.99, df = 1 (P = 0.32); I² = 0%
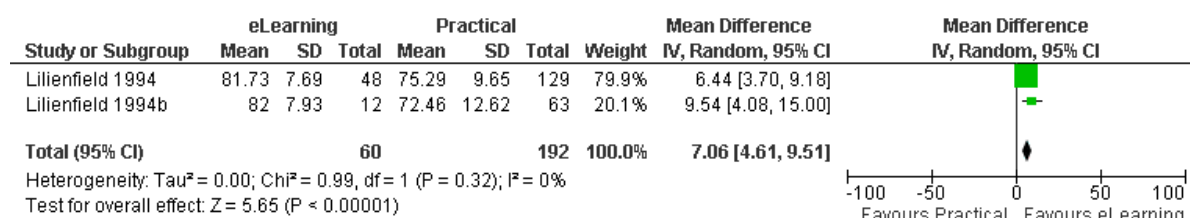Test for overall effect: Z = 5.65 (P < 0.00001)

Figure 2B Effects of Multimedia Interventions — Forest plot of Comparison 1: Multimedia versus Practical. Outcome: 1.2 Exam results.

Four studies replaced a wet laboratory practical experiment with an online multimedia version. Figures 2A and 2B represent outcomes of post-test results and end of year exam results respectively. The analyses indicate that the multimedia version of the practical improved

student's knowledge gains, demonstrated by the recall of practical information when completing a post-test, (mean difference 23.21, confidence interval -1.56 to 47.97), However, this was not significant, (Figure 2A, p=0.07), although significant learning gains were achieved in the end of year examinations (Figure 2B, p<0.00001). The heterogeneity in Figure 2A was high with an I-squared value of 95% suggesting high levels of population and methodological diversity.

Comparison 2: *Multimedia versus traditional lecture.*

Study or Subgroup	eLearning Mean	SD	Total	Lecture Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Dewhurst 1998	52.12	16.67	19	48.18	15.76	17	18.1%	3.94 [-6.66, 14.54]
Dewhurst 1998b	53.79	8.97	20	47.93	10	18	24.4%	5.86 [-0.21, 11.93]
Gibbons 2004	53	10.6	13	45.6	10.9	14	21.5%	7.40 [-0.71, 15.51]
Gibbons 2004c	59.7	27.6	14	67.9	23.7	16	10.1%	-8.20 [-26.74, 10.34]
Starbek 2010	68.49	22.49	115	51	14.89	112	25.9%	17.49 [12.54, 22.44]
Total (95% CI)			181			177	100.0%	7.44 [0.19, 14.69]

Heterogeneity: Tau² = 46.49; Chi² = 15.85, df = 4 (P = 0.003); I² = 75%
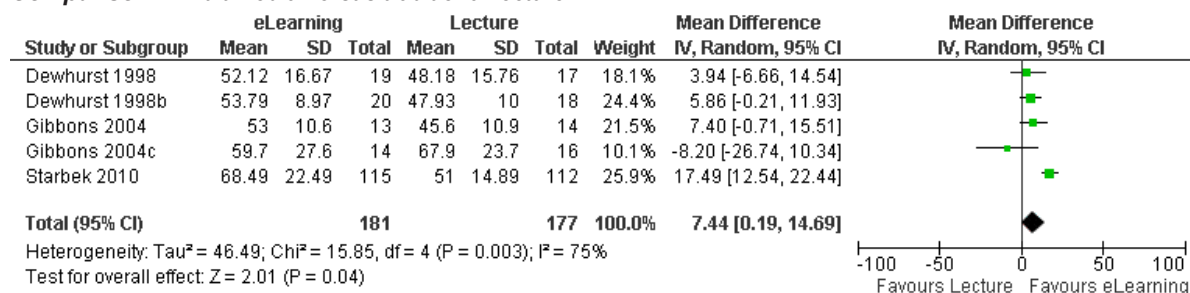Test for overall effect: Z = 2.01 (P = 0.04)

Figure 3 *Effects of Multimedia interventions — Forest plot of Comparison 2: eLearning versus Lecture. Outcome: 2.1 Post-Test results.*

When comparing multimedia substituted for a traditional lecture (Figure 3), there were significant enhancements in knowledge gains observed in a post-test (see table 3)

Table 3 *Summary of study comparisons. A positive (+) effect estimate indicates the outcome favours multimedia and a negative (-) effect shows a favourable control approach.*

Comparison / subgroup	No. Studies	No. Participants	Effects Estimate	I2	P
1 Multimedia vs practical	1.1 post-test	3	235	+23.21 [-1.56, 47.97]	95%	p=0.07
	1.2 exam	1	252	+7.06 [4.61, 9.51]	0%	p<0.00001
2 Multimedia vs lecture	2.1 post-test	3	358	+7.44 [0.19, 14.69]	75%	p=0.04
	2.2 exam	1	45	-0.90 [-3.97, 2.17]	0%	p=0.57
	2.3 exam % correct	1	73	-1.25 [-3.17, 0.67]	0%	p=0.20
3 Multimedia vs blended learning	3.1 exam	1	134	-13.00 [-17.92, -8.08]	0%	p<0.00001
4 Multimedia vs text book	4.1 post-test	2	261	+23.27 [19.80, 26.74]	0%	p<0.00001
5 Static vs interactive webpages	5.1 post-test	1	32	+21.40 [12.09, 30.71]	0%	p<0.00001
6 Graphic vs animation	6.1 short view, post-test	2	37	-11.80 [-20.48, -3.12]	0%	p=0.008
	6.2 repeat view, post-test	1	49	+13.10 [2.66, 23.54]	0%	p=0.01
	6.3 post-test	2	183	+7.10 [-17.79, 31.99]	100%	p=0.58
	6.4 retention test	2	151	+10.15 [4.37, 15.93]	99%	p=0.0006
7 Multimedia + lecture vs tutorial + lecture	7.1 post-test	1	179	+10.04 [0.13, 19.96]	58%	p=0.05
8 Multimedia vs multimedia + 3D	8.1 post-test	1	57	+18.00 [12.03, 23.97]	0%	p<0.00001
9 Lecture vs lecture + multimedia	9.1 retention test	1	76	+7.05 [-2.01, 16.11]	42%	p=0.13
	9.2 retention test (16 months)	1	128	+3.10 [-1.00, 7.20]	20%	p=0.14

It can be observed, that for studies substituting multimedia resources for lectures (comparison 2) that longer term knowledge gains were not observed in examination results suggesting that multimedia may only influence short-term understanding. In the remaining seven comparisons (Table 3), one study compared the use of multimedia alone compared to a blended learning approach (comparison 3); there was a significant favour toward blended learning ($p<0.00001$, 134 participants). Two studies compared the use of multimedia with that of a text book (comparison 4), and a positive impact on short-term learning was observed as indicated by improved post-test results ($p<0.00001$, 261 participants).

Adding levels of interactivity and animation to multimedia resources produced a positive outcome on learning gains observed in students. In comparison 5, improvements in post-test results were observed by adding interactive elements to web pages versus static web pages ($p<0.001$, 32 participants). In a series of studies looking at the effects of un-narrated animation (comparison 6), improvements in post-test and retention tests were observed in students viewing animations on cell signalling compared to graphics alone when viewed repeatedly ($p<0.01$, 49 participants). Viewing animations compared to a static graphic improved retention test results ($p<0.001$, 151 participants) but not more immediate post-test results ($p>0.05$, 183 participants).

Other studies observed that multimedia resources were a beneficial supplement to a lecture compared to a face-to-face tutorial accompanying a lecture (comparison 7, $p<0.05$, 179 participants). Significant enhancements to learning were noted through the use of 3D models compared to flat graphical equivalents (comparison 8, $p<0.001$, 57 participants). Finally, supplementing a traditional lecture with an animation did not enhance student learning above that of a lecture alone when considering a short-term retention test and a test several months later (comparison 9).

Discussion

In this systematic review, following an extensive literature search there were only 18 eligible studies dating back 26 years that addressed the question "are multimedia resources effective in life science education?" This did seem a surprisingly low number considering the advancements internet capability, and the increasing ease of authoring multimedia resources. The studies largely focused on biomedical science subjects including physiology, anatomy and biochemistry. Many studies were excluded because they did not evaluate a learning gain and focused user attitude toward using the resource. Many were excluded due to poor experimental design and study quality, and a disappointing number excluded due to the research being poorly reported with relevant data and information missing, and in total seven authors had to be contacted to supply missing data. In nearly all of the included studies, the evaluation was based on the learning gain of one year group or cohort rather than improving the robustness of the study by repeating it across centres or over a period of years.

This review data suggests that the use of multimedia resources generates a positive outcome as a substitute for wet laboratory practicals, and was effective in short-term understanding when used as a substitute for lectures. There is the need for further well designed studies with more accurate reporting of data in order to thoroughly evaluate the longer-term effectiveness of multimedia resources as a substitute for lectures. Further randomised controlled trials are required to understand more fully the benefits of multimedia resources in a wider range of educational scenarios.

The multimedia resources used for teaching physiology formed the majority of the studies identified, with the resources in the form of interactive online learning packages with animation and text. Many were developed in Macromedia Authorware. What was not clearly articulated

in the studies was the impact on staff preparation time and student time to complete the resources. The motivation to replace practicals included the ability to offer an alternative to animal experiments (Dewhurst, 1994), to save time during medical studies (Lilienfield, 1994) and to save faculty time and resources (Gibbons, 2004). In all these studies, students used the resources in scheduled timetabled sessions, and their impact when used as self-directed study would be an important question to ask. Whilst the primary outcome of interest in this review was learning gains, numerous studies had evaluated student satisfaction. Satisfaction is an important consideration as Bogacki suggests that students commented that they preferred not to use the resource in isolation but as part of an interactive class (Bogacki, 2004), and other student groups wanted to maintain a face-to-face component to their studies (Dewhurst, 1998), such user information would be important to determine when implementing new technology solutions.

Our research identified several studies that specifically compared the use of interactive elements and animations in comparison to static images. The use of interactive web pages compared to static pages enhanced post-test performance (Marsh, 2008), and animations improved student performance over static graphics (O'Day, 2006; O'Day, 2007). This provides important evidence to suggest that multimedia resources enhance learning processes in life sciences, supporting the design theories of Mayer that combining visual aids with text (in the form of printed text or audio), enhances learning processes in maths, engineering and a range of other academic subjects (Mayer, 2005). The theoretical basis for multimedia-enhanced learning is that a well designed resource (visual plus text) optimises the cognitive load. An excessive number of elements such as animation/audio as well as text could provide cognitive overload. Conversely, a single resource e.g. just text looses the impact of the visual imagery. Our meta-analysis results colludes with a previous meta-analysis of 26 studies (Hoffler and Leutner, 2007) in which advantages of instructional animation over static pictures are observed, particularly when the animation was graphically realistic as opposed to a more cartoon style.

The meta-analysis by Lin *et al.,* (2007) highlights the components within a multimedia resource that possibly contribute to an educationally effective outcome, with benefits observed from the addition of audio / narration; using chunking strategies to form "bite-size" chunks of learning to reduce cognitive load; using scaffolding and defining terminology throughout resources and by encouraging deeper learning and allowing the student to build on previous knowledge through comprehension exercises and tests. The authors concluded that the greatest impact on student learning was gained by encouraging meaningful learning, followed by chunking, audio/narration and scaffolding (Lin *et al.,* 2007). This provides a useful framework to apply to resource development approaches.

There is need for further research into the effectiveness of multimedia resources in life science education, not just focusing on a wider-range of pedagogical scenarios, but a wider-range of multimedia formats. Current technology and software developments greatly assist the production of multimedia in the form of presentations with narration and video. There is the need for those involved in pedagogic research to strive to adopt a randomised control study approach, or if limited by timetable and institutional restrictions, to adopt the best study design possible with the inclusion of a control in either a matched or longitudinal design (Slavin, 1986). In the present review only seven studies were randomised, and four were quasi-randomised with students selected based on their timetabled groups. A large number of studies were excluded from having no comparison group for control purposes, and for not performing a pre-test to determine the prior knowledge of the participants.

Many more studies were excluded due to poor data reporting with incomplete datasets, for example, studies that did not report the standard deviation or standard errors of their data.

Some studies reported the results of a knowledge test without stating the total marks awarded which is required before the meta-analysis can be performed. The poor reporting of study methodology and incomplete data reporting is well recognised, and both prevent the reliable assessment of quality of research studies and also introduce bias due to studies having to be un-necessarily excluded (Chan and Altman, 2005). The authors describe that in their review of 519 clinical trials, most failed to state the primary outcome of interest, and most fail to report on the presence of blinding and specifically which of the trial participants (patients, clinicians) were blinded.

In our review, the heterogeneity in the meta-analyses was high, occasionally over 75% reflecting the diversity of the study populations and methodologies employed. This is further evidence to suggest that guidelines for education research into learning technology could be established to ensure trials are conducted more consistently in the future, in terms of student selection, study design and nature of the outcomes measured. Such guidelines are available for reporting medical randomised trials, called the CONSORT statement (Consolidated Standards of Reporting Trials). This is a checklist to provide a useful reminder to authors on what to include in their publications in terms of details (Moher *et al.,* 2001), and this might provide a useful framework for the Bioscience community when not just reporting education research, but when planning and designing studies.

There are limitations to consider when conducting a systematic review. Reporting bias is likely to occur, hence in this review, meetings and published abstract databases were searched along with full journal articles to ensure that no studies were excluded on the basis that they had simply not been published. Experts in the field were contacted to identify on-going or as yet unpublished studies. No studies were excluded on the basis of being written in a non-English language.

Closing Comments

This systematic review highlights the need for randomised controlled trials to evaluate the effectiveness of multimedia in undergraduate life science education for a broader range of scenarios, not just replacing laboratory practicals and lectures, but as self-directed study aids and distance learning resources appropriate with modern educational models. Studies in the future would benefit from being rigorous in design with good quality reporting of all aspects of methodology and study results.

As one paper suggested "the preparation of such materials was not a trivial project" (Lilienfield, 1994), and indeed the acknowledgement that the development of such resources takes time and money is a common theme. It therefore is an exciting moment in time with multimedia resources becoming abundant through open education initiatives. In the UK alone, the Centre for Bioscience "Open Fieldwork Manual" comprises of numerous animations, video and resources to support life science teaching, including a laboratory skills resources (UK Centre for Bioscience, 2011).

With the future prospects of accessing good quality multimedia resources looking bright, the education community could consider adopting a more open approach to collaborate on enhancing educational research studies to enhance the evidence base and therefore decision making processes, not just on which type of resource to use but how to effectively deploy them in different educational scenarios.

Communicating Author: Dr Vivien Rolfe, School of Allied Health Sciences, Hawthorn Building, De Montfort University, Leicester, UK. LE1 9BH. vrolfe@dmu.ac.uk

General References

Chan, A-W. and Altman, D. G. (2005) Epidemiology and reporting of randomised trails published in PubMed journals. *Lancet* **365**, 1159–1162

Egger, M., Smith, G. D. and Phillips, A. N. (1997) Meta-analysis: Principles and procedures. *British Medical Journal,* **315**, 1533

Evans, J. and Benefield, P. (2001) Systematic reviews of education research: Does the medical model fit? *British Educational Research Journal,* **27** (5), 527–541

Higgins, J.P.T. and Green, S. (2011) Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0. www.cochrane-handbook.org (accessed 4th July 2011)

Hoffler, T.N. and Leutner, D. (2007) Instructional animation versus static pictures: a meta-analysis. *Learning and Instruction,* **17**, 722–738

Lin, H., Ching, Y.-H., Ke, F. and Dwyer, F. (2007) Effectiveness of various enhancement strategies to complement animated instruction: a meta-analytic assessment. *Journal of Educational Technology Systems,* **35** (2), 215–237

Mayer, R.E. (2005) The Cambridge Handbook of Multimedia Learning. Ed Mayer, R.E. Cambridge University Press, UK

Moher, D., Schultz, K. F. and Altman, D. G. (2001) The CONSORT statement: revised recommendations for improving the quality of reports of parallel group randomized trials. *BMC Medical Research Methodology*, **1** (2)

Pereira, J.A., Pleguezuelos, E., Meri, A., Molina-Ros, A., Molina-Tomas, M.C. and Masdeu, C. (2007) Effectiveness of using blended learning strategies for teaching and learning of human anatomy. *Medical Education* **41**, 189–195

Reddi, U.V. and Mishra, S. *Educational Multimedia. A handbook for teacher-developers.* A publication for the Commonwealth Educational Media Centre for Asia. www.cemca.org/emhandbook/edmul_full.pdf (accessed 4th July 2011)

Skinner, B. F. (1960) Teaching Machines. *The Review of Economics and Statistics,* **42** (3 part 2), 189–191

Slavin, R.E. (1986) Best-evidence synthesis: an alternative to meta-analytic and traditional reviews. *Education Research,* **15**, 5–11

Suppes, P. (1972) Computer-Assisted Instruction at Stanford. Technical Report No. 174. http://suppes-corpus.stanford.edu/techreports/IMSSS_174.pdf (accessed 4th July 2011)

UK Centre for Bioscience. (2011) Open Educational Resources. www.bioscience.heacademy.ac.uk/resources/oer/ (accessed 4th July 2011)

Studies Included in Meta-analysis

Bogacki, R.E., Best, A. and Abbey, L.M. (2004) Equivalence study of a dental anatomy computer-assisted learning program. *Journal of Dental Education,* **68** (8), 867–871

Dewhurst, D.G., Hardcastle, J., Hardcastle, P.T. and Stuart, E. (1994) Comparison of a computer simulation program and a traditional laboratory practical class for teaching the principles of intestinal absorption. *Advances in Physiology Education,* **267**, 95–104

Dewhurst, D.G. and Williams, A.D. (1998) An investigation of the potential for a computer-based tutorial program covering the cardiovascular system to replace traditional lectures. *Computers and Education,* **31**, 301–317

Evans, C., Gibbons, N.J., Shah, K. and Griffin, D.K. (2004) Virtual learning in the biological sciences: pitfalls of simply "putting notes on the web". *Computers and Education,* **43**, 49–61

Gibbons, N.J., Evans, C., Payne, A., Shah, K. and Griffin, D.K. (2004) Computer simulations improve university instructional laboratories. *Cell Biology Education,* **3**, 263–269

Glittenberg, C. and Binder, S. (2006) Using 3D computer simulations to enhance ophthalmic training. *Opthalmic and Physiological Optics,* **26**, 40–49

Jenkins, S., Goel, R. and Morrell, D.S. (2008) Computer-assisted instruction versus traditional lecture for medical student teaching of dermatology morphology: a randomized control trial. *Journal of the American Academy of Dermatology,* **59**, 255–259

Kronke, K.-D. (2010) Computer-based learning versus practical course in pre-clinical education: acceptance and knowledge retention. *Medical Teacher,* **32**, 408–413

Lilienfield, L.S. and Broering, N.C. (1994) Computers as teachers: learning from animations. *Advances in Physiology Education,* **266**, 47–54

Marsh, K.R., Griffin, B.F. and Lowrie, Jr D.J. (2008) Medical student retention of embryonic development: impact of the dimensions added by multimedia tutorials. *Anatomical Sciences Education,* **1**, 252–257

McGrath, P., Kucera, R. and Smith, W. (2003) Computer simulation of introductory neurophysiology. *Advances in Physiology Education,* **27**, 120–129

Nicholson, D.T., Chalk, C., Funnell, W.R.J. and Daniel, S.J. (2006) Can virtual reality improve anatomy education? A randomised controlled trial of a computer-generated three-dimensional anatomical ear model. *Medical Education,* **40**, 1081–1087

O'Day, D.H. (2006) Animated cell biology: a quick and easy method for making effective, high-quality teaching animations. *Life Science Education,* **5**, 255–263

O'Day, D.H. (2007) The value of animations in biology teaching: a study of long-term memory retention. *Life Science Education,* **6**, 217–223

Starbek, P., Erjavect, M.S. and Peklaj, C. (2010) Teaching genetics with multimedia results in better acquisition of knowledge and improvement of comprehension. *Journal of Computer Assisted Learning,* **26** (3), 214–224

Thatcher, J. (2006) Computer animation and improved student comprehension of basic science concepts. *Journal of the American Osteopathic Association,* **106** (1), 9–14

Excluded Studies

Adamczyk, C., Holzer, M., Putz, R. and Fischer, M.R. (2009) Student learning preferences and the impact of a multimedia learning tool in the dissection course at the University of Munich. *Annals of Anatomy,* **191**, 339–348

Blake, C.A., Lavoie, H.A. and Millette, C.F. (2003) Teaching medical histology at the University of South Carolina School of Medicine: Transition to virtual slides and virtual microscopes. *Anatomical Record (Part B: New Anatomy),* **275B**, 196–206

Buchowski, M.S., Plaisted, C., Fort, J. and Zeisel, S.H. (2010) Computer-assisted teaching of nutritional anemias and diabetes to first-year medical students. *American Journal of Clinical Nutrition,* **75**, 154–161

Corton, M.M., McIntire, D.D., Wai, C.Y., Ling, F.W. and Wendel, Jr. G.D. (2006) A comparison of an interactive computer-based method with a conventional reading approach for learning pelvic anatomy. *American Journal of Obstetrics and Gynecology,* **195**, 1438–1443

Dantas, A.M. and Kemm, R.E. (2008) A blended approach to active learning in a physiology laboratory-based subject facilitated by an e-learning component. *Advances in Physiology Education,* **32**, 65–75

Dewhurst, D.G., Macleod, H.A. and Norris, T.A.M. (2000) Independent student learning aided by computers: an acceptable alternative to lectures? *Computers in Education,* **35**, 223–241

Dunsworth, Q. and Atkinson, R.K. (2007) Fostering multimedia learning of science: exploring the role of an animated agent's image. *Computers in Education,* **48**, 677–690

Fawver, A.L., Branch, C.E., Trentham, L., Robertson, B.T. and Beckett, S.D. (1990) A comparison of interactive videodisc instruction with live animal laboratories. *Advances in Physiology Education,* **4**, S11–S14

Garg, A.X., Norman, G.R., Eva, K.W., Spero, L. and Sharan, S. (2002) Is there any real virtue of virtual reality? The minor role of multiple orientations in learning anatomy from computers. *Academic Medicine,* **77** (10), S97–S99

Goldberg, H.R. and McKhann, G.M. (2000) Student test scores are improved in a virtual learning environment. *Advances in Physiology Education,* **23** (1), 59–66

Guy, J.F. and Frisby, A.J. (1992) Using interactive videodiscs to teach gross anatomy to undergraduates at The Ohio State University. *Academic Medicine,* **67**, 132–133

Kohlmeier, M., Althouse, L., Stritter, F. and Zeisel, S.H. (2000) Introducing cancer nutrition to medical students: effectiveness of computer-based instruction. *American Journal of Clinical Nutrition,* **71**, 873–877

Kohlmeier, M., McConathy, W.J., Lindell, K.C. and Zeisel, S.H. (2003) Adapting the contents of computer-based instruction based on knowledge tests maintains effectiveness of nutrition education. *American Journal of Clinical Nutrition,* **77**(supplement), 1025S–1027S

McAteer, E., Neil, D., Barr, N., Brown, M., Draper, S. and Henderson, F. (1996) Simulation software in a life sciences practical laboratory. *Computers in Education,* **26** (1–3), 101–112

McFarlin, B.K. (2008) Hybrid lecture-online format increases student grades in an undergraduate exercise physiology course at a large urban university. *Advances in Physiology Education,* **32**, 86–91

McLean, P., Johnson, C., Rogers, R., Daniels, L., Reber, J., Slator, B.M., Terpstra, J. and White, A. (2005) Molecular and cellular biology animations: development and impact on student learning. *Cell Biology Education,* **4**, 169–179

McNulty, J.A., Halama, J., Dauzvardis, M.F. and Espiritu, B. (2000) Evaluation of web-based computer-aided instruction in a basic science course. *Academic Medicine,* **75** (1), 59–65

McNulty, J.A., Halama, J. and Espiritu, B. (2004) Evaluation of computer-aided instruction in the medical gross anatomy curriculum. *Medical Education,* **17**, 73–78

Petersson, H., Sinkvist, D., Wang, C. and Smedby, O. (2009) Web-based interactive 3D visualisation as a tool for improved anatomy learning. *Anatomical Sciences Education,* **2**, 61–68

Rawson, R.E. and Quinlan, K.M. (2002) Evaluation of a computer-based approach to teaching acid/base physiology. *Advances in Physiology Education,* **26** (2), 85–97

Stith, B.J. (2004) Use of animation in teaching cell biology. *Cell Biology Education,* **3**, 181–188