

This author's accepted manuscript may be used for non-commercial purposes in accordance with [Wiley Terms and Conditions for Self-Archiving](#).

The full details of the published version of the article are as follows:

TITLE: Anatomy Teaching, a “Model” Answer? Evaluating “Geoff”, a Painted Anatomical Horse, as a Tool for Enhancing Topographical Anatomy Learning

AUTHORS: Bietzk, E; Weller, R; Simons, V; Channon, S B

JOURNAL: Anatomical Sciences Education

PUBLISHER: Wiley

PUBLICATION DATE: 9 November 2018 (online)

DOI: <https://doi.org/10.1002/ase.1823>

ASE-18-0011.R2

Research Report

Anatomy teaching, a “model” answer? Evaluating "Geoff", a painted anatomical horse, as a tool for enhancing topographical anatomy learning

Edward Bietzk¹, Renate Weller¹, Victoria Simons¹, Sarah B Channon^{2*}

¹Department of Clinical Science and Services, The Royal Veterinary College, University of London, North Mymms, Hatfield, Hertfordshire, United Kingdom

²Department of Comparative Biomedical Sciences, The Royal Veterinary College, University of London, London, United Kingdom

Running Title: Models for topographical anatomy teaching

*Correspondence to: Dr. Sarah B. Channon, Department of Comparative Biomedical Sciences, The Royal Veterinary College, Royal College Street, London NW1 0TU, United Kingdom. E-mail: schannon@rvc.ac.uk

Grant Sponsor: Animal Care Trust, UK registered charity number 281571.

ABSTRACT

Development of new methods for anatomy teaching is increasingly important as we look to modernize and supplement traditional teaching methods. In this study, a life-size equine model, "Geoff", was painted with surface and deep anatomical structures with the aim of improving students' ability to convert theoretical knowledge into improved topographical anatomy knowledge on the live horse. Third and fourth year veterinary medicine students ($n = 45$) were randomly allocated into experimental (used "Geoff") and control (used textbook) groups. The efficacy of the model was evaluated through a structured oral exam using a live horse. Questionnaires gathered information on student confidence and enjoyment of the task. There was no significant difference in the performance of experimental and control groups either immediately ($44 \pm 20\%$ vs. $40 \pm 21\%$; $P = 0.504$) or nine weeks after the learning intervention ($55 \pm 17\%$ vs. $55 \pm 20\%$; $P = 0.980$). There were however specific questions on which the experimental group performed better than controls, and for which gender effects were apparent. The students using "Geoff" showed a transient gain in confidence following the session (Likert scale 2.7 to 3.6) however the initial increase was no longer present at the second test. There was a significant influence of gender on confidence with greater confidence gains in females in the Experimental group. The students found the model to be extremely useful and both groups found the sessions enjoyable. The model will be of benefit as a complementary learning tool for students.

Keywords: Veterinary education; undergraduate education, veterinary anatomy, topographical anatomy; models; live anatomy

INTRODUCTION

Sound anatomy knowledge and the ability to apply this to patients in a clinical context is essential in medical and veterinary practice (Sugand et al., 2010; Azer, 2013). Traditionally, anatomy teaching has been provided through lectures, dissections and the use of additional resources such as textbooks. Students still find dissections central in the understanding of deep anatomy, aiding recall and three-dimensional perception (Azer and Eisenburg, 2007). Time available in the curriculum (Waterston and Stewart, 2005; Fitzgerald et al., 2008; Yammine and Violato, 2015), as well as cadaver supply and availability for students have decreased however in recent times (Kumar et al., 2001; Martinsen and Jukes, 2005; Chen et al., 2018). Further issues surrounding cadaver use such as cost, and safety (McLachlan and Patten, 2006), as well as the moral and ethical concerns surrounding dissection are also a continued consideration for educators (Kinnison et al., 2009; Woodmansey, 2016). Consequentially, in many medical institutions, cadaver dissection is no longer the only, or main method, of practical anatomy instruction (Drake, 2002; Drake et al., 2009; Craig et al., 2010; Watmough et al., 2010; Cho and Hwang, 2013; Topping, 2014; Halliday et al., 2015).

Though some medical education establishments in modern times have moved away from dissection either partially or completely, such a dramatic shift has not yet been seen within veterinary education. Students are expected to be competent in elective surgeries, for example spays, castration and dental surgery, on graduation (RCVS, 2014; AAVMC, 2018). In the United Kingdom veterinary graduates obtain full licensure for independent practice on completion of undergraduate studies but in the United States licensure is obtained on sitting the North American Veterinary Licensing Examination (NAVLE). This sits in contrast to human

1
2
3 medicine students who, in both the UK and US, typically complete a residency or foundation
4 program prior to obtaining independent license. As such, cadavers retain an important role in
5 development of anatomical knowledge and surgical skills in many veterinary schools (Latorre et
6 al., 2006). Veterinary education providers still nevertheless face the time pressures of modern
7 curricula (Hall et al., 2013), and must balance the “3Rs” (reduction, replacement, and refinement
8 of animal use, as defined by Russell and Burch (1959)), with the need to provide authentic and
9 effective anatomy learning experiences for students (Hart et al., 2005; Martinsen and Jukes,
10 2005).
11
12
13
14
15
16
17
18
19
20
21
22
23

24 In order to achieve an effective balance between traditional and more ethically and economically
25 sustainable anatomy teaching methods, many novel teaching methods have been proposed and
26 developed. These often aim not to replace traditional methods, but to complement and enhance
27 the learning experience. At the study institution, haptic technology (Kinnison et al., 2009),
28 plastination (Latorre et al., 2007), near-peer teaching (Hall et al., 2013) and physical anatomical
29 models (Braid et al., 2012; Preece et al., 2013) have been successfully integrated into the
30 veterinary anatomy curriculum. Elsewhere, body painting (Finn et al., 2011; Senos et al., 2015),
31 ultrasound (Pawlina and Drake, 2015; Jamniczky et al., 2017), three-dimensional (3D) printing
32 (Smith et al., 2018), artistic approaches (Backhouse et al., 2017) and computer-based learning
33 systems (Khalil et al., 2010) have also been shown to be effective alternatives, and add diversity
34 to the curriculum. Such a multi-modal approach may be especially important in this era of
35 “personalized” and “self-directed” learning, where educators are increasingly aware of an ever
36 more diverse student body and the inevitable accompanying variety of learning styles and
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 preferences (Entwistle and Tait, 1990; Murphy et al., 2004; O'Mahoney et al., 2016; Hernández-
4
5 Torrano et al., 2017).
6
7
8
9

10 To successfully promote self-directed learning, provision of suitable, scaffolded (Rosenshine
11 and Meister, 1992), and accessible supplementary learning resources is key (Smith et al., 2018).
12 Historically in a pre-internet era, students will have turned to textbooks to aid their personal
13 study. Textbooks are still a popular resource choice by medical students (Choi-Lundberg et al.,
14 2015), but are potentially not optimal for learning as they can contain a level of detail that leads
15 to considerable extraneous (detrimental) cognitive load upon students (Khalil et al., 2010). They
16 also use two-dimensional photographs and drawings to depict three-dimensional anatomical
17 structures. Many students find interpretation of such images difficult and struggle with
18 translating this information onto living beings (Brenton et al., 2007). Resources that lower
19 cognitive load may be more beneficial for learning, particularly for novices (van Merriënboer
20 and Sweller, 2009). Low fidelity models present simple, representative information to the
21 learner, making it easier to remember while maintaining the dimensionality, form and spatial
22 relationships between the structures (Chan and Cheng, 2011). Meta-analyses illustrate that three
23 dimensional tools produce significantly better results in overall knowledge outcome, spatial
24 understanding, and long term knowledge retention when compared to all other teaching methods
25 (Yammine and Violato, 2015). Evidence suggests that manipulation of physical models in a 3D
26 space can significantly benefit a student's visuospatial understanding of structures, especially
27 those structures with complex spatial relationships (Krontiris-Litowitz, 2003; Waters et al., 2005;
28 Latorre et al., 2007; Jittivadhna et al., 2009; Motoike et al., 2009; Oh et al., 2009; Jittivadhna et
29 al., 2010; DeHoff et al., 2011; Preece et al., 2016; Smith et al., 2018). The use of 3D models is
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 therefore an exciting area of development in anatomy education with much potential for further
4
5 refinement and application.
6
7
8
9

10 Topographical anatomy (regional relations between structures and features), “live” anatomy
11 (anatomy as it relates to the living animal) and “surface” anatomy (concerning the surface and
12 palpable features of the body) are related dimensions of anatomy whereby a learner’s ability to
13
14 transition between two-dimensional (2D) and three-dimensional (3D) representations may be
15 particularly important. Good topographical anatomy knowledge is essential for the practicing
16
17 medic or veterinarian in order to accurately carry out clinical procedures (Azer, 2013). Locating
18
19 structures beneath the skin by recognizing their location and palpable characteristics (Chou et al.,
20
21 2010) as well as identifying anatomical landmarks on the patient during clinical procedures are
22
23 skills that must be acquired. Body painting has been used in human medicine (Op Den Akker et
24
25 al., 2002; Finn and McLachlan, 2010; Nanjundaiah and Chowdapurkar, 2012) and more recently
26
27 the veterinary field (Senos et al., 2015) to allow students to explore and learn surface anatomy.
28
29 Body painting however, despite its richness as a learning experience, is time consuming and
30
31 requires care and supervision to ensure the structures are marked correctly (Nanjundaiah and
32
33 Chowdapurkar, 2012). Body painting in animals is challenging, since live animals are often
34
35 reluctant to stay still for the required duration. Live animals are often not available in large
36
37 numbers to allow all students to participate in such an activity (whereby in medical education
38
39 students can participate as models themselves). In this study therefore, a permanently painted
40
41 life size anatomical model was proposed, created and evaluated, to assess if 2D anatomical
42
43 representations on a life sized 3D model could help veterinary students learn to identify the
44
45 location and spatial relationships of anatomical structures on a living equine.
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 It was hypothesized that studying the painted equine model would significantly enhance a
4
5 student's ability to apply anatomical knowledge onto a live horse, when compared to students
6
7 using traditional methods. Secondly, it was hypothesized that any learning benefits from using
8
9 the model horse would still be apparent several weeks after use of the model horse (enhance
10
11 medium term knowledge retention). Thirdly, it was hypothesized that use of the model horse
12
13 would increase student self-reported confidence and enjoyment during the learning experience
14
15 when compared to those learning topographical anatomy via traditional methods.
16
17
18
19
20

21 MATERIALS AND METHODS

22
23
24
25

26 Educational Context

27
28 The course studied is a five-year undergraduate degree in Veterinary Medicine at the Royal
29
30 Veterinary College, University of London, UK. The program follows an integrated (both
31
32 horizontally and vertically) body-systems based curriculum, with strong emphasis on self-
33
34 directed learning (Knowles, 1975) and professional skills development (Dale et al., 2008).
35
36 Anatomy is taught heavily integrated with physiology, biochemistry, and clinical science, and as
37
38 such lectures featuring solely anatomy are rare. Practical anatomy is an important part of the
39
40 program, and is taught principally through both prosection and dissection, dependent on the
41
42 region and experience of the students. In addition, students undertake Integrated Structure and
43
44 Function (ISF) tutorials with live animals which highlight topographical anatomy, and revisit
45
46 basic and clinical science concepts in the context of a live animal. Anatomy teaching is biased
47
48 towards the first two years of the course (though is revisited in clinical context in later years),
49
50 and during these two years students participate in approximately 13 hours of prosection
51
52
53
54
55
56
57
58
59
60

1
2
3 practicals, 42 hours of dissection, and 3 hours of ISF tutorials. These sessions cover all body
4
5 systems sequentially, and include the major veterinary species (dog, horse, and ruminant) as well
6
7 as a small ‘exotic’ species component (including birds, rabbits, fish, and reptiles). As anatomy
8
9 teaching is integrated throughout the program it is not examined as a separate entity – summative
10
11 assessments, which all take place at the end of the academic year, integrate all systems-based
12
13 teaching. Anatomy features in every assessment type (multiple choice questions [MCQs],
14
15 problem solving short-answer questions [PSQs], essays), however its most rigorous and valid
16
17 assessment is during ISF oral examinations – an end-of-year semi-structured oral examination,
18
19 which takes place in the dissection room utilizing specimens and live animals.
20
21
22
23
24
25

26 Model creation

27
28 A standard life size commercially available fiberglass horse (Shires Equestrian Products,
29
30 Leominster, Herefordshire, UK), affectionately named “Geoff” by veterinary students, was
31
32 painted using acrylic paint (Figure 1). The design depicted a 2D representation of both surface
33
34 and deep anatomy, using textbooks and photographs from whole horse dissections for reference.
35
36 Each side of the horse was painted to convey different aspects of anatomy creating as many
37
38 views and layers as possible within the space available. The model was photographed, and
39
40 labelled flashcards showing anatomical structures and auscultation landmarks were created in
41
42 Microsoft PowerPoint (Microsoft Corp, Redmond WA), printed, laminated and provided for
43
44 students to use during the learning exercise. After the study these flashcards were made available
45
46 for all students to use to deliver an opportunity for self-directed learning of equine anatomy at
47
48
49
50
51 any stage of the veterinary medicine course.
52
53
54
55
56
57
58
59
60

Initial Investigations.

Data were collected at the Royal Veterinary College, London, UK, following approval from the RVC Ethics and Welfare committee. An initial questionnaire was distributed to final year students on the Bachelor of Veterinary Medicine course which was used to highlight the areas of equine anatomy these advanced students found most challenging and to identify the resources that they commonly used to learn anatomy – the responses to the latter question were used to inform the textbook used for the control group in this study.

Participants.

Students were recruited to the study by invitation, sent via email to all third and fourth year students on the Bachelor of Veterinary Medicine course (270 students and 238 students respectively; 21 % male and 79 % female). These year groups were chosen since at this stage of the course students had undertaken anatomy teaching of all body systems, and completed all dissections and live animal tutorials. The two cohorts were therefore identical in the formal anatomy teaching they had received to date and differed only in the amount of clinical teaching they had received (with fourth year students having completed 10 weeks of extra mural studies placements within veterinary practices). None of the students had started structured rotational clinical teaching within the institution.

Forty-five volunteer students were randomly allocated into the experimental group ($n = 24$; year 3 = 10, year 4 = 14), or the control group ($n = 21$; year 3 = 13, year 4 = 8). Both groups were provided the same set of learning objectives: Identify the location of the bones and joints of the limbs and vertebral column; Show the position of the superficial musculature, and major tendons

of the thoracic and pelvic limbs; Identify the paths of the nerves of the distal limb; Identify the location of the heart and the individual valves within it; Identify the sites of venipuncture in the horse; Outline the topographical anatomy of the gastrointestinal system. The experimental group were given "Geoff", the model horse, accompanied by laminated cards depicting labelled images of "Geoff"; the control group were each given a copy of the textbook Veterinary Anatomy of Domestic Mammals (König and Liebich, 2007). Study groups of approximately 6 students were given 30 minutes to address the learning objectives with the resources provided. Students were requested to converse within their groups to simulate peer-to-peer learning, as expected in an authentic self-directed learning session.

Model Evaluation

Questionnaire

Prior to the learning opportunity, students were provided with a questionnaire (S1, Supplementary Information) to evaluate their learning preferences, their practical equine experience, and confidence in their equine anatomy knowledge/application. The questionnaire used a five-point Likert scale. Cronbach alpha for the questionnaire was 0.809, above the threshold of 0.70, indicating reliability.

Knowledge Test

Following the learning opportunity, students were individually tested on their learning using a 10 item test that was developed in conjunction with anatomy teaching staff and initially piloted on a small selection of final year students (See Table S2 for questions and associated scoring).

Cronbach alpha for the test was 0.751, above the threshold of 0.70, indicating reliability of the

1
2
3 test instrument. The Kendall's Tau B coefficient for the test was 0.376 which was significant (P
4 = 0.02), demonstrating validity. The test took the form of a structured oral examination using a
5
6 live horse, and was ten minutes long. Questions required students to identify anatomical
7
8 landmarks on the horse and in some cases to discuss clinical or functional significance of this
9
10 anatomy. The examiners were two members of teaching staff who were fully briefed on the
11
12 process and blinded to which group the students belonged.
13
14
15
16
17
18

19 The questionnaire was re-administered following testing. At this point, the questionnaire also
20
21 asked students to rate their enjoyment of the exercise. The experimental group were additionally
22
23 asked how useful they found the model horse in the context of the university teaching resources
24
25 already available. A free text comment box was provided for any further information that
26
27 participants wished to provide.
28
29
30
31
32

33 All volunteers were invited to return nine weeks later to repeat the live horse testing (same test)
34
35 to assess medium-term knowledge retention. 22 out of the original 45 returned with 10 students
36
37 from the experimental group, and 12 from the control group. Similar proportions of Year 3 and 4
38
39 students were retained at the returning session (12 Year Three students and 10 Year Four
40
41 students). The questionnaire to assess confidence, enjoyment and usefulness of the activity was
42
43 repeated at this second testing occasion. Students were not provided any feedback or informed of
44
45 their test scores until after the second testing occasion, and examiners were instructed to take
46
47 care not to give feedback or cues during either testing session.
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 Data analysis.

4
5 Factor Analysis

6
7
8 Factor analysis of both the test and questionnaire was undertaken in order to determine whether
9
10 there were distinct constructs within the data. Bartlett's Test of Sphericity and Kaiser-Meyer-
11
12 Olkin (KMO) measure of sampling adequacy (Beavers et al., 2013) were first computed to
13
14 ensure the data were suitable for factor analysis. Factor analysis was undertaken using the
15
16 Principal Components method with oblique rotation (direct oblimin) in IBM SPSS Statistics
17
18 Version 24 (IBM Corp. in Armonk, NY). The Kaiser Criterion Method in conjunction with a
19
20 Scree Plot (Baryla et al., 2012) were used to identify the number of factors to retain.
21
22
23
24
25

26
27 Factor analysis was justified for the test by a KMO measure of 0.611 and a Bartlett's
28
29 significance of $P < 0.0001$. Initial analysis identified four factors, explaining 66 % of the
30
31 variance in the data. However, one factor contained only one item (Question 4). This item was
32
33 removed from the analysis, however the following three factor solution was identical to the
34
35 previous iteration and only explained 58 % of the variance in the data. Therefore the four factor
36
37 solution was retained (Table S3, supplementary information). Follow up statistical analysis of the
38
39 four constructs as defined by the factor analysis solution was not undertaken as there was no
40
41 clear conceptual rationale for the identified solution.
42
43
44
45

46
47 Factor analysis was justified for the questionnaire by a KMO measure of 0.68 and Bartlett's test
48
49 significance of $P < 0.0001$. Two factors were identified which explained 91 % of the variance,
50
51 one which mapped strongly to the three questions pertaining to confidence levels; the other
52
53 grouped the measures of enjoyment and future use of the model. Factor loadings for each
54
55
56
57
58
59
60

component of the questionnaire are provided in Table 1. When considered as separate constructs, the Cronbach alpha statistic for the “confidence” and “enjoyment” items increased to 0.889 and 0.862 respectively. The Kendall’s Tau B coefficients for the items within each of these constructs ranged from 0.559 to 0.721 for the “confidence” items, and was 0.575 for the “enjoyment” items (all significant, $P < 0.0001$). This demonstrated validity of the survey instrument.

Test scores

The data were assessed by histogram and using a D’Agostino and Pearson normality test (D’Agostino, 1986). Test score was found to be normally distributed thus differences in test scores between treatment groups at each test occasion were analyzed using independent t tests. Data for individual questions were found to be a mixture of normally and non-normally distributed data and so differences in individual question scores between treatment groups, at each test occasion, were all assessed using a Mann-Whitney U test, accepting some loss of power on the three questions containing normal data. Following this, further analysis of test scores between test occasions was undertaken solely using the data from participants who returned for the second test occasion. For this, a paired students t-test and Wilcoxon Signed Rank test were used to compare the pre-post test scores, and the non-normally distributed individual question scores respectively. The Pearson Product Moment Correlation Coefficient was also calculated to assess the degree of association between test scores at each testing occasion.

A separate analysis was conducted to assess the influence of gender on learning outcomes. Test scores, as well as questionnaire item responses were compared at each test occasion between

male and female participants using independent t- tests. Gender influence on scores for individual questions (data not normally distributed), and when the data file was further split by experimental group (due to the small group sample sizes) was assessed using a Mann-Whitney U test.

Questionnaire Data

Likert scale responses on the questionnaire data were found to be normally distributed and were therefore treated as continuous data (Sullivan and Artino, 2013). Questionnaire responses were compared at each testing stage between the control and treatment groups by means of independent t-tests. Paired t tests were used to measure changes in questionnaire item responses from pre to post learning intervention, and between the first and second testing sessions. Analysis of questionnaire responses was conducted for individual questions, as well as for the combined data for each of the two identified individual constructs, “confidence” and “enjoyment” (combined data are presented as means across the construct component scores rather than as summed values). All statistical tests were conducted in IBM SPSS Statistics Version 24 (IBM Corp. in Armonk, NY) and the P-value was set at 0.05.

RESULTS

Participant demographics

Forty-five students volunteered for the study, of which 82 % (n = 37) identified as female and 18 % (n = 8) male. Gender distribution between study groups was broadly equivalent (83 % [n = 20] female within the intervention group, and 81 % [n = 17] female within the control group).

Topographical Anatomy Performance

Test performance was similar for both Experimental and Control groups immediately after the learning intervention. Mean total test scores for the Experimental group, (11 ± 5 or 44 ± 20 %; mean $\pm$ SD), were not significantly different ($P = 0.504$; Cohen's $d = 0.2$ [Confidence Interval $0.39 - 0.79$]) from those of the Control group (10 ± 5 ; 40 ± 21 % (Figure 2). Neither group exceeded the institutional pass mark for this level of 50 %. There was no effect of year of study on immediate post-test performance (Year three 12 ± 6 vs Year four 9 ± 2 ; Cohen's $d = -0.6$ [$-1.2 - -0.02$]; $P = 0.116$).

Nine weeks later, the performance of Experimental versus Control groups was similar ($P = 0.980$). Both groups improved upon their original test scores after nine weeks (Experimental group 14 ± 4 or 55 ± 17 %; Control group 14 ± 5 or 55 ± 20 %). The apparent increase in medium term knowledge was not significant for the experimental group ($P = 0.234$) despite a large effect size (Cohen's $d = 1.2$ [$0.25 - 2.15$]) however, the control group showed a significant and moderate increase in performance ($P = 0.038$; Cohen's $d = 0.7$ [$-0.1 - 1.6$]; Figure 2). There was a strong and significant correlation between first and second test scores for the experimental group ($r = 0.875$, $P < 0.0001$) but no significant correlation for the control group ($r = 0.359$, $P = 0.251$). There was no effect of year of study on medium term test performance (Year three 14 ± 5 vs Year four 13 ± 4 ; Cohen's $d = 0.2$ [$-1.0 - 0.6$]; $P = 0.737$).

Results by individual question can be found in Table S4, supplementary information.

Gender Influence on Anatomy Test Performance

There were no gender influences on overall test scores for either Experimental or Control groups (Figure 2). Both Question 5 (initial test occasion) and Question 3 (second test occasion) showed an effect of gender. For question 5, with all participants combined, male participants performed significantly better than female participants (2 ± 1 vs 1 ± 1 ; $P = 0.044$; Cohen's $d = 0.97$ [$0.18 - 1.76$]), however when the participants were grouped by study method, or when only returning participants were considered, there was no longer a significant gender effect. For question 3 (medium term), female participants performed significantly better than males (1 ± 1 vs 0 ± 1 ; $P = 0.023$; Cohen's $d = 1.3$ [$-0.2 - 2.3$]); when broken down by study method, this difference was still significant in the Experimental group ($P = 0.021$; Cohen's $d = 2.1$ [$-0.5 - 3.7$]) but not in the control group ($P = 0.423$), however only two male participants were present in the control group at this test occasion.

Self-reported questionnaire measures

Confidence

There was initially no significant difference in students' self-reported confidence in anatomy knowledge between Experimental and Control groups. This was evidenced by similar scores for the 'Confidence' construct for both experimental and control groups (2.7 ± 0.7 and 2.7 ± 0.6) respectively ($P = 0.899$; Figure 3).

Following the learning exercise, there was no difference in reported confidence between the experimental group and control group (3.1 ± 0.9 versus 2.7 ± 0.8 ; $P = 0.099$). However, when considering only the sample of students who returned for the follow up test nine weeks later, the

1
2
3 experimental group rated their confidence higher than the control group (3.6 ± 0.8 vs 2.8 ± 0.6 ;
4
5 $P = 0.031$; Cohen's $d = 1.1$ [0.1 - 2.0]; Figure 3). There was resultantly a significant but
6
7 moderate increase in reported confidence between pre- and post- test occasions in the
8
9 experimental group ($P = 0.024$; Cohen's $d = 0.5$ [0.04 - 1.0]). The control group did not report a
10
11 change in confidence between pre and post testing ($P = 0.596$; Figure 3). When answers to the
12
13 individual aspects of confidence were considered, the effects seen could be attributed principally
14
15 to differences in both confidence visualizing three-dimensional structures, and confidence
16
17 identifying structures on a live horse. These aspects of confidence improved from pre to post
18
19 exercise in the experimental group (3D structures: $P = 0.016$, Cohen's $d = 0.3$ [0.13 - 0.8]; Live
20
21 horse: $P = 0.025$, Cohen's $d = 0.5$ [0.1 - 1.0]), but not the control group ($P > 0.05$). This resulted
22
23 in significant differences between groups at the post exercise test occasion (3D, $P = 0.011$;
24
25 Cohen's $d = 0.34$ [0.5 - 1.3]; Live horse, $P = 0.041$, Cohen's $d = 1.1$ [0.2 - 2.1]). Conversely,
26
27 confidence in identifying structures on images did not change between test occasions, for either
28
29 group ($P > 0.05$).
30
31
32
33
34
35
36

37 After nine weeks, student confidence within the experimental group fell to within pre exercise
38
39 levels (Figure 3). 'Medium term' confidence of Experimental and Control groups was not
40
41 significantly different from pre exercise ($P = 0.343$, $P = 0.475$), or immediate post exercise ($P =$
42
43 0.148 , $P = 0.832$) values. Additionally, there was no difference between the reported confidence
44
45 of the two groups at this time point ($P = 0.457$; Figure 3).
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Enjoyment

Reported enjoyment of the learning exercise did not significantly differ between Experimental and Control groups ($P = 0.294$; Figure 4). Both Experimental and Control groups seemed to enjoy the exercise with 71 % and 62 % of students respectively selecting that they found the exercise “enjoyable”. Only 8 % of students that used the model horse responded that they did not enjoy the exercise; this proportion was higher (but not significantly so) in the Control group (19 %).

Students in the Experimental group suggested that “Geoff” was a useful addition to anatomy learning resources. 60% of students identified the model horse as “very useful” and a further 30% as “useful”. Despite this, when asked, no student from either group had used the model horse between the two testing sessions in their own time. Students who returned for follow up testing after nine weeks were asked to rate how useful the exercise was for improving long term anatomical knowledge. There was no significant difference in response to this question between the two groups ($P = 0.801$).

Gender influence on survey responses

Gender was found to influence participant self-reported confidence levels, but no other survey responses. In the whole cohort of students, mean self-reported confidence prior to the learning intervention differed between male and female participants (Figure 3). Females reported a confidence score of 2.6 ± 0.7 compared to 3.2 ± 0.6 in male participants ($P = 0.035$; Cohen’s $d = 0.8$ [0.1 – 1.6]). This was predominantly due to a higher initial confidence in visualization of 3D structures in the male participants (3.3 ± 0.9 versus 2.5 ± 0.7 ; $P = 0.013$, Cohen’s $d = 1.1$ [0.3 –

1.9)). Female participant confidence increased however after the learning activity ($P = 0.018$, Cohen's $d = 0.5$ [$0.1 - 0.9$]), whereas this was not the case for males ($P > 0.05$). This was particularly evident in the experimental group ($P = 0.016$, Cohen's $d = 0.6$ [$0.03 - 1.2$]) and not in the control group ($P > 0.05$). Again, this increase in confidence was predominantly attributable to female experimental group participants in the domain of visualizing 3D structures ($P = 0.019$, Cohen's $d = 0.8$ [$0.2 - 1.30$])

When considering only data from the cohort that returned nine weeks later there were no significant gender differences in pre- exercise confidence levels (Figure 3; $P > 0.05$). However after completing the exercise, males reported significantly higher confidence levels in visualizing 3D structures than females ($P = 0.031$, Cohen's $d = 1.1$ [$0.1 - 2.1$]). This gender difference was found in the control group only, where both general "confidence" levels, as well as confidence in 3D visualization were significantly higher in male participants with a strong effect size (Confidence: $P = 0.03$, Cohen's $d = 3.0$ [$1.1 - 5.0$]; 3D: $P = 0.015$, Cohen's $d = 3.2$ [$1.2 - 5.1$]). The significant pre-post increase in confidence levels in female experimental group participants was still present when considering only this returning cohort of students ($P = 0.045$).

DISCUSSION

Topographical Anatomy Performance

The results of this study highlight that the use of a painted model horse did not significantly enhance the topographical anatomy knowledge in veterinary medicine students when compared to a control study method. The model was however at least as beneficial as the textbook, since

1
2
3 student test scores post learning exercise were similar regardless of study method. The lack of
4
5 difference between Experimental and Control groups falls in contrast to a similar study of
6
7 comparable design (Braid et al., 2012) which found 16 % higher performance in students who
8
9 had used a model horse wearing an anatomical rug depicting anatomical structures compared to
10
11 those studying using textbooks alone. On initial inspection the learning exercises in the two
12
13 studies appear similar – using color and simple illustration on a life size model to aid learning.
14
15 However a key difference is that the “Anatorug” (Braid et al., 2012) was interactive, allowing
16
17 students to stick labels to the rug. This simple addition of an “active” component to the resource
18
19 may account for the apparent greater efficacy of the “Anatorug” over the painted model in the
20
21 current study. Indeed, literature shows that anatomical models requiring interaction from the
22
23 student have superior outcomes to traditional methods (e.g., Preece et al., 2013). The importance
24
25 of an active component in resource design is also illustrated by the success of other popular
26
27 modern methods for supplementing traditional anatomy teaching methods, e.g., clay modelling
28
29 (Krontiris-Litowitz, 2003; Motoike et al., 2009; Oh et al., 2009; Estevez et al., 2010; DeHoff et
30
31 al., 2011) and body painting (McLachlan and Regan De Bere, 2004; McMenamin, 2008; Finn et
32
33 al., 2011) and is likely one of the reasons why cadaver dissection is still considered, despite its
34
35 challenges, a “Gold standard” teaching methodology for many.
36
37
38
39
40
41
42
43

44
45 Another reason for the lack of improved performance in students in the Experimental group may
46
47 be the integrated nature of the test questions used for evaluation. Questions tested clinical
48
49 anatomy, topographical anatomy, as well as basic anatomical factual knowledge. The test as a
50
51 result required students to bring their existing knowledge, and combine it with topographical and
52
53 surface anatomy, on a live animal, in a clinical context. Factor analysis of the test was unable to
54
55
56
57
58
59
60

1
2
3 separate these constructs for in depth analysis. Future studies could look to modify the
4
5 assessment rubric to explicitly assess these elements separately and to facilitate better
6
7 exploration of how “Geoff” or similar models enrich teaching and learning outcomes.
8
9

10 11 12 Medium-term retention 13

14 Medium-term test scores were collected since immediate testing after a learning intervention
15
16 may not appropriately measure true learning (i.e. committal to long term memory), and instead
17
18 may better reflect short term recall from working memory. When a learner engages with a
19
20 resource, information initially passes into working memory (Miller, 1956). The working memory
21
22 provides temporary storage and processing of ‘to be learned’ information. Here, information is
23
24 packaged into manageable pieces, for integration with existing knowledge into long term
25
26 memory (Baddeley, 1997; Young et al., 2014). Working memory has a finite capacity when
27
28 dealing with new information (Miller, 1956) both in terms of amount and duration. Individuals
29
30 also have differing abilities to process information in working memory and these differences can
31
32 affect performance (Meinz and Hambrick, 2010). Long term memory however is a knowledge
33
34 base with, theoretically, a relatively unlimited capacity, and so once information is transferred to
35
36 long term memory, individual differences are of lesser consequence (Sweller, 2016).
37
38
39
40
41
42
43

44 To encode information into long-term memory, students usually try to understand new
45
46 information by processing it at a “deeper” level (Craig and Lockhart, 1972), by thinking about its
47
48 meaning and linking it to existing knowledge (elaborative rehearsal). Students who rely on rote
49
50 memorization (maintenance rehearsal), are instead trying to keep the information for longer in
51
52 their working memory. Learners can therefore influence the effectiveness of transferring
53
54
55
56
57
58
59
60

1
2
3 information from working to long term memory and instructional strategies that promote
4
5 successful processing in working memory will enhance learning and understanding (Khalil et al.,
6
7 2005). This phenomenon forms the basis of Cognitive Load Theory which suggests that effective
8
9 learning will occur when instructional conditions are aligned with a learner's cognitive
10
11 architecture (Sweller, 1988; Sweller and Chandler, 1991, 1994; Sweller, 1999; Paas et al., 2003).
12
13 Instructional design that produces extraneous cognitive load (e.g. through inappropriate
14
15 instructional design) can overload working memory and inhibit successful organization and
16
17 transfer of schemata to long term memory (Sweller, 2016; Pickering, 2014).
18
19
20
21
22
23

24 Mean test scores increased between test occasions, for both groups in this study, suggesting that
25
26 students were successful in transferring knowledge from that committed to working memory into
27
28 longer term memory. Further analysis however suggests that there may be differences in
29
30 outcomes of students using the physical model horse versus a textbook in terms of effective
31
32 learning. All participants using "Geoff" appeared to consistently successfully transfer
33
34 information into longer term memory: the strong and significant correlation between initial and
35
36 medium-term test scores in the Experimental group indicates that follow-up performance was
37
38 predictable and maintained across all individuals. The equal benefit between students in this
39
40 group infers that conditions (student motivation/ learning strategies, and instructional design)
41
42 encouraged successful learning to take place. There was no correlation between initial and
43
44 medium-term scores however in the Control group, highlighting inconsistent performance gains,
45
46 or perhaps specific gains only for some individuals. One interpretation of this is that using a
47
48 textbook may impose extraneous cognitive load, and therefore individuals with poorer working
49
50 memory may have underperformed at the initial testing session (and conversely those with good
51
52
53
54
55
56
57
58
59
60

working memory may have over performed). An alternative might be that individual study strategies were more varied in the control group, with some students using the text book relying on rote memorization instead of deeper learning strategies. Future studies should aim to explore this hypothesis, and consider working memory capacity, and student study strategies as variables in studies that evaluate educational interventions.

It is plausible that some test-retest effect may have accounted for some of the improvement seen at the later test occasion (Bjork, 1988; Karpicke and Blunt, 2011; Roediger and Butler, 2011), in particular since the form of assessment was highly valid (using a live horse) and potentially a rich learning experience in itself (oral examination) – despite the lack of immediate feedback.

General student performance

The purpose of this study was not to teach topographical anatomy as if encountered for the first time; moreover, the resources were used as a revision aid for third year veterinary students at the start of their clinical years of teaching. It is striking therefore that neither the experimental nor the control group succeeded in achieving an average score of above 50% in this study, highlighting how challenging many students find such material. This is significantly below the level of knowledge expected at this stage in the course and shows that, as has been documented in the medical and veterinary professions (Waterston and Stewart, 2005; Fitzgerald et al., 2008; Bhangu et al., 2010; Braid et al., 2012; Preece et al., 2013), the anatomy knowledge of students is often disappointing. This study took place in the first week of term, so one explanation for poor performance is that students may have forgotten some of their anatomy teaching over the extended summer period. The test may have been too challenging, though was designed to

1
2
3 reflect the learning objectives covered and already assessed during the first two years of the
4
5 course. The specific study task presented to students may also have been too large a volume to
6
7 cover during the time allocated (despite this being a revision exercise). This is substantiated by
8
9 free comments from one student asking for more time and two students who thought repeating
10
11 the exercise would be useful. It would be interesting to study the influence of increased time
12
13 with the resource, and over a prolonged period of study. Interaction at earlier stages of the
14
15 course may also provide better knowledge improvements (Sugand et al., 2010), as may the
16
17 provision of stricter guidance during teaching sessions; clear guidance has previously resulted in
18
19 enhanced learning over the self-directed method utilized here (Kooloos et al., 2012).
20
21
22
23
24
25

26 It requires significant cognitive effort to achieve a thorough knowledge of anatomical structures
27
28 and their relationships (Moxham et al., 2011). General notes made by the examiners in this study
29
30 found that a common theme across the student body was that students could name and identify
31
32 structures, however they failed to apply their anatomical knowledge correctly to the live horse.
33
34 For example, many appreciated the shape and orientation of the equine caecum, however they
35
36 were unsure on which side of the horse this lay. This is illustrated by the extremely low scores
37
38 for this question in every group on both occasions (Table S4, supplementary information). The
39
40 2D nature of textbooks makes them poor at illustrating 3D spatial relationships. The painted
41
42 model horse was created as a life size equine replica, with the aim of improving direct
43
44 application of knowledge to the clinical situation. As accurate as possible in the anatomy it
45
46 depicts, “Geoff” can only illustrate selected structures in each region. It is not truly a 3D model -
47
48 moreover a 2D representation painted onto a 3D structure, lacking multiple viewpoints of
49
50 individual structures, and without a real sense of depth. It is thought that the ability of the
51
52
53
54
55
56
57
58
59
60

1
2
3 student to control and alter the viewpoint is one of the major benefits of a 3D model for anatomy
4 learning (Garg et al., 2001) and this may account for the relative lack of efficacy of this model on
5 this occasion in aiding students form such spatial relationships. Moreover, Geoff does not strictly
6 depict surface landmarks, or allow palpation of surface structures, which might be key in
7 enabling students to form the link between 2D, 3D, topographical and surface anatomy.
8
9

10 11 12 13 14 15 16 17 Student Confidence

18
19 Following successful teaching, with improved competency student confidence should increase
20 (Butter et al., 2007). The data support the hypothesis that confidence would be improved by
21 using the model horse. Many studies have previously shown increases in student confidence
22 when learning from topographic models (McMenamin, 2008; Braid et al., 2012). However, the
23 increase in experimental group confidence alongside limited improvement in learning outcomes,
24 presents a conflict, and a cautious tale for researchers looking to use student-reported outcomes
25 alone in assessing the efficacy of learning interventions. It also highlights the phenomenon of
26 “unconscious incompetence” and the importance of developing metacognitive skills in students
27 on a professional program. In this study, student test scores were significantly below desirable,
28 and certainly below the threshold for a 'pass' grade in a summative examination at this
29 institution. Students yet considered themselves confident in their anatomy knowledge. Surgical
30 residents are one group that have previously been found to be falsely confident in their abilities
31 (Bowyer et al., 2015), and future studies should consider how educators can best promote
32 accurate and reflective self-assessment alongside self-directed teaching resources such as
33 “Geoff”. Interestingly, the confidence that the experimental group had gained during the first
34 testing had waned and returned to pre-learning levels by the second test occasion, suggesting that
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 the confidence gained from using “Geoff” was a short term emotion and that without the
4
5 concomitant enhancement in knowledge, this confidence did eventually return to match the
6
7 student’s own ability.
8
9

10 11 12 Gender Influence 13

14
15 Observed effects of gender on test scores were minimal, with the exception of question five and
16
17 three. Question five asked the student to identify the patella ligaments on the live horse, and was
18
19 answered better by male participants. The patella ligaments are palpable on the cranial aspect of
20
21 the stifle joint which is a highly sensitive region for some horses. Palpation of these structures
22
23 can cause horses to become agitated or kick out with their hind limbs. Therefore it is possible
24
25 that the better performance of male participants on this question may relate to gender typical
26
27 attitudes and perceptions towards danger (Harris and Miller, 2000). Educators should potentially
28
29 be aware of such gender bias when using assessment forms which integrate ‘hidden curriculum’
30
31 (Hafferty, 1998; Roder and May, 2017) elements, or non-assessed skills such as animal handling,
32
33 alongside assessment of basic or clinical science outcomes. Question three required the student
34
35 to indicate the location of the lumbar vertebrae. Females in the experimental group performed
36
37 better on this question, though the reason for this is unclear. The lumbar vertebrae in the model
38
39 were one of the few structures that were painted white against a dark background. It is possible
40
41 that such monochromatic contrast may favor female learning preferences. Color is known to
42
43 influence both attention (Pan, 2010) and the way in which learners remember both words and
44
45 pictures (Smilek et al., 2002; Spence et al., 2006). There is further evidence to suggest that the
46
47 context in which color is used is also an important feature of whether color aids memory. Colors
48
49 that are congruent with the image or word they are associated with have a positive impact on
50
51
52
53
54
55
56
57
58
59
60

memory, however incongruent colors do not (Boyatzis and Varghese, 1999; Olurinola et al., 2015). Contrast is also a factor that can influence the effectiveness of color on memory (Dzulkifli and Mustafar, 2013). Information on how gender may influence the effect of color on attention and memory however is negligible, and further work is recommended to allow educators to fully understand the effects of color on learning in the diverse student body.

As well as appearing to influence some aspects of anatomy learning, gender was also found to influence participant confidence, in particular in the area of 3D visualization of structures. Increases in confidence pre-post learning activity were most notable in female participants within the experimental group suggesting that use of “Geoff” was particularly beneficial for this subset of the student body. Spatial abilities have previously been linked to topographical anatomy performance (Rochford, 1985; Guillot et al., 2007; Hoyeck et al., 2009). There are also well documented effects of gender on spatial ability, with males consistently outperforming females on spatial ability tests (Maccoby and Jacklin, 1974; Peters et al., 2007; Langlois et al., 2013). Models such as “Geoff” may be fundamentally useful for providing female students with additional support for learning three dimensional anatomy. Providing additional resources and choice for a diverse student body may therefore enhance student outcomes, as well as offering improved inclusivity.

Student Enjoyment

In order for students to actively engage in a learning session and get the maximum from it, student motivation, engagement and enjoyment is essential (Cake, 2006). Learning from a textbook is far less stimulating than active learning and it was interesting that student enjoyment

of the task was not significantly different between the two groups in this study. This may be because all students enjoyed the test experience (with the live horse), and this part of the session influenced their questionnaire response - though this interpretation is based on the authors observations and was not directly assessed. A further possibility is a “volunteering effect” whereby the students in this study, who elected to participate, were excited about the learning opportunity; this intrinsic motivation may have resulted in an enjoyable experience for all students. This could be evaluated further once the model horse is embedded within the formal curriculum.

Despite the lack of statistical significance, a far greater number of students responded negatively to the Control exercise than the Experimental exercise. The Experimental group also stated overwhelmingly that “Geoff” was a useful addition to the anatomy resources at the university (90%). This reports similar findings to other studies (Kinnison et al., 2009; Preece et al., 2013) where models, in particular those that can be used for self-directed study, were looked upon favorably by students who appreciated the tactile and interactive nature of models.

Limitations of the study

A pre-test was not included in this study, as the learning situation and the students involved fit the criteria laid out by Hartley and Davies (1976), suggesting a pre-test could alter the post-test knowledge outcome potentially biasing results. In particular, the test situation using a live horse, could itself have been a rich learning experience, thus conducting this twice (pre and post intervention) would have been problematic. Instead a control group was used, alongside true random allocation of students to each study group. Though a difference in baseline knowledge

cannot be ruled out between the two groups, there is no reason to believe this the case, with both groups containing similar numbers of students from the two different year groups, and stating similar levels of self-reported confidence in equine anatomy. However, the results of this study highlight that individual differences such as knowledge, but also working memory capacity, student learning preferences, and spatial ability may be important measures to consider in future evaluations of this and other anatomy teaching interventions.

The sample size for this study was small, and of a self-selecting population. This has the potential to increase the likelihood of a type II error occurring; however, for the main non-significant result (total test score for experimental versus control groups) a power analysis suggests that at 80 % power, a sample size of > 700 would have been required to observe a significant difference. This is larger than the population of students available for study, and the sample appears sufficient on this occasion. Nevertheless, results by gender in particular should be interpreted with caution, given the relatively low proportion of male students in the student body. It is possible that voluntary recruitment of participants (and the relatively low uptake of this opportunity) may have biased the study participants toward very high, or very low achieving students. This is likely to represent the more highly motivated portion of the student body, and therefore the results should be interpreted in this context.

This study considered medium term effects on knowledge gain, which are often neglected in studies evaluating learning interventions. Retention rates were at 49 % which though resulting in significantly lower numbers of participants at the second test occasion, is still respectable for a voluntary teaching session. One reason for the low retention may be that oral exams make it

1
2
3 difficult to blind the assessors to student identity. Students may have chosen not to return due to
4
5 this 'exposed' style of assessment. The low numbers of participants returning should be noted
6
7 when interpreting the results: in assessing differences between study groups at the immediate test
8
9 stage, data is presented for both the full cohort of participants and the reduced cohort for balance.
10
11
12
13

14 CONCLUSION

15
16
17 With novel anatomical teaching methods in high demand to provide alternatives and adjuncts to
18
19 cadaver dissections for student learning, "Geoff" provides an exciting innovation for teaching
20
21 veterinary students. The results of this study showed that the painted model horse did not provide
22
23 significantly better test scores than the textbook when used for anatomy learning under the
24
25 conditions applied in this project. As hypothesized however, the test results were similar
26
27 suggesting the model was at least as useful as a textbook alone for learning topographical
28
29 anatomy. There were apparent gender effects that suggest that more study into the influence of
30
31 gender on learning outcomes of a range of teaching modes would be a useful addition to the
32
33 field. Students clearly found "Geoff" useful and engaging, yet not necessarily more enjoyable
34
35 than traditional methods. As such this addition to the anatomy teaching resources will form a
36
37 useful complementary tool alongside traditional teaching methods, and further research may help
38
39 find the area and timing in the course in which the model can best be utilized.
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

ACKNOWLEDGEMENTS

The authors would like to thank Chris Kench for his permission to use his photographs of “Geoff” as the basis for Figure 1 in this article.

NOTES ON CONTRIBUTORS

EDWARD BIETZK, B.VET. MED., was a final year veterinary medicine student at the Royal Veterinary College, University of London, North Mymms, Hatfield, Hertfordshire, United Kingdom. He is now working as a veterinary surgeon in a busy small animal practice in Bedfordshire.

RENATE WELLER, D.V.M., Ph.D., M.Sc., M.R.C.V.S., F.H.E.A., is a professor in diagnostic imaging in the Department of Clinical Science and Services at the Royal Veterinary College, University of London, North Mymms, Hatfield, Hertfordshire, United Kingdom. She splits her time between biomechanical and educational research, teaching in the B.Vet.Med. program and running the equine diagnostic imaging clinical service. She has recently been made a National Teaching Fellow of the Higher Education Academy.

VICKY SIMONS, B.VET.MED., was a final year veterinary medicine student at the Royal Veterinary College, University of London, North Mymms, Hatfield, Hertfordshire, United Kingdom. She is now a small animal veterinary surgeon working in practice in Wiltshire.

SARAH B. CHANNON, B.Sc. (Hons), Ph.D., M.Sc., F.H.E.A., is a senior lecturer in veterinary anatomy in the Department of Comparative Biomedical Sciences at the Royal Veterinary College, University of London, London, United Kingdom. She teaches anatomy to students in the B.Vet.Med. Degree and other programs, and her research interests are in functional anatomy, biomechanics, and education.

LITERATURE CITED

AAVME. 2018. Association of American Veterinary Medical Colleges. Competency-Based Veterinary Education: CBVE Framework. 1st Ed. Washington, DC: Association of American Veterinary Medical Colleges. 20 p. URL: <http://aavmc.org/data/files/cbve-publication-1-framework.pdf> [accessed 26 April 2018].

Azer SA, Eizenberg N. 2007. Do we need dissection in an integrated problem-based learning medical course? Perceptions of first-and second-year students. *Surg Radiol Anat* 29:173–180.

Azer SA. 2013. The place of surface anatomy in the medical literature and undergraduate anatomy textbooks. *Anat Sci Educ* 6:415–432.

Backhouse M, Fitzpatrick M, Hutchinson J, Thandi CS, Keenan ID. 2017. Improvements in anatomy knowledge when utilizing a novel cyclical “observe-reflect-draw-edit-repeat” learning process. *Anat Sci Educ* 10:7–22.

Baddeley AD. 1997. *Human Memory: Theory and Practice*. 1st Rev. Ed. Hove, East Sussex, UK: Psychology Press Ltd., Publishers. 423 p.

Baryla E, Shelley G, Trainor W. 2012. Transforming rubrics using factor analysis. *Pract Assess Res Eval* 17:1–7.

1
2
3 Beavers AS, Lounsbury JW, Richards JK, Huck SW, Skolits GJ, Esquivel SL. 2013. Practical
4 considerations for using exploratory factor analysis in educational research. *Pract Assess Res*
5
6
7 Eval 18:1–13.
8
9

10
11
12 Bhangu A, Boutefnouchet T, Yong X, Abrahams P, Joplin R. 2010. A three-year prospective
13 longitudinal cohort study of medical students' attitudes toward anatomy teaching and their career
14 aspirations. *Anat Sci Educ* 3:184–190.
15
16
17
18

19
20
21 Bjork RA. 1988. Retrieval practice and the maintenance of knowledge. In: Gruneberg MM,
22 Morris PE, Sykes RN (Editors). *Practical Aspects of Memory: Current Research and Issues*.
23
24 Volume 1. *Memory in Everyday Life*. 1st Ed. Oxford, UK: John Wiley & Sons. p 396–401.
25
26
27
28

29
30
31 Bowyer MW, Shackelford SA, Garofalo E, Pugh K, Mackenzie CF. 2015. Perception does not
32 equal reality for resident vascular trauma skills. *J Surg Res* 198:280–288.
33
34
35
36

37
38 Boyatzis CJ, Varghese R. 1994. Children's emotional associations with colors. *J Gen Psychol*
39 155:77–85.
40
41
42
43

44
45 Braid F, Williams SB, Weller R. 2012. Design and validation of a novel learning tool, the
46 “Anato-Rug,” for teaching equine topographical anatomy. *Anat Sci Educ* 5:256–263.
47
48
49

50
51 Brenton H, Hernandez J, Bello F, Strutton P, Purkayastha S, Firth T, Darzi A. 2007. Using
52 multimedia and Web3D to enhance anatomy teaching. *Comput Educ* 49:32–53.
53
54
55
56
57
58
59
60

1
2
3
4
5 Butter J, Grant TH, Egan M, Kaye M, Wayne DB, Carrión-Carire V, McGaghie WC. 2007. Does
6
7 ultrasound training boost Year 1 medical student competence and confidence when learning
8
9 abdominal examination? Med Educ 41:843–848.
10
11
12

13
14 Cake MA. 2006. Deep dissection: Motivating students beyond rote learning in veterinary
15
16 anatomy. J Vet Med Educ 33:266–271.
17
18
19
20

21
22 Chan LK, Cheng MM. 2011. An analysis of the educational value of low-fidelity anatomy
23
24 models as external representations. Anat Sci Educ 4:256–263.
25
26
27

28
29 Chen D, Zhang Q, Deng J, Cai Y, Huang J, Li F, Xiong K. 2018. A shortage of cadavers: The
30
31 predicament of regional anatomy education in mainland China. Anat Sci Educ 11:397–402.
32
33
34

35
36 Cho MJ, Hwang YI. 2013. Students' perception of anatomy education at a Korean medical
37
38 college with respect to time and contents. Anat Cell Biol 46:157–162.
39
40
41

42
43 Choi-Lundberg DL, Low TF, Patman P, Turner P, Sinha SN. 2016. Medical student preferences
44
45 for self-directed study resources in gross anatomy. Anat Sci Educ 9:150–156.
46
47
48

49
50 Chou CL, Topp KS, Kruidering-Hall M. 2010. Integrating surface anatomy learning with clinical
51
52 skills training. Med Educ 44:1127–1128.
53
54
55
56
57
58
59
60

1
2
3 Craik FI, Lockhart RS. 1972. Levels of processing: A framework for memory research. *J Verb*
4
5 *Learn Verb Behav* 11:671–684.
6
7
8
9

10 Craig S, Tait N, Boers D, McAndrew D. 2010. Review of anatomy education in Australian and
11
12 New Zealand medical schools. *ANZ J Surg* 80:212–216.
13
14
15
16

17 D'Agostino RB. 1986. Tests for normal distribution. In: D'Agostino RB, Stephens MA (Editors).
18
19 *Goodness-Of-Fit Techniques*. 1st Ed. Boca Raton, FL: CRC Press. p 367–420.
20
21
22
23

24 Dale VH, Head SD, May SA. 2008. Students' first impressions of a new, more integrated
25
26 curriculum with increased self-directed study time. In: Gourlay L, Saxby-Smith S (Editors).
27
28 *Excellence in Teaching Conference 2008 Annual Proceedings*. 1st Ed. London, UK: King's
29
30 Learning Institute, King's College London. p 21–33.
31
32
33
34

35 DeHoff ME, Clark KL, Meganathan K. 2011. Learning outcomes and studentperceived value of
36
37 clay modeling and cat dissection in undergraduate human anatomy and physiology. *Adv Physiol*
38
39 *Educ* 35:68–75.
40
41
42
43

44 Drake RL. 2002. Meeting the challenge: The future of the anatomical sciences in medical school
45
46 curricula. *Anat Rec* 269:68.
47
48
49
50

51 Drake RL, McBride JM, Lachman N, Pawlina W. 2009. Medical education in the anatomical
52
53 sciences: The winds of change continue to blow. *Anat Sci Educ* 2:253–259.
54
55
56
57
58
59
60

1
2
3
4
5 Dzulkipli MA, Mustafar MF. 2013. The influence of colour on memory performance: A
6 review. Malays J Med Sci 20:3–9.
7
8
9

10
11
12 Entwistle N, Tait H. 1990. Approaches to learning, evaluations of teaching, and preferences for
13 contrasting academic environments. High Educ 19:169–194.
14
15
16

17
18
19 Estevez ME, Lindgren KA, Bergethon PR. 2010. A novel three-dimensional tool for teaching
20 human neuroanatomy. Anat Sci Educ 3:309–317.
21
22
23

24
25
26 Finn GM, McLachlan JC. 2010. A qualitative study of student responses to body painting. Anat
27 Sci Educ 3:33–38.
28
29
30

31
32
33 Finn GM, White PM, Abdelbagi I. 2011. The impact of color and role on retention of
34 knowledge: A body-painting study within undergraduate medicine. Anat Sci Educ 4:311–317.
35
36
37

38
39
40 Fitzgerald JE, White MJ, Tang SW, Maxwell-Armstrong CA, James DK. 2008. Are we teaching
41 sufficient anatomy at medical school? The opinions of newly qualified doctors. Clin Anat
42 21:718–724.
43
44
45

46
47
48 Garg AX, Norman GR, Eva KW, Spero L, Sharan S. 2002. Is there any real virtue of virtual
49 reality?: The minor role of multiple orientations in learning anatomy from computers. Acad Med
50 77:S97–S99.
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5 Guillot A, Champely S, Batier C, Thiriet P, Collet C. 2007. Relationship between spatial
6 abilities, mental rotation and functional anatomy learning. *Adv Health Sci Educ Theory Pract*
7
8 12:491–507.
9
10

11
12
13
14 Hafferty FW. 1998. Beyond curriculum reform: Confronting medicine's hidden curriculum. *Acad*
15
16 *Med* 73:403–407.
17
18

19
20
21 Hall ER, Davis RC, Weller R, Powney S, Williams SB. 2013. Doing dissections differently: A
22 structured, peer-assisted learning approach to maximizing learning in dissections. *Anat Sci Educ*
23
24 6:56–66.
25
26
27

28
29
30 Halliday N, O'Donoghue D, Klump KE, Thompson B. 2015. Human structure in six and one-half
31 weeks: One approach to providing foundational anatomical competency in an era of compressed
32
33 medical school anatomy curricula. *Anat Sci Educ* 8:149–157.
34
35
36

37
38
39 Harris MB, Miller KC. 2000. Gender and perceptions of danger. *Sex Roles* 43:843–863.
40
41
42

43
44 Hart LA, Wood MW, Weng HY. 2005. Mainstreaming alternatives in veterinary medical
45 education: Resource development and curricular reform. *J Vet Med Educ* 32:473–480.
46
47
48

49
50
51 Hartley J, Davies IK. 1976. Preinstructional strategies: The role of pretests, behavioral
52 objectives, overviews and advance organizers. *Rev Educ Res* 46:239–265.
53
54
55
56
57
58
59
60

1
2
3
4
5 Jamniczky HA, Cotton D, Paget M, Ramji Q, Lenz R, McLaughlin K, Coderre S, Ma, IW. 2017.

6
7
8 Cognitive load imposed by ultrasound-facilitated teaching does not adversely affect gross
9
10 anatomy learning outcomes. *Anat Sci Educ* 10:144–151.
11
12

13
14 Jittivadhna K, Ruenwongsa P, Panijpan B. 2009. Hand-held model of a sarcomere to illustrate
15
16 the sliding filament mechanism in muscle contraction. *Adv Physiol Educ* 33:297–301.
17
18
19

20
21 Jittivadhna K, Ruenwongsa P, Panijpan B. 2010. Beyond textbook illustrations: Hand-held
22
23 models of ordered DNA and protein structures as 3D supplements to enhance student learning of
24
25 helical biopolymers. *Biochem Mol Biol Educ* 38:359–364.
26
27
28
29

30
31 Karpicke JD, Blunt JR. 2011. Retrieval practice produces more learning than elaborative
32
33 studying with concept mapping. *Science* 331:772–775.
34
35
36

37
38 Khalil MK, Paas F, Johnson TE, Payer AF. 2005. Interactive and dynamic visualization in
39
40 teaching and learning of anatomy: A cognitive load perspective. *Anat Rec* 286B:8–14.
41
42
43

44
45 Khalil MK, Mansour MM, Wilhite DR. 2010. Evaluation of cognitive loads imposed by
46
47 traditional paper-based and innovative computer-based instructional strategies. *J Vet Med Educ*
48
49 37:353–357.
50
51
52
53
54
55
56
57
58
59
60

1
2
3 Kinnison T, Forrest ND, Frean SP, Baillie S. 2009. Teaching bovine abdominal anatomy: Use of
4 a haptic simulator. *Anat Sci Educ* 2:280–285.
5
6
7
8
9

10 Knowles MS. 1975. *Self-Directed Learning: A Guide for Learners and Teachers*. 1st Ed. New
11 York: Association Press. 135 p.
12
13
14
15
16

17 König HE, Liebich HG (Editors). 2007. *Veterinary Anatomy of Domestic Mammals: Textbook*
18 *and Colour Atlas*. 4th Ed. Stuttgart, Germany: Schattauer GmbH Verlag. 787 p.
19
20
21
22
23

24 Kooloos JG, de Waal Malefijt MC, Ruiter DJ, Vorstenbosch MA. 2012. Loosely-guided, self-
25 directed learning versus strictly-guided, station-based learning in gross anatomy laboratory
26 sessions. *Anat Sci Educ* 5:340–346.
27
28
29
30
31
32

33 Krontiris-Litowitz J. 2003. Using manipulatives to improve learning in the undergraduate
34 neurophysiology curriculum. *Adv Physiol Educ* 27:109–119.
35
36
37
38
39

40 Kumar AM, Murtaugh R, Brown D, Ballas T, Clancy E, Patronek G. 2001. Client donation
41 program for acquiring dogs and cats to teach veterinary gross anatomy. *J Vet Med Educ* 28:73–
42 77.
43
44
45
46
47
48

49 Langlois J, Wells GA, Lecourtois M, Bergeron G, Yetisir E, Martin M. 2013. Sex differences in
50 spatial abilities of medical graduates entering residency programs. *Anat Sci Educ* 6:368–375.
51
52
53
54
55
56
57
58
59
60

1
2
3 Latorre R, Rodríguez MJ. 2007. In search of clinical truths: Equine and comparative studies of
4 anatomy. Equine Vet J 39:263–268.
5
6
7
8
9

10 Latorre RM, García-Sanz MP, Moreno M, Hernández F, Gil F, López O, Ayala MD, Ramírez G,
11 Vázquez JM, Arencibia A, Henry RW. 2007. How useful is plastination in learning anatomy? J
12 Vet Med Educ 34:172–176.
13
14
15
16
17
18

19 Latorre R, Bainbridge D, Tavernor A, López Albors O. 2016. Plastination in anatomy learning:
20 An experience at Cambridge University. J Vet Med Educ 43:226–234.
21
22
23
24
25

26 Maccoby EE, Jacklin CN. 1974. The Psychology of Sex Differences. 1st Ed. Stanford, CA:
27 Stanford University Press. 634 p.
28
29
30
31
32

33 Martinsen S, Jukes, N. 2005. Towards a humane veterinary education. J Vet Med Educ 32:454–
34 460.
35
36
37
38
39

40 McLachlan JC, Regan De Bere S. 2004. How we teach anatomy without cadavers. Clin Teach
41 1:49–52.
42
43
44
45
46

47 McLachlan JC, Patten D. 2006. Anatomy teaching: Ghosts of the past, present and future. Med
48 Educ 40:243–253.
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 McMenamin PG. 2008. Body painting as a tool in clinical anatomy teaching. *Anat Sci Educ*
4
5 1:139–144.
6
7

8
9
10 Meinz E, Hambrick D. 2010. Deliberate practice is necessary but not sufficient to explain
11 individual differences in piano sight-reading skill: The role of working memory capacity.
12
13
14
15 *Psychol Sci* 21:914–919.
16
17

18
19 Miller GA. 1956. The magical number seven, plus or minus two: Some limits on our capacity for
20 processing information. *Psychol Rev* 63:81–97.
21
22
23

24
25
26 Motoike HK, O’Kane RL, Lenchner E, Haspel C. 2009. Clay modeling as a method to learn
27 human muscles: A community college study. *Anat Sci Educ* 2:19–23.
28
29
30

31
32
33 Moxham BJ, Shaw H, Crowson R, Plaisant O. 2011. The future of clinical anatomy. *Eur J Anat*
34 15:29–46.
35
36
37

38
39
40 Murphy RJ, Gray SA, Straja SR, Bogert MC. 2004. Student learning preferences and teaching
41 implications. *J Dent Educ* 68:859–866.
42
43
44

45
46
47 Nanjundaiah K, Chowdapurkar S. 2012. Body-painting: A tool which can be used to teach
48 surface anatomy. *J Clin Diagn Res* 6:1405–1408.
49
50
51

52
53
54

55
56
57

58
59
60

1
2
3 Oh CS, Kim JY, Choe YH. 2009. Learning of cross-sectional anatomy using clay models. *Anat*
4
5 *Sci Educ* 2:156–159.
6
7
8
9

10 Olurinola O, Tayo O. 2015. Colour in learning: Its effect on the retention rate of graduate
11
12 students. *J Educ Pract* 6:1–6.
13
14
15
16

17 O'Mahony SM, Horgan M, O'Flynn S, O'Tuathaigh CM. 2016. Association between learning
18
19 style preferences and anatomy assessment outcomes in graduate-entry and undergraduate
20
21 medical students. *Anat Sci Educ* 9:391–399.
22
23
24
25

26 Op Den Akker JW, Bohnen A, Oudegeest WJ, Hillen B. 2002. Giving color to a new curriculum:
27
28 Bodypaint as a tool in medical education. *Clin Anat* 15:356–362.
29
30
31
32

33 Paas F, Renkl A, Sweller J. 2003. Cognitive load theory and instructional design: Recent
34
35 developments. *Educ Psychol* 38:1–4.
36
37
38
39

40 Pan Y. 2010. Attentional capture by working memory contents. *Can J Exp Psych* 64:124–128.
41
42
43
44

45 Pawlina W, Drake RL. 2015. New (or not-so-new) tricks for old dogs: Ultrasound imaging in
46
47 anatomy laboratories. *Anat Sci Educ* 8:195–196.
48
49
50
51

52 Peters M, Manning JT, Reimers S. 2007. The effects of sex, sexual orientation, and digit ratio
53
54 (2D:4D) on mental rotation performance. *Arch Sex Behav* 36:251–260.
55
56
57
58
59
60

Peterson RA. 1994. A meta-analysis of Cronbach's coefficient alpha. *J Consum Res* 21:381–391.

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

Preece D, Williams SB, Lam R, Weller R. 2013. “Let's get physical”: Advantages of a physical model over 3D computer models and textbooks in learning imaging anatomy. *Anat Sci Educ* 6:216–224.

RCVS. 2014. Royal College of Veterinary Surgeons. RCVS Day One Competencies. 1st Ed. London, UK: Royal College of Veterinary Surgeons. 12 p. URL: <https://www.rcvs.org.uk/document-library/day-one-competences/1day-one-competences-updated-26-march-2014.pdf> [accessed 26 April 2018].

Rochford K. 1985. Spatial learning disabilities and underachievement among university anatomy students. *Med Educ* 19:13–26.

Roder CA, May SA. 2017. The hidden curriculum of veterinary education: Mediators and moderators of its effects. *J Vet Med Educ* 44:542–551.

Roediger HL III, Butler AC. 2011. The critical role of retrieval practice in long-term retention. *Trends Cognit Sci* 15:20–27.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Rosenshine B, Meister C. 1992. The use of scaffolds for teaching higher-level cognitive strategies. *Educ Leader* 49:26–33.

Russell WMS, Burch RL. 1959. *The Principles of Humane Experimental Technique*. 1st Ed. London, UK: Methuen & Co. Ltd. 238 p.

Santos JR. 1999. Cronbach's alpha: A tool for assessing the reliability of scales. *J Extension* 37:1–5.

Senos R, Ribeiro MS, Martins Kde S, Pereira LV, Mattos MF, Kfoury Júnior JR, Rodrigues MR. 2015. Acceptance of the bodypainting as supportive method to learn the surface locomotor apparatus anatomy of the horse. *Folia Morphol (Warsz)* 74:503–507.

Smilek D, Dixon MJ, Cudahy C, Merikle PM. 2002. Synesthetic color experiences influence memory. *Psychol Sci* 13:548–552.

Smith CF, Tollemache N, Covill D, Johnston M. 2018. Take away body parts! An investigation into the use of 3DDprinted anatomical models in undergraduate anatomy education. *Anat Sci Educ* 11:44–53.

Sokal RR, Rohlf FJ. 1981. *Biometry: The Principles and Practice of Statistics in Biological Research*. 2nd Ed. San Francisco, CA: W.H. Freeman & Company. 859 p.

1
2
3
4
5 Spence I, Wong P, Rusan, M, Rastegar N. 2006. How color enhances visual memory for natural
6 scenes. Psychol Sci 17:1–6.
7
8
9

10
11
12 Sugand K, Abrahams P, Khurana A. 2010. The anatomy of anatomy: A review for its
13 modernization. Anat Sci Educ 3:83–93.
14
15
16
17

18
19 Sullivan GM, Artino AR Jr. 2013. Analyzing and interpreting data from Likert-type scales. J
20 Grad Med Educ 5:541–542.
21
22
23
24

25
26 Sweller J. 1988. Cognitive load during problem solving. Cognit Sci 12:257–285.
27
28
29

30
31 Sweller J. 1999. Instructional Design in Technical Areas. 1st Ed. Melbourne, Australia:
32 Australian Council for Educational Research (ACER) Press. 192 p.
33
34
35
36

37
38 Sweller J. 2016. Working memory, long-term memory, and instructional design. J Appl Res
39 Mem Cognit 5:360–367.
40
41
42
43

44
45 Sweller J, Chandler P. 1991. Evidence for cognitive load theory. Cognit Instruct 8:351–362.
46
47
48

49 Sweller J, Chandler P. 1994. Why some material is difficult to learn. Cognit Instruct 12:185–233.
50
51
52
53
54
55
56
57
58
59
60

1
2
3 Topping DB. 2014. Gross anatomy videos: Student satisfaction, usage, and effect on student
4 performance in a condensed curriculum. *Anat Sci Educ* 7:273–279.
5
6
7
8
9

10 Hernández-Torrano D, Ali S, Chan CK. 2017. First year medical students' learning style
11 preferences and their correlation with performance in different subjects within the medical
12 course. *BMC Med Educ* 17:131.
13
14
15
16
17
18

19 van Merriënboer JJ, Sweller J. 2010. Cognitive load theory in health professional education:
20 Design principles and strategies. *Med Educ* 44:85–93.
21
22
23
24
25

26 Waters JR, Van Meter P, Perrotti W, Drogo S, Cyr RJ. 2005. Cat dissection vs. sculpting human
27 structures in clay: An analysis of two approaches to undergraduate human anatomy laboratory
28 education. *Adv Physiol Educ* 29:27–34.
29
30
31
32
33
34

35 Waterston SW, Stewart IJ. 2005. Survey of clinicians' attitudes to the anatomical teaching and
36 knowledge of medical students. *ClinAnat* 18:380–384.
37
38
39
40
41

42 Watmough SD, O'Sullivan H, Taylor DC. 2010. Graduates from a reformed undergraduate
43 medical curriculum based on Tomorrow's Doctors evaluate the effectiveness of their curriculum
44 6 years after graduation through interviews. *BMC Med Educ* 10:65.
45
46
47
48
49
50

51 Woodmansey D. 2016. Vet schools defend using dead imported dogs for dissection. *Vet Times*.
52 9 December 2016. Veterinary Business Development Ltd., Peterborough, UK. URL:
53
54
55
56
57
58
59
60

1
2
3 <https://www.vettimes.co.uk/news/vet-schools-defend-using-dead-imported-dogs-for-dissection/>
4
5 [accessed 11 January 2018].
6
7
8
9

10 Yammine K, Violato C. 2015. The effectiveness of physical models in teaching anatomy: A
11
12 meta-analysis of comparative studies. *Adv Health Sci Educ Theory Pract* 21:883–895.
13
14
15

16
17 Young JQ, Van Merriënboer J, Durning S, Ten Cate O. 2014. Cognitive load theory:
18
19 Implications for medical education: AMEE guide no. 86. *Med Teach* 36:371–384.
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

FIGURE LEGENDS

Figure 1: Images of “Geoff”, the painted model horse for topographical anatomy teaching. A, Left lateral view; B, Cranial view; C, Right lateral view showing superficial musculature, thoracic and abdominal organs, muscles and tendons of the distal limb; D, Caudal view.

Figure 2: Test performance for Experimental (“Geoff”) and Control groups at each testing session, by gender. There was a significant increase in test score for the experimental group between test occasions. There were no significant gender effects on overall test score. Pale bars indicate first testing session; dark bars show testing session nine weeks later. Plain bars indicate Control group, patterned bars indicate Experimental group. Grey bars indicate data for all students combined; blue and purple bars indicate data for male and female students respectively. Mean and standard deviation shown.

Figure 3: Mean student confidence scores pre and post initial testing and post second testing session, by gender. After teaching males reported higher confidence than females. Female confidence increased from pre to post test in the experimental group only. After nine weeks, males in the control group reported significantly higher confidence than females in the same group. Plain bars indicate Control group and patterned bars indicate the Experimental group. Blue and purple bars indicate data for male and female students respectively. Shade of bar (pale through dark) indicates time point of assessment. A five-point Likert scale where 1 = very poor, 5 = excellent. Error bars represent standard deviation.

Figure 4: Mean student enjoyment ratings for both experimental and control groups, by gender, from the first testing session. Female students in the experimental group enjoyed the session more than those in the control group. The reverse trend was true for male students. Plain bars indicate Control group and patterned bars indicate the Experimental group. Blue and purple bars indicate data for male and female students respectively. A five-point Likert scale where 1 = not enjoyable at all, 5 = very enjoyable. Error bars represent standard deviation.

Table 1. Factor loading scores from principal component analysis of survey instrument.

Item	Factor 1	Factor 2
Confidence identifying anatomical structures on images	-0.055	0.953
Confidence visualising anatomy in 3D	0.069	0.896
Confidence identifying anatomical structures on a live horse	-0.003	0.928
Usefulness	0.773	0.153
Enjoyment	1.018	-0.139

The factor loading scores relates to the relationship between the variable (survey item) and its underlying factor. The higher the factor loading score, the stronger the item's relationship to the underlying factor.

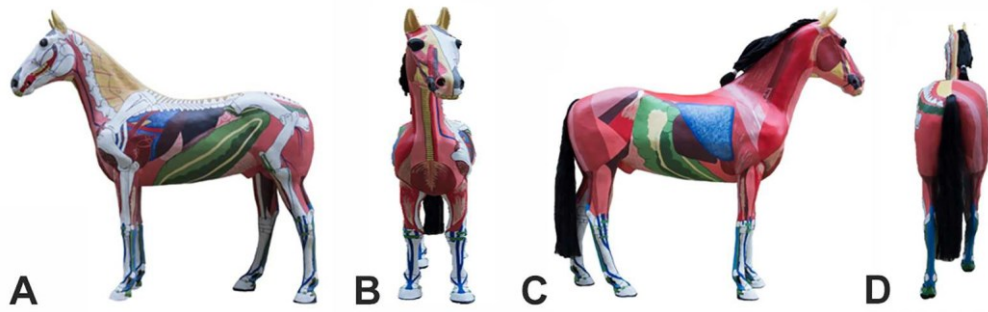


Figure 1: Images of "Geoff", the painted model horse for topographical anatomy teaching. A, Left lateral view; B, Cranial view; C, Right lateral view showing superficial musculature, thoracic and abdominal organs, muscles and tendons of the distal limb; D, Caudal view.

194x63mm (300 x 300 DPI)

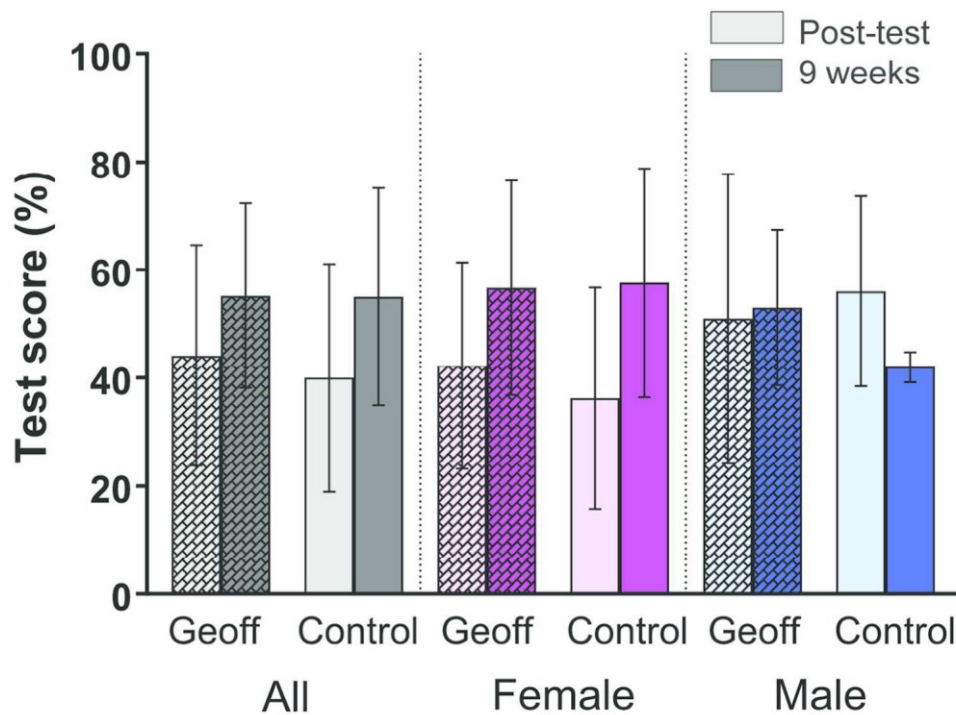


Figure 2: Test performance for Experimental ("Geoff") and Control groups at each testing session, by gender. There was a significant increase in test score for the experimental group between test occasions. There were no significant gender effects on overall test score. Pale bars indicate first testing session; dark bars show testing session nine weeks later. Plain bars indicate Control group, patterned bars indicate Experimental group. Grey bars indicate data for all students combined; blue and purple bars indicate data for male and female students respectively. Mean and standard deviation shown.

127x97mm (300 x 300 DPI)

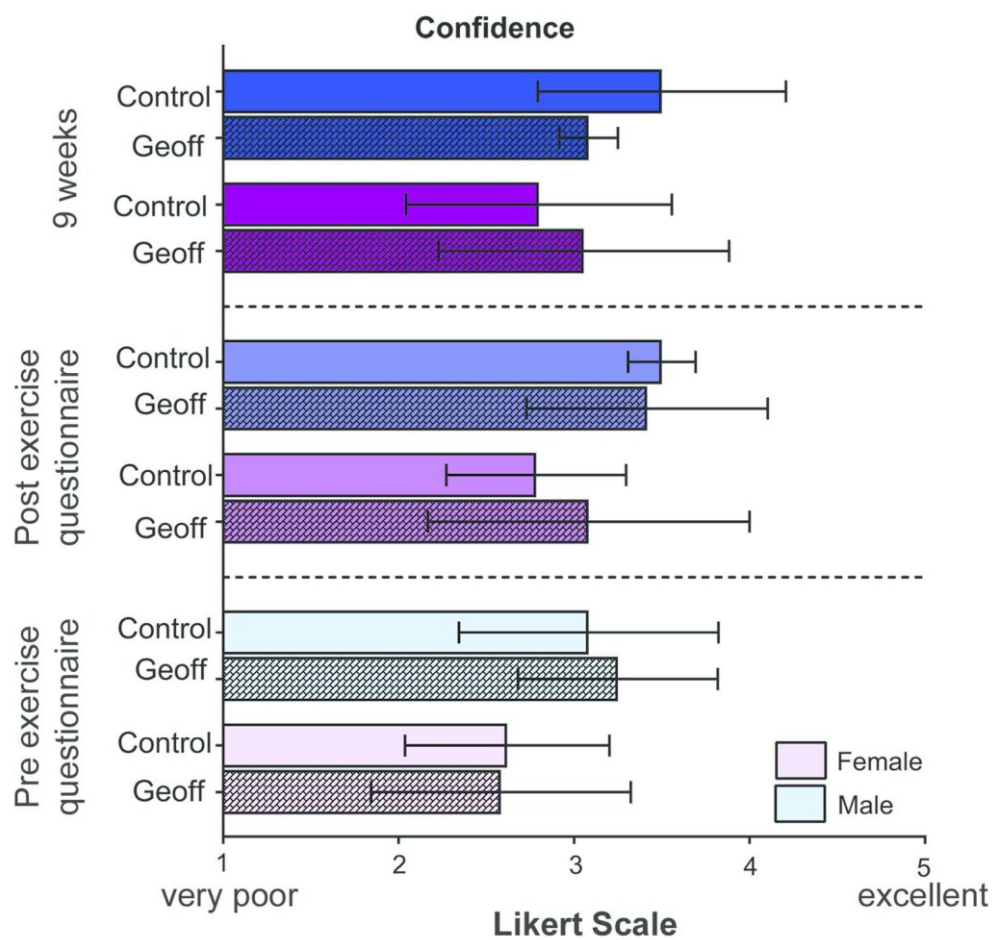


Figure 3: Mean student confidence scores pre and post initial testing and post second testing session, by gender. After teaching males reported higher confidence than females. Female confidence increased from pre to post test in the experimental group only. After nine weeks, males in the control group reported significantly higher confidence than females in the same group. Plain bars indicate Control group and patterned bars indicate the Experimental group. Blue and purple bars indicate data for male and female students respectively. Shade of bar (pale through dark) indicates time point of assessment. A five-point Likert scale where 1 = very poor, 5 = excellent. Error bars represent standard deviation.

128x122mm (300 x 300 DPI)

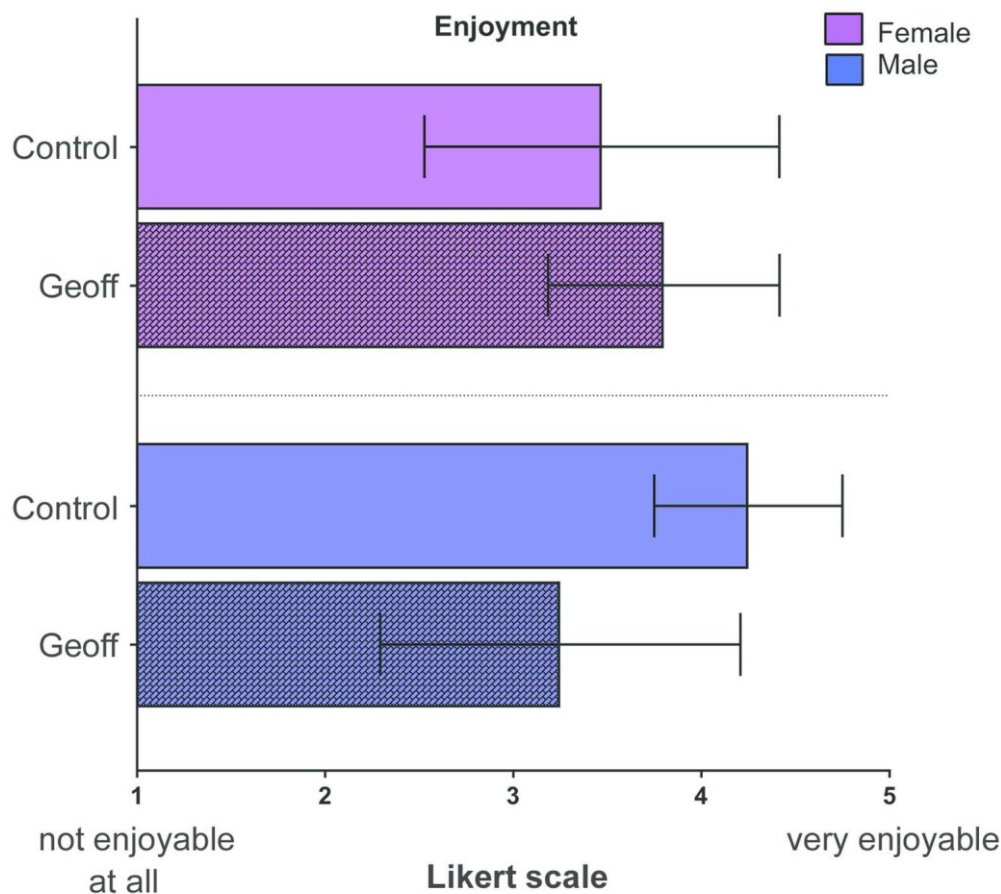


Figure 4: Mean student enjoyment ratings for both experimental and control groups, by gender, from the first testing session. Female students in the experimental group enjoyed the session more than those in the control group. The reverse trend was true for male students. Plain bars indicate Control group and patterned bars indicate the Experimental group. Blue and purple bars indicate data for male and female students respectively. A five-point Likert scale where 1 = not enjoyable at all, 5 = very enjoyable. Error bars represent standard deviation.

128x119mm (300 x 300 DPI)