ASE-16-0081.R2

Research Report

Effects of image-based and text-based active learning exercises on student examination performance in a musculoskeletal anatomy course

M. Melissa Gross[1], Mary C. Wright[2], Olivia S. Anderson[3]

[1]Department of Movement Science, School of Kinesiology, University of Michigan, Ann Arbor, Michigan

[2]Harriet W. Sheridan Center for Teaching and Learning, Brown University, Providence, Rhode Island

[3]Department of Nutritional Sciences, School of Public Health, University of Michigan, Ann Arbor, Michigan.

Running title: Effects of image- and text-based active learning on performance

Correspondence to: Dr. M. Melissa Gross, Department of Movement Science, University of Michigan, 401 Washtenaw Avenue, Ann Arbor, MI 48109-2214. USA. E-mail: mgross@umich.edu

Grant sponsor: Echo^{360}, Inc.

ABSTRACT

Research on the benefits of visual learning has relied primarily on lecture-based pedagogy, but the potential benefits of combining active learning strategies with visual and verbal materials on learning anatomy has not yet been explored. In this study, the differential effects of text-based and image-based active learning exercises on examination performance were investigated in a functional anatomy course. Each class session was punctuated with an average of 12 text-based and image-based active learning exercises. Participation data from 231 students were compared with their examination performance on 262 questions associated with the in-class exercises. Students also rated the helpfulness and difficulty of the in-class exercises on a survey. Participation in the active learning exercises was positively correlated with examination performance ($r = 0.63$, $P < 0.001$). When controlling for other key demographics (gender, underrepresented minority status) and prior grade point average, participation in the image-based exercises was significantly correlated with performance on examination questions associated with image-based exercises ($P < 0.001$) and text-based exercises ($P < 0.01$), while participation in text-based exercises was not. Additionally, students reported that the active learning exercises were helpful for seeing images of key ideas (94%) and clarifying key course concepts (80%), and that the image-based exercises were significantly less demanding, less hard and required less effort than text-based exercises ($P < 0.05$). The findings confirm the positive effect of using images and active learning strategies on student learning, and suggest that integrating them may be especially beneficial for learning anatomy.

Keywords: Gross anatomy education, musculoskeletal anatomy, undergraduate education, kinesiology education; active learning; visual learning; image-based learning, text-based learning, functional anatomy; effectiveness of anatomy education

INTRODUCTION

Anatomical Sciences in Kinesiology Education

The study of human anatomy is fundamental to the field of kinesiology. Kinesiology is the multi-disciplinary study of movement and physical activity that draws from many disciplinary areas and includes multiple sub-disciplinary areas (Gill, 2007). A description of the field is provided on the National Academy of Kinesiology website: "Kinesiology is a common name for college and university academic departments that include many specialized areas of study in which the causes, processes, and consequences, and contexts of physical activity are examined from different perspectives" (NAK, 2016).

The American Association of Kinesiology has established a curriculum for kinesiology programs that includes the scientific foundations of physical activity as a core element (Chodzko-Zajko, 2014; AKA, 2015a). Anatomy is fundamental to the scientific foundation core and has been specifically identified as a required course in many kinesiology programs (AKA, 2015b). The number of majors in kinesiology programs is growing, increasing by more than 50% from 2003-2008 (Wojciechowska, 2010; AKA, 2015c). This growth has been attributed to the popularity of kinesiology as preparation for allied health graduate programs (Thomas, 2014).

The Movement Science program in the School of Kinesiology at the University of Michigan emphasizes the study of human movement from biological and behavioral perspectives across the lifespan. A fundamental requirement for Movement Science students is to understand the relationship between body structure and body movement, and how the relationship changes as a function of age, injury or disease. The Human Musculoskeletal Anatomy is a required course that is designed to provide students with an in-depth knowledge of human musculoskeletal anatomy and its relationship to body movement. Students learn bone names and landmarks, joint structure and kinematic characteristics, and the names, attachments and functions of the major muscles in the human body.

Two important goals of the Human Musculoskeletal Anatomy course are to help students learn anatomy deeply rather than just memorize the material for examinations, and to help students form mental models of musculoskeletal anatomy that persist so that they can apply anatomical knowledge in more advanced courses as they progress in their program. Using images and active learning strategies may help instructors achieve these instructional goals.

Learning with Images in Traditional Environments

Most university introductory science, technology, engineering and mathematics (STEM) courses are lecture-based (Astin and Astin, 1992; FSSE, 2015), with instructors typically using slide presentation software to supplement the lecture (Hurtado et al., 2012), often to add a visual dimension to the presentation. Like in other STEM disciplines, images are routinely included in lecture slides in anatomy courses (Vorstenbosch et al., 2013). Imaging anatomy has been integrated into gross anatomy instruction worldwide, including in traditional lectures (Grignon et al., 2016). Because anatomy is the study of the physical structures of the body and their spatial relationships, the use of images in anatomical education is of particular importance (Vorstenbosch et al., 2013; 2014). Anatomy students need to develop a mental model of body structures and their spatial relationships, and images used by anatomy instructors may help students achieve that goal (Gyselinck, 1996; Butcher, 2006).

At one time, it was thought that a mix of verbal and visual formats would allow students to better encode information in long-term memory (Paivio and Csapo, 1973; Clark and Paivio, 1991), thus increasing learning. Surprisingly, there is mixed evidence for this approach (Mayer and Massa, 2003; Kollöffel, 2012). On the one hand, Mayer (2011) suggests that the presentation of both visual and verbal cues – rather than verbal cues alone – promotes generative processing, or the building of connections between new and prior knowledge. This idea, known as the multimedia principle, supports the use of both text-based and image-based materials during instruction. Multiple studies (Mayer and Gallini, 1990; Mayer and Anderson, 1992; Mayer et al., 1996; Moreno and Mayer, 2002) support this principle by showing that students' performance on transfer tests was higher when presented with both visual and verbal materials. Butcher (2006) demonstrated that students improved their mental models more when learning anatomy with text and diagrams compared to text-only materials, providing support for the multimedia principle in anatomical education.

However, other research indicates that in certain instructional contexts, there can be limited learning gains from the use of visuals, possibly because they increase cognitive load. For example, Tangen et al. (2011) compared student quiz scores following three different types of lectures – with visuals relevant to the content, with visuals not relevant, and with text-based bullet points – and found no significant difference in accuracy among them. Lin and Atkinson (2011) found similar results, albeit for animated vs. static graphics. Tangen and colleagues suggest that the finding of no difference between text and congruent image conditions might be due to insufficient cognitive resources to allow the elaborative processing needed. They hypothesize that "there was simply not enough time or resources for elaborative processing to occur" (Tangen et al., 2011). It may be that the deeper learning expected with image-based slides could emerge if learners were given an opportunity to interact with the material more extensively during a class session to enable the needed elaborative processing.

Cognitive processes associated with text and anatomical images may also differ. In a gross anatomy course, medical students seemed to use different cognitive processes when responding to assessment questions posed as labeled images or answer lists (Vorstenbosch et al., 2014). The type of anatomical image used can also affect item difficulty (Vorstenbosch et al., 2013). Hunt (1978) found that item difficulty increased when questions were presented to medical students in the original data format (image) compared to radiological reports (text), and suggested that students be given more time to complete assessments with images. Butcher (2006) suggested that the most effective type of anatomical image for learning may depend on the level of cognitive support required by the learner.

Learning with Images in Active Learning Environments

Much of the previous work on visual and verbal formats have investigated the effects on learning in traditional, knowledge transmission settings (Butcher, 2006; Bartholomé and Bromme, 2009) or in assessments of learning (Hunt, 1978; Crisp and Sweiry, 2006; Berends and van Lieshout, 2009; Vorstenbosch et al., 2013, 2014; Holland et al., 2015). In this study, however, the effects of verbal and visual formats on learning were investigated in an active learning setting, in which students had time to think about their answers and discuss their answers with peers. Active learning exercises provide instructors with opportunities to support student learning in the classroom. Recently, the use of audience response system technology (ARS; "clickers") and other active learning exercises have been shown to increase learning and performance in STEM courses (Knight and Wood, 2005; Bruff, 2009; Barr, 2014; Freeman et al., 2014). Although active learning practices vary among STEM disciplines, instructors in undergraduate STEM courses reported using clickers regularly to engage students in actively answering questions or interacting with peers, and to assess students' conceptual understanding (Hora and Holden, 2013). Audience response systems have been used in anatomy courses with mixed results, with some studies finding little or no change in examination scores with ARS use (Stein et al., 2006; Hoyt et al., 2010) and others finding a positive correlation between ARS and examination scores (Alexander et al., 2009). Regardless of the potential influence on examination performance, anatomy students perceived ARS as valuable for their learning, reporting that ARS was a useful feedback tool (Alexander et al., 2009), ARS reviews were more beneficial than traditional, lecture-style reviews (Stein et al., 2006), and ARS tended to enhance understanding of the material (Hoyt et al., 2010).

An increasing number of instructors have been using these active learning pedagogies, also sometimes in conjunction with images. For instance, peer instruction is an active learning practice that asks students to respond to Conceptest questions before and after discussing them with classmates (Mazur, 1997; Crouch and Mazur, 2001; Mazur, 2009). Peer Instruction and Conceptests have been used in STEM courses including physics (Crouch and Mazur, 2001; Lasry et al., 2008), chemistry (Donovan, 2008; Miller et al., 2014), mathematics (Schlatter, 2002), and geosciences (McConnell et al., 2006). Conceptests are short conceptual questions that are designed to reveal common misconceptions (Fagen et al. 2002) and to assess student learning at the comprehension, application or analysis level of Bloom's taxonomy (Bloom, 1956) rather than just memorization (McConnell et al., 2006). For example, students might be asked to interpret or analyze an image, rather than just identify structures within the image. In mathematics, Conceptest questions have been designed specifically to develop students' ability to think in three dimensions (Schlatter, 2002). However, the role of images in enhancing student learning in STEM classrooms – particularly in an active learning environment – is unclear.

Therefore, research on visuals may be helpfully contextualized by two key factors. First, visual learning formats may be more beneficial in course contexts requiring the interpretation of complex visual material (Kotze et al., 2012). In classes such as these visual presentation formats may allow for increased cognitive processing. Given the visual nature of much STEM course content, including anatomy, pictorial representations can be especially helpful for students who struggle with spatial visualization in such courses (Kolloffel, 2012).

Second, research on learning through visual-verbal presentations has relied primarily on lecture-based pedagogy, not accounting for the processing time students may need to make sense of both pieces of information (Lin and Atkinson, 2011; Tangen et al., 2011). Given increasing emphasis on the importance of active learning pedagogy in STEM classes (Freeman et al., 2014), research that examines the role of visuals in active learning is needed. Indeed, it remains unknown the degree to which particular modes of active learning are beneficial to student learning (Freeman et al., 2014). In the context of anatomical education, it would be useful to know whether verbal or visual formats are more effective for student learning in active learning settings.

Purpose of the Study

The purpose of this study was to investigate the potential differential effects of text- and image-based active learning exercises on students' examination performances in the context of an undergraduate musculoskeletal anatomy course. Specifically, the hypothesis that participation in image-based activities will result in higher examination scores than participation in text-based activities was tested. Previous studies have treated active learning as a separate condition from visual- and text-based learning (e.g., Bockoven, 2004; Tangen et al., 2011), but their integration was investigated in this study.

The key research questions examined in this study were: (1) Is student participation in text- and image-based active learning exercises, as measured by course-based analytics data, positively correlated with examination performance? If so, which is a better predictor of examination performance and do any of these trends vary by key students' demographics?; (2) Do students report that text- and image-based active learning exercises assist their learning? If so, do students perceive image-based exercises to require more cognitive load than text-based engagement and do students perceive both types of active learning to be equally useful for examination preparation?

METHODS

Course Content and Structure

Human Musculoskeletal Anatomy is a three credit, second-year course offered in the School of Kinesiology at the University of Michigan. The course is required for students in the Movement Science program in the School of Kinesiology and is open to students from other programs at the

University of Michigan, with enrollments in the course divided approximately 50% between majors and non-majors. Most commonly, students take the course as sophomores. Students attend bi-weekly 80-minute didactic lecture sessions (36 hours total), punctuated with active learning exercises. A systems approach is used to cover musculoskeletal anatomy and function, including the skeletal system (11 hours), the muscular system (15 hours) and musculoskeletal function (4 hours); remaining sessions are used for examinations. Screencasts of the lecture sessions with audio are recorded and made available to the students immediately following each session. To encourage active participation in class, students complete an online quiz based on assigned reading that is due before each lecture session begins. Students also complete an online quiz after each class session that reviews the lecture material. Students' grades in the course are based on four examinations (60%), online quizzes (10%), participation in active learning activities during lecture sessions (15%) and an independent project (10%). Assigned readings for the course are drawn from a textbook (Behnke, 2012). Examinations in the course are multiple choice format and emphasize knowledge of musculoskeletal structures and their relationships, and concepts, rather than recognition of musculoskeletal structures.

Active Learning Exercises

Each class session in Human Musculoskeletal Anatomy consisted of lecture segments punctuated with active learning exercises in which students worked independently or in pair-share dyads to answer interactive questions asked through interactive presentation software (Echo^{360} Inc., Dulles, VA). In this project, the interactive presentation software was used to provide both course content (i.e., slides) and active learning exercises, but other technology platforms or non-technology-based approaches could be used instead. Both text- and image-based active learning exercises were included in each class session (Figure 1). Text-based exercises consisted of a written question and multiple-choice answers. Image-based exercises consisted of either a question associated with a particular image (and multiple-choice answers) or an image map where students were asked to click on the image in order to answer the question. Individual course topics were practiced via a text- or image-based active learning question, but not both. Each 80-minute class session included between 4 and 25 active learning questions, with the proportion of image- and text-based exercises dependent upon the number of course topics covered in each session. The average number of active learning questions per class session was 12, with 75% of class sessions having 14 or fewer active learning exercises. A few class sessions (20%) were designed to reinforce muscle attachments and relationships, and these sessions had the most active learning questions (17 to 25). Active learning questions varied in difficulty according to Bloom's (1956) taxonomy with 70% at the knowledge level and 30% at the comprehension level. The proportion of comprehension-level questions was slightly greater for image-based (33%) than for text-based active learning exercises (23%).

In-Class Participation Data

Student participation in all learning exercises was tracked using analytics data embedded in the interactive lecture presentation software. The presentation software system collected information about participation in all exercises, so if a student entered in an answer for a question, the student was counted as having participated in the exercise. Each individual question was flagged by the instructor as either text- or image-based, so that participation for each category could be calculated. Additionally, demographic information including gender, underrepresented minority (URM) status, and cumulative grade point average (GPA) were obtained for each student from university registrar records. The study protocol was approved by the Institutional Review Board at the University of Michigan.

Overall, the analyses examined students' participation and performance in 259 in-class active learning exercises over the two terms, with roughly similar numbers of exercises in the fall (123) and winter (136). (There were 383 activities total, but only questions associated with examination questions were analyzed for this article.) Most of the exercises were repeated in the winter term, but some were deleted or replaced. Given the nature of the content taught in this course, a majority (60%) of these exercises were image-based, with comparable proportions in the fall (62%) and winter (59%).

Examination Performance Data

Each term, students took four examinations, all in multiple-choice format. The examination questions were designed to assess knowledge of musculoskeletal structure and function, and conceptual relationships between structure and function. Some of the multiple-choice questions referred to images, but most did not (80%). Of the multiple-choice questions referring to images, half of the questions used images to illustrate a body movement. Fewer than 2% of examination items required students to identify musculoskeletal structures on images; among these, only cross-sectional images were used for muscle identification.

Examinations included questions directly associated with both text- and image-based in-class active learning exercises, as well as questions not associated with active learning exercises (Appendix A). The level of difficulty of all examination questions was similar to the text-based and image-based learning activities. The examination questions were the same in each term except that an additional five examination questions were added in the winter term. Overall, there were 431 examination questions (213 in the fall and 218 in the winter), and a majority (61%) of assessment items corresponded with in-class exercises. Of the 262 examination questions assessing material that was reinforced through active learning, a slight majority (60%) was associated with image-based exercises. The percent correct and point serial correlation for each examination question were obtained from the university examination scoring service and were used to indicate examination item difficulty and discrimination (Holland et al., 2015).

Survey Data

Students were asked to participate in an end-of-semester survey, which asked them to reflect on the helpfulness of in-class exercises, the difficulty of the in-class learning activities, and how helpful the in-class learning activities were for preparing for examinations. First, students rated the helpfulness of the in-class exercises related to five aspects of the course on a five-point Likert scale (anchor items were 1 = strongly disagree and 5 = strongly agree). Specifically, students were asked to rate how helpful the activities were for clarifying key course concepts, collaborating with other students, increasing interest in the subject, seeing images of key ideas, and asking questions.

Second, students were asked to rate the difficulty of the in-class exercises. Survey questions about the difficulty of in-class learning activities were modeled on the instrument for measuring cognitive load described by Lin and Atkinson (2011), which itself was adapted from a modified version of the NASA Task Load Index (NASA-TLX), developed and validated by Hart and

Staveland (1988) and Hart (2006). Although bio-measures are sometimes used to measure cognitive load, they are challenging to use in a classroom context and subjective rating scales like the NASA-TLX are frequently used instead (Leppink and van den Heuvel, 2015). Paas and colleagues write that "although self-ratings may appear questionable, it has been demonstrated that people are quite capable of giving a numerical indication of their perceived mental burden" (Paas et al., 2003).

Students were asked three questions in which they rated their level of mental activity, how hard they had to work, and how much effort they invested. The set of questions was repeated for activities generally, for activities using images, and for activities using text. Students rated how much mental activity was required by responding to the question, "Thinking specifically of exercises using [activities generally/images/text], how much mental activity was required when participating in activities? That is, were [activities/activities with images/activities with text] easy or demanding?" using an eight-point Likert scale (anchor items were 1 = easy, 8 = demanding). Students rated how hard they had to work by responding to the question "Thinking specifically of exercises using [activities generally/images/text], how hard did you have to work to understand the contents of exercises?" using an eight-point scale (anchor items were 1 = not very hard, 8 = very hard). Finally, students rated how much effort they invested by responding to the question "Thinking specifically of exercises using [activities generally/images/text], how much effort did you have to invest in order to participate in exercises during class?" using an eight-point Likert scale (anchor items were 1 = low effort, 8 = high effort).

Scores for each difficulty question were averaged for each of the in-class exercise types. Then, mean scores across the three difficulty questions were averaged to create a composite difficulty score for each of the in-class exercise types. Thus, one composite difficulty score was created to estimate cognitive load for activities generally, one for image-based activities, and a third for text-based activities.

Finally, students were asked to rate the helpfulness of in-class activities for examination preparation. Again, the question was asked separately for activities generally, for activities using images, and for activities using text. For each question, students rated how helpful activities were on an eight-point Likert scale (anchor items were 1 = not at all helpful, 8 = very helpful).

Statistical Analysis

To test for differences in student demographics between semesters, an unpaired t-test was used to test for significant differences in GPA and a chi-square test was used to evaluate differences among gender and underrepresented minority (URM) status. An unpaired t-test was used to test for differences between image-based and text-based learning activities in student scores on the survey questions. A simple Pearson's correlation was used to examine the relationship between participation in active-learning exercises and student performance on examinations. The SPSS statistical package, version 22, (IBM Corp., Armonk, NY) was used for all statistical analyses.

Two linear multivariate regression analyses were used to look for correlations between performance on examination questions associated with in-class active learning exercises and student participation in these exercises, using SPSS statistical package. Linear multivariate regression was used to account also for student demographics (gender and URM status) and prior GPA (grades in prior courses prior to the term of interest), because they have been found to be key predictors of STEM performance and educational achievement generally (Tobias, 1990; Hazari et al., 2007; Bettinger, 2010; Ost, 2010). Examination performance was measured as accuracy, i.e., the percent of examination questions answered correctly out of the total possible. The dependent variable for one regression was performance on examination questions associated with image-based exercises, and the dependent variable for the second regression was performance on examination questions associated with text-based exercises.

RESULTS

Student Demographics

Enrollment in the course averaged 116 students (96 in the fall and 135 in the winter). The course was comprised of approximately half majors and non-majors (54% Movement Science and 46% other concentrations over the two terms), with a slightly greater proportion of majors in the winter (59%) compared to the fall (50%). Nearly two-thirds of students (64%) were female, and 11% were underrepresented minorities. Most commonly, students took the course as sophomores (49%). Overall, students enrolled in the course had a mean incoming GPA of 3.2 (±SD 0.46) for all prior courses. Comparing fall and winter term students, there was no statistically significant difference in GPA, gender composition or URM status. Analyses included all students who completed the class and engaged in active learning exercises via a laptop. (Six students responded via paper to in-class exercises, and these students' analytics were not included in the analysis.)

Participation in Active Learning Exercises and Examination Performance

Overall, participation rates were similar for both types of active learning exercises, with about three fourths of all students participating in text-based (75%) and image-based (74%) exercises. Individual students participated in more image-based (mean = 58 exercises, ±SD 14) than text-based (mean = 39 exercises, ±SD 11) exercises, likely because there were more image-based than text-based activities offered to students.

Overall, students correctly answered an average of 77% (±SD 15%) of the examination questions associated with in-class exercises. The validity of the examination items was good, with average point biserial values of 0.384 for examination questions associated with active learning exercises.

There was a positive and strong relationship between participation (measured by the mean number of active learning exercises) and examination performance (based on the mean number of associated examination questions answered correctly) ($r = 0.63$, $P < 0.001$). However, participation was only moderately correlated with examination performance on unrelated questions ($r = 0.32$, $P < 0.001$). As displayed in Figure 2, students in the top quartile of examination performance on exercise-related questions tended to also be at the top for participation rate. Similarly, students receiving the lowest 25% of examination grades tended to be below the median for participation rate.

Insert Figure 2 about here

When predicting performance on examination questions associated with image-based exercises, participation in image-based active learning exercises and prior GPA had a statistically significant impact (Table 1). Although it might be expected that prior GPA would have a significant relationship with examination performance, interestingly, image-based active learning had a stronger effect ($\beta = 0.51$, $P < 0.001$) compared to prior achievement ($\beta = 0.24$, $P < 0.001$) on examination questions associated with image-based exercises. There were no statistically significant differences for text-based participation, gender nor URM status.

Alternatively, when predicting performance on examination questions associated with text-based exercises, again participation in image-based active learning exercises and prior GPA had a statistically significant relationship (Table 1). Participation in image-based learning exercises had a stronger effect ($\beta = 0.33$; $P = 0.01$) than prior GPA ($\beta = 0.20$; $P = 0.002$) on examination questions associated with text-based exercises. Text-based participation, gender and URM status were again not significant.

Insert Table 1 about here

In summary, results for both regressions showed that participation in image-based learning exercises was associated with higher examination scores, while participation in text-based active learning was not. As might be expected, prior GPA was also associated with higher examination scores. Underrepresented minority status and gender were not significantly correlated with examination performance in this class, holding other covariates constant, suggesting that this benefit holds for a diversity of students.

Helpfulness and Difficulty of Active Learning Exercises

Response rate on the end-of-semester survey was 75% (N = 158). The proportions of survey respondents that were Movement Science majors (58%), female (66%), underrepresented minorities (9%), and sophomores (53%) were similar to the respective proportions in the overall class composition. Cronbach's alpha for the survey instrument used in this study was 0.94, signifying a high degree of internal consistency.

Overall, students found the active learning exercises to be very helpful to their learning experiences, particularly for seeing images of key ideas and clarifying course concepts. Nearly all students (94%) agreed or strongly agreed that the active learning exercises were helpful for "seeing images of key ideas in the course" (mean = 4.55, ±SD 0.58, on a five-point Likert scale). Most students (80%) agreed or strongly agreed that active learning exercises were helpful for "clarifying key course concepts" (mean = 4.11, ±SD 0.71). Student ratings for helpfulness of active learning exercises on other aspects of their learning, including "increasing my interest in the subject" (mean = 3.63, ±SD 0.89), "collaborating with other students" (mean = 3.26, ±SD 1.12), and "asking questions" (mean = 3.30, ±SD 1.16), were less highly rated but were all above the neutral midpoint, indicating that overall students perceived the active learning exercises to be helpful to their learning.

Overall, based on the composite score, students reported that the cognitive load demanded by the active learning exercises was neither easy nor demanding, i.e., a neutral rating on an eight-point Likert scale (mean score = 4.01, ±SD 1.71). However, students rated image-based (mean score =

3.88, ±SD 1.71) exercises as requiring significantly less load compared to text-based exercises (mean score = 4.35, ±SD 1.78) ($P < 0.001$). The results for each individual component of the composite score showed similar results, with students rating image-based activities as less demanding, less hard and requiring less effort than text-based activities (Table 2).

Insert Table 2 about here

Students reported that the in-class activities generally were quite useful for preparing for examinations (mean score = 6.04, ±SD 1.61, on an eight-point Likert scale). Students rated the image-based exercises (mean score = 5.90, ±SD 1.63) and text-based exercises (mean score = 5.94, ±SD 1.55) as similarly helpful in preparing for examinations.

DISCUSSION

In our first set of research questions, we asked if student participation in clicker-based exercises was positively associated with exam performance, if image- or text-based exercises were more strongly correlated, and if there were any differences by demographics. We found that student participation in active learning exercises in a musculoskeletal anatomy class was positively correlated with student performance on examinations. Further, participation in image-based learning exercises was more strongly correlated with examination performance than participation in text-based exercises, when controlling for key background characteristics. Additionally, there were no significant differences in these effects by demographics. These findings not only confirm the benefits of active learning strategies and the use of images for student learning that have been described in other studies, but also suggest that integrating images into active learning exercises may be particularly beneficial for all students' learning in anatomy courses.

Our second research question inquired if students perceived that image-based questions require more cognitive load than text-based engagement, and the degree to which students perceive both types of active learning to be useful. Like in other anatomy and kinesiology courses using active learning strategies (Alexander et al., 2009; Stein et al., 2006; Hoyt et al., 2010; Barr et al., 2014), students in this study had positive responses to the active learning exercises. The majority of students agreed that the in-class exercises were helpful for seeing images of key ideas and clarifying course concepts, and that participation in the learning exercises was quite useful for examination preparation. Similarly, geoscience and chemistry students in courses using Conceptests perceived a positive impact on their learning, reporting that the Conceptests were useful for helping them learn, and giving them the opportunity to test their understanding and discuss their answers with others (McConnell et al., 2006; Donovan, 2008). Considering the positive impact of active learning on examination scores compared to traditional lectures in STEM courses (Freeman et al., 2014) and the positive perceptions of students towards active learning exercises, the time required for active learning exercises during class sessions is merited.

Students found the exercises helpful but students rated image-based exercises as less demanding, less hard and requiring less effort than the text-based activities. We suggest that the images used in the active learning exercises reduced cognitive load in students with limited knowledge of anatomy by assisting in the visualization of complex, three-dimensional structures. Students were freed from having to produce visualizations of anatomy *de novo* and were instead able to connect new information to prior knowledge. In-class active learning exercises using images may have been particularly valuable in helping students discern critical features in an image, learning how to “see" important elements in the image, and beginning to construct the mental models that are necessary for more advanced anatomy applications. The use of visual materials along with verbal cues has been shown to be particularly effective for low-knowledge learners (Kalyuga et al., 1998, 2000; Mayer, 2011), perhaps because students with less experience in a subject require additional information before generative processing can occur. Human Musculoskeletal Anatomy is an introductory course, implying that, without knowing the anatomy background of each student, the majority of the students likely had low knowledge of the course content, even if they were otherwise good students (i.e., had a high GPA in other courses). Therefore, the use of visual material in the active learning exercises may have been particularly beneficial to student learning in this introductory course.

Findings of this study support previous research on the benefits of active learning strategies in anatomical education. Alexander et al. (2009) demonstrated a positive effect of ARS questions on student examination performance for medical students in an anatomy curriculum. In that study, students answered daily ARS questions written in the National Board of Medical Examiners® (NBME®) style at the end of laboratory exercises. Students’ cumulative ARS scores predicted examination scores on the anatomy written final examination ($r = 0.37$, $P < 0.001$) and NBME gross anatomy subject examination ($r = 0.53$, $P < 0.001$). Student participation in active learning exercises in this study also predicted examination performance, with participation in the active learning exercises ($r = 0.63$, $P < 0.001$) having an even stronger effect on examination performance than ARS scores had on the anatomy final written examination or NBME gross anatomy subject examination in the previous study. Taken together, both studies suggest that incorporating active learning exercises into daily classroom sessions improves student learning.

However, other studies suggest that use of ARS does not always result in improved examination scores. In one case, the active learning strategy was deployed just before the examination, serving as a check of student knowledge, rather than earlier in the learning process when the ARS questions may have been used by students to help develop anatomical knowledge (Stein et al., 2006). The contrast of our findings with these other studies suggests that it is important to embed active learning exercises in lectures as low-stakes formative assessments throughout the entire term. This idea is consistent with studies that have shown a large positive effect of repeated testing on student performance (Butler and Roediger, 2007; Karpicke and Roediger, 2008; Soderstrom et al., 2016), including in anatomy contexts (Logan et al., 2011; Dobson and Linderholm, 2015).

Most importantly, in contrast to other studies, the format of the active learning exercises in our study was explicitly distributed between image-based and text-based questions. Formulating the active learning exercises as images or as text may have helped the students develop different skills. Vorstenbosch et al. (2014) suggested that questions with images test students' ability to interpret and extract cues from visual information and questions without images test the quality of their mental images. Since learning anatomy requires the construction of mental models of anatomical structures and their spatial relationships, constructing active learning exercises with images may be particularly effective in helping students develop mental models. In this study, the examination questions were mostly textual multiple choice questions, likely testing the quality of the students' mental images. Thus, the finding that participation in image-based but not text-based exercises was positively correlated with examination performance (holding other variables constant) supports the importance of using image-based active learning exercises in learning anatomy.

The benefit of using images in active learning exercises for student learning was demonstrated in this study, but it is not clear which type of images might have been most helpful, or whether cues provided with the images were useful to support students' learning. Leppink and van den Heuvel (2015) proposed strategies to help optimize intrinsic cognitive load that suggest instructors gradually increase task complexity and task fidelity, and provide less instructional support as learning progresses. According to these strategies, instructors should select images for active learning exercises beginning with more simple and schematic images and then progressing to more complex and realistic images such as those derived from medical imaging modalities. Butcher (2006) directly addressed the issue of image complexity for learning, and demonstrated strong support for the benefit of simpler diagrams designed to highlight the representation of critical relationships, particularly for low knowledge learners. At a given level of image complexity, Fenesi et al. (2017) found that image quality itself seemed to have only a limited effect on anatomy learning, and the advantage of higher quality images did not apply with additional learning opportunities. The images in the active learning exercises used in this study varied in complexity, from schematic representations to radiological images; if the images had been systematically ordered by complexity or fidelity, or had been designed specifically to represent spatial or conceptual relationships, it is possible that the positive effect of the image-based active learning exercises might have been even stronger.

The images in the active learning exercises in this study also differed in the amount and type of cues provided. For example, some images in the active learning exercises were paired with visual cues such as labels or highlights or with instructional prompts to help guide students' attention and understanding, and others were not. The visual relationship between images and corresponding instructional text can affect learning, depending on whether the text is redundant with image representation or integrated with the image (Chandler and Sweller, 1991). Bartholomé and Bromme (2009) also investigated the impact of visual cues and prompting on learning with images and text, and found that the best learning results were achieved when minimal support was given. Interestingly, spatial ability interacted with prompting, so that learners with low spatial abilities benefited most from prompting and learners with high spatial abilities learned less with prompting. These findings suggest that instructors need to consider the learners' abilities and needs when selecting the most effective images for active learning exercises, and modify image complexity, image fidelity and instructional cueing accordingly. As Butcher (2006, p. 194) states, "It is not enough to simply encourage active engagement with learning materials; rather, good multimedia materials support the learner's effective performance of essential cognitive processes. Effective impact on learning requires that representations can successfully support the cognitive processes necessary for deep comprehension."

Limitations of the Study

One limitation of this study was the survey approach to measuring cognitive load, which relied on students' self-reports. Although indirect, subjective assessments are a common way to measure cognitive load (Paas et al., 2003; Leppink et al., 2013; Young et al., 2016), direct measures exist to assess this condition, ranging from physiological measures to experimental approaches (e.g., asking subjects to perform two tasks simultaneously) (Brünken et al., 2010). Future researchers may wish to utilize more sophisticated methodologies to measure cognitive load.

Further, the survey was conducted at the end of the term and asked students to rate the difficulty of the learning exercises that were embedded in multiple class sessions. It is not known if students rated the peak load, average load or cumulative load for each of the types of active learning exercises (Paas et al., 2003), or if their ratings of the learning exercises were influenced more by recent class sessions rather than by earlier ones. Thus, it may have been useful to ask students to complete the survey questions more frequently during the term to understand the relationship between their ratings at the end of the term and their ratings of load imposed by the learning exercises in a single class session as sampled throughout the term. Regardless, the students' ratings of load on all three measures were consistently lower for the image-based compared to the text-based exercises at the end of the term.

Many factors affect student learning, including those investigated in this study and others that were not. The multiple regression analyses in this study demonstrated a positive effect of image-based exercises on student learning compared to text-based exercises, holding other important variables (prior GPA, gender, URM status) constant. The adjusted R^2 values for the regression analyses were moderate (0.40 and 0.37), indicating that factors not included in the analyses were also important. These factors might have included study habits, participation in study groups, or non-cognitive characteristics. Further, the relative effects of image-based exercises on examination performance were also moderate as indicated by the β coefficients in the regression analyses (0.51 and 0.33). Even so, the relative effects of image-based exercises were greater than for prior GPA ($\beta = 0.24$ and 0.20), which has been shown to be a strong predictor of student learning outcomes in STEM courses (Tobias, 1990; Hazari et al., 2007; Bettinger, 2010; Ost, 2010). Although other factors that might have affected student performance on examinations were not included in the study, the finding that there was a differential effect of image-based and text-based exercises in an active learning environment on examination performance while accounting for other key factors suggests that visualization should be included as a factor to explain student learning in future studies of anatomical education practices.

Consistent with another study of ARS (Hoyt et al., 2010), we found more poorly performing students did not engage in the exercises at the same rate as more successful students. Although a significant portion of the grade (15%) was awarded for participation, it may be that a higher proportion would encourage greater engagement. However, in their study of active learning in a Human Anatomy and Physiology course, Cavanagh et al. (2016) found that student "buy-in" to active learning involves a complex web of factors, such as student–instructor relationships, inclusive pedagogy, and student perceptions of course content and learning activities.

CONCLUSIONS

Image-based and text-based active learning exercises were embedded in lectures in an undergraduate musculoskeletal anatomy course. Overall, participation in the active learning exercises was positively correlated with student examination performance. A differential effect of in-class exercise format on examination performance was observed, with a positive relationship for image-based exercises but not text-based exercises, holding other key factors constant. Cognitive load may have been lower with the image-based exercises in the active learning setting, with students reporting that the image-based exercises were less demanding, less hard and required less effort than the text-based activities. The findings suggest that active learning exercises using images positively impacts student learning when implemented in anatomy courses.

The potential advantages of visualizations combined with active learning strategies for student learning in anatomical education are not limited to embedding images in lecture presentations. In the context of medical education in the United Kingdom, Backhouse et al. (2017) demonstrated an improvement in test scores when students completed an interactive tutorial that emphasized visualizing anatomical structures by guiding students through an observe-reflect-draw-edit-repeat (ORDER) cycle of learning, compared to a control interactive tutorial in which students were presented with text and image descriptions of the same anatomy. The difference in test scores was not related to students' verbal or visual learning preference; this is perhaps not surprising because both groups were presented with images, and the preference for learning with visuals may not be accompanied by a greater ability to learn from visual information (Willingham et al., 2015). Rather, the results support the notion that interacting with images may be especially helpful to students as they are constructing mental representations of anatomical structures.

ACKNOWLEDGMENTS

Financial support for this project was provided by a grant from Echo^{360}, Inc. The authors have no conflicts of interest with the funding sponsor. Thanks to Dr. Anne Greenberg for analysis of student participation and examination performance data, Dr. Ronit Greenberg for assisting with literature reviews on visually-based learning and to Ms. Melinda Thompson for her assistance with data entry.

NOTES ON CONTRIBUTORS

M. MELISSA GROSS, Ph.D., is an Arthur F. Thurnau professor and associate professor in the Department of Movement Science in the School of Kinesiology at the University of Michigan in Ann Arbor, Michigan. She teaches anatomy and biomechanics to undergraduate students and directs research activities in the Behavioral Biomechanics Laboratory. Her research interest is musculoskeletal structure and biomechanics of movement behavior especially as related to bodily expression of emotion in health and disease.

MARY C. WRIGHT, Ph.D., is Director of the Harriet W. Sheridan Center for Teaching and Learning at Brown University in Providence, Rhode Island. Previously, she was Director of Assessment at the University of Michigan's Center for Research on Learning and Teaching. Her research interests include evaluation of teaching and learning innovations, educational development, graduate student professional development, and curricular assessment of student learning.

OLIVIA S. ANDERSON, Ph.D., R.D., is a clinical assistant professor in the Department of Nutritional Sciences in the School of Public Health at the University of Michigan in Ann Arbor, Michigan. Previously, she was a postdoctoral fellow at the University of Michigan's Center for Research on Learning and Teaching. She teaches undergraduate and graduate courses in nutrition as well as pedagogy and professional development. Her research interests include developing, implementing, and assessing pedagogy that creates an engaged learning environment for the

nutritional sciences students and studying early-life environmental exposures and the impact they have on metabolism and the epigenome.

LITERATURE CITED

AKA. 2015a. American Kinesiology Association. The undergraduate core in kinesiology. American Kinesiology Association, Champaign, IL. URL: http://www.americankinesiology.org/the-undergraduate-core-in-kinesiology [accessed 1 November 2016].

AKA. 2015b. American Kinesiology Association. AKA undergraduate core elements. Section one: The AKA core undergraduate elements. American Kinesiology Association, Champaign, IL. URL: http://www.americankinesiology.org/the-undergraduate-core-in-kinesiology/section-one-the-aka-undergraduate-core-elements [accessed 1 November 2016].

AKA. 2015c. American Kinesiology Association. White papers. Kinesiology on the move: One of the fastest growing (but often misunderstood) majors in academia. American Kinesiology Association, Champaign, IL. URL: http://www.americankinesiology.org/white-papers/white-papers/kinesiology-on-the-move--one-of-the-fastest-growing-but-often-misunderstood-majors-in-academia [accessed 1 November 2016].

Alexander CJ, Crescini WM, Juskewitch JE, Lachman N, Pawlina W. 2010. Assessing the integration of audience response system technology in teaching of anatomical sciences. Anat Sci Educ 2:160–166.

Astin AW, Astin HS. 1992. Undergraduate Science Education: The Impact of Different College Environments on the Educational Pipeline in the Sciences. 1st Ed. Los Angeles, CA: Higher

Education Research Institute, University of California Los Angeles. 377 p. URL: http://eric.ed.gov/?id=ED362404 [accessed 23 May 2016].

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.

Bartholomé T, Bromme R. 2009. Coherence formation when learning from text and pictures: What kind of support for whom? J Educ Psychol 101:282–293.

Barr ML. 2014. Encouraging college student active engagement in learning: The influence of response methods. Innovat High Educ 39:307–319.

Behnke RS. 2012. Kinetic Anatomy. 3rd Ed. Champaign, IL: Human Kinetics, Inc. 329 p.

Berends IE, van Lieshout EC. 2009. The effect of illustrations in arithmetic problem-solving: Effects of increased cognitive load. Learn Instruct 19:345–353.

Bettinger E. 2010. To be or not to be: Major choices in budding scientists. In: Clotfelter CT (Editor). American Universities in a Global Market. 1st Ed. Chicago, IL: University of Chicago Press. p 69–98.

Bloom BS (Editor). 1956. Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook 1: Cognitive Domain. 1st Ed. New York, NY: David McKay Co, Inc. 201 p.

Bockoven J. 2004. The pedagogical toolbox: Computer-generated visual displays, classroom demonstration, and lecture. Psychol Rep 94:967–975.

Bruff D. 2009. Teaching with Classroom Response Systems: Creating Active Learning Environments. 1st Ed. San Francisco, CA: Jossey-Bass. 240 p.

Brünken R, Plass JL, Leutner D. 2010. Direct measurement of cognitive load in active learning. Educ Psychol 38:53–61.

Butcher KR. 2006. Learning from text with diagrams: Promoting mental model development and inference generation. J Educ Psychol 98:182–197.

Butler AC, Roediger III HL. 2007. Testing improves long-term retention in a simulated classroom setting. Eur J Cognit Psychol 19:514–527.

Cavanagh AJ, Aragón OR, Chen X, Couch BA, Durham MF, Bobrownicki A, Hanauer DI, Graham MJ. 2016. Student buy-in to active learning in a college science course. CBE Life Sci Educ 15:ar76.

Chandler P, Sweller J. 1991. Cognitive load theory and the format of instruction. Cognit Instruct 8:293–332.

Chodzko-Zajko W. 2014. The American Kinesiology Association undergraduate core curriculum©. Quest 66:288–294.

Clark JM, Paivio A. 1991. Dual coding theory and education. Educ Psychol Rev 3:149–210.

Crisp V, Sweiry E. 2006. Can a picture ruin a thousand words? The effects of visual resources in exam questions. Educ Res 48:139–154.

Crouch CH, Mazur E. 2001. Peer instruction: Ten years of experience and results. Am J Phys 69:970–977.

Dobson JL, Linderholm T. 2015. The effect of selected "desirable difficulties" on the ability to recall anatomy information. Anat Sci Educ 8:395–403.

Donovan W. 2008. An electronic response system and conceptests in general chemistry courses. J Comput Math Sci Teach 27:369–389.

Fagen AP, Crouch CH, Mazur E. 2002. Peer instruction: Results from a range of classrooms. Phys Teach 40:206–209.

Fenesi B, Mackinnon C, Cheng L, Kim JA, Wainman BC. 2017. The effect of image quality, repeated study, and assessment method on anatomy learning. Anat Sci Educ (in press; doi: 10.1002/ase.1657).

Freeman S, Eddy SL, McDonough M, Smith MK, Okoroafor N, Jordt H, Wenderoth MP. 2014. Active learning increase student performance in science, engineering, and mathematics. Proc Natl Acad Sci U S A 111:8410–8415.

FSSE. 2015. Faculty Survey of Student Engagement. FSSE 2015 Aggregate Frequencies Frequency Distributions by Disciplinary Area. Faculty Who Teach Lower-Division Courses. Bloomington, IN: Indiana University School of Education. 23 p. URL: http://fsse.indiana.edu/pdf/FSSE_IR_2015/summary_tables/summary_tables_Lower_Division_by_Discipline.pdf [accessed 10 June 2016].

Gill DL. 2007. Integration: The key to sustaining kinesiology in higher education. Quest 59:269–286.

Grignon B, Oldrini G, Walter F. 2016. Teaching medical anatomy: What is the role of imaging today? Surg Radiol Anat 38:253–260.

Gyselinck V. 1996. Illustrations and mental models in text comprehension. Ann Psychol 96:495–516.

Hart SG. 2006. NASA-task load index (NASA-TLX); 20 years later. In: Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting (HFES 2006); San Francisco, CA, 2006 October 16-20. p 904–908. Human Factors and Ergonomics Society, Santa Monica, CA.

Hart SG, Staveland LE. 1988. Development of NASA-TLX (task load index): Results of experimental and theoretical research. In: Hancock PA, Meshkati N (Editors), Human Mental Workload. 1st Ed. Amsterdam, The Netherlands: North-Holland Press. p 139–183.

Hazari Z, Tai R, Sadler P. 2007. Gender differences in introductory university physics performance: The influence of high school physics preparation and affective factors. Sci Educ 91:847–876.

Holland J, O'Sullivan R, Arnett R. 2015. Is a picture worth a thousand words: An analysis of the difficulty and discrimination parameters of illustrated vs. text-only vignettes in histology multiple choice questions. BMC Med Educ 15:184.

Hora MT, Holden J. 2013. Exploring the role of instructional technology in course planning and classroom teaching: Implications for pedagogical reform. J Comput High Educ 25:68–92.

Hoyt A, McNulty JA, Gruener G, Chandrasekhar A, Espiritu B, Ensminger D, Price R Jr, Naheedy R. 2010. An audience response system may influence student performance on anatomy examination questions. Anat Sci Educ 3:295–299.

Hunt D. 1978. Illustrated multiple choice examinations. Med Educ 12:417–420.

Hurtado S, Eagan MK, Pryor JH, Whang H, Tran S. 2012. Undergraduate Teaching Faculty: The 2010–2011 HERI Faculty Survey. 1st Ed. Los Angeles, CA: Higher Education Research Institute, University of California Los Angeles. 104 p.

Kalyuga S, Chandler P, Sweller J. 1998. Levels of expertise and instructional design. Hum Factors 40:1–17.

Kalyuga S, Chandler P, Sweller J. 2000. Incorporating learner experience into the design of multimedia instruction. J Educ Psychol 92:126–136.

Karpicke JD, Roediger HL 3rd. 2008. The critical importance of retrieval for learning. Science 319:966–968.

Knight JK, Wood WB. 2005. Teaching more by lecturing less. Cell Biol Educ 4:298–310.

Kollöffel B. 2012. Exploring the relation between visualizer–verbalizer cognitive styles and performance with visual or verbal learning material. Comput Educ 58:697–706.

Kotze SH, Mole CG, Greyling LM. 2012. The translucent cadaver: An evaluation of the use of full body digital X-ray images and drawings in surface anatomy education. Anat Sci Educ 5:287–294.

Lasry N, Mazur E, Watkins J. 2008. Peer instruction: From Harvard to the two-year college. Am J Phys 76:1066–1069.

Leppink J, Paas F, Van der Vleuten CP, Van Gog T, Van Merriënboer JJ. 2013. Development of an instrument for measuring different types of cognitive load. Behav Res Methods 45:1058–1072.

Leppink J, van den Heuvel A. 2015. The evolution of cognitive load theory and its application to medical education. Perspect Med Educ 4:119–127.

Lin L, Atkinson RK. 2011. Using animations and visual cueing to support learning of scientific concepts and processes. Comput Educ 56:650–658.

Logan JM, Thompson AJ, Marshak DW. 2011. Testing to enhance retention in human anatomy. Anat Sci Educ 4:243–248.

Mayer RE. 2011. Applying the science of learning to multimedia instruction. Psychol Learn Motiv 55:77–108.

Mayer RE, Anderson RB. 1992. The instructive animation: Helping students build connections between words and pictures in multimedia learning. J Educ Psychol 84:444–452.

Mayer RE, Bove W, Bryman A, Mars R, Tapangco L. 1996. When less is more: Meaningful learning from visual and verbal summaries of science textbook lessons. J Educ Psychol 88:64–73.

Mayer RE, Gallini JK. 1990. When is an illustration worth ten thousand words? J Educ Psychol 82:715–726.

Mayer RE, Massa LJ. 2003. Three facets of visual and verbal learners: Cognitive ability, cognitive style, and learning preference. J Educ Psychol 95:833–846.

Mazur E. 1997. Peer instruction: Getting students to think in class. In: Redish EF, Rigden JS (Editors). The Changing Role of Physics Departments in Modern Universities, Part Two: Sample Classes. Proceedings of International Conference on Undergraduate Physics Education, College Park, Maryland, August 1996. 1st Ed. Woodbury, NY: American Institute of Physics. p 981–988.

Mazur E. 2009. Education. Farewell, lecture? Science 323:50–51.

McConnell DA, Steer DN, Owens KD, Knott JR, Van Horn S, Borowski W, Dick J, Foos A, Malone M, McGrew H, Greer L, Heaney PJ. 2006. Using conceptests to assess and improve student conceptual understanding in introductory geoscience courses. J Geosci Educ 54:61–68.

Miller K, Lasry N, Lukoff B, Schell J, Mazur E. 2014. Conceptual question response times in peer instruction classrooms. Phys Rev ST Phys Educ Res 10:020113.

Moreno R, Mayer RE. 2002. Learning science in virtual reality multimedia environments: Role of methods and media. J Educ Psychol 94:598–610.

NAK. 2016. National Academy of Kinesiology. Kinesiology: The discipline and related professions. National Academy of Kinesiology, Champaign, IL. URL: http://www.nationalacademyofkinesiology.org/what-is-kinesiology [accessed 1 November 2016].

Ost B. 2010. The role of peers and grades in determining major persistence in the sciences. Econ Educ Rev 29:923–934.

Paas F, Tuovinen JE, Tabbers H, Van Gerven PW. 2003. Cognitive load measurement as a means to advance cognitive load theory. Educ Psychol 38:63–71.

Paivio A, Csapo K. 1973. Picture superiority in free recall: Imagery or dual coding? Cognit Psychol 5:176–206.

Schlatter MD. 2002. Writing conceptests for a multivariate calculus class. Probl Resour Issues Math Undergrad Stud 12:305–314.

Soderstrom NC, Kerr TK, Bjork RA. 2016. The critical importance of retrieval--and spacing--for learning. Psychol Sci 27:223–230.

Stein PS, Challman SD, Brueckner JK. 2006. Using audience response technology for pretest reviews in an undergraduate nursing course. J Nurs Educ 45:469–473.

Tangen JM, Constable MD, Durrant E, Teeter C, Beston BR, Kim JA. 2011. The role of interest and images in slideware presentations. Comput Educ 65:856–872.

Thomas JR. 2014. The public face of kinesiology in the 21st century. Quest 66:313–321.

Tobias S. 1990. They're Not Dumb, They're Different: Stalking the Second Tier. 1st Ed. Tucson, AZ: Research Corporation. 96 p.

Vorstenbosch MA, Klaassen TP, Kooloos JG, Bolhuis SM, Laan RF. 2013. Do images influence assessment in anatomy? Exploring the effect of images on item difficulty and item discrimination. Anat Sci Educ 6:29–41.

Vorstenbosch MA, Bouter ST, van den Hurk MM, Kooloos JG, Bolhuis SM, Laan RF. 2014. Exploring the validity of assessment in anatomy: do images influence cognitive processes used in answering extended matching questions? Anat Sci Educ 7:107–116.

Willingham DT, Hughes EM, Dobolyi DG. 2015. The scientific status of learning styles theories. Teach Psychol 42:266–271.

Wojciechowska I. 2010. A quickly growing major. Inside Higher Ed. 11 August 2010. Inside Higher Ed., Washington, DC. URL: https://www.insidehighered.com/news/2010/08/11/kinesiology [accessed 1 November 2016].

Young JQ, Irby DM, Barilla-LaBarca ML, Ten Cate O, O'Sullivan PS. 2016. Measuring cognitive load: Mixed results from a handover simulation for medical students. Perspect Med Educ 5:24–32.

APPENDIX A

Example examination questions are listed below. The examination questions were associated with image-based or text-based active learning exercises or were not associated with any active learning exercises. All of the examination questions were multiple choice. None of the examples questions referred to a figure.

Example examination questions associated with image-based active learning exercises:

Which landmark is found only on a thoracic vertebra?

Shoulder separations are most associated with damage to which ligament?

What is the best way to stretch rectus femoris?

Which of the following landmarks is most anterior, medial and proximal on the humerus?

Which of these carpal bones is included in the distal row of carpals?

Example examination questions associated with text-based active learning exercises:

Which muscle shortens during rotation of the trunk to the left?

Which of the following muscles is most lateral?

Which of the following muscles medially rotate the hip?

Which muscle inserts on the base of the first metacarpal?

Lateral epicondylitis might be associated with injury to which muscle?

Example examination questions not associated with in-class active learning exercises:

Osteoporosis and aging can lead to exaggeration of which spinal curvature?

Which structure is most likely to be injured with excessive valgus loading at the knee?

Which of the following muscles are included in the prevertebral group?

Which muscles elevate the temporomandibular joint?

Winging of the scapula is associated with loss of function in which muscle?

FIGURE LEGENDS

Figure 1. Examples of image-based and text-based active learning exercises. At left, a schematic drawing of a transverse section of the left lower leg demonstrates the position of muscles as they cross the ankle joint (medial-lateral axis) and the subtalar joint (anterior-posterior axis). Yellow dots indicate individual student responses to the image-based question; darker yellow dots indicate correct responses. At right, colored bars indicate the number of student responses to each of the possible answers to the text-based question. The green highlight indicates the correct response. The number of students that participated in each learning exercise are indicated at the bottom of each screen.

Figure 2. The number of students attempting in-class exercises, grouped by examination performance and level of participation. Bar color indicates level of participation in active learning exercises, with students attempting a high number of exercises represented in dark blue, students attempting an average number of exercises represented in blue, and students attempting a low number of exercises represented in light blue. Bar length indicates the number of students attempting in-class activities at high, middle or low levels for each level of examination performance. For example, students with an average examination performance (middle graph) that attempted a high number of in-class exercises (dark blue bar) are represented by the fourth bar from the top; there were 18 students who fell into this category.

Table 1. Regression Models Predicting Examination Performance on Associated Image-Based Questions and Text-Based Questions

Model	**Associated Image-Based Questions**				**Associated Text-Based Questions**			
	Estimate	**SE**	**β**	***P*-value**	**Estimate**	**SE**	**β**	***P*-value**
Intercept	0.18	0.07		0.01	0.16	0.08		0.05
Image-based exercise participation	0.42	0.11	0.51	<0.001	0.32	0.13	0.33	0.01
Text-based exercise participation	0.05	0.10	0.07	0.62	0.20	0.11	0.24	0.07
Underrepresented minority status (Non-URM coded 0)	0.03	0.03	0.05	0.37	0.01	0.03	0.01	0.84
Gender (Male coded 0)	-0.02	0.18	-0.73	0.23	-0.03	0.02	-0.10	0.12
Grade Point Average	0.75	0.02	0.24	<0.001	0.08	0.02	0.20	0.002
Adjusted R^2	0.40				0.37			

Number of participants N = 176; Estimate, or unstandardized coefficient, indicates how much the dependent variable (examination performance) varies when all other predictor variables are held constant; SE, standard error of the estimate; β, or standardized coefficient, removes the effect of units by standardizing variances so that comparisons among predictor variables measured with different scales is easier.

Table 2. Student Ratings for the Learning Activities

Learning Activity Questions	**Mean (±SD)**
Were learning activities….	
easy or demanding?	4.28 (±1.66)[a]
with images easy or demanding?	4.16 (±1.79)[b]
with text easy or demanding?	4.70 (±1.82)
How hard did you work to understand the….	
content of learning activities?	3.70 (±1.69)[c]
content of learning activities with images?	3.54 (±1.64)[d]
content of learning activities with text?	4.00 (±1.80)
How much effort was invested to participate in….	
learning activities?	4.06 (±1.68)[e]
learning activities with images?	3.95 (±1.64)[d]
learning activities with text?	4.34 (±1.66)

[a]Likert scale: 1 = easy, 8 = demanding; [b]Image-based mean is significantly different from text-based mean ($P < 0.001$); [c]Likert scale: 1 = not very hard, 8 = very hard; [d]Image-based mean is significantly different from text-based mean ($P < 0.05$); [e]Likert scale: 1 = low effort, 8 = high effort.

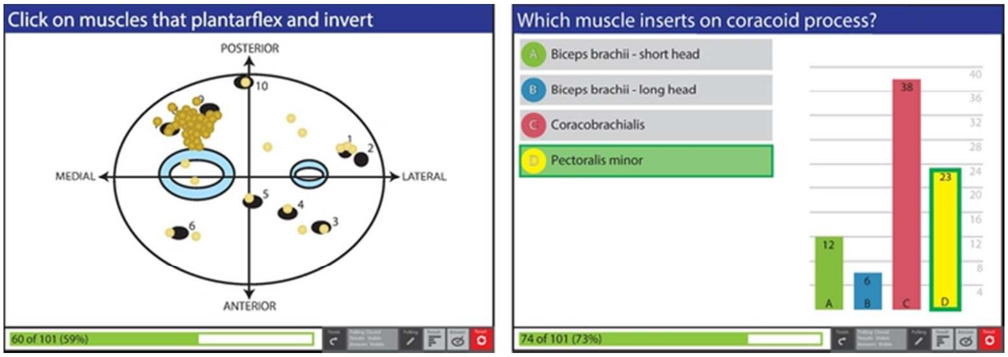

Figure 1. Examples of image-based and text-based active learning exercises. At left, a schematic drawing of a transverse section of the left lower leg demonstrates the position of muscles as they cross the ankle joint (medial-lateral axis) and the subtalar joint (anterior-posterior axis). Yellow dots indicate individual student responses to the image-based question; darker yellow dots indicate correct responses. At right, colored bars indicate the number of student responses to each of the possible answers to the text-based question. The green highlight indicates the correct response. The number of students that participated in each learning exercise are indicated at the bottom of each screen.

60x21mm (300 x 300 DPI)

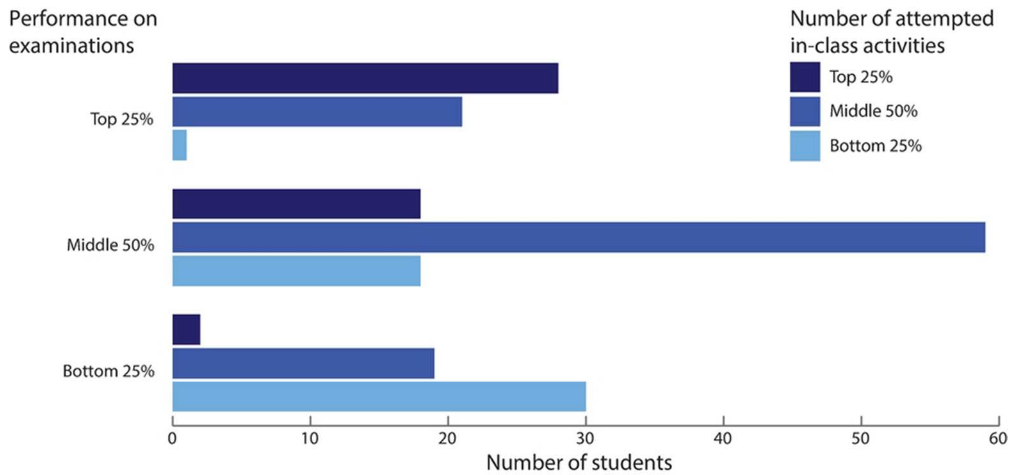

Figure 2. The number of students attempting in-class exercises, grouped by examination performance and level of participation. Bar color indicates level of participation in active learning exercises, with students attempting a high number of exercises represented in dark blue, students attempting an average number of exercises represented in blue, and students attempting a low number of exercises represented in light blue. Bar length indicates the number of students attempting in-class activities at high, middle or low levels for each level of examination performance. For example, students with an average examination performance (middle graph) that attempted a high number of in-class exercises (dark blue bar) are represented by the fourth bar from the top; there were 18 students who fell into this category.

80x37mm (300 x 300 DPI)