DISSERTATION

EVALUATING LEARNING RESOURCE SELECTION TO SUPPORT GROSS ANATOMY EDUCATION

Submitted by
Jason Martin
Department of Biomedical Sciences

In partial fulfillment of the requirements
For the Degree of Doctor of Philosophy
Colorado State University
Fort Collins, Colorado
Fall 2023

Doctoral Committee:

Advisor: Christianne Magee

Andrew West
Carolyn Meyer
Leann Kaiser
Quinton Winger

ABSTRACT

EVALUATING LEARNING RESOURCE SELECTION TO SUPPORT GROSS ANATOMY EDUCATION

Anatomy educators are tasked with developing, maintaining, and revising comprehensive curricula that strengthen the foundation of a growing body of scientific knowledge necessary to be successful medical professionals. This work sought to evaluate the impact of instructional timing, spatial ability, self-efficacy, and resource preference in the animal anatomy classroom for undergraduate and graduate students at Colorado State University. It identified factors relevant to optimizing student performance on animal gross anatomy examinations. The first chapter provides a background on anatomy learning resources in the context of self-efficacy and spatial ability research. Chapter two found that following a transition to remote instruction, students value resources that assist with navigating their learning ecology and assist with content mastery. Chapter three was a five-year retrospective study that identified a correlation between prior examination experience and dissection examination scores. Chapter four compared two measures of visuospatial ability with atlas- and specimen-based animal anatomy assessments following experimental single-resource and real-world multi-resource instruction. Chapter five was a mixed-methods analysis that developed an experimental framework used to describe the relationship between learner self-efficacy and animal gross anatomy assessment scores. Using this framework, increased time attending in-person didactic lectures and teaching assistants mediated open laboratories was found to be beneficial for low self-efficacy students while independent exploration of open-laboratory was beneficial for high self-efficacy students. An experimental study reported in chapter five failed to find a relationship between resource preference and performance on experimental anatomy assessments. When taken together, the research provides suggestions for anatomy educators seeking to use the selective distribution of learning resources to improve the instruction of animal anatomy students.

ACKNOWLEDGEMENTS

Rarely one to be at a loss for words, I struggle to adequately express my gratitude for the patience, feedback, guidance, and support of my advisor and mentor, Dr. Christianne Magee. Thank you, Dr. Magee. The generosity you have provided me over the past ten years extends well beyond funding and feedback and I thank you for supplying me with the freedom to explore my interests, helping me to build my network, and never shying away from holding me accountable. Thank you for showing me how to navigate disappointing results, complex student needs, improve my presentation style, work through grief, and how to celebrate success.

I would also like to extend my gratitude to my committee members Drs. Andrew West, Carolyn Meyer, Leann Kaiser, and Quint Winger. Your feedback and advice were exceptional, and I always left our meetings with a clearer understanding of where I was headed. Dr. Winger, you always provided unambiguous insight necessary to improve my work. Dr. Kaiser, as a direct result of your teaching and feedback on my committee, my educational methodology and teaching has been strengthened substantially. Dr. Meyer, your anatomy teaching and research experience always provided a relevant and helpful perspective. Dr. West, as a veterinary student I was immensely thankful for the support you provided me and my classmates, and as a graduate student I saw how closely you worked with faculty across universities. Thank you all.

I would also like to thank Dr. Ray Whalen for his dedication to veterinary anatomy education and his commitment to improving the effectiveness of anatomy learning resources. Next, I want to thank Andrea Linton for reading the many, many drafts that preceded this document and all of her support which has improved my writing over the past few years. I want to thank Gwen Cahill for helping me with the operation of my experimental sessions and the

hours of data entry and analysis necessary for these studies. To Dr. Anna Fails and Kenny Ivie, your passion for teaching has been an inspiration. To Brittany Runyan, for helping to manage important things such as money, paperwork, and deadlines. I want to thank Andrew Garrett, Paulina Svec, Damon Mango, Dr. Kenneth Compton and the rest of the animal anatomy teaching and Virtual Veterinary Education Tools team for all their support building resources used as part of my studies and for helping me collect student data necessary for analysis.

I want to share my utmost appreciation for my parents, Fred and Cheryl. As well as to my family Megan, Amy and Bruce, Hannah, Kathryn, and Elizabeth, and my Nana Petrozzi. Your support, generosity, and love never faded. Finally, to Tippy and Smudge, you were, without question, the reason this I started and finished this journey.

DEDICATION

I would like to dedicate my dissertation to my daughters, Tippy Marie and Smudge Anne.

May all our curiosities be as resolute and compassionate as yours were.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I – REVIEW OF LITERATURE

1.1. PURPOSE

In the twenty century, biomedical education became burdened with disseminating an ever-growing body of knowledge to the next generation of professionals.[1,2] The golden age of modern medicine is frequently used to describe the first half of the twentieth century, culminating in 1955 with the release of the polio vaccine and the importance of the placebo.[3–5] Golden ages are often attributed to a period of application following technological innovations, and the placebo provided medicine a methodological means of critically evaluating centuries of knowledge. While the general population initially benefited from rapid systematic evaluations of healthcare, medical educators remain concerned with addressing the growing body of knowledge[1], demand for additional healthcare workers[6], and increased enrollment within professional degree programs.[7–9]

The two-pillar model has been attributed to Abraham Flexner and William Osler, who advocated for the inclusion of both foundational and clinical knowledge within medical education.[9–11] Flexner called for medical schools to be tied to established universities that required prerequisites for students, research experience of instructors, and a curriculum that presented timely medical knowledge.[9,10] Osler insisted that students needed a medical education that included clinician-guided patient contact.[9,10] Research-focused instructors struggled to connect the growing body of scientific knowledge to upcoming clinical experiences while the increased student-to-clinician ratio impinged upon the availability and value of bedside experiences.[9]

A shift to an organ-system-based approach to instruction in the 1950s provided some integration between the foundational and clinical curricula[1,9], of which gross anatomy is frequently described as part of the former.[12–14] Anatomists noted that this transition occurred in parallel with a decrease in instructional time dedicated to morphological studies.[2,15–21] The change has been perceived as substantial, trending from a reported 549 hours at the start of the twentieth century to fewer than 200 hours during the twenty-first century.[2,15,22–24] However, reports of reduced instructional time often focused exclusively on dedicated anatomical instruction did not include vertical approaches to clinical proficiency that use spaced repetition within other courses to reinforce anatomical knowledge throughout a degree program. While vertical approaches to anatomy education are most associated with instruction centered around interpretations of radiographs and ultrasounds, it is also included in neurologic, orthopedic, surgical, and reproductive instruction.[2] Therefore, anatomical resources provide substantial value beyond gross anatomy coursework.

The COVID-19 pandemic necessitated the replacement of face-to-face learning with independent, asynchronous, and remote instruction. The loss of in-person didactic lectures and practical cadaveric laboratories has been described as "disruptive innovation" as it forced anatomy educators to adopt unfamiliar solutions and promote resources previously used to support a primary methodology.[25] The pandemic provided an unprecedented opportunity to examine learning resources in modern anatomy education.[26] Three research questions were developed: How do we measure the contribution of a learning resource to anatomy education? How does the relative timing of anatomy instruction influence assessment outcomes? How do we use learner-specific variables to guide the selection of learning resources? These questions will be addressed throughout this work.

1.2. METHODOLOGY

This work includes a mixed-method research design that combines qualitative and quantitative approaches into a shared body of knowledge. Occasionally, there is a misconception that these terms describe the absence or presence of numerical data. However, such a definition inappropriately assigns value to the technique used rather than the intention of the research.[27] Here, research is defined as qualitative if the analysis intends to describe data or generate a hypothesis. By contrast, the intention of quantitative research analyses is to use data to test a hypothesis.

Mixed-method designs have seen increased popularity in healthcare research.[27–36] Justifications for the methodology are numerous and include, but are not limited to, intentions to provide a broader depiction of the topic, increase the credibility of the analysis, highlight a diversity of perspectives, develop or revise frameworks, and evaluate processes.[37] Independent of the researcher's justification, the study must describe a method for integrating qualitative and quantitative analyses.[28,37] Frequently, there is an order of operations applied where one analysis is used to explain another. An example of this can be seen in Chapter Five, whereby an exploratory approach used the results of a qualitative data analysis to generate a framework that was examined using subsequent quantitative analysis. Alternatively, qualitative and quantitative data can be analyzed concurrently. Although chapter two does not self-identify as a mixed-method analysis, examining its methodology reveals a convergent approach was used before analysis to convert survey comments that described resource use to numerical resource counts.

Fields that traditionally employ quantitative analysis have identified strategies such as triangulation, reflexivity, member checking, prolonged engagement, audit trail, peer debriefing, and thick description to increase the trustworthiness of qualitative analysis.[38] Triangulation uses two data sources or multiple researchers to analyze the data independently. Member checking is a methodological process where participants can confirm or deny conclusions developed from the data associated with them. Prolonged engagement describes methods where the researcher and participants have time to build rapport. Qualitative studies provide an audit trail to help the reader connect experimental data with the study's conclusions. Peer debriefing involves sharing study details with collaborators not directly involved with the research project to provide an opportunity for alternative interpretations of the data. Thick description involves providing sufficient methodological details to allow readers to evaluate how well the results can be transferred to other populations. Chapter Five employed each of these techniques. Reflexivity was partially addressed by the details provided within the contributor's notes and is supplemented within this chapter by including a researcher's perspective and teaching philosophy.

1.2.1. Researcher's Perspective

My career in anatomy education began in 2014 as an undergraduate teaching assistant at Colorado State University (CSU). Without prior training in pedagogy or andragogy, my early teaching experiences borrowed heavily from my undergraduate study in behavior analysis and eight years of experience managing salespeople. Prior to receiving my Master of Biomedical Science degree, I was privileged to develop programs for the Fort Collins Museum of Discovery, be a course instructor for an introductory biology course at Aims Community College in Greeley, Colorado, provide laboratory instruction for both the Biology and Biomedical Science departments as a CSU undergraduate and graduate teaching assistant, and assist with anatomy education research as a CSU graduate research assistant. While enrolled in CSU's Doctor of Veterinary Medicine, I completed night courses in education. After receiving my DVM, I returned to the anatomy laboratory to instruct undergraduate, graduate, and professional students in animal gross anatomy.

What began as single-task operant behavior training has expanded into a vertically integrated, data-driven teaching philosophy. This change was greatly influenced by the courses completed as part of a Facilitating Adult Learning graduate certificate. The opportunity to partner with faculty and graduate students within the School of Education allowed me to critically evaluate established instructional methods of gross anatomy instruction and identify how the results of mixed methods analyses could be used to improve teaching. In partnership with faculty mentors, I have evaluated data-driven approaches to increasing early exploration of veterinary medicine through elementary and high school outreach, expanded undergraduate and graduate experiences, and identified ways to support professional development.

1.2.2. Teaching Philosophy

Education is a method of developing content experts at a rate faster than curiosity. The educator's ability to curate relevant content and organize an approach to mastery differentiates independent exploration from education. When I started teaching, I measured my effectiveness in content delivered per hour; if I had time to say it, they had time to learn it. I would explain forty-five slides in twenty minutes and then expect the students to use the remaining hour to complete a laboratory assignment efficiently. My experience with conventional approaches to teaching taught me that introductory courses served as gatekeepers to challenging professions. This perspective was supported when my student evaluations exceeded the mean for graduate teaching assistants, and observation rubrics completed by the department rated my instruction highly. Ignoring the countless glazed eyes that predominated my lectures, my numbers looked great.

Due to enrollment sizes, those early instructional opportunities were highly structured; as someone who would go on to specialize in anatomy education, I frequently found control to be at odds with engagement. It was not until I became a teaching assistant for animal anatomy that I finally felt a sense of instructional autonomy. I started developing my teaching philosophy after exploring the path students would take through the course by working backward from the exam. What did learners need to know daily to prepare them for what was to follow? I spoke regularly with the course coordinator, who would become my mentor, and the other instructors to understand the expectations and clarify any ambiguity of the learning objectives.

As an anatomy educator, I rarely teach anatomists. My students enroll in my courses to strengthen their content expertise in medicine, art, and athletics. While I encourage students to attend open laboratory sessions to increase their exposure to gross anatomy concepts, I attend these sessions to understand better what these students value and hope to obtain from anatomical study. A consequence of this rapport is that I can identify opportunities for intervention that would have otherwise gone unnoticed. For example, some students regularly attend laboratory sessions but remain uncomfortable engaging with specimens, peers, or teaching assistants. My solutions involve counterintuitive experiences. Biosecurity protocols require gloves when handling specimens, and by removing my gloves, I force students resistant to touching specimens to engage directly with materials. Through conversations with students, I learned that those who routinely failed to speak up during study periods routinely lacked confidence in their knowledge, even if their understanding was sufficient to participate. Students report enjoying when I ask them to make up anatomical names that sound like they could be correct but which they know are wrong. This activity is a fun diversion from their study that provides me with a less threatening situation to address the volume of their voice and frequency of speaking before inviting them into a stressful conversation.

My educational philosophy is to support student motivations for individual learning. To foster an environment where a student's failure is perceived as a consequence of content exposure, I encourage them to evaluate the methods they use to learn critically. I avoid practices that question their ability to achieve mastery but do call out behaviors that restrict learning. I recognize that while their personal goals may not directly mirror the course objectives, their success results from their engagement.

1.3. LITERATURE

1.3.1. Key Terminology

Several key words and phrases are used to describe specific conditions associated with learning resources. Due to a lack of a unified lexicon, anatomy education literature occasionally applies definitions that do not align with other studies. Rather than redefine the terms as they arise or attempt to explain the language used in each of the selected works, the following reference is provided.

Table 1.1. Key Terminology, Definitions, and Examples.

Term	Definition	Example
Specimen	An anatomical learning resource derived from a previously living organism.	1. Dissected Cadaver. 2. Preserved bones.
Learning Resource	Any source of information accessible to a student that aids in the understanding of gross anatomy.	1. Anatomical atlas. 2. Teaching Assistant. 3. Lecture attendance.
Two-Dimensional (2D)	Lacking sufficient detail to generate a stereoscopic image.	1. A single photograph. 2. A rotatable object is presented on a screen, but that has not been observed to rotate.
Monoscopic	Visual observation where both eyes perceive identical images.	1. Two or more photographs taken from different angles. 3. A rotatable object that has been observed rotated.
Mental Rotation Test (MRT)	A spatial ability test that requests participants to	1. Vandenberg and Kuse[39]

	visualize a 2D image in 3D and mentally rotate it before they compare it with another 2D image.	
Stereoscopic	Visual observation where each eye perceives a different image.	1. An object whose height, length, and width are perceived independently by the left and right eye.
Three-Dimensional (3D)	Containing sufficient detail to generate a stereoscopic image.	1. Two or more photographs taken from different angles. 2. A rotatable object that has been observed rotated. 3. A physical object with height, length, and width is perceived with stereoscopic vision.
Three-dimensional Visualization Technologies (3DVT)	Any learning resource that presents three dimensions to any component of the	1. Virtual Reality Headset. 2. 3D movie.

	learning resource.	3. Laparoscopic Surgical Simulator
Two-Dimensional (2D)	Lacking sufficient detail to generate a stereoscopic image.	4. A single photograph. 5. A rotatable object presented on a screen, but that has not been observed to rotate.

1.3.2. Overview

This literature review intends to present a representative sample of the literature on learning resources that support gross anatomy instruction. The section begins with a description of Carroll's Model of School Learning and is supported by Bloom's Learning for Mastery. Relevant literature on gross anatomy learning resources and their relationship to self-efficacy and spatial ability are included. Due to the rate of technological advancement and the multitude of applications in which variations can be applied to instruction, this chapter is not intended to be an exhaustive review. Literature published over a decade ago has been excluded unless justified through substantial contributions. Due to the volume of literature published, only anatomy education research made available via the National Library of Medicines online portal PubMed is described in detail. Understanding the role of alternatives to cadaveric instruction will further the field of anatomy education.

1.3.3. A Model of School Learning

Carroll's A Model of School Learning (1963) was an early attempt to apply a mathematical analysis to learning.[40] Carroll observed that many cognitive variables that influence learning within a structured educational experience, including maturity and motivation, contain sufficient overlap to make identifying those factors' influence challenging.[41] In his initial proposal, he stated that he was searching for a model that was as comprehensive as possible without sacrificing simplicity. Carroll self-identified that his model was an oversimplification of learning, and the consequence of this had implications on the extent to which the model could be applied; he deferred to B. F. Skinner's work on operant behavior for any task that was predominantly emotional or the consequence of values.[41] Carroll used the term task quite broadly and included any change from ignorance to either recall of knowledge or capacity to perform a skill.

The model states that the extent to which one will learn a task depends on the time needed to learn the task. Carroll clarified that this statement came with three qualifications: that time spent learning only accrued when the student was trying to learn, that there were learner-specific variables that influenced how much time the student spent learning, and that there were variables that influenced how much time it will take a student to learn the task.[41] In total, Carroll proposed that five variables defined the degree of learning: Three variables were expressed in units of time (opportunity to learn, perseverance, and aptitude) and two as assumptions of achievement (quality of instruction and ability to understand instruction).

The degree of learning is represented as a ratio of the time available and the time needed to learn.[40,41] The time available (numerator) is the amount of time available for learning and is exclusively the variable with the shortest duration. Of the three, the opportunity to learn is when learning could feasibly occur. Perseverance[40], or motivation[42], is the time a learner assigns to learning. The third, aptitude, is the amount of time it takes a learner to achieve a set level of competency, and it represents the time it takes to learn a task and the time saved due to prerequisite knowledge.[40,42]

The time needed (denominator) is calculated by modifying aptitude with the previously named assumptions of achievement. Increases in the quality of instruction or ability to understand instruction increase aptitude and decrease the time it takes a student to learn. In the model, the ability to understand instruction is typically considered a measure of general intelligence[42] and is therefore considered static within the defined instructional period. By contrast, the quality of instruction is influenced by the educators' skills and can increase or decrease the total time needed to learn.[40,42]

Initial investigations of the model attempted to understand better how variables intersected. Charles Spearman's seminal work on calculating general intelligence culminated in various intelligence tests by the middle of the twentieth century[43,44] and these tests were regularly being administered to school children.[45] Limitations of general intelligence assessments remain a concern[46,47], but correlation analysis derived from elementary school students[45] and adult learners[48] overlap with outcomes of computer modeling.[42] A learner's intelligence is the dominant factor when instructional quality is low, aptitude dominates in environments with high-quality instruction and restricted opportunity to learn, and perseverance dominates with high-quality instruction and extended opportunity to learn.[42,45]

Carroll has been criticized for focusing exclusively on individual learners.[49–51] Harnischfeger and Wiley, 1978 raised concern that focus on the individual learner fails to account for conditions of the curriculum within the classroom.[49] Arlin, 1982 noted that course management necessitates prioritizing instruction to rate of learning.[51] Real-world application of the model's linear relationship to learning has been questioned. A meta-analysis of over 800 school achievement variables identified the classroom as a complex dynamic system composed of many interdependent, non-linear relationships.[50] Two and a half decades after its original publication, Carroll addressed some of these concerns, praising the pursuit of understanding complex conditions of social dynamics and emotions but noting that a simplified model remains helpful when working with complex phenomena.[40]

The research presented in the following chapters is relevant to Carroll's A Model of School Learning because this work looks at how the availability of learning resources influences student performance; by provisioning learning resources for student use, their presence and absence are expected to influence the quality of instruction.

1.3.4. Learning for Mastery

Learning for Mastery is a learning strategy derived from Carroll's model, which presumes that all learners can learn any task or skill within an unconstrained time.[52] By removing the burden of instructional duration from teaching, attention shifts to variables that influence the rate of learning.[52,53] The consequence is that when instruction quality is adequately addressed, learning gains increase.[40]

Carroll describes the difference between his model and Bloom's mastery as a difference in educational philosophies, whereby Carroll was concerned about opportunity equality, not parity of performance.[40] Indeed, while Carroll's model challenged the educator to modify the quality of instruction to maximize the number of learners who achieved a sufficient degree of learning[40,41,45], Bloom questioned the practice of students matriculating at 70% mastery and proposed a requirement where they must demonstrate an understanding of 90% or more to pass assessments.[52]

Criticism of Learning for Mastery notes that prioritizing the learning process over instructional time does not guarantee that additional learning time will produce changes in student achievement.[54] By prioritizing performance, discrepancies in traditional classroom learning rates are magnified by Learning for Mastery[51], with slow learners necessitating additional financial support and fast learners squandering learning opportunities.[55]

The anatomy courses included in the following studies provide more opportunities for content engagement than most enrolled students regularly use. Between available open laboratory sessions and online tools, the opportunity to learn is expected to be greater than perseverance or aptitude. For this reason, the study of anatomy resources may advance Bloom's Learning for Mastery by adding to the body of knowledge about learning resources. The literature review will describe how select anatomical learning resources help students achieve content mastery.

1.3.5. Measuring Achievement

A strong understanding of theoretical knowledge and practical application must be demonstrated in medicine before allowing students to perform a procedure with live patients.[56] Within existing biomedical curriculums, students have multiple opportunities to demonstrate their knowledge. At the undergraduate level, the anatomy curriculum leans towards establishing sufficient competency in anatomic identification, while graduate and professional-level anatomy are expected to demonstrate an understanding of anatomical relationships and function.[23,57] Consequently, undergraduate anatomy examinations place greater emphasis on low-order, identification-based questions, while postgraduate anatomy courses dedicate more attention to higher-order, relationship- or clinical-based questions.[57]

Comparisons of undergraduate and postgraduate anatomy students are challenging.[57–59] Thompson and Griffin, 2021 emphasize differentiating between the results of computer-based tests and laboratory-based practicals. Their study investigated how the order level of questions impacted undergraduate and medical student anatomy examination performance.[57] Both populations took both a computer-based test and a laboratory practical. They noted that all students took longer to answer higher-order questions than lower-order ones, independent of test type, and they associated increased time to answer a question with increased difficulty. However, while medical students correctly answered more lower-order questions than higher-order ones, regardless of test format, comparisons of order level for undergraduates were associated with question presentation; for undergraduates, there was no difference between order levels for computer-based tests but a significant difference for laboratory practicals.[57]

Professional degree programs are notoriously competitive. Students pursuing medical careers prioritize professional school entrance requirements, and the consequence of this behavior is that non-academic characteristics become increasingly relevant to identifying applicants who will be successful students.[60] With respect to failure rates, anatomy is one of the most challenging subjects in the sciences.[61] Even in the face of data-driven intervention strategies that encourage more engaging learning experiences, the quality of instruction remains relatively unchanged.[62] It has been proposed that increased engagement with cadaveric resources leads to more knowledgeable students.[63] The following explores two factors associated with anatomy performance: visuospatial ability and self-efficacy.

1.3.6. Visuospatial Ability

Visuospatial (spatial) ability is the mental capacity for non-linguistical manipulation of visual stimuli. Its prominence in educational research can be traced back to measures of intelligence, G, and was used as a summative measure to describe one's ability to generate images, place the image in memory, retrieve the image, and transform it.[64] Anatomy is occasionally described as a visual science, and the applications of spatial ability to predict academic success or explain performance deficits have helped maintain its use as a prominent measure within anatomy education literature.[65] The regular inclusion of various spatial ability measures has resulted in the publication of four comprehensive literature reviews since 2012.[66–69]

A 2012 meta-analysis of 217 studies sought to explore if spatial ability skills could be changed through training, if the effects of training could be transferred to other spatial tasks, and then evaluated if select variables were related to training effects.[69] Training spatial ability resulted in nearly half a standard deviation improvement in mean effect size (Effect Size = .47, Standard Erro*r* = .04, total studies = 206, total effect sizes evaluated = 1038). This effect was also computed when evaluating the transfer of training to new spatial ability measures (Effect Size = .48, Standard Erro*r* = .04, total studies = 170, total effect sizes evaluated = 764). Neither age of the participant nor the type of training influenced effect sizes. Participants with higher levels of initial spatial ability were less likely to see improvements with training, possibly due to an upper limit of ability (although the effect was inconsistent across studies). A 2020 systematic review of the effects of spatial ability training, while enrolled in a semester of gross anatomy, reported similar results (Total Effect = 0.49, 95% confidence interval [0.17; 0.82], $n = 11$).[66]

A 2021 meta-analysis examined studies to determine if the spatial ability measure Mental Rotation Task (MRT) correlated with anatomy examination performance.[67] The selection of this test reflects the popularity of the MRT spatial ability-related anatomy education research.[66,67,70–74] The task presents multiple 2D images drawn to represent 3D structures, and participants were asked to indicate if the drawings were rotated but otherwise the same or if they represent objects of different shapes.[39,75] Three assessment measures were included to evaluate anatomy knowledge: written examination, laboratory practical, and or an (undefined) drawing test. While a statically significant correlation was reported for mental rotation task scores and anatomy performance (Effect Size = .24, 95% confidence interval [0..09; 0.38], *n* = 15), the small effect size was attributed to the nine studies with professional students, seven studies with practical examinations, and two studies with drawing tests.

A 2015 meta-analysis described the process of evaluating thirty-six studies that compared the effect of three-dimensional visualization technologies (3DVT) on an empirical outcome of anatomy knowledge.[68] It's important to note that 3DVT is not synonymous with stereoscopic 3D visuals. Any learning resource which contains a 3D component can qualify as 3DVT, even if the visual component of instruction is in 2D.[76] By contrast, stereoscopic 3D describes only those resources where binocular vision provides access to a 3D image. Confusion over implementations and interpretations of 3D is not unprecedented.[77–80]

1.3.7. Self-Efficacy

When deciding to engage in a particular behavior, individuals make two assessments: what will the outcome of a behavior be, and can they expect to engage in said behavior successfully?[81] The former assesses a cause-effect relationship, but the latter is a cognitive self-evaluation and a functional definition of self-efficacy. A more formal definition of self-efficacy is a belief held by an individual that they can accomplish a task.[81] Knowledge of a person's self-efficacy provides insight into the binary decision to initiate behavior and the relative expenditure and sustainability of effort.[81]

Measures of self-efficacy have been reported to predict gross anatomy performance.[63] Burgoon, Meece, and Granger, 2012 compared the results of 157 first-year medical students self-efficacy surveys with performance on eight anatomy tests. Besides referencing student dissections, the study provides little detail about the resources available to the students. The tests were distributed over four units, and students completed both a written examination and a laboratory practical for each unit. When controlling for prior experience (physical and biological scores on the MCAT), a statistically significant relationship was found from a linear regression analysis between self-efficacy survey scores and all four laboratory practicals and two written examinations. While all laboratory practicals were significant to less than .001, only laboratory practical three had an effect size high enough to be interpreted as typical.[82]

A three-university study on motivation to learn anatomy included a correlation between self-efficacy scores to examination performance.[83] Participants included first- and second-year undergraduate chiropractic science students, first-year dental students, and second-year medical students. Descriptions of available learning resources were more robust for experienced students and professional degree programs. First-year chiropractic science students received a combination of lecture, laboratory, and tutorial sessions. Second-year chiropractic science students were provided prosected cadavers and radiographic anatomy outside course lectures. Dental students completed a dissection of the human head and neck in conjunction with lectures, problem-based learning activities, clinical cases, and practical studies. Medical students received lectures, tutorials, live human ultrasounds, and dissection. While these descriptions suggest had access to different resources, the authors note that instructional methodologies were similar. A statistically significant relationship was reported between self-efficacy and anatomy grade (N = 251, R^2 = .109, $p < .001$), however, results were not provided for different education levels. These results align with the previously described study, statistically significant but less than typical in effect size.[82]

Not all education research has found a relationship between self-efficacy and anatomy examination performance. One study reported a non-significant, low effect-size relationship between self-efficacy and human gross anatomy practical examination performance in first- and second-year medical students in the United Kingdom (N = 428, R = .10, p = .38).[84] Students received 30 hours of workbook-guided anatomy instruction per module from various methods, including prosected cadavers, osteological specimens, anatomical demonstrations, models (of unknown form), radiographs, and a digital resource involving touch screen computers. However, it should be noted that the study reported all examination finding in normalized z-scores. While it is reasonable to use normalized data for comparing multiple tests that may contain variations in difficulty, in the absence of best-fit data and non-normalized examination performance, it is difficult to differentiate between a student population who scored highly on the examination and a measure that was not associated with exam performance.

Another study involving students (N = 131) enrolled in an undergraduate human gross anatomy course in Australia found that providing supplemental online anatomy videos did not influence examination performance compared to mean scores of prior semesters who did not have access to the videos.[85] However, they did note low self-efficacy learners (based on a survey presented by Burgoon et al., 2012) were statistically more likely to watch the videos. The results presented in this study should give anatomy educators pause if this population performed as well as they did because they watched the supplemental videos or if they invested study time in a resource that did not help them achieve content mastery.

Even when studies directly compare resource use, the relationship between reports of self-efficacy surveys and examination scores is inconsistent. A qualitative study involving gynecologists (N = 12) found that daily anatomy practice was associated with increased self-efficacy.[86] A study of active learning using ultrasonography to teach human gross anatomy to 140 third-year medical students did not improve their examination or practical scores, but self-efficacy scores did increase (*M* = 3.68, *SD* = ±0.56; $p < 0.001$).[87] However, no change in self-efficacy was reported in second-year medical students (N = 402) who supplemented their flipped-classroom lecture-laboratory anatomy course with either ultrasound or videos.[88] One explanation for the difference in effect between the two ultrasound interventions could be content presented, as the former was used to support hepatobiliary, urinary, and cardiac anatomy, and the latter was used to support musculoskeletal limb anatomy.

1.4. SUMMARY

The literature review demonstrated the need to evaluate learning resources in the context of the learner-specific factor's spatial ability and self-efficacy. A detailed meta-analysis demonstrated the malleability of spatial ability[69], and this finding was replicated in the context of anatomy education.[66] The MRT was reported to be correlated with gross anatomy performance, but the assessment format and population variables can influence the strength of the relationship.[67] The difficulty of evaluating 3D representations was demonstrated.[68] Two studies that reported a relationship between self-efficacy and anatomy performance[63,83] were critically evaluated and compared with two studies that failed to find a relationship.[84,85] The inconsistent reports of anatomical studies influencing self-efficacy are demonstrated.[86–88] Attention to anatomy learning resources, emphasizing spatial ability and self-efficacy, can be supported by Carroll's A Model of School Learning and Bloom's Learning for Mastery. The study will add to our understanding of educational learning resource management in anatomy education.

REFERENCES

1. Densen, P. Challenges and opportunities facing medical education. Trans. Am. Clin. Clim. Assoc. 122, 48–58 (2010).

2. Chan, A. Y.-C. C., Stapper, C. P. M., Bleys, R. L. A. W., Leeuwen, M. van & Cate, O. ten. Are We Facing the End of Gross Anatomy Teaching as We Have Known It for Centuries? Adv Medical Educ Pract 13, 1243–1250 (2022).

3. Logan, A. C., Prescott, S. L. & Katz, D. L. Golden Age of Medicine 2.0: Lifestyle Medicine and Planetary Health Prioritized. J Lifestyle Medicine 9, 75–91 (2019).

4. Varga, S. The criticism of medicine at the end of its "golden age." Theor. Med. Bioeth. 43, 401–419 (2022).

5. Walsh, K. The golden age of medical education? Clin Teach 2, 2–4 (2005).

6. Brändle, T. & Colombier, C. Healthcare expenditure projections up to2050: ageing and the COVID-19 crisis. (2022).

7. AAMC. 2021 Fall Applicant, Matriculant, and Enrollment Data Tables. Preprint at https://www.aamc.org/media/57761/download (2021).

8. AAVMC. 2021-2022 Annual Data Report. Preprint at https://www.aavmc.org/about-aavmc/public-data/ (2022).

9. Norman, G. Medical education: past, present and future. Perspect. Méd. Educ. 1, 6–14 (2012).

10. Shelton, P. G., Corral, I. & Kyle, B. Advancements in Undergraduate Medical Education: Meeting the Challenges of an Evolving World of Education, Healthcare, and Technology. Psychiatr. Q. 88, 225–234 (2017).

11. Buja, L. M. Medical education today: all that glitters is not gold. BMC Méd. Educ. 19, 110 (2019).

12. Abrahams, A., Pienaar, L., Bugarith, K., Gunston, G. & Badenhorst, E. A foundational knowledge assessment tool to predict academic performance of medical students in first-year anatomy and physiology. Adv Physiol Educ 46, 598–605 (2022).

13. Almizani, M. S. et al. Clinicians' and Students' Perceptions and Attitudes Regarding the Anatomical Knowledge of Medical Students. Adv Medical Educ Pract 13, 1251–1259 (2022).

14. Bhattarai, L., KC, S., Shrivastava, A. K. & Sah, R. P. Importance of Cadaveric Dissection in Learning Anatomy For Medical Students: A Review. Med Phoenix 6, 68–72 (2022).

15. Leung, K., Lu, K.-S., Huang, T.-S. & Hsieh, B.-S. Anatomy Instruction in Medical Schools: Connecting the Past and the Future. Adv Health Sci Educ 11, 209–215 (2006).

16. Cottam, W. W. Adequacy of medical school gross anatomy education as perceived by certain postgraduate residency programs and anatomy course directors. Clin Anat 12, 55–65 (1999).

17. Bergman, E. M., Prince, K. J. A. H., Drukker, J., Vleuten, C. P. M. van der & Scherpbier, A. J. J. A. How much anatomy is enough? Anat Sci Educ 1, 184–188 (2008).

18. Memon, I. K. Anatomy education faces challenges in Pakistan. Anat Sci Educ 2, 193–194 (2009).

19. Farrokhi, A., Soleymaninejad, M., Ghorbanlou, M., Fallah, R. & Nejatbakhsh, R. Applied anatomy, today's requirement for clinical medicine courses. Anat Cell Biology 50, 175–179 (2017).

20. Singal, A., Bansal, A. & Chaudhary, P. Cadaverless anatomy: Darkness in the times of pandemic Covid-19. Morphol 104, 147–150 (2020).

21. Lunn-Collier, R., Layman-Lemphane, J. I., Baatjes, K. J. & Correia, J. Clinicians' opinions on the clinical relevance of anatomy education at Stellenbosch University. Anat Sci Educ (2022) doi:10.1002/ase.2175.

22. Linton, A. et al. Enhancing Anatomical Instruction: Impact of a Virtual Canine Anatomy Program on Student Outcomes. Anat Sci Educ 15, 330–340 (2022).

23. Martin, J. F. et al. How Virtual Animal Anatomy facilitated a successful transition to online instruction and supported student learning during the coronavirus pandemic. Anatomia Histologia Embryologia (2022) doi:10.1111/ahe.12799.

24. Lopez, E. K. N., Johnson, J. H., Cunningham, E. P., Warshaw, J. & Baker, E. W. A unique and effective method of anatomy education: Small-group learning with prosected plastinated specimens. J Dent Educ 86, 425–436 (2022).

25. Lachman, N. & Pawlina, W. Reconsidering laboratory-based anatomy within the backdrop of digital transformation: Bringing an old practice into a new world. Anat Sci Educ 15, 439–446 (2022).

26. Gaur, U. et al. Challenges and Opportunities of Preclinical Medical Education: COVID-19 Crisis and Beyond. Sn Compr Clin Medicine 2, 1992–1997 (2020).

27. Maudsley, G. Mixing it but not mixed-up: Mixed methods research in medical education (a critical narrative review). Méd. Teach. 33, e92–e104 (2011).

28. Bressan, V. et al. Mixed-methods research in nursing – a critical review. J. Clin. Nurs. 26, 2878–2890 (2017).

29. Janssen, A., Shaw, T., Goodyear, P., Kerfoot, B. P. & Bryce, D. A little healthy competition: using mixed methods to pilot a team-based digital game for boosting medical student engagement with anatomy and histology content. Bmc Med Educ 15, 173 (2015).

30. Low, E. Z. et al. Assessing medical students' perception and educational experience during COVID-19 pandemic. Ir J Medical Sci 1971 - 1–6 (2022) doi:10.1007/s11845-022-03118-3.

31. Xu, X., Mangina, E. & Campbell, A. G. HMD-Based Virtual and Augmented Reality in Medical Education: A Systematic Review. Frontiers Virtual Real 2, 692103 (2021).

32. Valanci-Aroesty, S. et al. Implementation and Effectiveness of Coaching for Surgeons in Practice – A Mixed Studies Systematic Review. J Surg Educ 77, 837–853 (2020).

33. Awadh, A. B., Clark, J., Clowry, G. & Keenan, I. D. Multimodal Three-Dimensional Visualization Enhances Novice Learner Interpretation of Basic Cross-Sectional Anatomy. Anat Sci Educ 15, 127–142 (2022).

34. Alharbi, Y., Al-Mansour, M., Al-Saffar, R., Garman, A. & Alraddadi, A. Three-dimensional Virtual Reality as an Innovative Teaching and Learning Tool for Human Anatomy Courses in Medical Education: A Mixed Methods Study. Cureus 12, e7085 (2020).

35. Fillmore, E. P., Brokaw, J. J., Kochhar, K. & Nalin, P. M. Understanding the current anatomical competence landscape: Comparing perceptions of program directors, residents, and fourth-year medical students. Anat Sci Educ 9, 307–318 (2016).

36. Messerer, D. A. C. et al. What factors motivate male and female Generation Z students to become engaged as peer teachers? A mixed-method study among medical and dental students in the gross anatomy course. Anat Sci Educ (2021) doi:10.1002/ase.2114.

37. Hong, Q. N., Rees, R., Sutcliffe, K. & Thomas, J. Variations of mixed methods reviews approaches: A case study. Res. Synth. Methods 11, 795–811 (2020).

38. Hadi, M. A. & Closs, S. J. Ensuring rigour and trustworthiness of qualitative research in clinical pharmacy. Int. J. Clin. Pharm. 38, 641–646 (2016).

39. Peters, M. et al. A Redrawn Vandenberg and Kuse Mental Rotations Test - Different Versions and Factors That Affect Performance. Brain Cognition 28, 39–58 (1995).

40. Carroll, J. B. The Carroll Model. Educ Researcher 18, 26–31 (1989).

41. Carroll, J. B. A Model of School Learning. Teach Coll Rec Voice Scholarsh Educ 64, 1–9 (1963).

42. Essa, A. & Mojarad, S. Does Time Matter in Learning? A Computer Simulation of Carroll's Model of Learning. Lect Notes Comput Sc 458–474 (2020) doi:10.1007/978-3-030-50788-6_34.

43. Spearman, C. "General Intelligence," Objectively Determined and Measured. Am. J. Psychol. 15, 201 (1904).

44. Kaufman, A. S., Choi, D., Kapoor, H. & Kaufman, J. C. Intelligence in Context, The Cultural and Historical Foundations of Human Intelligence. 59–92 (2022) doi:10.1007/978-3-030-92798-1_4.

45. Carroll, J. B. & Spearritt, D. A Study of a "Model of School Learning." Preprint at https://eric.ed.gov/?id=ED045477 (1967).

46. Sternberg, R. J. The Intelligent Attitude: What Is Missing from Intelligence Tests. J. Intell. 10, 116 (2022).

47. Anderson, W. F. Relation of Lorge-Thorndike Intelligence Test Scores of Public School Pupils to the Socio-Economic Status of their Parents. J. Exp. Educ. 31, 73–76 (1962).

48. Sjogren, D. D. Achievement as a Function of Study Time. Am. Educ. Res. J. 4, 337–343 (1967).

49. Harnischfeger, A. & Wiley, D. E. Conceptual Issues in Models of School Learning. J. Curric. Stud. 10, 215–231 (1978).

50. Huitt, W. G., Huitt, M. A., Monetti, D. M. & Hummel, & J. H. A Systems-based Synthesis of Research Related to Improving Students' Academic Performance. (2009).

51. Arlin, M. Teacher Responses to Student Time Differences in Mastery Learning. Am. J. Educ. 90, 334–352 (1982).

52. Bloom & S., B. Learning for Mastery. Instruction and Curriculum. Regional Education Laboratory for the Carolinas and Virginia, Topical Papers and Reprints, Number 1. Evaluation Comment 12 (1968).

53. Huyar, D. A. & Esin, M. N. Effects of Bloom's mastery learning model based on CBRN preparedness program on learning skills of nursing students: A randomized controlled trial. Nurse Educ. Pr. 67, 103568 (2023).

54. Slavin, R. E. A Model of Effective Instruction. Educ Forum 59, 166–176 (1995).

55. Slavin, R. E. Mastery Learning Reconsidered. Rev Educ Res 57, 175–213 (1987).

56. Berger, M. et al. Simulation-Based Mastery Learning Course for Tube Thoracostomy. Mededportal 18, 11266 (2022).

57. Thompson, A. R. & Giffin, B. F. Higher-Order Assessment in Gross Anatomy: A Comparison of Performance on Higher- Versus Lower-Order Anatomy Questions between Undergraduate and First-Year Medical Students. Anat Sci Educ 14, 306–316 (2021).

58. Robertson, E. M., Thompson, K. L. & Notebaert, A. J. Perceived Benefits of Anatomy Coursework Prior to Medical and Dental School. Anat Sci Educ 13, 168–181 (2020).

59. Bergman, E. M., Prince, K. J. A. H., Drukker, J., Vleuten, C. P. M. van der & Scherpbier, A. J. J. A. How much anatomy is enough? Anat Sci Educ 1, 184–188 (2008).

60. Ranasinghe, P., Ellawela, A. & Gunatilake, S. B. Non-cognitive characteristics predicting academic success among medical students in Sri Lanka. BMC Méd. Educ. 12, 66 (2012).

61. Higgins-Opitz, S. B. & Tufts, M. Performance of first-year health sciences students in a large, diverse, multidisciplinary, first-semester, physiology service module. Adv. Physiol. Educ. 38, 161–169 (2014).

62. Vitali, J., Blackmore, C., Mortazavi, S. & Anderton, R. Tertiary Anatomy and Physiology, A Barrier for Student Success. Int. J. High. Educ. 9, 289 (2020).

63. Burgoon, J. M., Meece, J. L. & Granger, N. A. Self-efficacy's influence on student academic achievement in the medical anatomy curriculum. Anat Sci Educ 5, 249–255 (2012).

64. Lohman, D. F. Spatial Ability and G. Preprint at (1993).

65. Lufler, R. S., Zumwalt, A. C., Romney, C. A. & Hoagland, T. M. Effect of visual–spatial ability on medical students' performance in a gross anatomy course. Anat Sci Educ 5, 3–9 (2012).

66. Langlois, J., Bellemare, C., Toulouse, J. & Wells, G. A. Spatial Abilities Training in Anatomy Education: A Systematic Review. Anat Sci Educ 13, 71–79 (2020).

67. Roach, V. A. et al. Correlating Spatial Ability With Anatomy Assessment Performance: A Meta-Analysis. Anat Sci Educ 14, 317–329 (2021).

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

69. Uttal, D. H. et al. The Malleability of Spatial Skills: A Meta-Analysis of Training Studies. Psychol Bull 139, 352–402 (2013).

70. Na, Y. et al. Spatial Visualization of Human Anatomy through Art Using Technical Drawing Exercises. Anat Sci Educ 15, 587–598 (2022).

71. Labranche, L., Wilson, T. D., Terrell, M. & Kulesza, R. J. Learning in Stereo: The Relationship Between Spatial Ability and 3D Digital Anatomy Models. Anat Sci Educ 15, 291–303 (2022).

72. Huk, T. Who benefits from learning with 3D models? the case of spatial ability. J Comput Assist Lear 22, 392–404 (2006).

73. Bogomolova, K. et al. Stereoscopic three-dimensional visualisation technology in anatomy learning: A meta-analysis. Med Educ 55, 317–327 (2021).

74. Dickson, J., Gardiner, A. & Rhind, S. Veterinary Anatomy Education and Spatial Ability: Where Now and Where Next? J Vet Med Educ 49, 297–305 (2022).

75. Vandenberg, S. G. & Kuse, A. R. Mental Rotations, a Group Test of Three-Dimensional Spatial Visualization. Percept Motor Skill 47, 599–604 (1978).

76. Hariri, S., Rawn, C., Srivastava, S., Youngblood, P. & Ladd, A. Evaluation of a surgical simulator for learning clinical anatomy. Méd. Educ. 38, 896–902 (2004).

77. Jacquesson, T. et al. Stereoscopic three-dimensional visualization: interest for neuroanatomy teaching in medical school. Surg Radiol Anat 42, 719–727 (2020).

78. Chytas, D., Piagkou, M. & Natsis, K. Stereoscopic three-dimensional visualization: interest for neuroanatomy teaching in medical school. Surg. Radiol. Anat. 42, 1381–1382 (2020).

79. Weeks, J. K. et al. Harnessing Augmented Reality and CT to Teach First-Year Medical Students Head and Neck Anatomy. Acad. Radiol. 28, 871–876 (2021).

80. Chytas, D., Salmas, M., Skandalakis, G. P. & Troupis, T. G. Augmented and virtual reality in anatomy education: Can they be effective if they do not provide immersive experience? Anat Sci Educ 15, 431–433 (2022).

81. Bandura, A. Self-efficacy: Toward a unifying theory of behavioral change. Psychol. Rev. 84, 191–215 (1977).

82. Morgan, G. A., Barrett, K. C., Leech, N. L. & Gloeckner, G. W. IBM SPSS for Introductory Statistics: Use and Interpretation. 106–134 (2019) doi:10.4324/9780429287657-6.

83. Meguid, E. M. A., Smith, C. F. & Meyer, A. J. Examining the Motivation of Health Profession Students to Study Human Anatomy. Anat Sci Educ 13, 343–352 (2020).

84. Finn, G. M. et al. Exploring relationships between personality and anatomy performance. Anat Sci Educ 8, 547–554 (2015).

85. Langfield, T., Colthorpe, K. & Ainscough, L. Online instructional anatomy videos: Student usage, self-efficacy, and performance in upper limb regional anatomy assessment. Anat Sci Educ 11, 461–470 (2018).

86. Koppes, D. M. et al. Anatomy in the daily practice of the gynecologist, essential or just window dressing? Anat. Sci. Educ. 16, 497–503 (2023).

87. Chen, W.-T. et al. Does ultrasound education improve anatomy learning? Effects of the Parallel Ultrasound Hands-on (PUSH) undergraduate medicine course. Bmc Med Educ 22, 207 (2022).

88. Vandenbossche, V. et al. Ultrasound versus videos: A comparative study on the effectiveness of musculoskeletal anatomy education and student cognition. Anat. Sci. Educ. (2023) doi:10.1002/ase.2311.

CHAPTER II – HOW VIRTUAL ANIMAL ANATOMY FACILITATED A SUCCESSFUL TRANSITION TO ONLINE INSTRUCTION AND SUPPORTED STUDENT LEARNING DURING THE CORONAVIRUS PANDEMIC[1][2][3]

2.1. INTRODUCTION

The COVID-19 pandemic significantly disrupted education at all ages and stages worldwide (Ray & Srivastava, 2020; United Nations Educational, Scientific and Cultural Organization [UNESCO], 2021). The rapid shift to remote learning in March 2020 presented a challenge to educators who still needed to provide students with quality learning experiences without in-person learning opportunities. While psychologists and sociologists will continue to evaluate the long-term impact of the pandemic on learners (Armstrong-Mensah, Ramsey-White, Yankey, & Self-Brown, 2020), educators must critically evaluate the effectiveness of their curricular design in response to the pandemic. Although online teaching occurred well before the start of the COVID-19 pandemic and there are evidence-based best practices for designing remote curricula (Evans et al., 2020; Said & Schwartz, 2021), anatomy is a subject that both educators and students agree is best taught through hands-on cadaveric learning (Mohamed, 2020; Onigbinde, Chia, Oyeniran, & Ajagbe, in press). Therefore, anatomy faculty with cadaver-

---

[1] This chapter is a modified version of *How Virtual Animal Anatomy Facilitated A Successful Transition To Online Instruction And Supported Student Learning During The Coronavirus Pandemic* published in *Anatomia, Histologia, Embryologia* and has been reproduced here with the permission of the copyright holder.

[2] Reference: Martin, J. F., Arnold, O. R., Linton, A., Jones, J. D., Garrett, A. C., Mango, D. W., Juarez, K. A., Gloeckner, G., & Magee, C. (2023). How Virtual Animal Anatomy facilitated a successful transition to online instruction and supported student learning during the coronavirus pandemic. *Anatomia, Histologia, Embryologia*, 52(1), 36–49.

[3] Link: https://doi.org/10.1111/ahe.12799

based laboratory courses were presented with a significant challenge in March 2020 to create equivalent learning experiences without cadaveric access.

Remote learning resources available to anatomy faculty at the start of the pandemic included a variety of virtual meeting software (e.g., Zoom, Microsoft Teams, etc.), learning management systems (e.g., Canvas, Blackboard, Doodle, etc.), and software programs to facilitate computer-assisted learning (CAL) of anatomy. Anatomy faculty worldwide shared digital and e-learning resources, coming together as communities of practice to support one another. In response to the pandemic, one CAL program that was made freely available from March to June 2020 was Colorado State University's (CSU) Virtual Animal Anatomy (VAA) software, which was being used at more than 100 schools and by more than 12,000 new student users by the end of March 2020 (College of Veterinary Medicine & Biomedical Sciences, 2021). This software suite includes the Virtual Canine Anatomy (VCA) program and is a cadaver-based atlas with additional equine, feline, and bovine anatomical content. While there is strong evidence that VCA and other CAL software can enhance in-person anatomy learning (Durosaro, Lachman, & Pawlina, 2008; Garrett et al., 2019; Linton, Schoenfeld-Tacher, & Whalen, 2005; Linton et al., 2021; Magee et al., 2015; Tam et al., 2010) by increasing student excitement and assessment outcomes (Yang, Yang, Yang, & Fan, 2020), before March 2020, the VAA had never been evaluated as a replacement for cadaveric learning. The anatomical specimens within the VAA are professionally dissected and photorealistic that, combined with the ability to manipulate and rotate the virtual anatomical specimens, provides VAA users with a virtual laboratory experience. Virtual reality is known to facilitate anatomy learning when students do not have access to cadavers for study. Implementation of effective virtual resources can provide the flexibility and personalization needed for students to best support themselves (Evans et al., 2020; Linton et al., 2005) and promote a more active learning environment since students are able to view and manipulate full-body models using virtual cadavers (Custer & Michael, 2015).

Another challenge for anatomy educators in the online environment is conducting assessments. Summative assessments, which evaluate students' overall progress, are an important part of their overall advancement and have been the focus of more traditional curricula. Formative assessments help develop students' intellectual capabilities and identify their strengths and weaknesses as well as help students determine which learning strategy works best for them. Active learning strategies embedded in a scaffolded curriculum are most effective when students combine activities that provide them with feedback on their ability to utilize their newly gained knowledge (May & Silva-Fletcher, 2015). In an in-person anatomy learning environment, weekly specimen-based quizzes or dissection evaluations on the prior week's laboratory objectives can be effective formative assessments to ensure that learners are progressing through a scaffolded curriculum. In an online environment, the perceived effectiveness of formative assessments, among other factors such as institutional support, is positively associated with the impact and effectiveness of the online curricula (Tartavulea, Albu, Albu, Dieaconescu, & Petre, 2020). Adaptive e-Learning (AEL) platforms, such as Firecracker, can be used as a comprehensive assessment tool to support both face-to-face and remote learning by helping instructors prepare students for summative assessments and identifying a student's progress relative to that of their peers (Gupta et al., 2020). The human anatomy AEL from Firecracker has a well-established question bank, generates new content every week, and continuously reviews item characteristics with an appropriate level of difficulty to prepare students for summative assessments (Gupta et al., 2020). Given that this type of AEL software may increase learning efficacy and efficiency and improve student preparedness for summative assessments (Gupta et al., 2020), it is important to note that no such software is widely available for animal anatomy curricula. Therefore, instructors wishing to provide formative assessments in a remote environment must develop tools within their learning management system to support student learning.

In March of 2020, the domestic animal anatomy courses at CSU were 8 weeks into a 16-week semester when lockdown orders and the transition to remote instruction began. The teaching team had 7 days over the natural semester break ("Spring Break") to redesign the remainder of the semester. Our underlying hypothesis was that the VAA program would effectively support student learning in both the in-person and online environments. Recognizing the importance of critically evaluating the new course curriculum, we used our historical knowledge of the student population and curriculum, as well as student surveys at the end of the semester, to determine the following: (a) if students considered the transition to online learning to be a success, and (b) if the VAA supported animal anatomy learning in a fully remote learning environment.

2.2. MATERIALS AND METHODS

2.2.1. Animal And Human Subjects

The strictest humane treatment of animals has always been used for the development of the VAA software programs and all anatomy-related curricula at CSU. All animals were euthanized prior to dissection. All animal work was pre-approved by the CSU Institutional Animal Care and Use Committee, which is in compliance with the Animal Welfare Act as enforced by the U.S. Department of Agriculture, and the Public Health Service Policy on Humane Care and Use of Laboratory Animals, which is administered by the Office for

Laboratory Animal Welfare of the National Institutes of Health. CSU is also accredited by the Association for Assessment and Accreditation of Laboratory Animal Care International. All work related to human subjects was approved under Protocol ID 20-10330H by the CSU Institutional Review Board, in compliance with the Common Rule [45 CFR 46.101(b)] and was considered "exempt" on the basis that the research involved the collection or study of existing data, documents, and records in such a manner that subjects could not be identified, either directly or through identifiers linked to the subjects.

A total of 346 students over 3 years (n = 108 in 2018, n = 120 in 2019, n = 118 in 2020) participated in one semester of undergraduate domestic animal anatomy at CSU. Course performance data from the university's learning management system (Canvas) was de-identified and used for analysis. The VAA was freely provided to the students via the learning management system (LMS), and information about VAA use was also de-identified and used for analysis. No gender, ethnicity, or socioeconomic data were collected about the students enrolled during this period.

2.2.2. Domestic Animal Gross Anatomy: A Scaffolded Undergraduate Anatomy Curriculum

Domestic Animal Gross Anatomy is an in-person course offered to students who have previously completed an undergraduate-level animal biology course. Canine gross anatomy is used as the foundation for comparative anatomy instruction of equids, bovids, camelids, felids, and suids via 3 hours per week of didactic lectures. Students attend 2 hours per week of a laboratory session exploring prosected fresh and preserved cadavers with the guidance of faculty, graduate teaching assistants (TAs), and undergraduate TAs as part of the regular course curriculum. All cadaver-based formative assessments and all summative assessments are given during the 2-hour lab. In addition, 9 hours per week of scheduled open laboratory access staffed with TAs are offered to students seeking additional guided or independent hands-on study of the anatomical resources available in the laboratory.

The university LMS is always available to students via a standard web browser on a computer, tablet, or smartphone application and acts as a central hub for distributing course materials and communications. Students may access their syllabi, grades, download copies of lecture presentations or a complete digital copy of their laboratory manual, watch recordings of lectures presented so far within the course, and directly and privately communicate with faculty and TAs. The LMS is also used for administering some formative assessments and the computer-based portion of all summative assessments.

Course content is divided into four units correlated to topics associated with regional dissection in a 15-week semester. The first half of the semester begins with the pelvic limb followed by the thoracic limb, and the second half of the semester historically includes a unit on the thorax, abdomen, and pelvis, followed by a unit on the head and neck region. Each unit receives 4 weeks of study except for the head and neck component, which receives 3 weeks. Summative assessments are a significant part of student course performance, with one 100-point examination at the conclusion of each unit (for a total of four such 100-point assessments), comprising two 50-point sections, one of which is a computer-based assessment of lecture material and the other of which is a 50-point cadaver-based laboratory assessment that integrates lecture material. Then, the course also includes multiple low-stakes, high-frequency formative assessments that are collectively worth 50 points, for a total of 450 points.

During the 2020 Spring semester, all on-campus CSU courses were notified of the transition to online remote study following the thoracic limb examination (Week 8). Initially, COVID-19 lockdowns were not anticipated to last for longer than 3 weeks and course faculty prioritized the value of in-person teaching of comparative gastrointestinal anatomy over head and neck anatomy in the event that in-person learning was to resume. Therefore, in 2020, in contrast to the 2018 and 2019 semesters, the head and neck unit was transitioned to the third position, followed by the unit on the thorax, abdomen, and pelvis. Moreover, lectures on both of these units were provided as 50-minute asynchronous recordings captured from the 2019 semester. The in-person laboratory sessions were replaced with 42 hours per week of optional online open laboratory sessions hosted by TAs on Zoom, guiding students as they interacted with digital content such as the VAA, lecture materials, and other image-based resources. Subsequently, as is now common knowledge, in-person teaching did not resume and lockdowns remained in place until the end of the semester. Additional relevant changes for 2020 are noted in the following sections.

2.2.2.1. Formative in-class assessments: iClicker questions.

Historically, in some in-person lectures, students have been given an on-screen question to which they were given approximately 1 min to submit their answers via the iClicker platform (see Table 2.1). iClicker uses a handheld remote that allows students the opportunity to respond in real-time to a presented multiple-choice question. Instructors use responses to monitor attendance, participation, and content engagement without drawing critical public attention to individual students. While in 2018 and 2019, iClickers contributed 10 points to the total course points with 0.5 points for participation and 0.5 points for accuracy, iClicker questions were only offered in 2020 prior to the shift to online learning; therefore, these 10 points were automatically awarded to all students and designated as “participation” in 2020.

2.2.2.2. Formative online assessments: Check-in quizzes.

Starting in 2019, students were given 1-week periods within the LMS to answer 11 open-resource formative assessments ("check-in quizzes") on lecture content (see Table 2.1). The first check-in quiz focused on course dynamics, requiring students to be familiar with their syllabus and laboratory policies. In Week 9 of the 2020 semester, an additional course dynamics check-in quiz was added to ensure students had reviewed the new syllabus and online policies following the transition to remote learning, resulting in a total of 12 check-in quizzes for the semester. All other check-in quizzes asked questions of comparable complexity to lecture examination assessments. Each check-in quiz was worth 2.5 points.

2.2.2.3. Formative in-person assessments: Laboratory quizzes. Each course unit introduces approximately 500 anatomical terms correlated to structures that may be assessed on that unit's summative examination (see Table 2.1). These terms may include directional terms, layperson terminology, and the *Nomina Anatomica Veterinaria* (NAV) nomenclature associated with specific anatomical structures. For each weekly laboratory exercise, a "short objectives" list comprising 10 to 30 terms is provided in the laboratory manual to direct student learning. During non-examination weeks, students simulate an abbreviated closed-resource formative laboratory assessment ("laboratory quizzes") where they are asked to answer five questions regarding tagged cadaveric structures using terms from the previous week’s short objectives list. These questions may involve having students identify tagged structures or describe relationships to the tagged structures. Each of the laboratory quizzes contributes up to 5 points to the final course grade.

Students who score less than 2 out of 5 possible points on a laboratory quiz are contacted by the course instructor to investigate opportunities for assistance. Interventions begin at Week

2, after the student has completed their first laboratory quiz. All students are provided the opportunity to receive assistance from the teaching team with study strategies, including VAA and laboratory resource utilization.

Seven of these formative, cadaver-based assessments (two for each 4-week unit and one for the 3-week unit) were presented to students in 2018 and 2019. After the transition to remote study in 2020, the remaining cadaver-based laboratory quizzes were replaced with two computer-based quizzes for each of the remaining units delivered via the course LMS. Each of the four new quizzes was a timed, open-resource assessment that correlated to that week's laboratory exercises and was worth 5 points. Although short objectives were still provided to encourage students to study, the testable material was not limited to these objectives.

2.2.2.4. Formative hybrid lecture-laboratory assessments: Guided quizzes.

After the transition to remote teaching in 2020, students were provided opportunities to assess their ability to integrate lecture and laboratory content via "guided quizzes" (see Table 2.1). These open-resource assessments followed a linear progression such that answers to previous questions directed responses to follow-up questions and were due the day before each unit examination. Guided quizzes were created and delivered via LMS; question types included "hot spot" structure identification, categorization or sorting, sequential ordering, and multiple answer selection. A total of six guided quizzes (three for each unit) were provided. Quizzes ranged from 12 to 15 questions for a total of 7.5 points per guided quiz. Quizzes were modeled after adaptive e-learning platforms even though no adaptive qualities were embedded in the course LMS platform. After each question, students were provided with detailed answer feedback and additional context or links to course resources to integrate various topics across laboratory and lecture material. After completing each section, students were asked if they were ready to proceed to the next topic within the guided quiz. Students were not timed for these open-resource formative assessments and were given two opportunities to complete them.

2.2.3. Analysis of Virtual Animal Anatomy

All domestic animal anatomy students are provided access to CSU's VAA application through the university LMS. Previously distributed as Virtual Canine Anatomy, the VAA contains equine, feline, and bovine anatomy content in various stages of program development. The VAA suite of anatomy software functions as an online atlas, providing "layers" of interactive photographs of prosected specimens and photorealistic 3D models to guide users through the progression of regional dissections. Students are encouraged to use the VAA to prepare for upcoming laboratory sessions or independent review by directly selecting anatomical structures within photos, searching for laboratory objectives by common or NAV nomenclature, reviewing animated or video dissection instructions to help them understand how the cadaver was dissected to achieve each layer, reading descriptions of anatomical structures, and/or by electing to take the embedded (non-graded) gross anatomy quizzes.

Student utilization of the VAA in the undergraduate Domestic Animal Gross Anatomy course previously had been collected as part of a series of learning ecology assessments conducted by The Institute of Learning and Teaching (TILT) at CSU and a significant course redesign that was implemented from 2014–2016. From 2014–2018, the VCA was distributed to CSU students as an Adobe Flash encoded software loaded on a USB drive; until 2019, students also had access to the virtual program on the desktop of computers in the laboratory. At the start of the second week of the 2018 Spring semester, the VAA software was converted to HTML and student VAA access was directly integrated into the LMS with entry to the VAA through a portal separate from the course page. This modification provided an opportunity for quantitative analysis of VAA usage throughout the semester without relying on student surveys. Beginning in 2020, access to VAA was only available through the LMS.

Utilization of the VAA software was measured by the number of times per week a student logged into the VAA site via the LMS. Due to the 2020 reversal of the thorax, abdomen, and pelvis unit with the head and neck unit during the second half of the semester, utilization was categorized as occurring during the first half (Examinations 1 and 2) or second half (Examinations 3 and 4) of a given year. Any VAA utilization during the 1-week break between the two parts of the semester was excluded from the analysis. The first week of VAA access was not captured during the 2018 semester due to delayed LMS integration at the start of that term.

2.2.4. End-of-Semester Student Survey 2020

Students were offered a 2-point extra credit opportunity during the week before the 2020 final examination if they completed an optional, anonymous two-part course survey via the course LMS. Students were provided 1 week to complete the survey and were not required to complete it in a single session. Students could elect to complete some or all of the survey questions before submitting it. In Part I, a 7-point Likert scale (1 = strongly disagree to 7 = strongly agree) was used to determine their agreement with statements about their experience with the course during the semester. There were no negatively-phrased statements (see Table 2.2).

Part II of the survey presented a list of course resources, and students were asked to select the resources they believed were most helpful to them before and after the COVID-19 transition. Students were prompted to provide comments in a text field to share any additional thoughts about the resources they had used. Before submitting the survey, students were also given the opportunity to provide additional comments about the course, the resources provided with the course, and the COVID-19 transition.

2.2.5. Statistical Analysis

All student data (i.e., grades, VAA access frequency, and survey responses), were curated in Microsoft Excel and made anonymous by a third party. Student examination performance as a grade represented as a percentage correlating to the first (Examinations 1 and 2) or second half (Examinations 3 and 4) of the semester were averaged to generate a mean examination score for the respective portion of the semester. VAA frequency access for the correlating part of the semester each year was also summed. In 2020, one student accessed the VAA 418 times (median access 27 times/student) prior to transitioning to remote learning and was excluded as an outlier. For each variable, the skewness was computed to assist with appropriate test selection. If the absolute value of skewness was greater than 1.0, then data did not sufficiently meet the assumptions of normality and nonparametric statistical methods were used and descriptive data were reported using median and interquartile range. Levene's test was computed for normally distributed data to test the assumption that variances were equal, and a t-test of independent samples was performed to assess differences between group means. For metrics that did not meet the conditions of normality, the Mann–Whitney U test was used to assess differences in central tendency. Pearson's correlation coefficient (r) was computed to assess relationships between two normally distributed variables, while Spearman's rank rho correlation coefficient (r_s) was used to assess relationships if at least one variable was not normally distributed. Effect sizes associated with r and rho with absolute values of s less than .3 were considered less than typical, of greater or equal to .3 but less than .5 were considered typical, and of greater or equal to .5 were considered larger than typical (Morgan, Barrett, Leech, & Gloeckner, 2020). For all results, $p <$

0.05 was considered significant. IBM SPSS Statistics (Version 26) was used for all computations.

2.3. Results

2.3.1. Results of the Student Survey

2.3.1.1. Students perceived the 2020 COVID-19 transition as successful.

Of the 118 students who completed the 2020 semester, 107 students submitted the optional, extra credit end-of-semester survey. In Part I, most students (92.5%) agreed that the COVID-19 transition was a success, indicated that they enjoyed animal anatomy (98.1%, *n* = 105), and responded that their interest in animal anatomy had increased (96.3%, *n* = 103) after having taken the course. Students who completed the course survey mostly agreed (91.6%, *n* = 98) that they were interested in taking additional animal anatomy courses. Tables 2.2 and 2.4 present this data.

In Part II, 92.5% (*n* = 99) of students agreed that the COVID-19 transition was a success, but students were more likely to agree that the course time commitment, available resources, and their course performance were better before the COVID-19 transition as compared to after the transition (see Table 2.3). There was a moderate association between pre- and post-COVID-19 time commitment, and while both were associated with perceived transition success, only the post-COVID-19 transition would be considered a larger than typical effect size. Additionally, only belief in personal performance post-COVID-19 was correlated with perceived course success. However, there was no observed association between pre- and post-COVID-19 responses regarding the perceived value of the available resources. Student agreement that post-

COVID-19 resources facilitated learning was correlated with reports that the transition was a success.

2.3.1.2. Perceived resource helpfulness in COVID-19 transition success.

Resources perceived as helpful changed after the COVID-19 transition (see Table 2.4). The resources ranked most frequently as helpful before the COVID-19 transition were face-to-face lectures (ranked helpful by 86.0% of students) and open laboratory sessions (ranked helpful by 85.1% of students). After the transition, the resources ranked as helpful most frequently were guided quizzes (ranked helpful by 84.1% of students) and VAA (ranked helpful by 82.2% of students).

Not all resources were associated with the agreement that the transition was a success. (Please see Table 2.5 for a presentation of the correlations.) Students who reported that they perceived the face-to-face lecture as helpful before the COVID-19 transition were significantly less likely to agree that the transition was a success. No other pre-COVID-19 resources were associated with perceived transition success. However, four resources post-COVID-19 were associated with reports of a successful transition. Weekly check-in quizzes, reported as helpful by 48.6% (n = 52 of 107) of students after the transition, were correlated with transition success. Three other items associated with transition success were online open laboratory sessions, course instructors, and the written laboratory exercise; however, less than a quarter of the students reported each of these resources as helpful after the transition (as shown on Table 2.4).

2.3.1.3. Students provided comments on COVID-19 transition success and course resources.

Of the 107 students who completed the survey, 68.2% (n = 73) provided comments about the helpfulness pre-COVID-19 of the course resources; and 52.3% (n = 56) provided comments about the helpfulness post-COVID-19 of the course resources. A qualitative analysis of student comments was performed to identify perspectives that may have influenced how they reported perceived helpfulness. In the comments provided, comparing the perceived helpfulness of pre-COVID-19 resources to those still available to students after the online transition away from in-person, cadaver-based laboratories, open laboratory dropped from 29% (31 out 73) to 17.9% (10 out of 56) of respondents, as did the helpfulness of peer teaching assistants from 16.8% (18 out of 73) to 10.7% (6 out of 56). There was also a decrease in perceived helpfulness of the course instructors from 10.9% (8 out of 73) to 7.1% (4 out of 56) and a decrease in the perceived helpfulness of classmates from 15.1% (11 out of 73) to 8.9% (5 out of 56) in the post-COVID-19 environment. Major themes observed in the qualitative analysis included the use of VAA to support other course resources (e.g., lecture, open laboratory), the value of guided quizzes, and the importance of group learning to guide students' study and evaluate students' knowledge.

In the comments provided regarding post-COVID-19 (n = 56), students most frequently (28.6%; 16 out of 56) mentioned the guided quizzes as a helpful resource after the COVID-19 transition, while others (21.4%; 12 out of 56) mentioned the VAA program. Of the 16 students who mentioned guided quizzes as helpful, 11 of these did not mention any other resources as helpful, while five of them referenced the VAA.

After March 2020, students were still required to apply the didactic course material, but without access to physical anatomical specimens. For six students, the VAA was perceived as a supplemental lecture resource, whereas two others mentioned in their comments using it to directly assess their knowledge. Two students commented on the inability to replicate a cadaveric experience with an online resource, indicating that the manner in which a resource is utilized may influence if it is perceived as helpful. Example comments from students included the following:

- “The diagrams we filled out in lectures were extremely helpful being online and not being able to interact with the cadavers! It helped me visualize things a lot better when applying this to the VCA.”
- “The main resources I facilitated [sic] before spring break were open laboratory, the virtual program, and lecture notes. [After spring break] the virtual programs and Echo 360 lectures were extremely helpful!”
- "I don't feel as though I got as much out of the class after spring break as I did before. This is due in part to not being able to attend open laboratory and seeing cadavers up close. There is quite a difference seeing an image and being able to move around from section to section on a specimen for yourself."

The guided quizzes were the most frequently reported helpful resource after the COVID-19 transition, as shown on Table 2.4, but were not associated with success in the COVID-19 transition (see Table 2.5). Student comments regarding the guided quizzes suggest that these quizzes may have helped direct student focus and aid in VAA utilization. Examples of these comments include the following:

- “The new guided quizzes were extremely helpful and helped me to understand what I needed to know for the exam. I would strongly suggest keeping the guided quizzes even after classes go back to normal.”
- “The guided quizzes and lab quizzes helped a lot. The VCA was also my best friend through this whole online time and I don't know how I would have done without it. Its [sic] a fantastic learning tool.”

Although the online laboratory sessions, teaching assistants, and course faculty were not frequently ranked as helpful resources post-COVID-19, 57.1% of students (32 out of 56) described the value of these resources in terms of guiding their study and evaluating their knowledge. In the post-COVID-19 environment, students described quizzes (16 out of 56; 28.6%), teaching assistants (7 out of 56; 12.5%), classmates (5 out of 56; 8.9%), and faculty (4 out of 56; 7.1%) as valuable resources for feedback on study direction, suggesting that students may perceive guidance and evaluation as more helpful than content presentation. Student comments regarding resource utilization included the following:

- "Before spring break, something that helped me a lot was having a study group. We would go over the information, and it was really useful to catch things I might have missed otherwise. It was especially useful because I overloaded my schedule and couldn't go to as much open laboratory as I should have. Something that was really helpful to me after spring break was the TA open laboratory sections [in Zoom] because it was actually more one on one that the lab sessions [that] were in person. I actually had a whole hour to myself with a TA several times during those sessions which really helped me nail down what I needed to focus on and what I didn't know yet."
- "The availability of open lab [after the transition] was amazing; I was really impressed with how helpful and supportive my classmates were and I miss getting to work with them; the instructors [sic] knowledge and support and the opportunities to check our understanding and preparedness through various quizzes."
- "[Prior to the transition] … forming [a] study group with classmates is also very helpful because we could teach others and during that process we are all learning. [After the

> transition] the VAA seemed like the most important and helpful tool during virtual classes."

Asynchronous lectures after the COVID-19 transition were not correlated with transition success (see Table 2.5), despite 69.2% of students reporting they were helpful (see Table 2.4). Ten students commented on the positive impact of the asynchronous lectures following the COVID-19 transition. However, four students noted that the archived lectures from previous semesters that had been provided as the asynchronous post-COVID transition lectures were not ideal for the remote learner. This sentiment is expressed in the following quote from a student:

> The video recordings from previous years were very limiting, as we could not clearly see where the lecturer was referring on the slide, especially if they were pointing directly on the slide. I would have loved to have new lecture recordings that would point out what is being talked about in the slides.

2.3.2. Virtual Animal Anatomy Access From 2018 to 2020

Examination performance data in 2018 and VAA access data in 2019 were skewed; therefore, all student performance and VAA access data were treated as nonparametric. Wilcoxon signed-rank tests were used to compare VAA access frequency differences between the first and second half of the semester for each year; in all three years, students accessed the VAA significantly more often in the first half of the semester than the second half of the semester (2018: N = 112, $z = -4.42$, $p < 0.01$, $r = -.42$; 2019: N = 123, $z = -8.23$, $p < 0.01$, $r = -.75$; 2020: N = 117, $z = -8.51$, $p < 0.01$, $r = -.79$). A Mann-Whitney U test was calculated to assess if there were differences in VAA access in either part of the semester between years. The VAA access mean rank for 2018 students was higher than that of 2019 or 2020 students for both halves of the semester. However, the difference in mean ranks for VAA access in 2020 was only significantly higher than 2019 for the first half of the semester (see Figure 2.1).

Examination scores increased significantly in the second half of the 2018 semester ($z = -2.85$, $p < 0.01$, $r = -.27$) and although they increased in 2019, this observation was not significant ($z = -0.15$, $p = 0.88$, $r = -0.1$). This relationship was not assessed in 2020 due to the modifications to the examinations in the second half of the semester. From 2018 to 2020 there was no difference in student examination performance within the first half of the semester.

A scatterplot was used to demonstrate the relationship between VAA access and corresponding examination performance; positive linear relationships were observed for 2018 and 2019. No notable correlations were found between VAA access during the first and second halves of the 2020 semester (see Figure 2.2). Using Spearman's rank correlation coefficient (Spearman's rho), a typical, positive correlation was also found between VAA use and the average of the examination scores when 2018 and 2019 data were combined for both the first half of the semester, r_s (231) = .31, $p < 0.01$, and the second half of the semester, r_s (231) = .32, $p < 0.01$.

2.4. DISCUSSION

The COVID-19 pandemic created health and economic crises around the world that inevitably impacted higher education practices. Specifically, accommodations required to maintain instruction through the pandemic necessitated adjustments to traditional anatomy instruction. However, if "necessity is the mother of invention," then this period has been characterized by creativity, collaboration, and innovation. Anatomy instruction has faced numerous challenges during the pandemic (Gummery, Cobb, Mossop, & Cobb, 2018; McBride

& Drake, 2018; Turney, 2007; Wilson, Kaza, Singpurwalla, & Brooks, 2020), and yet anatomy faculty are known for developing cutting-edge technologies and integrating them into novel teaching methods to improve student learning (Collins, 2008; Erickson, 2020; Linton et al., 2005; Reeves, Aschenbrenner, Wordinger, Roque, & Sheedlo, 2004; Shah, 2020; Zargaran, Turki, Bhaskar, Spiers, & Zargaran, 2020). The pandemic forced anatomy programs that had relied on in-person and hands-on cadaveric experiences to leverage existing—or create new—online and asynchronous learning tools (World Association of Veterinary Anatomists, 2020). More than a year after school closures began due to COVID-19, complete and partial school closures still impact more than 900 million learners worldwide (UNESCO, 2021). As in-person learning begins to resume some degree of normalcy, what we have learned from the pandemic will also forever change our approach to anatomy education.

CSU is internationally known for the development of cadaver-based virtual anatomy software. In addition to its high-ranking professional veterinary medicine program (U.S. News & World Report, 2019), the school also provides cadaver-based domestic animal anatomy instruction to undergraduate students and relies heavily on the VAA to supplement anatomy instruction. Compared with a weekly, in-person laboratory interaction, the remote learning environment following the March 2020 transition made connecting with students more difficult for the teaching team; nonetheless, students enrolled in the Domestic Animal Gross Anatomy course considered the COVID-19 online transition to be a success. Much of this success can be attributed to students and instructors using the VAA as a tool to support cadaveric learning in the online environment. Fortunately, CSU students and faculty were already familiar with the VAA and were able to readily leverage the 3D object rotation feature, in conjunction with interactive virtual meeting platforms, to replicate the laboratory environment for cadaver exploration and structure identification.

In the transition to online learning in March 2020, there was limited time for the creation of new content. As a result, course leadership prioritized developing guided quizzes over recording new lectures for the course. Students highly valued the guided quizzes and recorded lectures following the transition. Still, some indicated that the value of the asynchronous lecture recordings could have been improved with videos intentionally created for an online audience rather than relying on recordings of previous in-person lectures. This feedback shaped the 2021 semester, for which we created additional guided quizzes and developed new, intentionally recorded, asynchronous lectures for each unit.

The virtual open laboratory sessions were also highly valued by students following the online transition. It should be noted that students indicated that they were often the only person in the virtual space with the peer teaching assistant; however, this may also reflect the more than 40 hours of virtual open laboratory time that were created to support student learning each week. Teaching assistants staffed these periods with supervision from the course director. Early perceptions of these sessions were that small group learning with no more than two to three students per TA was better for student engagement. While visits to the open labs versus virtual study with peers was not evaluated, the virtual open labs provided students with flexibility and were a valued resource for learning in the post-COVID-19 transition environment. Although lectures for this course remain remote and asynchronous in delivery in 2021, in-person cadaveric-based laboratory learning has resumed. Nonetheless, the perceived value of the virtual sessions in 2020 resulted in the addition to the 2021 course of 10 hours per week of TA-staffed open laboratory and of 3 hours of faculty-staffed office hours.

Assessing which resources, strategies, and interventions are helpful to students can be challenging for instructors in a large course. The VAA has been repeatedly demonstrated to supplement student learning in the face-to-face environment (Garrett et al., 2019; Linton et al., 2005; Linton et al., 2021; Magee et al., 2015), and this study demonstrates that the VAA supports online anatomy learning as well. The VAA user analytics from 2018 and 2019 demonstrate that VAA use supports student success and that it can be used to identify students who are not yet engaging with the course resources in the online environment. However, a limitation of the quantitative analysis of VAA usage is that the quality of the user interaction with the VAA as a learning experience is not reflected by the metric that is captured to quantify program usage. For example, a student may open the VAA five times a week with the intention of studying but may aimlessly select anatomical structures from the cadaver images. In contrast, another student may use the program only once in conjunction with course materials—such as short objectives, laboratory exercises, and lecture content—to review material before attending a laboratory session. Therefore, the student who only used the program once may have experienced greater learning gains, but their frequency of use would be lower. This metric also does not account for group learning when login access to the VAA is shared in an in-person or online environment. Furthermore, students who used the VAA more frequently may have been doing so instead of hands-on cadaveric learning, which most will agree is the preferred learning tool. Therefore, while this quantitative analysis supports the value of VAA as an opportunity to capture data about student engagement with course resources, it is not recommended as the sole strategy for identifying students in need of early intervention.

Another challenge to large courses is that each learner brings a cohort of personal experiences that shapes their learning and may require a personalized approach that can be difficult to achieve in large, online courses. Knowing what learning strategies are effective for many students may be just as valuable for instructors as knowing how to guide an individual student through various alternate methods. Many have focused on creating student-centered curricula that leverage active learning (Moffett, 2015; Ramnanan & Pound, 2017). Dominant or preferred learning styles for anatomy learners typically involve active sensing, visual strategies, and sequential strategies that are used in conjunction with other tools to provide an optimal learning experience (Neel & Grindem, 2010; Quinn, Smith, Kalmar, & Burgoon, 2018). However, it may be more important to engage students in the exploration of learning strategies as an opportunity for growth and plasticity, dispelling the myth that students are capable of only one learning style. At the start of each semester, faculty provide incoming students in Domestic Animal Gross Anatomy with "best practices" for student success based on evidence gathered by the TILT Learning Ecologies program. Students are given instructions on how to use the VAA and active learning strategies to employ while using the VAA are promoted in the lecture and laboratory sessions. Blended learning environments that combine traditional face-to-face and online instructional methods allow students to participate in active learning, improving academic performance (Graham, 2006). This instruction methodology requires student participation and preparation outside of class, which means that failure to do so may lead to decreased in-class participation (Estai & Bunt, 2016). However, this student-focused approach aims to foster innovation and lifelong learning (Hwa, 2020).

The survey results reported herein provide evidence that students find the VAA program helpful for learning animal anatomy. A positive relationship was found between student access and summative examination performance. Future research will investigate how students utilized specific course resources, including the VAA program. Currently, the VAA includes a two-dimensional interactive atlas, recorded video tutorials, written descriptions, and non-interactive rotatable models. How these individual components of the VAA influence learning is poorly understood.

Spatial visualization of anatomic structures can be a time-consuming task that may not be ideal for a fast-paced online environment if students never have a chance to interact with actual specimens. Berry and Baker (2010) found that when biochemistry students used 3D structural representations in place of 2D images, they were more easily able to focus on areas of importance. Additionally, biochemistry students provided positive feedback regarding having the ability to manipulate the images outside of the lectures at their own pace rather than in class, since they are busy in class keeping up with lecture material (Berry & Baker, 2010). A better understanding of how spatial skills impact the use of the VAA and how anatomy students use existing 3D resources may help facilitate our understanding of anatomic learning. Virtual reality provides an extreme range of potential learning experiences, from immersive clinical applications to adaptive learning opportunities (Goh & Sandars, 2021). Anatomy instruction will continue to evolve; understanding the impact of virtual reality experiences on learning will be critical to knowing how to best develop these resources and integrate them into the curricula.

2.5. LIMITATIONS

As an opportunistic study at the start of a pandemic, there are inherent limitations to this study that would have been improved by surveying students at the start of the semester or having consistent teaching and learning methods throughout the semester. Trying to identify one parameter that contributed to student success is highly subjective since every student's learning strategy immediately changed and each one's experience was unique as a result of the pandemic.

Rather than focusing on a single factor that influenced student success, the nature of the pandemic required that the study design take a holistic approach to assessing student success.

Study design and data collection were limited by the predetermined institutional use of the LMS, which was also the mechanism by which students accessed the VAA software and all course materials. A limitation of this integration was that VAA access measured the number of times per week an individual student logged into the LMS and navigated to the VAA program. This data does not identify the duration of use, provide qualitative analysis of the VAA utilization, or track access when multiple users share a single session either in-person or when using the program during remote learning sessions. Ideally, weekly user data would be collated to coincide with the timing of assessments and would reflect all users of the program at any given time. In 2020, VAA access in the laboratory was captured because it required a user to log in, but dozens of students may have shared a single login. While 2020 VAA access was captured, it represents a pattern of usage that is inherently different from 2018 and 2019, even before the pandemic.

Ideally, a comparison of formative examination performance by year would have been conducted to assess student learning following the online transition as a result of the pandemic. All students are expected to adhere to a student code of conduct (Colorado State University, 2020), but the online examinations in 2020 were delivered without a proctor, making their outcomes less reliable. Furthermore, these two online assessments in 2020 were significantly modified from those delivered in 2018 and 2019; therefore, direct comparisons would be invalid. Fundamentally, while examination performance is a measure of learning, the online examinations may not have accurately reflected a student's complete understanding of the material, nor do examinations distinguish previous knowledge from that learned during the semester. Future studies may benefit from more granular analyses of specific course resources and intentional study design.

2.6. CONCLUSIONS

Anatomy is necessary to our understanding of physiology and is the foundation of many medical training programs. Cadaveric learning is considered an invaluable experience by students and faculty alike. As the pandemic has pushed computer-based programs to the forefront as necessary resources for remote learning, evaluating these resources becomes essential as anatomy curricula and teaching continue to evolve in the post-pandemic environment. The virtual environment created by the VAA can support online anatomy learning when used in conjunction with other best practices for online teaching.

REFERENCES

Armstrong-Mensah, E., Ramsey-White, K., Yankey, B., & Self-Brown, S. (2020). COVID-19 and distance learning: Effects on Georgia State University School of Public Health students. Frontiers in Public Health, 8. doi:10.3389/fpubh.2020.576227

Berry, C., & Baker, M. D. (2010). Inside protein structure: Teaching in three dimensions. Biochemistry and Molecular Biology Education, 38(6), 425–429. doi:10.1002/bmb.20434

College of Veterinary Medicine & Biomedical Sciences. Virtual animal anatomy (Version 1.4.9) [Computer software]. Colorado State University. Retrieved from https://virtualanimalanatomy.colostate.edu/vaanopcomm

Collins, J. P. (2008). Modern approaches to teaching and learning anatomy. BMJ, 337(7671), 665. doi:10.1136/bmj.a1310

Colorado State University. (2020). Student conduct code. Board of Governors. Retrieved from https://resolutioncenter.colostate.edu/wp-content/uploads/sites/32/2018/08/Student-Conduct-Code-v2018.pdf

Custer, T. M., & Michael, K. (2015). The utilization of the Anatomage Virtual Dissection Table in the education of imaging science students. Journal of Tomography & Simulation, 1(1), 102–106. doi:10.4172/jts.1000102

Durosaro, O., Lachman, N., & Pawlina, W. (2008). Use of knowledge-sharing web-based portal in gross and microscopic anatomy. ANNALS Academy of Medicine Singapore, 37(12), 998–1001. https://pubmed.ncbi.nlm.nih.gov/19159031/

Erickson, M. (2020, October 16). Educators will use virtual reality to teach anatomy. Stanford Medicine. Retrieved from https://med.stanford.edu/news/all-news/2020/10/educators-will-use-virtual-reality-to-teach-anatomy.html

Evans, D. J. R., Bay, B. H., Wilson, T. D., Smith, C. F., Lachman, N., & Pawlina, W. (2020) Going virtual to support anatomy education: A stopgap in the midst of the COVID-19 pandemic. Anatomical Scientific Education, 13(3), 279–283. doi:10.1002/ase.1963

Estai, M., & Bunt, S. (2016). Best teaching practices in anatomy education: A critical review. Annals of Anatomy – Anatomischer Anzeiger, 208, 151–157. doi:10.1016/j.aanat.2016.02.010

Garrett, A. C., Ivie, K. R., Linton, A., Martin, J. F., Delcambre, J. J., Whalen, L. W., & Magee, C. (2019, June 24-27). Use of virtual canine anatomy improves student outcomes in a domestic gross anatomy course [Paper presentation]. The 2019 Biennial Conference of the Teaching Academy of the Consortium of West Region Colleges of Veterinary Medicine (RTA), Davis, CA, United States. Retrieved from https://s3.wp.wsu.edu/uploads/sites/1358/2019/06/Garrett-Magee.pdf

Goh, P.-S., & Sandars, J. (2020). A vision of the use of technology in medical education after the COVID-19 pandemic. MedEdPublish, 8(2), 15. doi:10.15694/mep.2020.000049.1

Graham, C. R. (2006). Blended learning system: Definition, current trends and future directions . In C. J. Bonk & C. R. Graham (Eds.), The handbook of blended learning (pp. 3–21). San Francisco, CA: Pfeiffer Publishing. Retrieved from https://www.academia.edu/563281/Blended_learning_systems_Definition_current_trends_and_future_directions

Gupta, S., Ojeh, N., Sa, B., Majumder, M. A. A., Singh, K., & Adams, O. P. (2020). Use of an adaptive e-learning platform as a formative assessment tool in the cardiovascular system course component of an MBBS programme. Advances in Medical Education Practices, 11, 989–996. doi:10.2147/AMEP.S267834

Gummery, E., Cobb, K. A., Mossop, L. H., & Cobb, M. A. (2018). Student perceptions of veterinary anatomy practical classes: A longitudinal study. Journal of Veterinary Medical Education, 45(2), 163–76. doi:10.3138/jvme.0816-132r1

Hwa, S. P. (2020). Assessing university students' perceptions toward constructivist learning environments using the CLES. International Journal of Advanced Research in Education and Society, 2(2), 115–129. Retrieved from http://myjms.mohe.gov.my/index.php/ijares/article/view/10183

International Committee on Veterinary Gross Anatomical Nomenclature. (2017). Nomina anatomica veterinaria (6th ed.). Editorial Committee. Retrieved from http://www.wava-amav.org/downloads/nav_6_2017.zip

Linton, A., Garrett, A. C., Ivie , K. R., Jr., Jones, J. D., Martin, J. F., Delcambre, J. J., & Magee, C. (2021). Enhancing anatomical instruction: Impact of a virtual canine anatomy program on student outcomes. Anatomical Sciences Education. Advance online publication. doi:10.1002/ase.2087

Linton, A., Schoenfeld-Tacher, R., & Whalen, L. R. (2005). Developing and implementing an assessment method to evaluate a virtual canine anatomy program. Journal of Veterinary Medical Education, 32(2), 249–254. doi:10.3138/jvme.32.2.249

Magee, C., Burns, S., & Johnson, D. (2015). Assessing course quality [Workshop]. TILT Professional Development Institute, Colorado State University, Fort Collins, CO, United States.

May, S. A., & Silva-Fletcher, A. (2015). Scaffolded active learning: Nine pedagogical principles for building a modern veterinary curriculum. Journal of Veterinary Medical Education, 42(4), 332–339. doi:10.3138/jvme.0415-063R

McBride, J. M., & Drake, R. L. (2018). National survey on anatomical sciences in medical education. Anatomical Sciences Education, 11(1), 7–14. doi:10.1002/ase.1760

Moffett, J. (2015). Twelve tips for "flipping" the classroom. Medical Teacher, 37(4), 331–336. doi:10.3109/0142159X.2014.943710

Mohamed, R. (2020). Attitude of veterinary students to cadaveric dissection in teaching and learning veterinary anatomy in the Caribbean. International Research in Education, 8(1), 139–144. doi:10.5296/ire.v8i1.16761

Morgan, G. A., Barrett, K. C., Leech, N. L., & Gloeckner, G. W. (2020). IBM SPSS for introductory statistics: Use and interpretation (6th ed.). New York, NY: Routledge. doi:10.4324/9780429287657

Neel, J. A., & Grindem, C. B. (2010). Learning-style profiles of 150 veterinary medical students. Journal of Veterinary Medical Education, 37(4), 347–352. doi:10.3138/jvme.37.4.347

Onigbinde, O. A., Chia, T., Oyeniran, O. I., & Ajagbe, A.O. (in press). The place of cadaveric dissection in post-COVID-19 anatomy education. Morphologie. doi:10.1016/j.morpho.2020.12.004

Quinn, M. M., Smith, T., Kalmar, E. L., & Burgoon, J. M. (2018). What type of learner are your students? Preferred learning styles of undergraduate gross anatomy students according to the Index of Learning Styles Questionnaire. Anatomical Sciences Education, 11(4), 358–365. doi:10.1002/ase.1748

Ramnanan, C. J., & Pound, L. D. (2017). Advances in medical education and practice: Student perceptions of the flipped classroom. Advances in Medical Education and Practice, 8, 63–73. doi:10.2147/AMEP.S109037

Ray. S., & Srivastava, S. (2020). Virtualization of science education: A lesson from the COVID-19 pandemic. Journal of Proteins and Proteomics, 11, 77–80. doi:10.1007/s42485-020-00038-7

Reeves, R. E., Aschenbrenner, J. E., Wordinger, R. J., Roque, R. S., & Sheedlo, H. J. (2004). Improved dissection efficiency in the human gross anatomy laboratory by the integration of computers and modern technology. Clinical Anatomy, 17(4), 337–344. doi:10.1002/ca.10245

Said, J. T., & Schwartz, A. W. (2021). Remote medical education: Adapting Kern's Curriculum Design to tele-teaching. Medical Science Educator, 31, 805–812. doi:10.1007/s40670-020-01186-7

Shah, A. (2020, April 3). Universities get creative with technology due to coronavirus closures. The Wall Street Journal. Retrieved from https://www.wsj.com/articles/universities-get-creative-with-technology-due-to-coronavirus-closures-11585918801?mod=searchresults&page=1&pos=2

Tam, M. D. B. S., Hart, A. R., Williams, S. M., Holland, R., Heylings, D., & Leinster, S. (2010). Evaluation of a computer program ('disect') to consolidate anatomy knowledge: A randomized-controlled trial. Medical Teacher, 32(3), e138–e142. doi:10.3109/01421590903144110

Tartavulea, C. V., Albu, C. N., Albu, N., Dieaconescu, R. I., & Petre, S. (2020). Online teaching practices and the effectiveness of the educational process in the wake of the COVID-19 pandemic. Amfiteatru Economic, 22(55), 920–936. doi:10.24818/EA/2020/55/920

Turney, B. W. (2007). Anatomy in a modern medical curriculum. Annals of the Royal College of Surgeons of England, 89(2), 104–107. doi:10.1308/003588407X168244

United Nations Educational, Scientific and Cultural Organization. (2021). Education: From disruption to recovery. Retrieved from https://en.unesco.org/covid19/educationresponse

U.S. News & World Report. (2019). Best veterinary schools. Retrieved from https://www.usnews.com/best-graduate-schools/top-health-schools/veterinarian-rankings

World Association of Veterinary Anatomists. (2020). Abstracts of the oral presentations of the 1st online conference of WAVA in August 2020. Anatomia, Histologia, Embryologia, 50(1), 4–7. doi:10.1111/ahe.12639

Wilson, A. B., Kaza, N., Singpurwalla, D. J., & Brooks, W. S. (2020). Are anatomy PhDs nearing extinction or adapting to change? United States graduate education trends in the anatomical sciences. Anatomical Sciences Education, 14(4), 432–439. doi:10.1002/ase.2013

Yang, C., Yang, X., Yang, H., & Fan, Y. (2020). Flipped classroom combined with human anatomy web-based learning system shows promising effects in anatomy education. Medicine, 99(46), e23096. doi:10.1097/MD.0000000000023096

Zargaran, A., Turki, M. A., Bhaskar, J., Spiers, H. V. M., & Zargaran, D. The role of technology in anatomy teaching: Striking the right balance. Advances in Medical Education and Practice, 11, 259–266. doi:10.2147/AMEP.S240150

TABLE 2.1. Assessments Used in Outcomes Analysis

Methodology	Assessment Type	Years Given (# Given/Year)	Point Value
iClicker questions	Formative, in-class assessment	2018 (17) 2019 (15) 2020 (9)	1 pt. each for maximum 10 pts.
Check-in quizzes	Formative, online assessment	2019 (11) 2020 (12)	2.5 pts. each, lowest score dropped
Laboratory quizzes	Formative, in-person* assessment	2018 (7) 2019 (7) 2020 (8)	5 pts. each, lowest score dropped
Guided quizzes	Formative hybrid lecture-laboratory assessments	2020 (6)	7.5 pts. each, none dropped

Note. * The last four 2020 laboratory quizzes were delivered via LMS.

TABLE 2.2. End of Semester Student Survey Course Statement Agreement

Statement	Median	IQR	n
The COVID-19 transition in this course was a success.	6	2	99
I enjoy this subject.	7	1	105
This class has increased my interest in this subject.	7	1	103
I want to take other classes in this subject.	7	1	98
The average time per week required by this course before Spring Break was appropriate.	6	2	90
The average time per week required by this course after Spring Break was appropriate.	6	1	83
The resources or activities made available to me before Spring Break facilitated my learning in this course.	7	0	103
The resources or activities made available to me after Spring Break facilitated my learning in this course.	5	3	79
I believed in my ability to get a passing grade for this class before Spring Break.	7	1	94
I believed in my ability to get a passing grade for this class after Spring Break.	6	2	83
The Virtual Animal Anatomy program helps me learn anatomy	7	1	99

Note. Student responses (N = 107) used a 1-7 Likert scale where 1 = strongly disagree, 4 = neutral, 7 = strongly agree. *IQR* = Interquartile Range

TABLE 2.3. Intercorrelations (rho Values) for End of Semester Student Survey Statements

Statement	n	1	2	3	4	5	6	7
1. The COVID-19 transition in this course was a success.	99	-	$.19^{*}$	$.60^{**}$	$-.02$	$.46^{**}$	.01	$.61^{**}$
2. The average time per week required by this course before Spring Break was appropriate.	90	-	-	$.32^{**}$	$.41^{**}$	$.30^{**}$	$.47^{**}$	$.23^{*}$
3. The average time per week required by this course after Spring Break was appropriate.	83	-	-	-	.03	$.45^{**}$	.14	$.58^{**}$
4. I believed in my ability to get a passing grade for this class before Spring Break.	104	-	-	-	-	$.39^{**}$	$.48^{**}$	.02
5. I believed in my ability to get a passing grade for this class after Spring Break.	79	-	-	-	-	-	$.21^{*}$	$.41^{**}$
6. The resources or activities made available to me before Spring Break facilitated my learning in this course.	94	-	-	-	-	-	-	.01
7. The resources or activities made available to me after Spring Break facilitated my learning in this course.	83	-	-	-	-	-	-	-

Note. *n* = number of students who scored a statement as 5 (slightly agree) or higher on 7-point scale (i.e., 5 = slightly agree to 7 = strongly agree) in the course survey (N = 107). $^{*}p < 0.05$, $^{**}p < 0.01$

TABLE 2.4. Course Resources Reported to be Helpful Pre- and Post-COVID-19 Transition

Course Resource	Pre-COVID-19			Post-COVID-19		
	n	%	CI95	n	%	CI95
Face-to-face lecture	92	86%	78%, 92%			
Open laboratory sessions†	91	85%	77%, 91%	24	22%	15%, 32%
The teaching assistants	77	72%	62%, 80%	26	24%	17%, 34%
The VAA program	67	63%	53%, 72%	88	82%	74%, 89%
Attending regular laboratory sessions	65	61%	61%, 70%			
The course instructors	49	46%	36%, 57%	23	21%	13%, 28%
The laboratory quizzes	49	46%	36%, 57%	42	39%	30%, 49%
My fellow classmates	38	36%	27%, 45%	23	21%	13%, 28%
The short objectives for lab quizzes	35	33%	24%, 42%	11	10%	5%, 18%
The weekly check-in quizzes	31	29%	21%, 39%	52	49%	39%, 58%
In-class question, iClickers	23	21%	13%, 28%			
Asynchronous lecture recordings	22	21%	13%, 28%	74	69%	60%, 78%
The written exercises	21	20%	12%, 27%	24	22%	15%, 32%
Video tours in the VAA or in LMS	19	18%	11%, 26%	26	24%	17%, 34%
Guided quizzes				90	84%	76%, 90%

Note. *n* = number of students who identified a resource as helpful in the course survey (N = 107). CI95 = 95% Confidence Interval. LMS = Learning Management System. VAA = Virtual Animal Anatomy program.

[†] Pre-COVID-19 Open Laboratory provided face-to-face cadaver laboratories while Post-COVID-19 provided interactions via online meeting software.

TABLE 2.5. Intercorrelations (rho Values) for Student Perceptions of Resource Helpfulness and COVID-19 Transition Success

Course Resource	Pre-COVID-19	Post-COVID-19
Face-to-Face lecture	$-.30^{**}$	
Open laboratory sessions[†]	−.04	$.23^{*}$
The teaching assistants	.11	.13
The VAA program	.05	−.15
Attending regular laboratory sessions	−.14	
The course instructors	.18	$.28^{**}$
The laboratory quizzes	.11	.18
My fellow classmates	−.15	−.01
The short objectives for lab quizzes	.00	−.13
The weekly check-in quizzes	.00	$.27^{**}$
In-class question, iClickers	−.11	
Asynchronous lecture recordings	.12	−.02
The written exercises	.10	$.25^{**}$
Video Tours in the VAA or in LMS	.09	.19
Guided quizzes		.18

Note. N = 107 survey responses. VAA = Virtual Animal Anatomy program. LMS = Learning Management System.

$^{*}p < 0.05$, $^{**}p < 0.01$.

[†] Pre-COVID-19 Open Laboratory provided face-to-face, cadaver laboratories while Post-COVID-19 provided interactions via online meeting software.

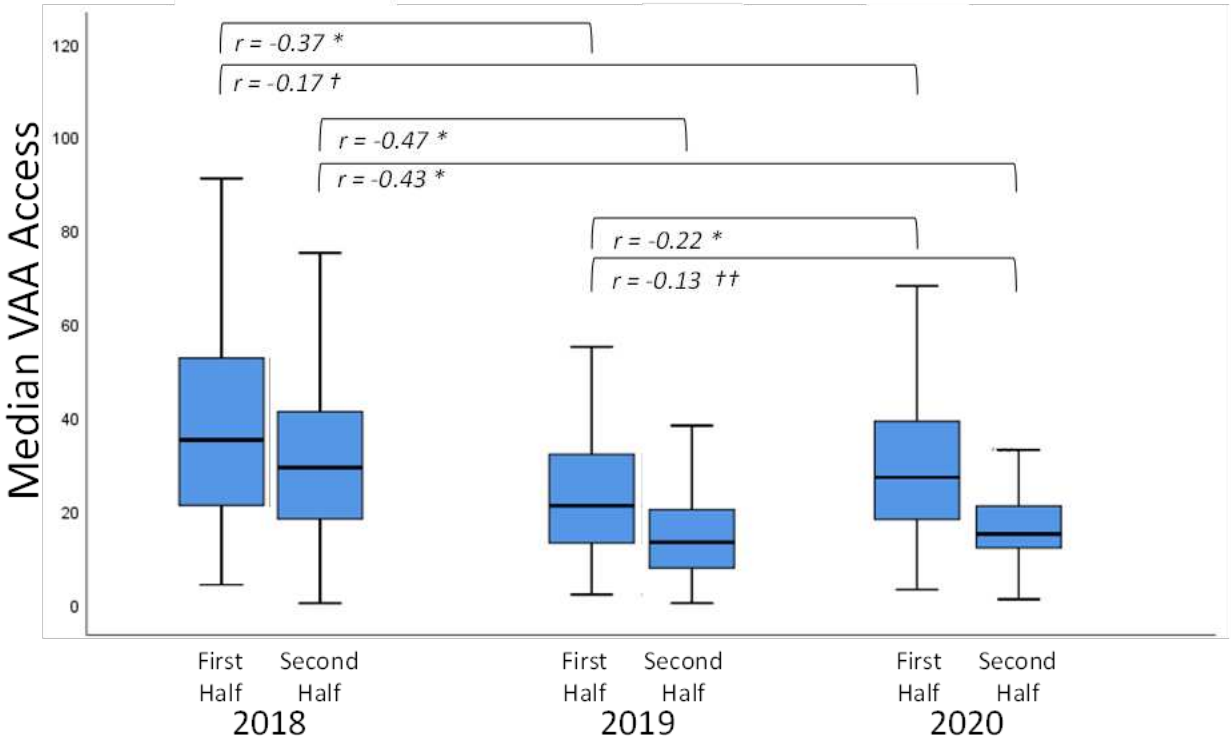


FIGURE 2.1. 2018 to 2020 Median, Interquartile Range, and Mann-Whitney U Test by Semester Half.

VAA = Virtual Animal Anatomy. Brackets represent Mann-Whitney U test.

Effect size reported as r value. 2018-2019 first half U = 4064.5 and second half U = 3140.0. 2018-2020 first half U = 5291.5 and second half U = 3310.5. 2019–2020 first half U = 5398.0 and second half U = 6136.0.

* $p < 0.01$, † = 0.01, †† = 0.05.

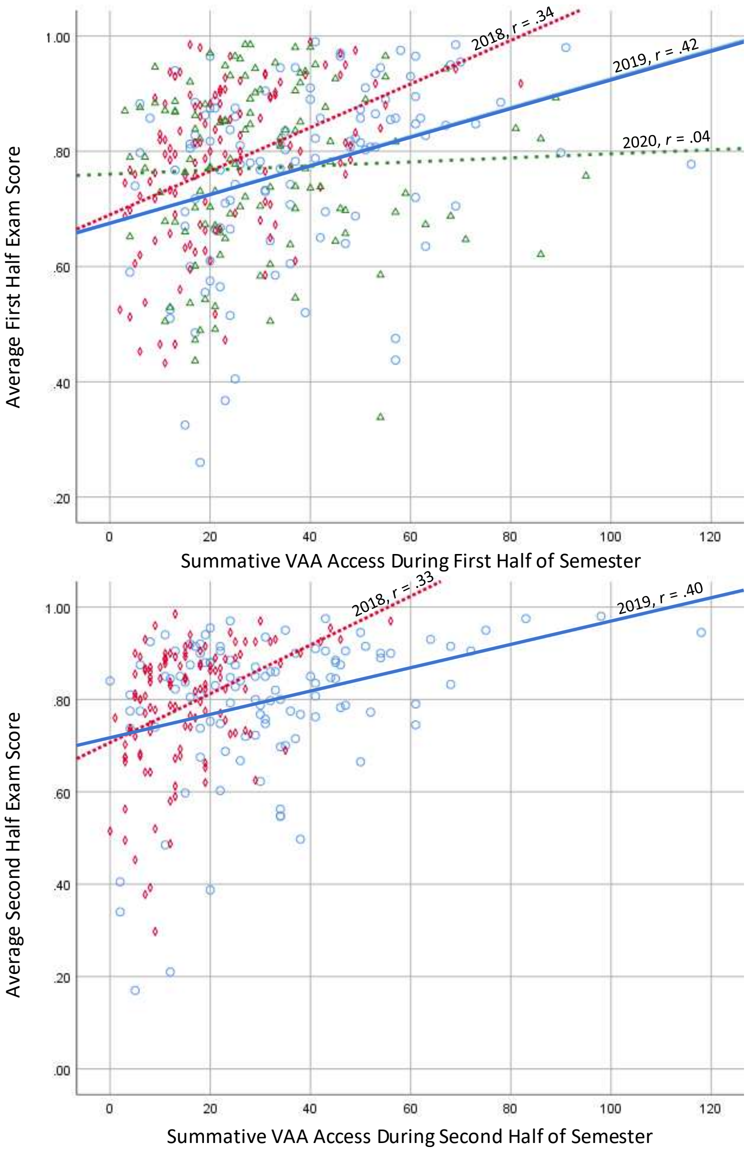


FIGURE 2.2. Scatterplots of 2018-2020 Summative Virtual Animal Anatomy Access vs. Examination Scores by Student.

Linear best fit lines for 2018 (N = 114, O), 2019 (N = 123, ◊), and 2020 (N = 117, Δ) for summative Virtual Animal Anatomy (VAA) access during select instructional periods and corresponding average examination score. As 2020 students completed different examinations and did not take laboratory-based examinations, their Examination 3 and Examination 4 performances were not included. Effect size, *r* = (.30) was considered typical.

CHAPTER III – RELATIONSHIP BETWEEN THE RELATIVE TIMING OF GROSS ANATOMY INSTRUCTION AND LABORATORY EXAMINATION PERFORMANCE

3.1. INTRODUCTION

Gross anatomy has been described as a foundational course within the biomedical sciences (Abrahams et al., 2022; Almizani et al., 2022; Bhattarai et al., 2022; Martin et al., 2022) and hundreds of hours of anatomical learning often precede medical students' client or patient interactions (Linton et al., 2022; Martin et al., 2022). While medical education in the nineteenth century consisted of two to three years of preclinical study of morphology, a biomedical revolution within the twentieth century necessitated a rapid reduction in anatomy instruction to make room for an expanding body of scientific knowledge (Chan et al., 2022). At the start of the twentieth century, a four-year medical degree averaged 549 hours of gross anatomy, but by the end of the century, students could expect to receive as few as 165 hours of instruction (Leung et al., 2006). While anatomy education literature frequently mentions a decrease in instructional time (Cottam, 1999; Leung et al., 2006; Bergman et al., 2008; Memon, 2009; Smith et al., 2014; Farrokhi et al., 2017; Singal et al., 2020; Chan et al., 2022; Lunn-Collier et al., 2022) it remains unclear if additional time dedicated to learning gross anatomy would offset reports of anatomy knowledge deficits (Cottam, 1999), or if veterinary medical education would see more value by adjusting the relative timing of anatomy instruction.

Carroll's A Model of School Learning (1963) assumes that any learner who is provided sufficient time can learn any skill and proposed five variables (opportunity to learn, perseverance, aptitude, quality of instruction, and ability to understand instruction) that could be used to calculate a student's degree of learning within a defined learning experience (Figure 3.1) (Carroll, 1989). By shifting emphasis from learning time to the learning process (Slavin, 1995) educators were challenged to redefine the relationship between the degree of learning and learning mastery (Bloom and S., 1968; Carroll, 1989). As additional learning time is provided, discrepancies in learning rates increase the educational costs of slower learners and the downtime of fast learners (Slavin, 1987). While it has been proposed that adaptive instruction provided by computer-assisted learning could avoid this outcome (Essa and Mojarad, 2020), such a feature has not been identified within available gross anatomy software.

In the absence of tools that can dynamically adjust instruction for each learner, veterinary medical educators are tasked to support or scaffold learners, using each learning opportunity to strengthen existing knowledge and support continual development (Dale et al., 2008). The application of scaffolding can be used to either guide the development of efficient independent learning strategies or transition the learner to higher levels of Bloom’s taxonomy (Dale et al., 2008). Providing complimentary resources that scaffold learning may increase time on task, increasing a learner's motivation to learn (Meehan and McCallig, 2019). Distinguishing between the opportunity to learn and motivation is essential to a successful academic intervention. An engaged student with insufficient instructional time requires a different intervention strategy than a disengaged student with abundant instructional opportunity.

Acknowledging a naturally dynamic learning environment promotes variations in instructional approaches, making comparing learning experiences challenging. The undergraduate anatomy curriculum is frequently introductory in design, while graduate and professional-level anatomy include additional functional perspectives (Thompson and Giffin, 2021; Martin et al., 2022). Consequently, undergraduate anatomy examinations place greater emphasis on low-order, identification-based questions, while postgraduate anatomy courses dedicate more attention to higher-order, relationship- or clinical-based questions (Thompson and Giffin, 2021).

Comparisons of undergraduate and postgraduate anatomy students are challenging (Bergman et al., 2008; Robertson et al., 2020; Thompson and Giffin, 2021). While the order level of examination questions can produce different outcomes for two populations of learners, the populations of interest are operationally defined by differences in opportunity to learn, demonstrations of academic perseverance, and prerequisite knowledge that influences aptitude. Existing anatomy education literature lacks an analysis that controls for these variables. This study seeks to understand better the relationship between available time to learn gross anatomy and practical examination performance.

3.2. MATERIALS AND METHODS

3.2.1. Course Designs

This retrospective study compared five years (2018 to 2022) of practical laboratory examinations completed as a component of two animal gross anatomy courses offered at Colorado State University. Both courses included comparative anatomy of equids, camelids, felids, and suids using canine gross anatomy as a foundational species. In addition, the same course coordinator and faculty provided instruction during the 15 weeks of undergraduate-level prosection- and graduate-level dissection-based animal anatomy. The first three units were four weeks in duration, while three weeks were provided for the fourth unit. At the end of each unit, students completed a summative practical examination.

3.2.1.1. Prosection Anatomy. A four-credit-hour prosection-based undergraduate animal gross anatomy course was provided to students who had previously completed an undergraduate-level animal biology course. Course enrollment ranged from 77 to 140 students per semester.

Formal instruction comprised three 50-minute didactic lectures per week and one 110-minute prosection-supported guided laboratory period. Students are also offered twelve hours per week of teaching assistant-supported open laboratory periods. The university learning management system (LMS), Canvas, acted as an online hub for distributing course documents, announcements, and didactic lecture recordings. In addition, the LMS provided students free access to an interactive comparative animal gross anatomy atlas, Virtual Animal Anatomy (Martin et al., 2022). When appropriate, lectures included images from this atlas. In addition, to allow LMS access while working with cadavers, course-provided computers or tablets were made available to students during guided and open laboratory sessions.

Each unit included multiple high-frequency, low-stakes formative assessments (50 of the 450 total course points) in the LMS and the laboratory to mimic the summative assessments. During non-examination weeks (11 total), there were formative (5-question) multiple-choice quizzes via the LMS. In addition, for each non-examination week after the first week of a unit (7 total), there were formative (5-question) fill-in-the-blank laboratory quizzes during their guided laboratory session.

At the end of each of the four units, during their scheduled 110-minute guided laboratory time, a summative 100-point assessment was divided into a 50-point multiple-choice examination in the LMS and a 50-point written fill-in-the-blank practical examination in the laboratory. Students received 50 minutes for each examination component, with a 10-minute break between each part.

3.2.1.2. Dissection Anatomy.

A three-credit-hour, graduate-level, dissection-based, animal gross anatomy course was offered to undergraduate students who either previously completed the prosection-based anatomy course or to graduate students as a requirement of a one-year Master of Science degree. Graduate students with no prior animal gross anatomy experience were required to enroll in both courses concurrently.

Dissection enrollment ranged from 20 to 40 students per semester. Groups of 2-3 students dissected an embalmed canine cadaver throughout the semester. Instruction was provided via two weekly 170-minute faculty-supported laboratory periods with one additional 170-minute faculty-supported open laboratory period. In addition, dissection students had after-hours access to the dissection laboratory and unlimited use of the LMS-hosted Virtual Animal Anatomy software.

Multiple high-frequency, low-stakes formative assessments of a student's anatomy knowledge were provided (100 of the course's 300 available points) in the LMS and laboratory. During non-examination weeks (11 total) and to help them prepare for their upcoming dissection, short multiple-choice and fill-in-the-blank quizzes were due in the LMS. For each non-examination week after the first week of the semester (10 total), students were orally assessed on their anatomy knowledge and thoroughness of their dissection with a brief (10-minute) faculty-led "table check." At the end of each unit, a summative assessment consisted of a 50-point fill-in-the-blank laboratory practical examination.

3.2.1.3. Differences in Unit Examinations.

The prosection- and dissection-based gross anatomy courses utilized the same laboratory manual. Students were provided the same list of anatomical structures in each course and asked to identify those structures on laboratory specimens and understand the structures' function or purpose. However, the prosection-based anatomy course emphasized identifying anatomical structures (lower-order knowledge-based questions), while the dissection-based course emphasized anatomical relationships (higher-order comprehension-based questions). This distribution resulted in prosection-based anatomy examinations comprising 75% lower-order and 25% higher-order questions. For dissection-based anatomy examinations, this ratio was reversed, presenting students with 25% lower-order and 75% higher-order questions.

3.2.2. Study Design

This five-year mixed-method retrospective study used records assessed from Colorado State University (CSU) between January 2018 to May 2022, approved by CSU's Internal Review Board (#3041). Following week 8 in 2020 and week 12 in 2021, laboratory instruction was replaced by remote instruction for both courses in response to a global pandemic. Those units were excluded from the analysis as practical examinations given during remote instruction could not use cadaveric prompts.

3.2.3. Examination Data.

Cadaveric instruction was divided into four units: pelvic limb (unit one), thoracic limb (unit two), thorax, abdomen, and pelvis (unit three), and head and neck (unit four). Each unit terminated with a non-cumulative, summative practical examination. Unit examination scores were used as a measure of learner performance.

During the study period, the total enrollment was 619 students; however, only the performance of 444 students was included in the study (Table 3.3). Some students enrolled in dissection more than a year after completing prosection-based anatomy (*n* = 8). Only the first enrollment was included for students who enrolled in prosection more than once (*n* = 12). Dissection students not concurrently enrolled in 2018 (*n* = 34) were excluded because there was insufficient data to confirm they completed prosection-based anatomy the prior year. The 121 prosection students enrolled in 2022 were excluded from cohort-specific comparisons because their participation in dissection the following year could not be verified at the time of analysis.

Students were grouped into one of three cohorts based on when they were enrolled in prosection-based anatomy relative to dissection-based gross anatomy. Eligible students who never enrolled in dissection-based anatomy were categorized as prosection-only. Concurrent students enrolled in both courses during the same semester, and consecutive students enrolled in the dissection course the Spring semester after completing prosection-based anatomy.

3.2.4. Statistical Methods

Consistent with other studies that evaluated gross anatomy examination performance (Thompson and Giffin, 2021; Martin et al., 2022), examination scores were not normally distributed, as defined by a skewness > |1|. Therefore, median (Med) and interquartile ranges (IQR) were used to report findings. Nonparametric Mann-Whitney U tests were used between cohorts to compare the outcomes of different learning experiences. Wilcoxon signed ranks tests were used for within-cohort analysis to compare individual learners' unit examination scores (Thompson and Giffin, 2021; Martin et al., 2022). Interpretations of the effect size r, as described by Cohen (1988), are regularly used with the selected analysis, and an r of |.30| was considered a typical effect size, |.20| represented a less-than-typical effect size, while |.50| and |.70|

represented a greater than and much greater than typical effect size, respectively (Morgan et al., 2019).

3.3. RESULTS

Using a Kruskal-Wallis nonparametric test to evaluate differences in examination performance due to the learners' year of enrollment, no difference was noted between each of the five years of prosection-based gross anatomy and total examination performance, χ^2 (4, N = 565) = 8.09, p = .088. When limiting the analysis to those students who at some point completed dissection-based anatomy, no difference was seen between years for prosection- nor dissection-based anatomy performance, χ^2 (4, n = 110) = 5.92, p = .205 and χ^2 (4, n = 110) = 6.89, p = .142, respectively. As no significant difference in examination performance was observed between years, the 5-year practical examination data for each course were pooled for further analysis.

3.3.1. Prosection-Based Anatomy

For each prosection-based anatomy unit, students who went on to take dissection significantly outperformed those who did not. In addition, concurrently enrolled students outperformed the other two cohorts, while prosection-only students consistently scored the lowest. Overall, a statistically significant performance increase was reported between units one and two, no significant differences between units two and three, and a statistically significant decrease from units three to four. Comparing unit examination performance per cohort revealed a similar pattern with one exception. The reported typical effect size observed for consecutively enrolled students' unit three and four examinations was not statistically significant.

3.3.1.1. Course Performance

3.3.2.1.1. Between Cohorts.

Total examination scores for prosection-based anatomy were greater for concurrent students or students who would later enroll in dissection-based gross anatomy (n = 110, *Med* = .90, *IQR* = .14) compared to students who did not take dissection (n = 334, *Med* = .74, *IQR* = .23). The prosection-only students have lower mean ranks (248.87) than their peers (424.16) on total examination score performance, U = 9497.0, $p < .001$, r = -.43. Compared with students who only completed prosection-based anatomy (n = 334, *Med* = 0.72, *IQR* = 0.23), total prosection-based practical examination scores were greater for both concurrently (n = 67, *Med* = 0.92, *IQR* = 0.08) and consecutively enrolled (n = 43, *Med* = 0.82, *IQR* = 0.15) students (Table 3.4). There was a significant difference in total prosection-based practical examination scores between each of the three cohorts (Table 3.1).

3.3.1.1.2. Between Units.

Of the 440 students who completed the first and second units, 66 scored higher on the first examination, 365 scored higher on the second examination, and 9 students scored the same. This difference indicates that prosection examination scores increase between the first and second examination, z = -14.67, $p < .001$, r = -.70. Of the 235 students who completed both the third and fourth unit examinations, 148 scored higher on the third exam, 78 scored higher on the fourth examination, and 9 students scored the same, n = 315, p = .431, r = -.02. Most students saw a decrease in examination scores between the third and fourth unit examinations, z = -5.46, $p < .001$, r = -.36.

3.3.1.2. Unit Performance.

A box and whisker plot was used to visualize prosection examination performance by course unit (Figure 3.2). Units are presented in order of course presentation, and for all three cohorts, unit one had the lowest median examination score. Unit-for-unit, concurrently enrolled students had the highest median unit score and smallest range, while prosection-only students had the lowest median score and the largest range.

3.3.1.2.1. Between Cohorts.

For each unit examination, the mean ranks of concurrently enrolled students were greater than prosection-only and consecutively enrolled students (Table 3.5). In addition, consecutively enrolled students had higher mean ranks than prosection-only students on all unit examinations. However, this was only deemed statistically significant for units one, two, and four.

3.3.1.2.2. Within Cohorts.

Of the 330 prosection-only students who completed unit examinations one and two, 49 scored higher on the unit one examination, 275 scored higher on the unit two examination, and 10 students scored the same (Table 3.6), $z = -12.67$, $p < .001$, $r = -.70$. Despite smaller populations of concurrently ($n = 67$) and consecutively enrolled students ($n = 43$) results were similar in effect size, $z = -5.49$, $p < .001$, $r = -.67$ and $z = -5.30$, $p < .001$, $r = -.81$, respectively.

For all three cohorts, the examination scores differences were not statistically significant between unit two and unit three examinations (Table 3.7). However, a statistically significant difference was calculated when comparing unit three and four examinations. Of the 176 prosection-only students who completed both the third and fourth unit examinations, 111 scored higher on the third examination, 60 scored higher on the fourth examination, and 5 students scored the same, $z = -4.45$, $p < .001$, $r = -.34$. This result was also observed with concurrently ($n = 35$) enrolled students but not consecutively ($n = 24$) enrolled, $z = -3.39$, $p < .001$, $r = -.57$ and $n = 24$, $p = .158$, $r = -.29$, respectively.

3.3.2. Dissection-Based Anatomy

While consecutively enrolled students outperformed concurrently enrolled students on units one, two, and three and concurrently enrolled students had a higher median unit four examination score (Figure 3.1), only units one and two saw statistically significant differences in performance. Within cohorts, a statistically significant performance increase was observed between units one and two. While no difference between units two and three was observed for concurrently enrolled students, consecutively enrolled students performed better on dissection unit examination two than examination three. Neither cohort showed a statistically significant change in performance between dissection examinations for units three and four.

3.3.2.1. Course Performance.

Within the dissection-based anatomy course, total examination scores were greater for consecutively enrolled students (*Med* = .85, *IQR* = .13) compared to concurrently enrolled students (*Med* = .82, *IQR* = .16) (Table 3.3). While consecutively enrolled students had a higher mean rank (63.73, $n = 43$) than concurrently enrolled students (50.22, $n = 67$), $U = 1086.5$, $p = .030$, $r = -.21$, the effect size was smaller than typical.

Of the 110 students who completed the first and second dissection units, four scored higher on the first unit examination, 102 scored higher on the second unit examination, and four scored the same. This difference indicates that the increase between the first and second dissection examination scores was significant, and the effect size was much larger than typical, $z = -8.65$, $p < .001$, $r = -.82$. There was not a statistically significant difference between examinations scores on the second and third or third and fourth units, $n = 80$, $p = .283$, $r = -.12$ and $n = 48$, $p = .612$, $r = -.07$, respectively.

3.3.2.2. Unit Performance.

A box and whisker plot was used to visually compare dissection unit examination performance for consecutively and concurrently enrolled students (Figure 3.3). Similar to what was observed with prosection-based anatomy, dissection unit examination one had the lowest median score of the four examinations (Figure 3.1). Consecutively enrolled students had greater median examination scores on the first three units, while concurrently enrolled students had greater scores on the last unit examination.

3.3.2.2.1. Between Cohorts.

For the first two dissection examinations, the mean ranks of consecutively enrolled students were greater than concurrently enrolled students (Table 3.2). For dissection examination unit one, consecutively enrolled students had a higher mean rank (66.58, $n = 43$) than concurrently enrolled students (48.39, $n = 67$), $U = 964.0$, $p = .003$, $r = -.28$. This relationship repeated with unit two, as consecutively enrolled students had a higher mean rank (65.40, $n = 43$) than concurrently enrolled students (49.15, $n = 67$), $U = 1015.0$, $p = .009$, $r = -.25$. Statistically significant differences were reported between consecutively and concurrently enrolled students' first two examination scores, and in both cases the effect size was less than typical. There was not a statistically significant difference between cohorts for the third and fourth units, $n = 80$, $p = .777$, $r = -.03$ and $n = 48$, $p = .280$, $r = -.14$, respectively.

Within Cohorts. For both concurrently and consecutively enrolled students, the differences between the first and second examinations were statistically significant (Table 3.6). Of the 67 concurrently enrolled students who completed both the first and second examinations, 2 scored higher on the first examination, 61 scored higher on the second examination, and 4 students scored the same, $z = -6.71$, $p < .001$, $r = -.82$. Of the 43 consecutively enrolled students who completed the first two examinations, 2 scored higher on the first examination, 37 scored higher on the second examination, and 4 students scored the same, $z = -5.27$, $p < .001$, $r = -.80$.

Of the 26 consecutively enrolled students who completed both the second and third examinations, 17 scored higher on the second, 8 scored higher on the third, and 1 student scored the same, $z = -2.20$, $p = .028$, $r = -.43$ (Table 3.6). A statistically significant difference between the second and third unit was not reported for concurrently enrolled students, $n = 54$, $p = .823$, $r = -.30$. Differences between the third and fourth examination was not statistically significant for either concurrent or consecutive students, $n = 35$, $p = .865$, $r = -.03$ and $n = 13$, $p = .534$, $r = -.17$, respectively.

Concurrently enrolled students consistently scored higher on prosection-based examinations than on the same dissection-based unit examination (Table 3.7). The differences were significant for concurrently enrolled students but much larger than typical. By contrast, the consecutively enrolled students had greater dissection-based examination scores than prosection-based first two units. However, the effect size of these differences (Unit 1: $n = 43$, $p = .007$, $r = -.41$; Unit 2: $n = 43$, $p = .004$, $r = -.44$) was almost half of what was seen with their concurrently enrolled students (Unit 1: $n = 67$, $p < .001$, $r = -.86$; Unit 2: $n = 67$, $p < .001$, $r = -.83$). While the third and fourth units had nonsignificant smaller than typical effect sizes (Unit 3: $n = 12$, $p = .790$, $r = -.08$; Unit 4: $n = 7$, $p = .735$, $r = -.13$), they were almost half of what was seen with the concurrently enrolled students (Unit 3: $n = 54$, $p < .001$, $r = -.81$; Unit 4: $n = 35$, $p = .002$, $r = -.52$).

3.4. DISCUSSION

This study analyzed the practical examination performance of a cadaver-based gross anatomy learning experience where the period of instruction was provided either over two semesters or within a single semester.

3.4.1. Is Additional Gross Anatomy Instruction Needed?

Carroll's Model of School Learning uses the shortest duration of three variables of time, opportunity to learn, motivation, and aptitude to calculate the degree of learning (Carroll, 1989). A consequence of the degree of learning being limited by a minimum variable is that interventions that target larger variables of time will not influence the degree of learning. This lack of influence is especially relevant to medical education because observations that total gross anatomy instruction has decreased (Cottam, 1999; Leung et al., 2006; Bergman et al., 2008; Memon, 2009; Smith et al., 2014; Farrokhi et al., 2017; Singal et al., 2020; Chan et al., 2022; Lunn-Collier et al., 2022), is only impactful if students do not have sufficient opportunity to learn what is expected of them.

This study's concurrently enrolled students outperformed their consecutively enrolled and prosection-only peers in all prosection-based unit examinations (Table 3.1). However, while consecutively enrolled students outperformed their prosection-only peers, the observed effect size was less than typical. Therefore, additional opportunities for learning during prosection-based anatomy had a larger effect on examination performance than factors that led to students enrolling in an additional gross anatomy course.

Concurrently enrolled students were provided additional instructional time and experiences to learn prosection-based gross anatomy. When their examination performance was compared with other students enrolled in the prosection-based gross anatomy course, concurrent students outperformed their peers (Table 3.3). An increase in opportunity to learn was associated with higher prosection examination scores. This relationship reveals that, at least for concurrent learners, the degree of learning was not limited by motivation because learners limited by motivation would not see improvements in performance with additional time.

The conclusion drawn from these results is that additional instructional time and experiences increase animal gross anatomy examination scores. While Bloom's learning for mastery proposed that 95% of learners should achieve 90% or more on learning assessments (Bloom and S., 1968), the threshold for matriculation from undergraduate prosection to the graduate dissection-based gross anatomy course is currently 70%, of which the median score was achieved by all three cohorts.

3.4.2. Should Curricular Redesign Account for Relative Timing of Instruction?

Carroll's model calculates the degree of learning as a ratio of actual learning time to aptitude, with no explicit consideration for the relative timing of learning. Instead, relative timing influences aptitude via prior experiences. Aptitude is generally defined as the time it takes to learn, but a student's prior knowledge changes what must be learned to achieve mastery. Therefore, aptitude is more completely defined as the sum of the time it will take to learn minus the time saved by previous learning mastery.

A consequence of concurrent enrollment is that dissection-based anatomy experiences necessitating higher-order understanding precede the knowledge acquired within prosection-based instruction. While scaffolding supports the development of efficient independent learning strategies for all learners, support for the transition between taxonomical levels is limited only to consecutive learners. Carroll's model would argue that with sufficient time for learning, there be no observable performance difference between an instructional model that uses scaffolded learning to move students from predominantly lower-order examination questions to one that is predominantly higher-order and an instructional model that asks learners to be prepared for lower and higher-order questions as long as both populations are provided the same opportunities for learning.

In this study, concurrent and consecutive students had equivalent opportunities to learn and instruction for both prosection- and dissection-based gross anatomy at each dissection examination. However, neither cohort maintained a performance advantage on all higher-order comprehension-based dissection examinations. Consecutively enrolled students outperformed concurrently enrolled students on the first two dissection examinations (Table 3.2). Due to the non-cumulative nature of the examinations, if prior prosection experience with a gross anatomy unit provided a performance benefit in dissection-based anatomy, consecutive students should have outperformed concurrent students on all four of the unit exams. Likewise, if dissection-based learning were associated exclusively with temporal proximity of prosection instruction, concurrent students would be expected to outperform consecutive students on all units. Convergence of performance between consecutive and concurrent learners is evidence that performance differences cannot be solely explained by prior knowledge or by the amount of temporal proximity of instruction to the examination.

Interestingly, performance convergences on dissection examinations mirror examination familiarity. When considering the gross anatomy examination history of concurrent and consecutive learners at the unit one dissection examination, consecutive learners have four prior prosection examinations to reflect upon, while concurrently enrolled students have none. For each sequential dissection examination, consecutive students increase their examination history by one, while concurrent students increase their examination history by two. Concurrently enrolled students had six gross anatomy examinations by the fourth dissection examination, while consecutively enrolled students had seven.

Three observations support a relationship between examination history and performance. First, neither cohort consistently outperformed the other on dissection examinations. Second, both prosection and dissection-based anatomy students performed significantly better on their second examination than their first (Table 3.6). Third, consecutive students score higher on unit one and two dissection examinations than on the same prosection examinations (Table 3.7).

3.4.3. Considerations for Veterinary Medical Educators

An instructor cannot presume that actual time spent learning is synonymous with time spent achieving content mastery. This study found that additional opportunities to learn increased gross anatomy examination performance. However, the median score for all three cohorts was sufficient for matriculation to more advanced coursework and those who completed the higher-level dissection course received passing marks. As students become more familiar with examinations, their scores increase despite no changes in content familiarity. These conclusions are consistent with studies investigating the benefit of prerequisite gross anatomy courses for medical anatomy classes (Peterson and Tucker, 2005) and the disruption in learning caused by the COVID-19 pandemic (Martin et al., 2022). Students adjust their approach to learning to align with expectations set by prior assessment experiences (Larsen et al., 2008; Watling and Ginsburg, 2019; Martin et al., 2022).

3.5. LIMITATIONS

Prior to the analysis of instructional time, examination performance was found to be consistent between years. This justified pooling of students into appropriate cohorts but provided no insight into the relative difficulty of any examinations. However, fluctuations in unit examination difficulty could influence initial student performance. A difficulty index and discrimination index would provide insight into each examination's difficulty. This analysis was not completed within the scope of this study because such an analysis would require a review of each student's response to every question, and laboratory examinations were returned to students after grading. Smaller sample sizes of dissection examinations three and four could explain nonsignificant relationships. Lastly, academic history remains a complex variable to control. Concurrently enrolled students are the product of enrollment in a one-year postgraduate degree program without prior gross anatomy experience. However, there was no way to discriminate between undergraduate students who elected to take dissection the following semester and those who took prosection-based anatomy, graduated, and then enrolled in a postgraduate degree program that included dissection-based anatomy as part of their course load. Academic experiences could explain variations in early learning strategies.

3.6. CONCLUSIONS

This study showed that additional instructional time and experiences provided by concurrently dissecting improved lower-order prosection-based gross anatomy scores. However, combining prosection- and dissection-based instruction into a single semester resulted in an initial disadvantage on higher-order examinations. The convergence of higher-order examination performance by concurrent and consecutive students may be better explained by examination familiarity as a component of a learner's aptitude. Therefore, instructors of subjects requiring foundational anatomical knowledge would be best served by working with anatomy instructors to prepare learners for the transition to the higher-order application of anatomy in their courses.

REFERENCES

Abrahams A, Pienaar L, Bugarith K, Gunston G, Badenhorst E. 2022. A foundational knowledge assessment tool to predict academic performance of medical students in first-year anatomy and physiology. Adv Physiol Educ 46:598–605.

Almizani MS, Alotaibi MA, Askar MFB, Albaqami NM, Alobaishi RS, Arafa MA, Jumaa MI. 2022. Clinicians' and Students' Perceptions and Attitudes Regarding the Anatomical Knowledge of Medical Students. Adv Medical Educ Pract 13:1251–1259.

Bergman EM, Prince KJAH, Drukker J, Vleuten CPM van der, Scherpbier AJJA. 2008. How much anatomy is enough? Anat Sci Educ 1:184–188.

Bhattarai L, KC S, Shrivastava AK, Sah RP. 2022. Importance of Cadaveric Dissection in Learning Anatomy For Medical Students: A Review. Med Phoenix 6:68–72.

Bloom, S. B. 1968. Learning for Mastery. Instruction and Curriculum. Regional Education Laboratory for the Carolinas and Virginia, Topical Papers and Reprints, Number 1. Evaluation Comment 12.

Carroll JB. 1989. The Carroll Model. Educ Researcher 18:26–31.

Chan AY-CC, Stapper CPM, Bleys RLAW, Leeuwen M van, Cate O ten. 2022. Are We Facing the End of Gross Anatomy Teaching as We Have Known It for Centuries? Adv Medical Educ Pract 13:1243–1250.

Cottam WW. 1999. Adequacy of medical school gross anatomy education as perceived by certain postgraduate residency programs and anatomy course directors. Clin Anat 12:55–65.

Dale VHM, Sullivan M, May SA. 2008. Adult Learning in Veterinary Education: Theory to Practice. J Vet Med Educ 35:581–588.

Essa A, Mojarad S. 2020. Does Time Matter in Learning? A Computer Simulation of Carroll's Model of Learning. Lect Notes Comput Sc 458–474.

Farrokhi A, Soleymaninejad M, Ghorbanlou M, Fallah R, Nejatbakhsh R. 2017. Applied anatomy, today's requirement for clinical medicine courses. Anat Cell Biology 50:175–179.

Larsen DP, Butler AC, III HLR. 2008. Test-enhanced learning in medical education. Med Educ 42:959–966.

Leung K, Lu K-S, Huang T-S, Hsieh B-S. 2006. Anatomy Instruction in Medical Schools: Connecting the Past and the Future. Adv Health Sci Educ 11:209–215.

Linton A, Garrett AC, Ivie KR, Jones JD, Martin JF, Delcambre JJ, Magee C. 2022. Enhancing Anatomical Instruction: Impact of a Virtual Canine Anatomy Program on Student Outcomes. Anat Sci Educ 15:330–340.

Lunn-Collier R, Layman-Lemphane JI, Baatjes KJ, Correia J. 2022. Clinicians' opinions on the clinical relevance of anatomy education at Stellenbosch University. Anat Sci Educ (in press; doi: 10.1002/ase.2175).

Martin JF, Arnold OR, Linton A, Jones JD, Garrett AC, Mango DW, Juarez KA, Gloeckner G, Magee C. 2022. How Virtual Animal Anatomy facilitated a successful transition to online instruction and supported student learning during the coronavirus pandemic. Anatomia Histologia Embryologia (in press; doi: 10.1111/ahe.12799).

Meehan M, McCallig J. 2019. Effects on learning of time spent by university students attending lectures and/or watching online videos. J Comput Assist Lear 35:283–293.

Memon IK. 2009. Anatomy education faces challenges in Pakistan. Anat Sci Educ 2:193–194.

Morgan GA, Barrett KC, Leech NL, Gloeckner GW. 2019. IBM SPSS for Introductory Statistics: Use and Interpretation. 106–134.

Peterson CA, Tucker RP. 2005. Undergraduate coursework in anatomy as a predictor of performance: Comparison between students taking a medical gross anatomy course of average length and a course shortened by curriculum reform. Clin Anat 18:540–547.

Robertson EM, Thompson KL, Notebaert AJ. 2020. Perceived Benefits of Anatomy Coursework Prior to Medical and Dental School. Anat Sci Educ 13:168–181.

Singal A, Bansal A, Chaudhary P. 2020. Cadaverless anatomy: Darkness in the times of pandemic Covid-19. Morphol 104:147–150.

Slavin RE. 1987. Mastery Learning Reconsidered. Rev Educ Res 57:175–213.

Slavin RE. 1995. A Model of Effective Instruction. Educ Forum 59:166–176.

Smith CF, Martinez-Álvarez C, McHanwell S. 2014. The context of learning anatomy: does it make a difference? J Anat 224:270–278.

Thompson AR, Giffin BF. 2021. Higher-Order Assessment in Gross Anatomy: A Comparison of Performance on Higher- Versus Lower-Order Anatomy Questions between Undergraduate and First-Year Medical Students. Anat Sci Educ 14:306–316.

Watling CJ, Ginsburg S. 2019. Assessment, feedback and the alchemy of learning. Med Educ 53:76–85.

TABLE 3.1. Comparison Of Total Prosection Laboratory Examination Scores By Cohort.

Cohorts	N	Mean Rank	U	p	r
			2662.5	< .001	-0.49
Prosection-Only[1]	334	328.26			
Concurrent[2]	67	175.47			
			4613.0	< .001	-0.20
Prosection-Only	334	181.31			
Consecutive[3]	43	248.72			
			630.5	< .001	-0.47
Concurrent	67	67.59			
Consecutive	43	36.66			

Mann-Whitney U test for total prosection practical examination scores between prosection-only, concurrent, and consecutive students from 2018 to 2022.

[1] Students who did not go on to take dissection within a year.

[2] Students who completed prosection- and dissection-based anatomy within the same semester.

[3] Students who completed prosection-based anatomy one year prior to dissection-based anatomy.

TABLE 3.2. Comparison Of Total Dissection Examination Scores By Cohort.

Exam	Cohorts	N	Mean Rank	U	p	r
				964.0	0.003	-0.28
1	Concurrent	67	48.39			
	Consecutive	43	66.58			
				1015.0	0.009	-0.25
2	Concurrent	67	49.15			
	Consecutive	43	65.40			
				674.5	0.777	-0.03
3	Concurrent	54	39.99			
	Consecutive	29	41.56			
				181.0	0.280	-0.14
4	Concurrent	35	25.83			
	Consecutive	24	20.92			

Mann-Whitney U test for dissection unit practical examination scores between concurrent, and consecutive students from 2018 to 2022.

TABLE 3.3. Descriptive Statistics Of Cohort Enrollment And Laboratory Format Between 2018 And 2022.

		Prosection[1]	Concurrent[2]	Consecutive[3]	Unit[4]			
Year	N	n	n	n	1	2	3	4
2018	105	99	6	34*	IP	IP	IP	IP
2019	107	87	13	7	IP	IP	IP	IP
2020	125	96	12	17	IP	IP	R	R
2021	85	52	19	14	IP	IP	IP	R
2022	22	121**	17	5	IP	IP	IP	IP
N	444	334	67	43				

Number of enrolled students per Spring semester. For students with multiple enrollments within the same course only the first enrollment was counted and included in the study. Numbers represent post exclusion criterion.

[1] Students who did not go on to take dissection within a year.

[2] Students who completed prosection- and dissection-based anatomy within the same semester.

[3] Students who completed prosection-based anatomy one year prior to dissection-based anatomy.

[4] IP: Face-to-face instruction occurred during that unit. R: Face-to-face instruction did not occur during that unit. * Enrollment prior to 2017 was not collected due to changes in instruction and therefore prosection enrollment in 2017 could not be confirmed for 2018 dissection students.

** 2022 Prosection was excluded because enrollment in dissection the following year could not be confirmed.

TABLE 3.4. Laboratory Examination Performance By Cohort And Course For Each Unit.

		Prosection-Only[1]			Concurrent[2]			Consecutive[3]		
		n	Median	IQR	n	Median	IQR	n	Median	IQR
Prosection Examination	1	334	0.65	0.27	67	0.90	0.15	43	0.78	0.15
	2	330	0.79	0.23	67	0.94	0.08	43	0.85	0.19
	3	232	0.82	0.22	54	0.96	0.08	29	0.86	0.14
	4	176	0.78	0.19	35	0.92	0.08	24	0.87	0.10
	Total	334	0.74	0.23	67	0.92	0.08	43	0.82	0.15
Dissection Examination	1				67	0.76	0.20	43	0.83	0.16
	2				67	0.86	0.15	43	0.90	0.08
	3				54	0.84	0.12	26	0.88	0.18
	4				35	0.88	0.19	13	0.86	0.14
	Total				67	0.82	0.16	43	0.85	0.13

Median and interquartile range for total course and unit prosection and dissection examination scores by enrollment type from 2018 to 2022.

[1] Students who did not go on to take dissection within a year.

[2] Students who completed prosection- and dissection-based anatomy within the same semester.

[3] Students who completed prosection-based anatomy one year prior to dissection-based anatomy.

TABLE 3.5. Comparison Of Unit Prosection Examination Scores By Cohort

Exam	Cohorts	N	Mean Rank	U	p	r
Unit 1				3068.0	< .001	-0.47
	Prosection-Only	334	176.69			
	Concurrent	67	323.21			
				4714.5	< .001	-0.19
	Prosection-Only	334	181.62			
	Consecutive	43	246.36			
				664.5	< .001	-0.45
	Concurrent	67	67.08			
	Consecutive	43	37.45			
Unit 2				3650.5	< .001	-0.43
	Prosection-Only	330	176.56			
	Concurrent	67	309.51			
				5079.0	.002	-0.16
	Prosection-Only	330	180.89			
	Consecutive	43	233.88			
				764.0	< .001	-0.40
	Concurrent	67	65.60			
	Consecutive	43	39.77			
Unit 3				1932.0	< .001	-0.47
	Prosection-Only	232	124.83			
	Concurrent	54	223.72			
				2703.5	.085	-0.11
	Prosection-Only	232	128.15			
	Consecutive	29	153.78			
				267.0	.001	-0.54

	Concurrent	54	51.56			
	Consecutive	29	24.21			
Unit 4				1070.0	< .001	-0.42
	Prosection-Only	176	94.58			
	Concurrent	35	163.43			
				1372.0	.005	-0.20
	Prosection-Only	176	96.30			
	Consecutive	24	131.33			
				248.5	.008	-0.34
	Concurrent	35	34.90			
	Consecutive	24	22.85			

Mann-Whitney U Test for prosection unit practical examination scores between prosection-only, concurrent, and consecutive students from 2018 to 2022.

TABLE 3.6. Comparison Of Sequential Prosection And Dissection Unit Examination Scores By Cohort.

	Concurrently Enrolled				Consecutively Enrolled				Prosection-Only			
Unit Score	N	z	p	r	N	z	p	r	N	z	p	r
		-5.49	< .001	-0.67		-5.30	< .001	-0.81		-12.67	< .001	-0.70
1 Higher	12				5				49			
2 Higher	52				38				275			
Same Score	3				0				6			
		-1.00	0.318	-0.14		-0.62	0.536	-0.11		-0.02	0.558	0.00
2 Higher	22				13				115			
3 Higher	28				12				104			
Same Score	4				4				13			
		-3.39	< .001	-0.57		-1.41	0.158	-0.29		-4.45	< .001	-0.34
3 Higher	24				13				111			
4 Higher	9				9				60			
Same Score	2				2				5			
		-6.71	< .001	-0.82		-5.27	< .001	-0.80		-	-	-
1 Higher	2				2				-			
2 Higher	61				37				-			
Same Score	4				4				-			
		-2.23	0.823	-0.30		-2.20	0.028	-0.43		-	-	-
2 Higher	26				17				-			
3 Higher	26				8				-			
Same Score	2				1				-			
		-0.17	0.865	-0.03		-0.62	0.534	-0.17		-	-	-
3 Higher	16				6				-			
4 Higher	17				5				-			
Same Score	2				2				-			

Wilcoxon Signed Rank Test for consecutive prosection unit examination scores for prosection-only, concurrently enrolled, and consecutively enrolled students. N represents the number of students who scored higher or the same on sequential unit examinations.

TABLE 3.7. Comparison Of Prosection And Dissection Examination Scores For Each Examination Unit By Cohort.

			Concurrently Enrolled				Consecutively Enrolled		
Exam	Unit Score	N	z	p	r	N	z	p	r
1			-7.05	< .001	-0.86		-2.72	0.007	-0.41
	Prosection	65				17			
	Dissection	2				26			
	Same Score	0				0			
2			-6.80	< .001	-0.83		-2.86	0.004	-0.44
	Prosection	60				14			
	Dissection	6				28			
	Same Score	1				1			
3			-5.94	< .001	-0.81		-0.27	0.790	-0.08
	Prosection	48				5			
	Dissection	5				6			
	Same Score	1				1			
4			-3.08	0.002	-0.52		-0.34	0.735	-0.13
	Prosection	24				5			
	Dissection	11				2			
	Same Score	0				0			

Wilcoxon Signed Rank Test for comparing prosection and dissection unit examination scores for concurrent and consecutive enrolled students. N represents the number of students who scored higher on the prosection or dissection examination or received the same score on both exams.

$$\text{Degree of Learning} = f\left(\frac{\text{Time actually spent learning}}{\text{Time needed to learn}}\right)$$

FIGURE 3.1. Equation representing Carroll's Model of School Learning

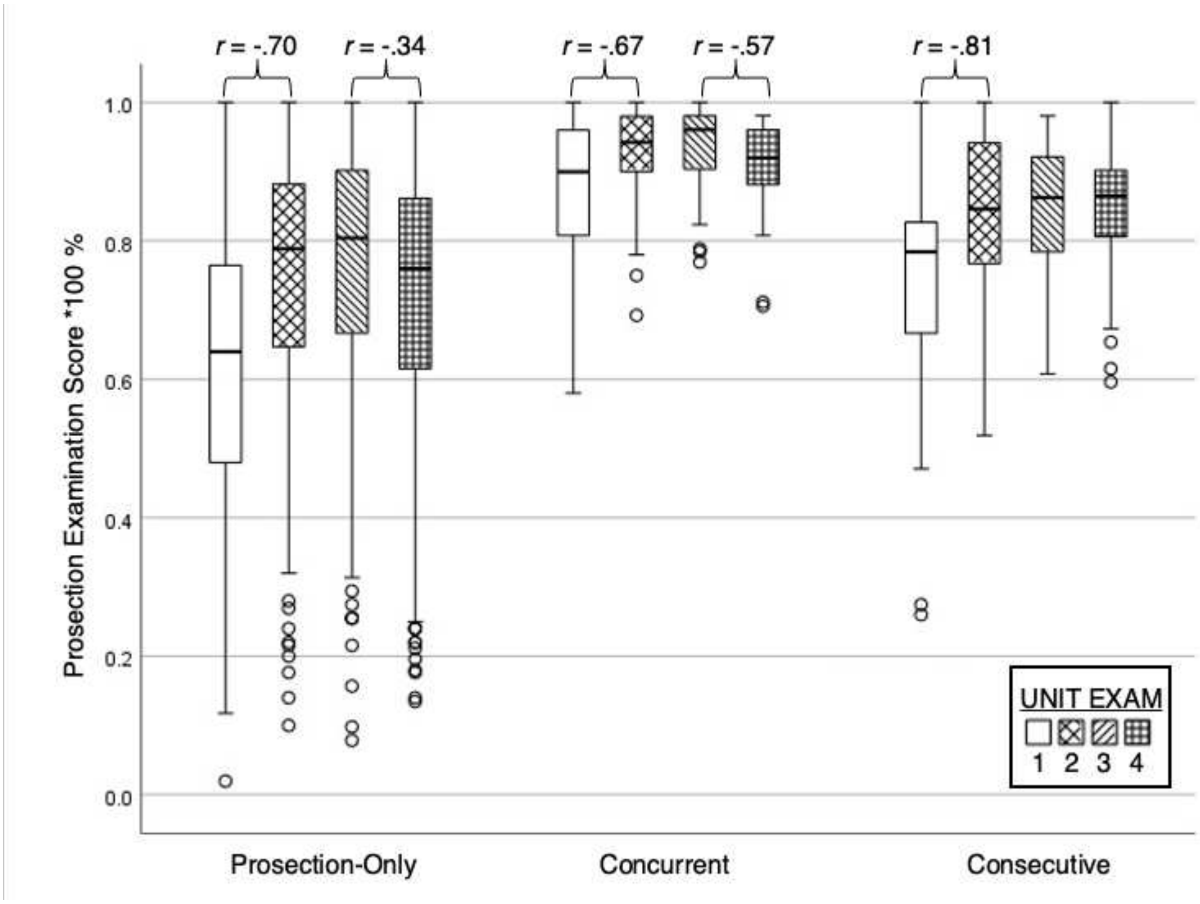


FIGURE 3.2. Prosection Laboratory Examination Scores By Cohort And Unit.

Legend: Boxplot representation of prosection unit examination scores for prosection-only, concurrently, and consecutively enrolled students with effect sizes (r value) for significant differences (*p* < .001) between examinations.

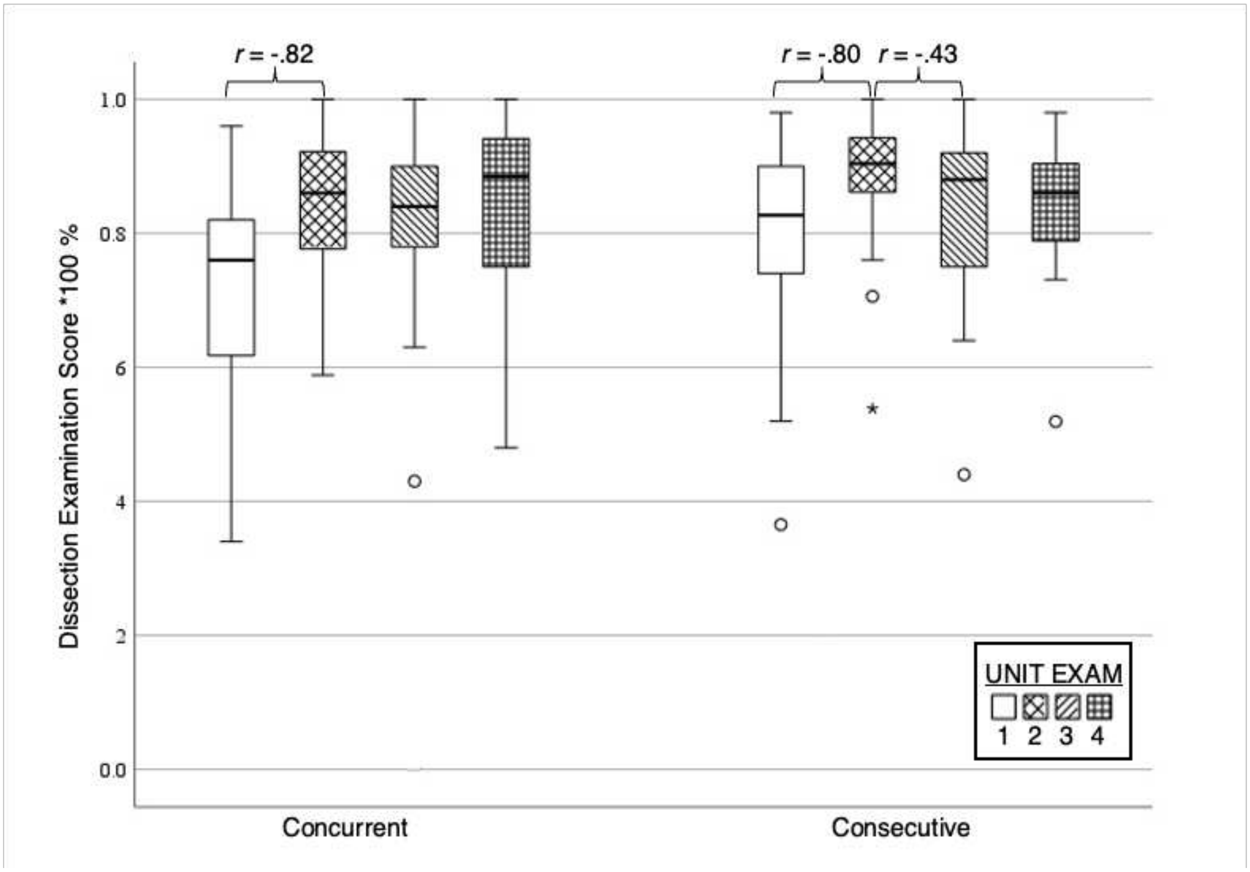


FIGURE 3.3. Dissection Laboratory Examination Scores By Cohort And Unit.

Legend: Boxplot representation of dissection unit examination scores for prosection-only, concurrently, and consecutively enrolled students with effect sizes (r value) for significant differences ($p < .001$) between examinations.

CHAPTER IV – HOW TWO MEASURES OF SPATIAL ABILITY AND THREE-DIMENSIONAL PRESENTATIONS OF LEARNING RESOURCES INFLUENCE GROSS ANATOMY ASSESSMENT OUTCOMES

4.1. INTRODUCTION

Veterinary medical education has been exposed to multiple generations of computer-based education and interactive instructional technologies. Alongside an increase in the digital distribution of course documents and provisions for remote instruction, interactive gross anatomy software such as the Virtual Animal Anatomy (VAA) software suite have undergone multiple revisions to provide educators with optimized tools to assist the varying needs of their learners.[1] The integration of cadaveric alternatives into gross anatomy curriculum has not only provided students with safer, lower cost, and more accessible learning resources[2], but the practice of mentally converting between two-dimensional (2D) and three-dimensional (3D) visuals increase learners' confidence in their anatomical understanding and clinical image interpretation.[3]

Visuospatial ability (spatial ability) is a broad term for any cognitive skill that involves non-linguistical mental manipulation of visual stimuli. Variations in a learner's spatial ability have been proposed as a factor contributing to gross anatomy learning[4], and outcomes of spatial ability assessments may serve as predictors of academic performance in a gross anatomy medical curriculum.[5]

The Mental Rotation Task (MRT) is frequently used as a measure of spatial ability in anatomy education research.[4,6–11] The task presents multiple 2D images drawn to represent 3D structures. The task involves rigid manipulation of the mentally visualized object such that the object's shape and size are preserved[12] before the participants attempt to identify if the drawings are rotated but otherwise the same or if they represent objects of different shapes.[13,14] The task has two primary components: manipulating the object through mental rotation and mental evaluation of spatial relations by comparing landmark locations. In gross anatomy research, MRT has been used to quantify cognitive processing associated with whole object orientation and 2D-3D conversion.[15] The MRT has repeatedly been associated with performance on gross anatomy assessments[4,8–11], but its role as a predictor of future performance has been questioned.[16]

Identifying complementary spatial ability tasks correlated with gross anatomy performance is necessary because some MRT participants never engage in mental rotation during spatial ability tasks.[17] The Landmark Position on Map (LPM) task is an alternative, non-rotation measure of spatial ability which prompts participants to indicate where a previously observed landmark would be located on a 2D map given two other landmarks as references.[18–20]

Before low-cost stereoscopic virtual reality (VR) headsets were available, computer-based learning resources primarily provided monoscopic experiences. For example, the traditional anatomy atlas, printed on paper or displayed on a screen, offers multiple images in 90-degree intervals (Figure 4.1a). The popularization of Apple Computer's QuickTime VR file format in the 1990s presented numerous photographs captured at set degrees intervals around an object to generate a rotatable 3D (r3D) object that could be viewed on a computer monitor (Figure 4.1b).[21] Modern implementations of r3D use unique features of photographs to generate a 3D object file with near-infinite degrees of rotation.

Anatomy atlases and r3D are monoscopic experiences in which the left and right eyes are presented with the same image. Presenting objects in these formats introduces a delay associated with the time it takes to shift attention between images to create a mental construction of a 3D object. Functionally, the difference between an atlas and an r3D is the inclusion of additional oblique views within r3D models. By comparison, stereoscopic vision of cadaveric preparations allows each eye to see a different image simultaneously, and a 3D perception can be processed without an additional delay.[22]

For monoscopic learning experiences, increased spatial ability is associated with greater anatomy assessment scores.[8,23] Huk reported that low spatial ability learners had higher scores on anatomy assessments when they studied with an atlas, and high spatial ability learners performed better when provided with r3D.[8] Levinson and colleagues replicated these findings and reported that when all spatial ability levels were provided an atlas, there was less variability in assessment performance compared to providing all learners with an r3D model.[23]

Stereoscopic 3D (s3D) shows promise as an instructional intervention.[10,24,25] A recent meta-analysis of s3D anatomy instruction stated that when low spatial ability learners were provided s3D, they outperformed low spatial ability learners who were provided an atlas or r3D instruction.[10] When s3D was provided for vascular anatomy instruction, low MRT score participants generated assessment performances comparable with their high MRT peers.[24] Jang and colleagues found that low spatial ability learners saw an additional benefit to assessment outcomes when s3D learning experiences included the ability to manipulate anatomical models directly.[25] However, due to the limited number of available s3D studies, it remains unclear if relationships between instructional methods, spatial ability measures, and gross anatomy assessment performance are restricted to specific anatomical regions or assessment formats.[9]

This study aimed to use two measures of spatial ability (MRT and LPM) as cofactors for describing how interactive computer-based learning (atlas, r3D, and s3D) and cadaveric instruction influence pre-professional gross anatomy course performance.

4.2. MATERIALS AND METHODS

This mixed-method study was completed over fifteen weeks in the spring of 2022. Students enrolled in a Colorado State University (CSU) undergraduate prosection-based domestic animal gross anatomy course were invited to participate (N = 96, 67% of course enrollment). The anatomy course consisted of a total of 455 possible points, and students could receive up to 16 points (3.5% of the course) by participating in all three sessions or an alternative activity. Course teaching staff were not involved in the experimental sessions. In addition to session data, all participants agreed to allow the researchers access to course performance and demographic data previously collected by the university. Course outcomes, research session outcomes, and demographic data were aligned and de-identified for analysis by a third party. Sessions occurred within private rooms in the basement of the CSU Anatomy Zoology building. Participants were provided Bose QuietComfort 45 wireless noise-canceling headphones. Session tasks were presented within CSU's learning management system, Canvas, on a 27-inch 2013 Apple iMac display with a USB mouse and keyboard. Before beginning each task, participants were provided a written description of the task, shown example question(s), and informed as to how much time they would be provided for the task and how much time they would have per question. The Institutional Review Board approved all procedures for Human Subjects (protocol #3041).

4.2.1. Learning Experiences

Prior to the start of the study, cadaveric preparations of an articulated bovine carpus (Session A) and a preserved canine heart (Session B) were digitally captured (Figure 4.2). Hundreds of images of each cadaveric preparation were converted into a digital object file using photogrammetry that was used to present each of the computer-based learning experiences. The cadaveric preparations were placed on a turntable that participants could rotate, and participants were instructed that due to biosecurity concerns, they were to avoid touching the specimen directly. An interactive atlas (atlas) was generated by placing dorsal, ventral, cranial, caudal, left lateral, and right lateral views into a Microsoft PowerPoint presentation, and participants could change the presented perspectives on a 27-inch digital display by pressing the keyboard's left or right arrow keys. The rotatable 3D (r3D) experience was embedded within an HTML file and presented on the same 27-inch display, and participants rotated the r3D in 10-degree intervals by moving their mouse. A Meta Quest 2 virtual reality headset was used to present two different stereoscopic 3D (s3D) learning experiences: sitting in a chair and using controllers (controller-s3D) or no controllers but sufficient space to freely walk about the room to view the s3D object (walking-s3D).

For each session, students were provided a list of available times during a two-week period and were free to sign up based on their availability. The order of signup during Session A determined which learning experience they received. The order of the first eight participants was cadaveric preparation (CP), atlas, CP, controller-s3D, CP, r3D, CP, and walking-s3D. This order was repeated for all Session A participants. Participants who received a digital experience in Session A were provided a CP in Session B, and participants who received a CP in Session A received a digital experience (atlas, controller-s3D, r3D, walking-s3D) based on the order they signed up for Session B.

4.2.2. Experimental Design

This experimental study was performed concurrently with a fifteen-week undergraduate-level gross anatomy course (Session A: weeks 2-3, Session B: weeks 8-9, Session C: weeks 14-15). Only enrolled students were recruited to participate. Specimens involved with Sessions A (bovine carpus) and B (canine heart) were not presented within the course curriculum until the corresponding session ended. Table 4.1 outlines the unit order. Session C was part of a follow-up study that recruited from the same population and included the same spatial ability measures.

4.2.2.1. Memory Tasks

The first session included two memory tasks to confirm that participants could remember visual and auditory information.[26] A sample audio file was provided before starting the auditory memory presentation unit, and participants were instructed how to adjust the headset's volume until the narrator's voice could be clearly heard. Before moving on, participants were informed that they would only be presented with the images and words once, and following the mid-session break, they would be asked to recognize the presented images and words. The visual memory task displayed 21 images, one at a time, for three seconds per image. The auditory memory presentation unit lasted for 2 minutes and 34 seconds and consisted of 12 spoken words repeated three times, first said on its own, then in a sentence, and again on its own.

Following a required five-minute break, participants were presented with two memory tests. The visual memory test gave participants three minutes to distinguish the 21 images they previously memorized from a pre-selected distractor. Three groups of distractors, representing seven questions each, were included. Novel distractors were different in shape and function, exemplar distractors were visually different but similar in appearance, and state distractors were the same object presented in different positions.[26] The order of the images and distractor were randomized between questions and subjects.

Following the visual memory test, the 12 words previously heard were presented individually alongside a distractor word. Distractors were classified as either unique, alliteration, or rhyme. Unique distractors had a different pronunciation and meaning than the previously presented word (e.g., popcorn or child). Alliteration distractors (e.g., power or pollen) were words with a similar phonetic start but a different ending. Finally, rhyme distractors started differently but shared phonetic endings (e.g., muddy or study). The word and distractor order was randomized between questions and subjects. Participants had two minutes to answer all 12 questions.

4.2.2.2. Landmark Position with Map

A Landmark Position with Map (LPM) task asked participants to accurately indicate the relative location of eight different local landmarks given the location of two other landmarks (Figure 4.3).[18–20,27] A map of the CSU campus was traced over from a satellite image acquired by Google Maps, and all points indicated the relative center of the landmark. Two anchor landmarks were labeled 1 and 2, and eight target landmarks were labeled with the letters A through H. Participants were provided a map and a key with landmark names and photographs in the study phase. After one minute of review, access to the map was removed, and participants were prompted to begin the LPM assessment immediately.

The LPM assessment presented the name and image of a target landmark. Participants were asked to use the two anchor landmarks to identify the location of the presented target landmark as accurately as possible. Then, using the mouse cursor, they could place a point to register where they perceived the target landmark was located. Unlimited adjustments could be made, and participants could return to any question within the two-minute assessment phase. All sessions used the same landmarks and maps, but the target landmark presentation order was randomized across participants and sessions.

4.2.2.3. Mental Rotation Task

A mental rotation task (MRT) was adopted from The Redrawn Vandenberg & Kuse Mental Rotation Test.[14] Participants were asked to identify if 40 pairs of monoscopic 3D drawings were the same object but drawn as if they had been rotated or if the drawings represented objects of different shapes. Pairs were presented individually, and participants could return to previous questions during the assessment. Participants had six minutes to answer all questions. All sessions used the same pairs, and the presentation order was randomized across participants and sessions.

4.2.2.4. Anatomy Presentation

Using the provided wireless noise-canceling headphones, participants listened to a brief audio narration (mean narration time across sessions: 477 seconds) in the presence of their provided learning experience (SP, atlas, r3D, controller-s3D, or walking-s3D). Before beginning the narration, all participants were provided instructions for how to view and interact with their provided learning experience. All participants started with a cranial view (Figure 4.2) but were free to change their perspective of the visual references to assist with learning. During instruction the narrator provided reorientation prompts. After the narration ended, participants received one minute to freely explore the visual in silence.

4.2.2.5. Anatomy Tests

Participants received three minutes to complete each of the four anatomy assessments. Two assessments referenced atlas images with graphical overlays, and two referenced cadaveric preparations with numbered tags. The atlas-based assessment included perspectives provided in the included atlas and oblique views. For both atlas and cadaveric preparations questions, participants were first asked to recall the answer by typing their response in a blank text box, with no provided word bank. Immediately following a recall question set, participants were asked the same questions but were provided a word bank to assist with answer recognition.

4.2.2.6. Domestic Animal Gross Anatomy Course Design

Domestic Animal Gross Anatomy is a four-credit-hour, prosection-based, undergraduate animal gross anatomy elective offered by CSU's Department of Biomedical Sciences. Undergraduate enrollment requires completing a 100-level animal biology course with a 70% or higher grade or acceptance into a biomedical science graduate degree program. In the spring of 2022, there were 143 students enrolled.

The course uses canine gross anatomy as a foundational species during the three 50-minute didactic lectures and one 110-minute prosection-based laboratory experience each week. As part of the 15 weeks of instruction, students are also exposed to cadaveric preparations detailing the unique gross anatomy of equids, bovids, camelids, felids, and suids. Twelve hours per week of teaching assistant-supported open laboratory periods were provided. Course documents, announcements, and didactic lecture recordings were available via Canvas. Course enrollment included free access to an interactive monoscopic comparative animal gross anatomy atlas, Virtual Animal Anatomy.[28] Both Canvas and the Virtual Animal Anatomy software were accessible anytime via internet-enabled devices or via provided iPad tablets while in the cadaveric laboratory.

Students receive four weeks of region-based instruction for three units (pelvic limb, thoracic limb, and thorax, abdomen, and pelvis) and three weeks for one unit (head and neck). During each unit, students receive high-frequency, low-stakes quizzes (55 of the course's 455 available points) designed to prepare them for the structure and format of the unit examinations (400 of the course's 455 available points). Each unit examination consists of a 50-minute, 50-point computerized multiple-choice test (recognition examination) and a 50-minute, 50-point short answer laboratory practical (recall examination).

4.2.2.7. Statistical Methods

Quantitative measures were visualized with histograms, and skewness was computed for each, with skewness < |1| being the cutoff for normal distribution. Effect size interpretations are based on suggestions originally provided by Cohen (1988).[29] The effect size d is regularly used with the selected analysis, and a d of |.50| was considered a typical effect size, |.20| represented a less than typical effect size, while |.80| and |1.00| represented a greater than and much greater than typical effect size, respectively. Likewise, Cohen's interpretations of Pearson's correlation coefficient were used where r = |.30| was considered typical while |.10| was considered small, and |.50| and |.70| were considered large and very large, respectively.[29]

Mental rotation task (MRT) scores were calculated by dividing the number of correct answers by 40, the number of questions asked. This method assigns equal value to all questions in the set and does not discriminate between incorrectly selected questions and those that did not receive a response. The decision to pool questions that lacked a response with incorrect answers aligns the task's output with the standard method for scoring course examinations. The differentiation of high and low MRT scores was determined by calculating z-scores, and those greater than 0 were classified as high MRT.

For the Landmark Position on Map test, the distance between landmarks 1 and 2 was used to calibrate each map. The difference between a landmark's actual location and the participant's reported location was measured in centimeters on the X and Y axes. The distance of the offset was then calculated using the Pythagorean theorem, hypotenuse = $\sqrt{(x2 + y2)}$.[27] This was repeated for landmarks A to H and reduced to a mean offset. Results were normalized with a z -score and multiplied by -1 for each mean offset. A larger positive number indicated increased map location accuracy. Differentiation of high and low LPM scores was determined by having a z-score, and those greater than 0 were classified as high LPM.

Experimental anatomy assessment scores were differentiated by atlas and specimen assessments. Scores for each were calculated by combining recall and recognition questions and dividing by 20, the number of questions asked. Two researchers independently evaluated questions, and all score disagreements were resolved before analysis. Course examinations included a recognition-based multiple-choice test presented within Canvas and a recall-based written laboratory practical. The course faculty graded all course examinations. However, only course examinations 1, 2, and 4 were analyzed due to their temporal proximity to experimental sessions. All anatomy assessments were normalized by converting to z-scores prior to analysis.

4.1. RESULTS

4.3.1. Demographic and Academic Data

Of the 96 students participating in the study, 86 completed Session A, 92 completed Session B, 84 completed Session C, and 81 completed all three sessions. Demographic data were available for seventy-nine participants. Sixty-five (68%) participants were undergraduates, and 14 (15%) were graduate students. Of the participants who reported their biological sex to the university, 72 (75%) were female, and 7 (7%) were male. While 49 (51%) received a scholarship of some kind compared with 30 (31%) who did not, only 9 participants (9%) were reported to have grown up in a limited-income household compared with 56 (58%) who reported they did not. Eight students (8%) were first-generation, while 57 (59%) were not.

For the 96 participants, the mean CSU grade point average was 3.47 (*SD* = .41), and the mean credits completed before the enrolled semester were 65.67 (*SD* = 33.12). The mean high school grade point average was 3.90 (*n* = 85, *SD* = .52), and the mean credits at entry to the university were 25.89 (*n* = 66, *SD* = 21.14). Sixty-one participants had combined SAT scores on file (*M* =1232.95, *SD* = 136.95), and 37 participants had their composite ACT score available (*M* =27.27, *SD* = 4.05).

4.3.2. Memory Tests

There was no statistically significant correlation between visual or auditory memory test scores and MRT, LPM, or experimental anatomy assessment scores. Of the 21 images participants were asked to recall as part of the visual memory test, the median amount students correctly identified was 19 (*IQR* = 2). Participants performed better identifying previously seen images with a novel image (*Med* = 1.0, *IQR* = .00) than exemplar (*Med* = .86, *IQR* = .14) or state images (*Med* = .86, *IQR* = .29). All participants were able to correctly identify the 12 words presented as part of the audio memory test (*Med* = 12.0, *IQR* = .00).

4.3.3. Changes in Spatial Ability Performance

For all sessions, MRT was found to be normally distributed, and LPM was normally distributed for sessions A and B. A paired sample t-test indicated no statistically significant changes in MRT scores between Sessions A and B, B and C, or A and C, $p = .637$, $p = .500$, $p = .103$, respectively. Likewise, no statistically significant difference was noted in LPM scores between Sessions A and B, $p = .770$. Wilcoxon signed ranks tests were used to compare LPM scores between Sessions A and C, and Sessions B and C, and no statistically significant difference was noted, $p = .365$ and $p = .836$, respectively. Additionally, there was no statistically difference in MRT or LPM scores between the learning experience groups.

4.3.4. Experimental Anatomy Performance

For Sessions A and B, linear regression analysis was computed for each learning experience to compare spatial ability measures within experimental anatomy assessments (Table 4.2). Positive r values indicate a correlation between higher spatial ability scores and increased experimental anatomy performance, and negative r values indicate a correlation between lower spatial ability and increased experimental anatomy performance. The MRT measure was correlated with sixteen of twenty (80%), and LPM was correlated with ten out of twenty (50%) experimental anatomy assessment scores.

For the MRT measure, the effect size was most consistent across Session A's atlas-based experimental anatomy assessment scores (Table 4.2). In Session A, the largest effect size for

MRT scores was associated with participants learning from the atlas (*r* = .91). However, a similar effect size between MRT and anatomy assessment scores was not reported for atlas-based instruction following Session A specimen-based (*r* = .01) assessments, or either Session B atlas-based (*r* = .11) or specimen-based (*r* = -.04) assessments. In Session B, negative correlations were reported between MRT and atlas-based and specimen-based assessments for participants who received controller-3D instruction, *r* = -.56 and *r* = -.29, respectively.

In seven out of eight experimental conditions, participants who received either controller-3D or walking-3D saw a negative correlation between LPM and assessment scores (Table 4.2). In Session A, a positive correlation between LPM and atlas-based assessment scores was reported for controller-3D participants (*r* = .39), but a negative correlation was reported for specimen-based assessments. In Session B, a negative correlation between LPM and atlas-based assessment scores was reported for controller-3D participants (*r* = -.41). In Session A, a positive correlation was reported between LPM scores and atlas- (*r* = .21) and specimen-based (*r* = -.43) assessments for r3D participants. However, a negative correlation was reported between LPM scores and atlas- (*r* = -.47) and specimen-based (*r* = -.34) assessments in Session B.

An analysis of covariance was used to assess if participants with low MRT or low LPM had higher experimental anatomy performance than participants with high MRT or high LPM test scores when controlling for their experimental learning experience. Results indicate that after controlling for the learning experience, there was no statistically significant difference between the experimental anatomy performance of low MRT and high MRT participants. Likewise, after controlling for the learning experience, there was no statistically significant difference between the experimental anatomy performance of low LPM and high LPM participants.

4.3.5. Course Examination Performance

A correlation was computed to investigate if there was a statistically significant association between the two spatial ability tests and the lecture and laboratory components of their temporally associated course examinations (Table 4.3). A statistically significant positive correlation was noted for all MRT relationships indicating that both lecture examinations and laboratory practical scores increased as MRT scores increased. A statistically significant correlation was not found between LPM measures and lecture examinations or laboratory practicals.

4.4. DISCUSSION

The results highlight how spatial ability is inconsistently associated with assessment outcomes following single-resource gross anatomy instruction. While no learning experience included in this study resulted in greater assessment performance than any other included resource, correlations were identified between spatial ability measures and experimental assessment outcomes among participants who used the same resource. The direction and effect sizes of reported correlations guide recommendations for resource selection when designing initial instruction and learner-specific intervention strategies.

When assessing spatial ability at weeks 2 and 3 (Session A), atlas-based assessments were consistently higher for participants with greater MRT scores. This is particularly true for participants who received an atlas during instruction. One explanation is that learners with higher MRT scores are better at interpreting previously unseen oblique views. Alternatively, the low effect size of atlas-based learners during the specimen-based assessment could be explained by participants reorienting themselves to match what they studied with the atlas. While these reorientation behaviors were observed during specimen-based assessments, the behaviors were not documented as part of the study and were not exclusive to those who received instruction with an atlas.

It has been previously reported that stereoscopic digital experiences, particularly those that offer direct manipulation of anatomical models, benefit low spatial ability learners more than high spatial ability learners[25,30] The negative correlation reported between LPM and experimental assessment outcomes for VR participants (controller-s3D and walking-s3D) is consistent with this idea. For single-resource learning experiences, the linear regression analysis supports an interpretation that high spatial ability learners benefit from monoscopic learning experiences and low spatial ability learners benefit from stereoscopic experiences. Why low-LPM learners outperform high-LPM learners when instruction is provided with a VR learning experience remains to be determined.

Like VR, Cadaveric preparations are a form of stereoscopic instruction. However, cadaveric preparations did not result in greater assessment scores for low LPM, as found with VR-based instruction. From an instructional equality perspective, cadaver-based instruction provides the benefits of stereoscopic instruction via methods that are less influenced by spatial ability.

While MRT scores have been reported to increase while learning gross anatomy[33], this study did not observe a significant change in MRT or LPM scores between sessions. However, MRT scores were positively correlated with lecture and laboratory practical scores, a finding consistent with previous literature.[5,8,23] The lack of a correlation between LPM and course performance may reflect the absence of VR within the student's anatomy curriculum. If LPM is exclusively relevant to VR-based experiences, the measure may be useful for identifying students who benefit most from VR-based instruction. As VR learning experiences become more accessible, instructors may be tempted to turn to it for introducing difficult-to-acquire content or providing instructional intervention. Anatomy students overwhelmingly support the inclusion of 3D instruction[31,32], and anatomy educators need to be prepared to discuss its use within their curriculum.

4.5. LIMITATIONS

This study contains a few methodological limitations. First, as participants were limited to learning with one anatomical model per session, variations in the models beyond monoscopic and stereoscopic perception may have contributed to differences in assessment outcomes. Secondly, the experimental learning experiences may not generalize to the way students approach studying for their gross anatomy course. Only audio narration was provided to all students; the outcomes may change if learners are provided both audio and written narration. Thirdly, multiple students selected for the VR experience noted that this was the first time they had used the technology, and the novel experience may have influenced participant engagement. Fourthly, the lack of anatomical pre-tests means that anatomy assessments measure knowledge, not learning. Although self-reported anatomy experience was not associated with experimental assessment performance, the time since prior anatomy experience was not investigated.

4.6. CONCLUSION

This study was the first to use LPM to compare the outcomes of monoscopic and stereoscopic 3D learning experiences in gross anatomy. In learning gross anatomy, monoscopic learning benefits high spatial ability learners, while stereoscopic learning benefits low spatial ability learners. Available resources and spatial ability may influence assessment performance.

Therefore, when developing new learning experiences or designing research, care should be taken to consider which population is being targeted. Based on the results of this study, MRT and LPM differentiate learners differently. The consequence is that LPM may be better suited for identifying learners who will benefit from stereoscopic learning.

REFERENCES

1. Linton, A., Schoenfeld-Tacher, R. & Whalen, L. R. Developing and Implementing an Assessment Method to Evaluate a Virtual Canine Anatomy Program. J Vet Med Educ 32, 249–254 (2005).

2. Little, W. B., Artemiou, E., Conan, A. & Sparks, C. Computer Assisted Learning: Assessment of the Veterinary Virtual Anatomy Education Software IVALA^{TM}. Vet Sci 5, 58 (2018).

3. Keenan, I. D. & Powell, M. Biomedical Visualisation, Volume 6. Adv Exp Med Biol 1235, 103–116 (2020).

4. Langlois, J., Bellemare, C., Toulouse, J. & Wells, G. A. Spatial Abilities Training in Anatomy Education: A Systematic Review. Anat Sci Educ 13, 71–79 (2020).

5. Lufler, R. S., Zumwalt, A. C., Romney, C. A. & Hoagland, T. M. Effect of visual–spatial ability on medical students' performance in a gross anatomy course. Anat Sci Educ 5, 3–9 (2012).

6. Na, Y. et al. Spatial Visualization of Human Anatomy through Art Using Technical Drawing Exercises. Anat Sci Educ 15, 587–598 (2022).

7. Labranche, L., Wilson, T. D., Terrell, M. & Kulesza, R. J. Learning in Stereo: The Relationship Between Spatial Ability and 3D Digital Anatomy Models. Anat Sci Educ 15, 291–303 (2022).

8. Huk, T. Who benefits from learning with 3D models? the case of spatial ability. J Comput Assist Lear 22, 392–404 (2006).

9. Roach, V. A. et al. Correlating Spatial Ability With Anatomy Assessment Performance: A Meta-Analysis. Anat Sci Educ 14, 317–329 (2021).

10. Bogomolova, K. et al. Stereoscopic three-dimensional visualisation technology in anatomy learning: A meta-analysis. Med Educ 55, 317–327 (2021).

11. Dickson, J., Gardiner, A. & Rhind, S. Veterinary Anatomy Education and Spatial Ability: Where Now and Where Next? J Vet Med Educ 49, 297–305 (2022).

12. Castro-Alonso, J. C. & Uttal, D. H. Spatial ability for university biology education. in Advances in Human Factors in Training, Education, and Learning Sciences: Proceedings of the AHFE 2018 International Conference on Human Factors in Training, Education, and Learning Sciences 283–291 (Springer International Publishing, 2019). doi:10.1007/978-3-319-93882-0_28.

13. Vandenberg, S. G. & Kuse, A. R. Mental Rotations, a Group Test of Three-Dimensional Spatial Visualization. Percept Motor Skill 47, 599–604 (1978).

14. Peters, M. et al. A Redrawn Vandenberg and Kuse Mental Rotations Test - Different Versions and Factors That Affect Performance. Brain Cognition 28, 39–58 (1995).

15. Bork, F. et al. The Benefits of an Augmented Reality Magic Mirror System for Integrated Radiology Teaching in Gross Anatomy. Anat. Sci. Educ. 12, 585–598 (2019).

16. Sweeney, K., Hayes, J. A. & Chiavaroli, N. Does spatial ability help the learning of anatomy in a biomedical science course? Anat Sci Educ 7, 289–294 (2014).

17. Geiser, C., Lehmann, W. & Eid, M. Separating "Rotators" From "Nonrotators" in the Mental Rotations Test: A Multigroup Latent Class Analysis. Multivar. Behav. Res. 41, 261–293 (2006).

18. Lopez, A., Caffò, A. O., Spano, G. & Bosco, A. The Effect of Aging on Memory for Recent and Remote Egocentric and Allocentric Information. Exp Aging Res 45, 1–17 (2018).

19. Lopez, A. et al. The Road More Travelled: The Differential Effects of Spatial Experience in Young and Elderly Participants. Int J Environ Res Pu 18, 709 (2021).

20. Lopez, A., Caffò, A. O. & Bosco, A. Topographical disorientation in aging. Familiarity with the environment does matter. Neurol Sci 39, 1519–1528 (2018).

21. Garg, A. X., Norman, G. & Sperotable, L. How medical students learn spatial anatomy. Lancet 357, 363–364 (2001).

22. Chytas, D., Piagkou, M. & Natsis, K. Does 3D stereoscopy support anatomical education? Surg Radiol Anat 43, 545–546 (2021).

23. Levinson, A. J., Weaver, B., Garside, S., McGinn, H. & Norman, G. R. Virtual reality and brain anatomy: a randomised trial of e-learning instructional designs. Med Educ 41, 495–501 (2007).

24. Cui, D., Wilson, T. D., Rockhold, R. W., Lehman, M. N. & Lynch, J. C. Evaluation of the effectiveness of 3D vascular stereoscopic models in anatomy instruction for first year medical students. Anat Sci Educ 10, 34–45 (2017).

25. Jang, S., Vitale, J. M., Jyung, R. W. & Black, J. B. Direct manipulation is better than passive viewing for learning anatomy in a three-dimensional virtual reality environment. Comput Educ 106, 150–165 (2017).

26. Brady, T. F., Konkle, T., Alvarez, G. A. & Oliva, A. Visual long-term memory has a massive storage capacity for object details. Proc National Acad Sci 105, 14325–14329 (2008).

27. Lopez, A., Caffò, A. O., Postma, A. & Bosco, A. How to separate coordinate and categorical spatial relation components in integrated spatial representations: A new methodology for analysing sketch maps. Scand J Psychol 61, 607–615 (2020).

28. Martin, J. F. et al. How Virtual Animal Anatomy facilitated a successful transition to online instruction and supported student learning during the coronavirus pandemic. Anatomia Histologia Embryologia (2022) doi:10.1111/ahe.12799.

29. Morgan, G. A., Barrett, K. C., Leech, N. L. & Gloeckner, G. W. IBM SPSS for Introductory Statistics: Use and Interpretation. 106–134 (2019) doi:10.4324/9780429287657-6.

30. Luursema, J.-M., Verwey, W. B., Kommers, P. A. M., Geelkerken, R. H. & Vos, H. J. Optimizing conditions for computer-assisted anatomical learning. Interact Comput 18, 1123–1138 (2006).

31. Meyer, E. R. & Cui, D. Anatomy Visualizations Using Stereopsis: Assessment and Implication of Stereoscopic Virtual Models in Anatomical Education. Adv Exp Med Biol 1235, 117–130 (2020).

32. Jacquesson, T. et al. Stereoscopic three-dimensional visualization: interest for neuroanatomy teaching in medical school. Surg Radiol Anat 42, 719–727 (2020).

33. Vorstenbosch, M. A. T. M. et al. Learning anatomy enhances spatial ability. Anat Sci Educ 6, 257–262 (2013).

4.1. TABLE 4.1. Outline of Session Units

Order	Task	Session Included	Task Time Limit (mm:ss)
1	Consent Form	A, B,	Unlimited
2	Survey 1	A, B,	10:00
3	Visual Memory Presentation	A, B*	01:03
4	Audio Memory Presentation	A, B*	02:34
5	Spatial Relations Task	A, B,	02:00
6	Mental Rotation Task	A, B,	06:00
7	Anatomy Instruction	A, B,	07:57 **
8	Visual Memory Test	A, B*	03:00
9	Audio Memory Test	A, B*	02:00
10	Anatomy Test 1	A, B,	03:00
11	Anatomy Test 2	A, B,	03:00
12	Anatomy Test 3	A, B,	03:00
13	Anatomy Test 4	A, B,	03:00
14	Survey 2	A, B,	10:00

* Included in Session B for participants who did not complete the task during Session A

** Mean time across three sessions.

4.1. TABLE 4.2. Linear Regression Analysis Effect size r for the Relationship Between LPM and MRT z-score and Experimental Atlas and Specimen Anatomy Tests in Sessions A and B

Spatial Assessment	Atlas-Based Assessment		Specimen-Based Assessment	
Resource	Session A	Session B	Session A	Session B
MRT				
CP	0.31	0.15	0.12	0.28
Atlas	0.91	0.11	0.01	-0.04
r3D	0.35	0.62	0.54	0.29
controller-s3D	0.36	-0.56	0.48	-0.29
walking-s3D	0.20	0.11	-0.06	0.29
LPM				
CP	0.08	-0.09	0.09	0.08
Atlas	0.18	0.10	0.37	0.11
r3D	0.21	-0.47	0.43	-0.34
controller-s3D	0.39	-0.41	-0.71	-0.19
walking-s3D	-0.42	-0.12	-0.55	-0.50

LP*M* = Landmark Position on Map; MRT = Mental Rotation Task; C*P* = Cadaveric Preparation; r3D = rotatable three-dimensional learning resource; controller-s3D = stereoscopic three-dimensional learning experience manipulated with virtual reality controller; walking-s3D = stereoscopic three-dimensional learning experience viewed by walking around a room wearing a virtual reality headset.

4.1. TABLE 4.3. Correlations between Spatial Ability MRT and LPM Scores and Course Examination Performance

			Lecture Examination		Laboratory Practical	
Spatial Assessment	Session	Course Unit	r	p	r	p
MRT	A	1	0.26	0.030	0.28	0.020
	B	2	0.35	0.003	0.42	< .001
	C	4	0.25	0.048	0.29	0.019
LPM	A	1	0.01	0.956	0.01	0.941
	B	2	0.13	0.241	-0.02	0.886
	C *	4	0.29	0.011	0.12	0.113

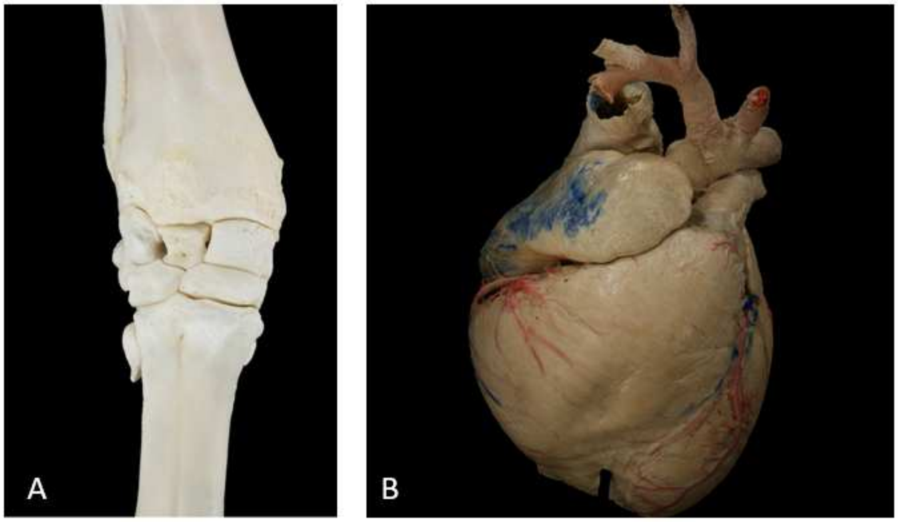
MRT = Mental Rotation Task

LP*M* = Landmark Position on Map

r = Pearson correlation, * Spearman rho

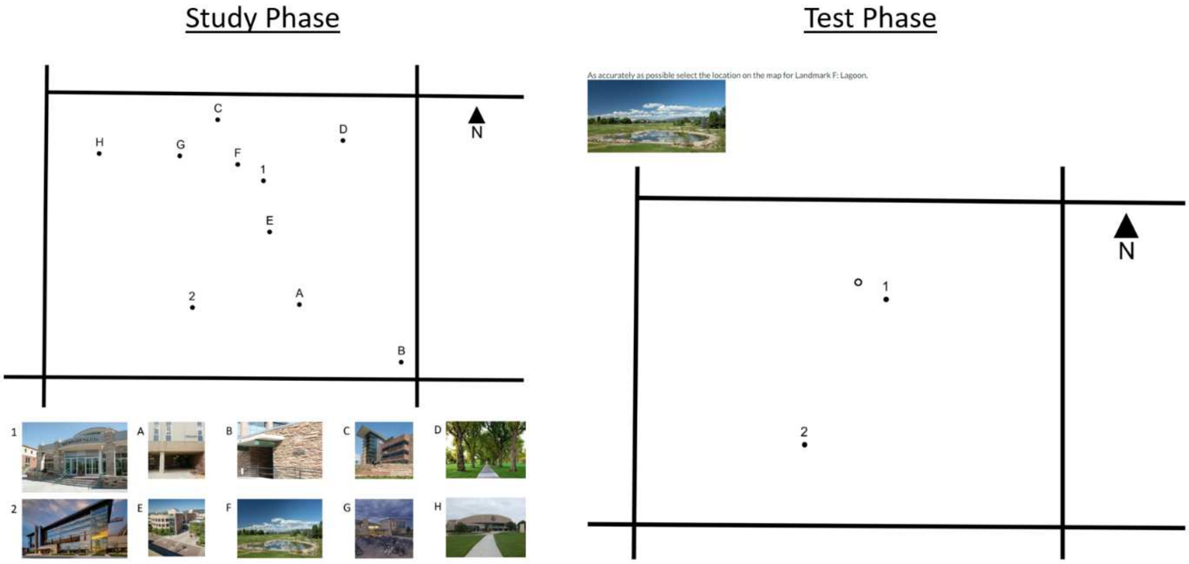
4.1. FIGURE 4.1. Comparison of Two Monoscopic Experiences of a Bovine Carpus: Atlas and r3D Object

Both the atlas and rotatable 3D object (r3D) present one perspective at a time resulting in the left and right eyes perceiving the same image. When the user requests a new perspective of the anatomical structure, the atlas rotates 90 degrees while the r3D rotates 10 degrees.



4.1. FIGURE 4.2. Images of Cadaveric Preparations of a Bovine Carpus and Canine Heart, Presented in Initial Orientation for Sessions A, B, and C.

A. Cranial view of an articulated bovine carpus. B. Cranial view of a persevered canine heart.



Study Phase

Test Phase

As accurately as possible select the location on the map for Landmark F: Lagoon.

4.1. FIGURE 4.3. Study Phase and Example Question for The Landmark Position with Map Task

In the Study Phase, participants had 60 seconds to review eight nearby landmarks. Locations 1 and 2 were anchor landmarks used for reference in all questions. Landmarks A-H were locations to be identified. Landmark A was the location of all three experimental sessions. In the test phase, Target Landmarks A-H were removed from the map. Then, participants were prompted to indicate where a named landmark was located on the map. In the example, the open circle was the correct location for Landmark F, Lagoon.

CHAPTER V – SELF-EFFICACY AND LEARNING RESOURCE SELECTION IN THE MODERN ANATOMICAL LEARNER: A MIXED METHOD ANALYSIS

5.1. INTRODUCTION

Various physical and digital educational resources have been developed to support gross anatomy instruction in response to the historical and perpetual challenges of maintaining a complete cadaveric collection. Despite the general availability of alternative instructional resources, cadaveric instruction has persisted as a preferred tool for educators and students alike.[3] In the spring of 2020, the COVID-19 pandemic forced gross anatomy courses worldwide to shift from a reliance on in-person instruction and physical models to remote instruction with virtual tools.[4–6] As the pandemic persisted, the number of digital resources available for gross anatomy instruction increased.[7] Forced resource diversity demanded revisions to instructional methodologies.

While not all changes were perceived as helpful, well-received instructional methodologies and learning resources remained a component of course curriculums following the return to in-person instruction.[8–10] However, new students entering a revised gross anatomy course e no context to differentiate between effective, tested resources, and untested novel resources. When provided with an abundance of resources and learning experiences, professional students prioritize resource selection in the context of their own learning environments.[11]

Self-efficacy is a factor associated with academic motivation[12], and increased self-efficacy has been reported to predict performance on gross anatomy examinations.[1] For this reason, self-efficacy measures have become more common in anatomy education research. However, the evaluation of low self-efficacy learners remains equally important. Compared to high self-efficacy learners, learners with low self-efficacy are more willing to engage with supplemental resources that do not improve assessment outcomes[13], and are more likely to report higher levels of burnout.[14]

It has been suggested that increased rates of cadaveric engagement have been linked with increased summative assessment scores in high self-efficacy learners.[1] However, most research involving gross anatomy self-efficacy predominately investigated in-person learning experiences. With anatomy courses integrating more digital resources implications of online self-efficacy should concern anatomy educators. This study examines a population of learners following three phases within the same semester that included some online instruction in the context of a course designed to merge hands-on experiences and best practices with digital resources. The following research questions are addressed within this study:

1. How does the self-regulated learner select a gross anatomy learning resource?
2. Do high and low self-efficacy learners differ in resource utilization within a hybrid lecture-laboratory gross anatomy course?
3. Is there a difference in assessment outcomes if a preferred resource is unavailable?

5.2. METHODS

5.2.1. Design

There were three phases in this mixed-method analysis. In phase 1, interviews and surveys asked participants to describe their experiences learning gross anatomy. In phase 2, the relationship between the participant's performance on gross anatomy examinations was compared with the results of a self-efficacy survey. In phase 3, in order to understand how resource preference influences assessment outcomes, participants were tested on the anatomy of a novel specimen following instruction with either a preferred or non-preferred learning resource.

5.2.2. Participants

This study was completed over fifteen weeks between January and April of 2022 at a public, four-year university in the Western United States with approval for Human Subjects research from the Institutional Review Board (IRB 3041). Two-thirds of students enrolled in an undergraduate domestic animal gross anatomy course elected to participate in the study (N = 96). Students could participate in up to three data collection sessions, which occurred during the two weeks prior to course examinations 1 (n = 85), 2 (n = 93), and 4 (n = 83). Participants received five extra credit points toward their final anatomy course score for each session in which they participated, and those who completed all three received one additional point (up to 16 total extra credit points; 3.5% of the course). An alternative community outreach activity was also offered for extra credit.

All participants were enrolled in the same Spring 2022 undergraduate animal anatomy course. Provided instruction includes three fifty-minute lectures per week, one 110-minute per week prosection-based instructional laboratory, 12 hours of optional open laboratory time staffed with teaching assistants, and unlimited access to Colorado State University's Virtual Animal Anatomy (VAA) software via the Learning Management System. Students receive four weeks of region-based anatomy instruction per unit to prepare them for a two-part non-cumulative examination covering approximately 500 anatomical structures. The exception is unit four, which is limited to three weeks of instruction due to the fifteen-week semester. Anatomy performance was measured using equally weighted computerized multiple-choice questions (lecture) and fill-in-the-blank laboratory practical questions (laboratory).

The university provided demographics and academic data. Participants' mean grade point average at the start of the semester was 3.47 (*SD* = .41), with a mean of 65.67 (*SD* = 33.12) credits completed. Most participants who reported their biological sex indicated female (*n* = 88; 91.7%), while 8.3% (*n* = 8) reported male. Most students were pursuing an undergraduate degree (*n* = 81; 84.4%), although 15.6% (*n* = 15) were registered as graduate students. Participants self-identified as White (*n* = 61; 63.5%), Hispanic or Latino (*n* = 16; 16.7%), Multiracial (*n* = 8; 8.3%), Asian (*n* = 5; 5.2%), Native American (*n* = 3; 3.1%), Black (*n* = 2; 2.1%), and International (*n* = 1; 1%). All participants reported familiarity with English. Eighty-five (89%) participants shared self-reflections on their understanding of non-English languages. Most participants (*n* = 50; 52.1%) reported that English was the only language they spoke in the past year, and 60 (62.5%) participants said they only spoke English before 2020. Only nineteen (22%) participants agreed they were fluent in a non-English language.

5.3. PHASE 1 – THE SELF-REGULATED LEARNER

Phase 1 aimed to answer the first research question, "How does the self-regulated learner select a gross anatomy learning resource?" Thematic analysis is a research method that identifies patterns within qualitative datasets.[15] Combining data-derived (inductive) and framework-derived coding (deductive coding) increases the rigor of thematic analysis.[16] In this present study, an inductive thematic analysis was performed on 40 interviews to explore how learning resource selection occurs within the novel context of a prosection-based animal gross anatomy course. A deductive thematic analysis was performed on eighty-three survey responses using an experimental framework generated from the themes identified during the inductive methodology.

5.3.1. Interviews

Participants were assigned an anonymous identifier and then asked to sign up for a one-hour block before each session. Three rooms were used for the study to accommodate all participants and provide sufficient privacy and social distancing. Participants were assigned to one of five sub-groups determined by order of study arrival and then asked to complete tasks that measured short-term memory, spatial ability, anatomy, and self-efficacy (described in more detail in phases 2 and 3). Upon completing the tasks, they were asked if they would be willing to participate in a post-session interview. Interview selection was randomized by interviewing the first participant from each sub-group to agree to complete an interview. Once all five sub-groups were interviewed, another round of interview selection began. Thirty interviews were conducted across the first (n = 12; interviews 1-12), second (n = 8; interviews 13-20), and third (n = 10; interviews 21-30) sessions.

The same room was used for all interviews. It was selected because it was large enough to allow more than six feet between the participant and interviewer, had a door with a window to enable the participant to see if anyone was in the hall, and was sufficiently insulated to prevent conversations from being overheard. At the start of each interview, participants consented to have the conversation audio and video recorded. The video was directed at the interviewer to confirm that body language was consistent across interviews. All interviews were conducted in English. At the start of each interview, participants were prompted to begin by sharing their experiences learning gross anatomy. Occasionally the interviewer would ask the participant to clarify or expand their statements. The participants determined their interview length, ranging from 2 minutes, 21 seconds to 59 minutes, 51 seconds (M = 26 mintues, 58 seconds; SD = 11m:36s; skewness = .45).

5.3.1.1. Interview Analysis After all 30 interviews were completed, the captured audio was transcribed. In addition to dialogue, the transcription also included audible actions (examples: laughs, chortles, stammers, dental clicks, sighs, snorts) and utterances (examples: um, hmm, uh-huh, ugh). The pacing of speech was documented by the presence and duration of pauses and speech emphasis (for example, change in volume or rate).

Thematic analysis of the interviews began with the first two authors independently reading through all transcripts and noting overall impressions of the dataset.[15,17] This first step revealed that participants frequently compare learning resources. Instead of describing interacting with a resource as an independent experience, participants would compare it with another resource.

> Interview 19: I'll kind of read through notes, but not as much and I like to just draw things out.
>
> Interview 25: I think open lab is better in my opinion, just because like there's not as many people compared to [the scheduled] lab.

This initial interpretation of metacognitive processes is described in multiple self-regulated learning frameworks.[18] The first two authors completed each step independently throughout the analysis, meeting between steps to discuss how to interpret and apply codes and themes. For each step, multiple rounds were performed as deemed necessary.

It was recognized during the familiarization phase that participants often provided context as to why a resource was or was not valued. As this observation was relevant to the research question, care was taken to code statements of resource selection separately from those describing the perceived value of a resource. When discussing generated codes, the repeated review and comparison of interview excerpts resulted in existing codes being rephrased and new codes being created to fit the underlying excerpt better. To explain how the self-regulated learner selects a gross anatomy learning resource, the first two authors generated a visual representation of the relationships between emerged themes assessment outcomes (Figure 5.1). Following the initial analysis, the methodological process was shared with coauthors and other parties for evaluation and refinement. The process of refining codes and peer debriefing enhances the trustworthiness of a thematic analysis by providing increased credibility.[19]

5.3.1.2. Interview Findings The results of the inductive thematic analysis describe a relationship where adjustments to gross anatomy learning strategies by perceived disruptions to learning, such as crowded classrooms, result in learning outcomes that influence their preference for prioritizing a resource. Changes in resource preferences influence the inclusion of learning strategies. An initial experimental framework was developed from the three themes which emerged from the 40 analyzed interviews (Figure 5.1). Identified themes included: resource preferences, learning strategies, and recognition of disruptions. Themes and associated subthemes follow.

5.3.1.2.1. Learning Strategy Preferences. Participants described seven strategies used to learn gross anatomy. These were: learning in 2D and 3D, drawing-to-learn, repetition, mnemonics, prioritizing orientation, establishing an understanding with an interactive anatomy atlas (VAA), and observing experts. The number of sub-themes presented demonstrates the diversity of available strategies considered by students. Although the participants discussed similar resources, they sometimes disagreed that specific tools were helpful.

> Interview 3: I feel like I learn maybe a little better with the 2D [atlas]. Just 'cause I can see like, you can keep going through the muscles or the bones or whatever and see like, "okay, this is deep to that, this is superficial," and it's kind of easier for me to visualize it.

> Interview 13: It's, it's hard for me to do it with [the] 2D [atlas] just because it's going to look, in my opinion, completely different on like a specimen or like the 3D-, the virtual anatomy website.

5.3.1.2.2. Recognition of Disruptions Disruptions to learning gross anatomy was classified into three sub-themes: information overload, undesirable environments, and absence of content within the VAA software. Participants describing information overload described a situation where engagement resulted in mental tiredness or difficulty following instructions.

> Interview 4: …not like exhausting but definitely, I could notice that like my brain is getting a little more tired by the end of it.

> Interview 24: I don't know, 'cause I was like, "Oh, okay," but then more information was coming out and I was like, "Oh! Oh, wait! Now I'm lost.

The scheduled laboratory periods were described as having high student to teaching assistant (TA) ratios resulting in the perception of a crowded learning environment. Participants remarked that they had difficulty learning in that environment.

> Interview 1: I just think there's too many students, and not enough TAs. And not enough of that one-on-one, or like small group setting that I need.

> Interview 5: I feel like when there are other people there, it's more distracting and I have a little bit more difficulty focusing on what I'm trying to figure out.

Participants were sensitive to the absence of content and information available in VAA. Within the undergraduate course, the first half of each of the four units is primarily canine anatomy, with the large animal anatomy presented in the second half. Participants described two ways in which this disrupts learning. The Virtual Canine Anatomy (VCA) section of the VAA is more robust than other sections, and learning strategies that rely on the software may struggle when transitioning species. Other learners reported being overwhelmed by the content provided within the canine sections and needed help focusing their attention.

> Interview 23: I think it would have been useful to have more anatomy in the program for other things than the dog. Like, I felt like I would be looking for something in the horse, like leg muscles or something, and they wouldn't have that.

> Interview 30: The VCA is really nice, but I really wish like, especially when you're learning, if we had our long objectives list or a short objectives list and the VCA automatically pulled up the photo, and it had all the things in it, instead of flipping back and forth.

5.3.1.2.3. Resource Preferences Participants described three resource traits (learning feedback, tactile processing, and phonetic processing) and two specific learning resources (teaching assistants and VAA) influencing resource preference. Interactions with other classmates and teaching assistants were described as positive, either through clarifying course concepts or affirming learning.

> Interview 8: But I love going to lab and like talking to the TAs and having them actually like point it out and like quiz me on things like that.

Positive feedback increased self-confidence and supported learners' belief that prior learning was effective.

> Interview 25: And like, just like having that praise after I got like something right, like boosted my confidence a lot 'cause like, I-, I do know what I'm talking about.

> Interview 28: But I also think I learn better when I'm confident, and having a TA totally increases my confidence 'cause I can say, "I think this is this," and they either say yes or no right there. So, instantly in my mind I can either eliminate that idea or solidify it, if that makes sense.

Support for the importance of learning feedback is found in the criticism of teaching assistants.

> Interview 11: They wont really let us do it on our own, they'll kind of jump in.

> Interview 21: I got really frustrated at one TA because I asked her something, I-, I asked her if I was labeling this thing correctly, and she told me that she wasn't going to just like tell me.

> Interview 22: Sometimes a TA would be helping us and she'd be like, "Just look at the posters on the wall," and I'd look at it and then I would just go back and like (chortles), I don't know, I didn't use them very much.

However, the benefits of learning feedback were not restricted to social interactions. One participant also described how assessments provided them with learning feedback.

> Interview 17: I think the things that I did differently is I actually went through the, you know, the guided quizzes that we have. I went back through all of those {uhm}, that helped and then I went to open lab.

Participants described learning resources that combined anatomical information with visual representations. Managing these components appears to play a role in resource preference.

> Interview 13: I think just because my brain just goes a mile-a-minute usually and so-, and when someone's talking, sometimes my brain will drift off somewhere else, and I'm like, I have to draw myself back.

> Interview 18: Not having a list of words in front of me when I'm being taught something, I can't figure out what I'm hearing. And because I'm a visual person, like if I see the word and I hear the word while I'm looking at the thing, then I'll remember it. But like, when I'm just hearing, whether it's a physical specimen or VR, it was kind of hard sometimes to keep up.

Participants provided two justifications for why anatomy software is a preferred resource. First, the software's interactivity helps focus attention on objects of interest. This feature may be valuable when information overload is a concern. Second, software curates ideal visuals reducing the need to locate alternative references.

> Interview 15: The most useful portion of the virtual anatomy was when there were four images on the same picture. So, you would get like left lateral, caudal/ cranial, and right lateral all side by side. So, I could click on one [labeled] structure, and it would be highlighted in every single image on how it would look from that view. So, I had a-, I could really make a good 3D image of what that structure would look like.

> Interview 21: {Uhm} and I mean the overall-, the V-, the VAA, not [the] VR, but {uhm}, it's really helpful in general to be able to see all-, 'cause not all of the specimens have

> everything, so it's nice to be able to have all of it online that you can reference all the different things that you need to.

However, when structures are visually similar, participants value tactile processing.

> Interview 22: I find it [touch and texture] very important, especially when it comes to anatomy 'cause you can feel the difference between nerves, and arteries, and veins, and you can tell the differences by feeling them or you feel the different like notches and like fossae and these like-. I feel like touch is really important in anatomy.

> Interview 29: I would say where it's beneficial to be able touch the nerves, veins and arteries because theoretically the artery should be bigger, you know, but like feeling a difference of like, oh, this is what a nerve feels like. Oh, this is what, an artery versus a vein.

5.3.2. Surveys

Before leaving the third session, participants (N = 84; 100%) responded to three open-ended prompts in a survey hosted on their Canvas (Table 5.1). Participants could complete these prompts in any preferred order and switch between prompts at will. However, once the survey was submitted, responses could not be modified.

5.3.2.1. Survey Analysis Survey responses were exported into an Excel spreadsheet for deductive thematic analysis. The same six thematic analysis steps were performed: familiarization, code generation, searching for themes, reviewing themes, defining themes, and writing up the analysis. The first two authors reviewed the initial framework generated from the previously described interview data and discussed the main themes: Learning Strategy Preferences, Recognition of Disruption, and Resource Preferences. An experimental framework was generated from the analysis (Figure 5.2).

5.3.2.2. Survey Findings The findings from the deductive thematic analysis describe a process where confidence translates into assessed outcomes via a cyclical learning experience.

5.3.2.2.1. Learning Strategy Preferences. Thirty-seven percent (37%; *n* = 31/84) of survey responses discussed learning strategy. Participant responses (71%; *n* = 22/31) mainly described learning with laboratory specimens (osteological specimens and cadaveric preparations). Twenty-six percent (26%; *n* = 8/31) of participants described the strategy discussed using the Virtual Animal Anatomy software program. Six surveys (19%; *n* = 6/31) described laboratory specimens and the Virtual Animal Anatomy software. One survey (3%; *n* = 1/31) stated their learning depended entirely on working with teaching assistants to connect anatomy information with relevant visuals.

Participants described improved three-dimensional visualization and the ability to touch structures as reasons they prioritized learning from laboratory specimens. Nineteen surveys directly mentioned how seeing the three-dimensional structure of laboratory specimens supported their learning by improving anatomical interpretation.

> Survey 2212: Having the specimen allows for me to have a continuous field of view to connect different anatomical features.
>
> Survey 2321: Specimens are three dimensional and I can see how different features look from different angles and views, while from a photo I only learn the features from one angle/view when in reality they may look different.

Participants noted that while rotatable three-dimensional gross anatomy models were easier to learn from than atlas images, they still prioritized laboratory specimens.

> Survey 2343: 2D [atlas] images are sometime more difficult to visualize attachment/connection points. Rotatable specimens are definitely easier, but I prefer a physical specimen to look at.

The ability to touch and interact with laboratory specimens was referenced in five surveys. Participants remarked that handling laboratory specimens helped them remember anatomy objectives.

> Survey 2268: I need to touch, move, poke, etc. to grasp the material. Often when using 2D pictures, the overall specimen cannot be appreciated, or the touch of it cannot be performed. Touch provides the essential priming for memory recall that I require to be successful.

> Survey 2303: I enjoy specimens more because I can actually touch them, and I think this helps me to remember better. Pictures are also helpful, but I tend to use them as a way to orient myself better on a specimen.

Participants who prioritized interacting with VAA described its value as an alternative to learning from laboratory specimens.

> Survey 2229: Sometimes, it is easier to relate when you have real specimens or photos of real specimens.

> Survey 2332: I think both can be a good supplementation of learning. The physical specimen would always outweigh the technology because that is what you will be

> working with in the future as well. However, I think the implementation of the VAA, whether that be 2D or 3D can be incredibly beneficial to my learning.

Some surveys indicated that the order in which they utilized VAA or laboratory specimens impacted their learning.

> Survey 2214: I study mostly from photos, so I get all my knowledge from photos before bringing it to specimen. I can't do it the other way around.
>
> Survey 2273: It is easier for me to go from a specimen to a photo since I would have already seen it in person.

5.3.2.2.2. Recognition of Disruptions Learning

Thirty-four percent (34%; *n* = 29/84) of survey responses discussed disruptions to learning. Disruptions to learning were classified into two sub-themes: information overload and learning-assessment alignment. Information overload included concerns about anatomical orientation (65%; *n* = 19/29) and auditory processing of anatomical concepts (58%; *n* = 17/29). Learning-assessment alignment includes perspectives on the relationship between how a participant learned gross anatomy content and how their performance was assessed.

Orientation refers to understanding an object's positioning, and participants reported that difficulty orientating can influence learning. Survey responses identified that participants engage in orientation behavior when using anatomy resources and that confidence in orienting themselves influenced resource selection.

> Survey 2277: The orientation of the specimen affects what I am looking at.
>
> Survey 2353: Yes. Having the specimen is much better because it's 3D, and you can rotate it and orient yourself.
>
> Survey 2298: It's much easier to identify things off a photo especially if you have already studied that photo because things don't change. Whereas identifying things on a cadaver things can be rotated or differ from species to species.

Participants expressed difficulty learning when information is presented entirely in an auditory format. This difficulty was often described as a skill the participant lacked.

> Survey 2212: I do not learn or retain [sic] information via auditory teaching.
>
> Survey 2317: I can sometimes lose track of what they are saying and miss the important thing that was mentioned.

Some expressed that they believed they could learn the information they heard, but only if additional support was provided.

> Survey 2224: I'm not an auditory learner. Without someone pointing and showing what they are talking about, I struggle to pick anything up.
>
> Survey 2247: I really struggle with learning by listening if I don't have something to read to go along with it.

Alternatively, participants linked perceived auditory processing deficits with orientation concerns.

> Survey 2223: I got lost when new words started to show up and caught on when I heard words I recognized and found where on the specimen we were referring to.

> Survey 2273: I had a really hard time when I was told it verbally because I could not orient myself on my own.

5.3.2.2.3. Resource Preferences Forty-three percent (*n* = 36/84) of survey responses discussed resource preference. Participants described how preferences for learning resources were associated with developing anatomical orientation (83%; *n* = 30/36), understanding how anatomical features were related (44%; *n* = 16/36), and skill transitioning between images and laboratory specimens (25%; *n* = 9/36). Survey responses exclusively described learning using atlas images provided within the VAA and available laboratory specimens. Participants overwhelmingly preferred to use laboratory specimens (94%; *n* = 34/36) when studying gross anatomy. Most (72%; *n* = 26/36) surveys only described laboratory specimens, while nine (28%; *n* = 10/36) mentioned using laboratory specimens and the VAA.

Those with exclusive preference for laboratory specimens cited how the three-dimensional structure of anatomical models helps them with anatomical orientation.

> Survey 2212: It is easier to visualize the different components when there is a physical specimen in front of me, rather than a computerized specimen.I can hold, rotate. and have full 360 degree visual field of a physical specimen where as the pictures on the computer are one view at a time. This can make connected the different views into one cohesive object difficult for me.

> Survey 2321: Specimens are three dimensional and I can see how different features look from different angles and views, while from a photo I only learn the features from one angle/view when in reality they may look different.

Those who reported they learned with a specimen and anatomy software still preferred the laboratory specimens because they found the spatial relationships provided by specimens easier to interpret.

> Survey 2240: Yes, the real specimen and 3D image are identical to me in terms of learning the anatomy. The 2D images and flat photographs can be hard to establish the spatial relationship.

> Survey 2201: With a specimen you can rotate the object to see it from different sides. With a picture you cant do that so it takes more effort learning from pictures.

A successful anatomy student needs to learn to differentiate anatomical structures even when they pass under or through other anatomical structures. Again, participants described how the laboratory specimen assists with learning.

> Survey 2347: I think having a physical specimen causes greater performance than a photo. I think this is because it is easier to distinguish between features when you have a 3D model in front of you.

> Survey 2321: Real specimens help me conceptualize better and help me see where things are supposed to be as I learn. I can manipulate it more so I can learn anatomical features better and retain it in my head.

Those who used atlas images to interpret anatomical relationships commented that they are easier to interpret than laboratory specimens because they simplify transitioning between superficial and deep layers. Once the atlas images are understood, these participants reported the atlas images can serve as a guide for exploring the diversity presented in the laboratory.

> Survey 2249: Photos tend to have a more simple/pretty example that is constant, while the real specimens look different through handling and position
>
> Survey 2302: Pictures are also helpful, but I tend to use them as a way to orient myself better on a specimen.

Participants described how transitioning between two-dimensional images and three-dimensional laboratory specimens was challenging. Some solved this problem by using a preferred resource as leverage to master another learning resource.

> Survey 2214: I study mostly from photos, so I get all my knowledge from photos before bringing it to specimen. I can't do it the other way around.
>
> Survey 2332: It was helpful to have many ways to learn the material and see it different ways. For me I found it easiest to study real specimens and then go to technology if I still did not understand.

Others described how they perceived images and specimens as separate objectives.

> Survey 2268: Technology and real specimens are helpful only if they're presented and tested on in the same way.
>
> Survey 2332: It is easiest to identify something in a picture if studied in a picture or to identify it on a specimen when studied on a specimen.

5.3.3. Phase 1 – Conclusions

Phase 1 focused on the perceptions of learning resources by gross anatomy students. An inductive thematic analysis of 40 interviews generated an initial framework that identified three themes: learning strategy preferences, recognition of disruptions, and resource preferences. A deductive thematic analysis of 84 surveys indicated these themes represented a cyclical process resulting in an experimental framework. Participants who reported resource preferences established learning strategies that utilized preferred resources. Information overload, undesirable environments, and the absence of content via their preferred resources could disrupt learning strategies. Regardless of which resource was preferred, participants described how access to preferred resources increased confidence in their ability to be successful.

5.4. PHASE 2 – THE SELF-EFFICACIOUS LEARNER

Phase 2 aimed to answer the research question, "Do high and low self-efficacy learners differ in resource utilization within a hybrid lecture-laboratory gross anatomy course?" Participants completed self-efficacy surveys (Table 5.2) during all three sessions and learning resource utilization measures (Table 5.3) during sessions 2 and 3. Anatomy performance was measured as part of regular course examinations.

5.4.1. Self-Efficacy, Learning Resource, and Anatomy Performance Measures

The self-efficacy measure included eight 7-point Likert scale questions presented in Table 5.2. Questions one through four aimed to measure how confident they were that they would be successful in the enrolled anatomy course, while questions five through eight aimed to measure how confident they were that they could successfully answer various examination question formats. To verify that the self-efficacy survey had good internal consistency, Cronbach's alphas were computed for the three sessions, .84, .88, and .91, respectively. No question was excluded, and self-efficacy was calculated as the mean of all eight items. Participants were ranked as low or high self-efficacy based on their score relative to that session's mean. Temporally proximal measures were used for all comparisons.

In sessions 2 and 3, participants were asked to report the number of hours per week they used 20 resources (Table 5.3). Three measures were derived from this data: how many hours per week participants use resources (resource time), the percent of resources used (total resources), and time using third-quartile resources (preferred resources). Resource time was calculated by summing all reported weekly use. Total resources were calculated by dividing the number of reported resources used per week by twenty. Preferred resources were calculated by dividing the third-quartile median by the total number of resource hours reported by the participant.

5.4.2. Results

Sixty-seven percent (N = 96/143) of enrolled students participated in the study. Descriptive statistics are included in Table 5.4. Participation in the study was consistent across sessions, with a small increase during session 2. No course examination scores were available for fourteen participants; three participants only had scores for the first two examinations. All effect size interpretations were based on statistical methods originally described by Cohen (1988).[20]

To see if self-efficacy changed during the three sessions, a Wilcoxon signed ranks test compared self-efficacy between sessions 1 and 2, and 2 and 3. There was no statistically significant difference in self-efficacy between sessions 1 and 2, N = 82, $z = -.418$, $p = .676$, $r = -.05$. However, a statistically significant increase was reported between sessions 2 and 3, N = 83, $z = -2.191$, $p = .028$, $r = -.24$. Using session 1 self-efficacy scores to differentiate participants, only those with high self-efficacy increased their score between sessions 2 and 3. Of the 43 high self-efficacy participants who completed sessions 2 and 3, 12 had higher self-efficacy at session 2, 25 had higher self-efficacy at session 3, and 6 did not change, N = 43, $z = -1.987$, $p = .047$, $r = -.30$.

5.4.2.1. Self-Efficacy and Examination Performance

An independent samples t-test was performed to compare low and high self-efficacy (Table 5.5). High self-efficacy participants consistently scored higher on course examinations than those with low self-efficacy (mean delta = 7.1). While statistical significance was inconsistent, lecture examinations one and two saw typical effect sizes, $t(70) = -2.11$, $p = .039$, $d = .57$ and $t(79) = -2.33$, $p = .023$, $d = .52$, respectively.

5.4.2.2. Self-Efficacy and Resource Utilization

To better understand how self-efficacy measured during session 1 was related to future resource use, session 1 self-efficacy scores were associated with three measures of resource utilization measures completed during sessions 2 and 3.

5.4.2.2.1. Resource Time

A Wilcoxon signed ranks test was used to compare the total reported time using resources during sessions 2 and 3. Of the 43 participants with session 1 high self-efficacy who also completed sessions 2 and 3, 33 reported more hours during session 2, 8 reported more hours during session 3, and 2 reported no change in resource use. This difference indicating high self-efficacy participants used resources less during session 3 than session 2 was statistically significant, N = 43, $z = -3.287$, $p = .001$, $r = -.50$. However, there was not a statistically significant difference in hours for low self-efficacy participants, N = 31, $z = -1.451$, $p = .147$, $r = -.26$.

5.4.2.2.2. Total Resources.

A paired samples t-test indicated that high self-efficacy participants used, on average, fewer resources when studying for their gross anatomy course during session 3 than during session 2, $t(42) = 3.08$, $p = .004$, $d = -.47$. However, there was not a statistically significant difference in the number of reported resources used by low self-efficacy participants, $t(30) = -.312$, $p = .758$, $d = -.05$.

5.4.2.2.3. Preferred Resources.

A paired samples t-test indicated no statistically significant difference in the relative time spent with preferred resources for either high or low self-efficacy participants, $t(42) = 1.51$, $p = .138$, $d = -.23$, and $t(30) = -.262$, $p = .795$, $d = -.05$, respectively.

5.4.2.3. Resource Utilization and Examination Performance

Spearman's rho was used to examine intercorrelations between the number of hours each resource was reported to be used during sessions 2 and 3 with temporally proximal lecture and laboratory examination scores. As self-efficacy was found to change between sessions 2 and 3, participants were classified as high or low self-efficacy based on that session's self-efficacy survey.

In session 2, the number of hours low self-efficacy participants spent in the open laboratory with a TA was positively correlated with both lecture and laboratory examination 2 performance of, $r_s(41) = .50$, $p = .005$ and $r_s(41) = .47$, $p = .009$, respectively. A similar positive correlation was computed for high self-efficacy participants who used open laboratory time without a TA, $r_s(49) = .54$, $p < .001$ and $r_s(49) = .53$, $p < .001$, respectively.

In session 3, high self-efficacy participants continued to see a positive correlation between time spent in the open laboratory without a TA and lecture and laboratory examinations 4 scores, $r_s(41) = .37$, $p = .017$ and $r_s(41) = .31$ $p = .050$, respectively. However, the strongest correlations for examination 4 performance for low self-efficacy participants were associated with lecture-based instruction. Increased hours attending in-person lecture was positively correlated with increased lecture ($r_s(39) = .38$, $p = .038$) and laboratory ($r_s(39) = .37$, $p = .043$) examination scores. Additionally, a negative correlation was computed for hours spent using asynchronous lectures for both lecture ($r_s(32) = .-41$, $p = .024$) and laboratory ($r_s(32) = -.38$, $p = .037$) examination scores.

5.4.2.4. Disruptions to Learning and Examination Performance

Due to the allocation of four weeks per course unit and a fifteen-week semester, students receive three weeks to cover the same amount of content presented in the prior units. This abbreviated study period was expected to act as a shared disruptor to learning.

A paired samples t-test was computed to compare scores on examinations 3 and 4 (Table 5.6). A statistically significant decrease was computed for low and high self-efficacy participants on lecture and laboratory examinations. However, while low self-efficacy participants saw a larger total decrease in performance between examinations 3 and 4, results of an independent samples t-test did not find differences between low and high self-efficacy examination scores.

5.4.3. Phase 2 – Conclusions

Phase 2 provided a quantitative analysis of the relationship between self-efficacy and learning resources described by the Phase 1 experimental framework. A relationship between self-efficacy and examinations was found, but only lecture examinations were reported to be statistically significant. For participants with low self-efficacy, no significant changes were reported in the three measures of resource utilization. By contrast, while high self-efficacy participants decreased the total amount of time and number of resources used, they maintained a similar proportion of time spent using preferred and less preferred resources. Low self-efficacy learners saw performance benefits with increased use of structured resources (TA supported open laboratory time and in-person lecture attendance) while high self-efficacy learners saw performance benefits with increased use of unstructured resources (open laboratory time without TA support). These results indicate that low and high self-efficacy participants differ in resource selection strategies. However, this difference in strategy does not provide either cohort protection from decreases in examination scores following reduced time for study. While the results here indicate that resource preference does influence resource selection, additional data is needed to determine if resource preference contributes to assessment outcomes.

5.5. PHASE 3 – THE PREFERENTIAL LEARNER

Phase 3 aimed to answer the research question, “Does the presence or absence of a preferred resource influence gross anatomy assessment outcomes?” As part of the experimental framework defined in phase 1, resource preference influences the selection of learning strategies. In phase 2, it was concluded that while participants did develop preferred resources during gross anatomy study, a preference for resources did not prevent reductions in preferred resource use.

5.5.1. Methods

During sessions 1 and 2, participants used a cadaveric specimen and a digital anatomical resource (an anatomy atlas, a three-dimensional rotatable model, or virtual reality). At the end of session 2, participants were asked to indicate if they preferred learning from the specimen or the randomly selected technology they were provided. Participants were informed at this time that their choice would determine what they received should they participate in session 3.

As part of each session, participants also completed a Mental Rotation Task (MRT)[21,22], a spatial ability measure that has previously been associated with gross anatomy assessment performance.[23–27] This task prompts participants to compare two drawings representing three-dimensional shapes by mentally visualizing and rotating one of them. Participants are asked to decide if the two drawings represent the same object, but one has been rotated or if the objects are different shapes. Participants have six minutes to answer 40 questions. The MRT score is calculated by dividing the correct responses by 40.

After all participants completed the second session, they were ranked by their MRT score. The participant with the lowest MRT score was then scheduled to receive their preferred resource in session 3, and the second lowest was scheduled to receive the opposite of their preferred resource. This pattern repeated until all participants were assigned a resource. When participant scores were tied, those who completed the MRT with fewer unanswered questions were ranked higher.

During session 3, participants received seven minutes of auditory instruction alongside a single visual resource of an osteological example of an emu pelvis as determined by their MRT ranking (Image 1). Following the experimental gross anatomy lesson, participants received one minute to freely explore the specimen before taking a mandatory five-minute break. Upon return, participants completed two gross anatomy quizzes, each comprising 20 questions. Based on arrival time to the study, participants either received a quiz using an articulated emu pelvis with numbered tags or a set of emu pelvis images taken in dorsal, ventral, cranial, caudal, left-lateral, and right-lateral positioning. The first ten questions were recall questions requiring the anatomical term to be entered into a text box. When the first ten questions were submitted, participants were prompted to answer the same ten questions again, however, this time, they were provided a word bank to assist with recognition of the answer. After completing the first twenty questions, they repeated the process for the alternative test method. The specimen and atlas-based quizzes used a different set of questions.

The first two authors independently graded all anatomy assessments and discussed any disagreements to ensure consistent grading across participants. Phonetic spellings were accepted as correct, and participant atlas-based quizzes and anatomy scores were determined by dividing the number of correct answers by twenty. To account for potential differences in quiz difficulty, both quizzes were normalized with z-scores prior to analysis.

5.5.2. Results

Three of the 83 participants who completed session 3 were excluded due to technical issues. This resulted in 41 participants receiving their preferred resource and 40 receiving the alternative. Visual resource distribution included 37 participants receiving instruction with an osteological specimen and 44 with technology. Scores for the altas-based assessments (*M* = 17.01%, *SD* = 14.42, skewness = 1.004) were slightly lower than the specimen-based assessments (*M* = 17.14%, *SD* = 15.94, skewness = 1.201).

5.5.2.1. Preferred Resources and Resource Utilization

A Spearman's rho was used to examine intercorrelations between the atlas-based and specimen-based quiz scores and the three measures of resource utilization described in Phase 2, total resource time, total resources, and preferred resource time (Table 5.7). Three out of the six pairs were significantly correlated. The strongest negative correlation was between total resource time and specimen-based quiz scores, $r_s(81) = -.34$, $p = .017$. Additionally, preferred resource time was negatively correlated with atlas- and specimen-based quiz scores, $r_s(81) = -.26$, $p = .021$, and $r_s(81) = -.33$, $p = .003$, respectively. These results mean that the more time participants reported using their preferred resource or all the resources in general, the lower their experimental quiz scores.

An analysis of covariance was used to assess whether participants who received a preferred resource scored higher on an experimental anatomy quiz than those who did not receive their preferred resource after controlling for reported resource time or preferred resource time. The assumption of homogeneity was violated for the atlas-based quizzes, however, because the cell sizes were nearly identical (40 and 41), when the analysis was completed. All other assumptions, independence of observations, linear relationships, and normal distribution of the dependent variable, were met.

Results indicate that after controlling for total resource time, there was no significant difference between those who received their preferred resource (adjusted $M = 18.90$, SE = 2.43) and those who did not (adjusted $M = 15.26$, SE = 2.46) in an experimental single resource specimen-based quiz, $F(1, 77) = 1.11$, $p = .296$, $\text{eta}^2 = .01$. After accounting for preferred resource time, no significant difference was reported between participants who received their preferred resource (adjusted $M = 19.41$, SE = 2.42) and those who did not (adjusted $M = 14.81$, SE = 2.45) in an experimental single resource specimen-based quiz, $F(1, 77) = 1.78$, $p = .186$, $\text{eta}^2 = .02$. For the atlas-based quizzes neither covariant total resource time or preferred resource time were statistically significant, (observed powe$r = .24$, $p = .207$ and observed powe$r = .28$, $p = .168$), respectively.

5.5.2.2. Preferred Resources and Self-Efficacy

A Spearman's rho was used to identify positive intercorrelations between session 3 self-efficacy scores and performance on the atlas-based($r_s(80) = .22$, $p = .047$) and specimen-based ($r_s(80) = .25$, $p = .028$) experimental anatomy quizzes. A Mann-Whitney U test was computed to compare quiz performance between low and high self-efficacy participants. The 42 high self-efficacy participants had higher mean ranks (46.75 and 46.83) than low self-efficacy participants (33.59 and 33.50) on the atlas-based ($U = 535.5$, $p = .011$, $r = -.29$) and specimen-based quizzes ($U = 532.0$, $p = .010$, $r = -.29$), respectively.

Results indicate that after controlling for self-efficacy, there was no significant difference between those who received their preferred resource (adjusted $M = 19.29$, SE = 2.41) and those who did not (adjusted $M = 14.86$, SE = 2.47) in an experimental single resource specimen-based quiz, (1, 76) = 2.58, $p = .113$, $\text{eta}^2 = .03$. Similarly, there was no a significant difference between those who received their preferred resource (adjusted $M = 17.77$, SE = 2.26) and those who did not (adjusted $M = 16.12$, SE = 2.32) in an atlas-based quiz, (1, 76) = .261, $p = .611$, $\text{eta}^2 < .01$.

5.5.3. Phase 3 – Conclusions

The relationship between resource preferences and assessment outcomes was investigated using an experimental anatomy quiz. Spatial ability, a measure previously associated with gross anatomy assessment performance[23–27] was controlled methodologically. Self-reported total resource time and preferred resource time were negatively correlated with quiz performances. A positive correlation was reported between self-efficacy and quiz performance. As participants reported an increased time using resources to study for their course examinations, performance on the experimental quiz decreased, and as self-efficacy increased, quiz scores increased. When accounting for these relationships receiving a previously indicated preferred resource did not influence quiz assessment outcomes.

5.6. DISCUSSION

This three-phase study used a mixed-method approach to identify factors that might explain how self-efficacy influences assessment outcomes. In phase 1, a hybrid inductive-deductive thematic analysis was used to develop an experimental framework that describes a process where the belief that one can successfully learn gross anatomy is associated with assessment outcomes. The framework describes a cyclical process by which resource preference influences learning strategies, learning strategies are selected to respond to disruptions to learning, and self-assessment of how well a strategy managed the disruption influences resource preference.

Phase 2 identified a relationship between self-efficacy and lecture-based anatomy assessments, and a weak association was also reported between session 1 self-efficacy and laboratory exam 1. While low self-efficacy participants did not report a change in resource utilization, high self-efficacy participants reported fewer hours using learning resources and the number of resources used weekly at session 3 compared to session 2. Low self-efficacy participants saw increased lecture and laboratory examination scores with increased time in structured learning environments (TA-led open laboratory and attending in-person lectures). Only high self-efficacy participants reported changes to how they used learning resources, which may explain why increased time in an unstructured learning environment (TA-free open laboratory) was associated with increased examination performance. These results support a relationship between gross anatomy resource utilization behaviors and gross anatomy self-efficacy.

Phase 3 used an experimental condition that accounted for spatial ability and showed that receiving a preferred learning resource did not significantly influence assessment performance following instruction with a novel anatomical model. However, participants with higher self-efficacy and those who reported using learning resources for fewer hours had higher experimental quiz scores, suggesting that while resource preference may influence learning strategy selection, preference does not significantly influence one's ability to learn novel anatomy.

5.7. LIMITATIONS

This study had several limitations. First, the included population of undergraduate students was recruited during a global pandemic, and they may not have sufficiently adjusted their study strategies to accommodate the variety of resources provided as part of a laboratory-based course. In phase 1, participants were randomly selected for interviews from those who agreed to complete an interview following a nearly one-hour experimental session. Those willing to participate in an interview may share perspectives with each other that are not shared with students who declined the interview. While the interviewer was not directly associated with the enrolled anatomy course, participants may have been reserved in their responses. Additionally, metacognitive and introspective prompts provided during the sessions may have influenced students to adjust their learning stratagy. In phase 2, an error capturing resource utilization prevented the comparison of resource utilization during semester weeks 2-3. The self-efficacy measure was developed for this study and designed to identify self-efficacy associated with learning gross anatomy. While it had strong internal consistency, conclusions drawn from it would benefit from a comparison with other established self-efficacy measures. In phase 3, participants were assessed using a novel anatomical model.

REFERENCES

1. Burgoon JM, Meece JL, Granger NA. Self-efficacy’s influence on student academic achievement in the medical anatomy curriculum. Anat Sci Educ. 2012;5(5):249–55.

2. Ulfatun T, Septiyanti F, Lesmana AG. University Students’ Online Learning Self-efficacy and Self-regulated Learning during the COVID-19 Pandemic. Int J Information Educ Technology. 2021;11(12):597–602.

3. Patel KM, Moxham BJ. Attitudes of professional anatomists to curricular change. Clin Anat. 2006;19(2):132–41.

4. Pather N, Blyth P, Chapman JA, Dayal MR, Flack NAMS, Fogg QA, et al. Forced Disruption of Anatomy Education in Australia and New Zealand: An Acute Response to the Covid-19 Pandemic. Anat Sci Educ. 2020;13(3):284–300.

5. Cheng X, Chan LK, Pan S, Cai H, Li Y, Yang X. Gross Anatomy Education in China during the Covid-19 Pandemic: A National Survey. Anat Sci Educ. 2021;14(1):8–18.

6. Hielscher AC, Everse S. Virtual peer teaching in the gross anatomy lab: a format of peer teaching and learning during the COVID-19 pandemic. MedEdPublish. 2022;12:48.

7. Attardi SM, Harmon DJ, Barremkala M, Bentley DC, Brown KM, Dennis JF, et al. An Analysis of Anatomy Education Before and During Covid-19: August-December 2020. Anat Sci Educ. 2021;15(1):10.1002/ase.2152.

8. Martin JF, Arnold OR, Linton A, Jones JD, Garrett AC, Mango DW, et al. How Virtual Animal Anatomy facilitated a successful transition to online instruction and supported student learning during the coronavirus pandemic. Anatomia Histologia Embryologia. 2022;

9. Lin DC, Bunch B, Souza RZDD, Chen D, Zhou J, Zumwalt AC, et al. Effectiveness of Pedagogical Tools for Teaching Medical Gross Anatomy During the COVID-19 Pandemic. Medical Sci Educ. 2022;32(2):411–22.

10. Schurr AF, Burg BJ, Dickinson E, Granatosky MC. No cuts, no buts: Satisfaction of first-year medical students with a hybrid prosection-based model for learning gross anatomy during the Covid-19 pandemic. Anat Sci Educ. 2022;15(5):827–38.

11. Johnson S, Purkiss J, Holaday L, Selvig D, Hortsch M. Learning histology – dental and medical students’ study strategies. Eur J Dent Educ. 2015;19(2):65–73.

12. Schunk DH. Self-Regulated Learning: The Educational Legacy of Paul R. Pintrich. Educ Psychol. 2005;40(2):85–94.

13. Langfield T, Colthorpe K, Ainscough L. Online instructional anatomy videos: Student usage, self-efficacy, and performance in upper limb regional anatomy assessment. Anat Sci Educ. 2018;11(5):461–70.

14. Ferriby A, Schaefer AF. The Relationship Between Anatomical Self-efficacy and Feelings of Burnout in First-year Medical Students. Medical Sci Educ. 2022;32(2):437–46.

15. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol. 2006;3(2):77–101.

16. Fereday J, Muir-Cochrane E. Demonstrating Rigor Using Thematic Analysis: A Hybrid Approach of Inductive and Deductive Coding and Theme Development. Int J Qual Methods. 2006;5(1):80–92.

17. Maguire M, Delahunt B. Doing a thematic analysis: A practical, step-by-step guide for learning and teaching scholars. AISHE-J [Internet]. 2017 Oct 31;9(3):14. Available from: https://ojs.aishe.org/index.php/aishe-j/article/view/335

18. Panadero E. A Review of Self-regulated Learning: Six Models and Four Directions for Research. Front Psychol. 2017;8:422.

19. Nowell LS, Norris JM, White DE, Moules NJ. Thematic Analysis. Int J Qual Meth. 2017;16(1):1609406917733847.

20. Morgan GA, Barrett KC, Leech NL, Gloeckner GW. IBM SPSS for Introductory Statistics: Use and Interpretation. 2019;106–34.

21. Peters M, Laeng B, Latham K, Jackson M, Zaiyouna R, Richardson C. A Redrawn Vandenberg and Kuse Mental Rotations Test - Different Versions and Factors That Affect Performance. Brain Cognition. 1995;28(1):39–58.

22. Vandenberg SG, Kuse AR. Mental Rotations, a Group Test of Three-Dimensional Spatial Visualization. Percept Motor Skill. 1978;47(2):599–604.

23. Langlois J, Bellemare C, Toulouse J, Wells GA. Spatial Abilities Training in Anatomy Education: A Systematic Review. Anat Sci Educ. 2020;13(1):71–9.

24. Huk T. Who benefits from learning with 3D models? the case of spatial ability. J Comput Assist Lear. 2006;22(6):392–404.

25. Roach VA, Mi M, Mussell J, Nuland SEV, Lufler RS, DeVeau KM, et al. Correlating Spatial Ability With Anatomy Assessment Performance: A Meta-Analysis. Anat Sci Educ. 2021;14(3):317–29.

26. Bogomolova K, Hierck BP, Looijen AEM, Pilon JNM, Putter H, Wainman B, et al. Stereoscopic three-dimensional visualisation technology in anatomy learning: A meta-analysis. Med Educ. 2021;55(3):317–27.

27. Dickson J, Gardiner A, Rhind S. Veterinary Anatomy Education and Spatial Ability: Where Now and Where Next? J Vet Med Educ. 2022;49(3):297–305.

TABLE 5.1. Session 3 Survey Questions

1. Does the format (photos or specimen) of a gross anatomy assessment influence your performance?
2. Briefly describe how technology and real specimens impact your learning of gross anatomy.
3. A friend tells you they plan to take BMS 305 [this course] next semester. What advice would you give them?

TABLE 5.2. Self-Efficacy Survey: Seven-Point Likert Scale

1 I am confident that I can perform successfully in an anatomy course.
2 At this point in the course, I am confident that I can learn the material expected of me for the next course exam.
3 At this point in the semester, I consider my personal outcome in this course to be a success.
4 I am confident that I could tutor another student on the material presented so far in the laboratory section of my enrolled anatomy course.
5 I am confident that I can identify anatomical structures presented in a photo that I studied from a cadaver.
6 I am confident that I can identify anatomical structures presented on a cadaver that I studied from a cadaver.
7 I am confident that I can identify anatomical structures presented on a cadaver that I studied from a photo.
8 I am confident that I can identify anatomical structures presented in a photo that I studied from a photo.

Provided Responses: 1 = Strongly Disagree; 2 = Somewhat Disagree; 3 = Neither Agree nor Disagree; 4 = Somewhat Agree; 5 = Strongly Agree.

TABLE 5.3. Learning Resource Measure

Since the last exam, how many hours per week did you spend with the following?
1. Anatomy quizzes provided via Canvas.
2. Attending In-Person Lectures
3. Cadavers or Special Preparations WITH OUT TA assistance.
4. Cadavers or Special Preparations WITH TA or Instructor assistance.
5. Looking at animal anatomy photos not included in course lectures or part of the Virtual Animal Anatomy (VAA) 2D interactive atlas
6. Making or reviewing drawing you personally did.
7. Palpations done on living animals.
8. Palpations done on other humans.
9. Palpations done on stuffed animals.
10. Palpations done on yourself.
11. Reading physical or electronic textbooks (not looking at pictures).
12. Reviewing drawings done by someone other than you.
13. Skeletons or 3D models you personally own.
14. Viewing Online, Live Lecture (AKA Synchronous Lectures)
15. Viewing Previously Recorded Lectures (AKA Asynchronous Lectures)
16. Virtual Animal Anatomy (VAA) 2D interactive atlas.
17. Virtual Animal Anatomy (VAA) 3D rotatable objects.
18. Virtual Animal Anatomy (VAA) in VR.
19. Virtual Animal Anatomy (VAA) Quizzes.
20. Virtual Animal Anatomy (VAA) Recorded Videos.

TABLE 5.4. Descriptive Statistics for Self-Efficacy, Resource Utilization, and Examination Performance at Three Points During the Study by Total and Individual Anatomy Self-Efficacy Score

		Session 1			Session 2			Session 3		
		Total	Low SE	High SE	Total	Low SE	High SE	Total	Low SE	High SE
	n	85	36	49	93	44	49	83	40	43
SE	Median	5.38	4.94	5.75	5.63	4.82	6.00	5.75	5.25	6.38
	IQR	0.88	0.63	0.63	1.25	1.00	0.44	1.13	1.10	0.50
	Skewness	-1.25	-2.32	0.819	-1.17	-1.47	0.77	-1.38	0.86	-0.15
RT	Median	-	-	-	24.00	24.50	24.00	22.00	22.00	22.00
	IQR	-	-	-	22.63	11.75	13.88	12.99	12.98	13.00
	Skewness	-	-	-	1.44	1.34	1.59	0.95	1.14	0.74
TR	Mean	-	-	-	12.10	11.80	12.37	12.10	11.95	12.44
	SD	-	-	-	2.60	2.84	2.36	2.60	3.00	2.50
	Skewness	-	-	-	-0.58	-0.57	-0.47	-0.58	-0.59	-0.46
PR	Mean	-	-	-	10.09	10.41	9.91	9.72	10.15	9.30
	SD	-	-	-	3.14	3.52	2.71	3.14	3.48	2.82
	Skewness	-	-	-	0.71	1.10	0.19	0.28	0.22	0.19
	n	82	30	42	81	38	43	69	36	33
Lec	Mean	80.35	75.30	82.40	78.09	73.97	81.72	74.74	72.58	77.09
	SD	14.95	15.23	13.25	15.36	14.46	14.49	17.24	17.39	17.02
	Skewness	-0.79	-0.13	-1.16	-0.68	-0.51	-0.92	-0.57	-0.40	-0.80
Lab	Mean	64.15	57.60	67.43	76.49	73.32	79.30	70.61	67.67	73.82
	SD	20.52	22.96	19.55	18.13	17.56	18.37	20.88	20.66	20.66
	Skewness	-0.36	0.08	-0.48	-0.68	-0.42	-0.47	-0.48	-0.17	-0.17

SE = Self-Efficacy; RT = Resource Time; T*R* = Total Resources; P*R* = Preferred Resources; Lec = Lecture Examination; Lab = Laboratory Examination. Session 1 occurred during weeks 2-3, Session 2 during weeks 8-9, and Session 3 weeks 14-15. Examinations occurred at the end of weeks 4, 8, and 15.

TABLE 5.5. Independent Samples t-test for Anatomy Examination Performance by Low and High Self-Efficacy

Variable	n	M	SD	t	df	p	d
Lecture Examination 1				-2.11	70	0.039	0.57
Low SE	30	74.30	15.23				
High SE	42	82.40	13.25				
Lecture Examination 2				-2.33	79	0.023	0.52
Low SE	38	73.97	15.46				
High SE	43	81.72	14.49				
Lecture Examination 4				-1.09	67	0.281	0.26
Low SE	36	72.58	17.39				
High SE	33	77.09	17.02				
Laboratory Examination 1				-1.96	70	0.055	0.46
Low SE	30	57.60	22.96				
High SE	42	67.43	19.55				
Laboratory Examination 2				-1.49	79	0.139	0.33
Low SE	38	73.32	17.56				
High SE	43	79.30	18.37				
Laboratory Examination 4				-1.23	67	0.224	0.30
Low SE	36	67.67	20.66				
High SE	33	73.82	20.95				

SE = Self-Efficacy. Low/high self-efficacy was based on being less than or greater than the mean for that for the most temporally proximal session.

TABLE 5.6. Paired Samples t-Test of Lecture and Laboratory Examinations Three and Four for Low and High Participants in Session 1

	Lecture Examinations				Laboratory Examinations			
	3 Mean	4 Mean	*d*	*p*	3 Mean	4 Mean	*d*	*p*
Low SE	79.31	71.88	-0.44	< 0.001	79.23	66.62	-0.61	< 0.001
High SE	79.71	74.93	-0.33	0.005	79.57	69.10	-0.59	< 0.001

SE = Self-Efficacy.

TABLE 5.7. Spearman’s rho Intercorrelations for Session 3 Experimental Anatomy Quiz Scores and Session 3 Resource Utilization Measures

	Atlas-Based Quiz		Specimen-Based Quiz	
	r	p	r	p
Total Resource Time	-0.19	0.083	-0.34	0.002
Total Resources	-0.01	0.957	-0.05	0.681
Preferred Resource Time	-0.26	0.021	-0.33	0.003s

N = 81.

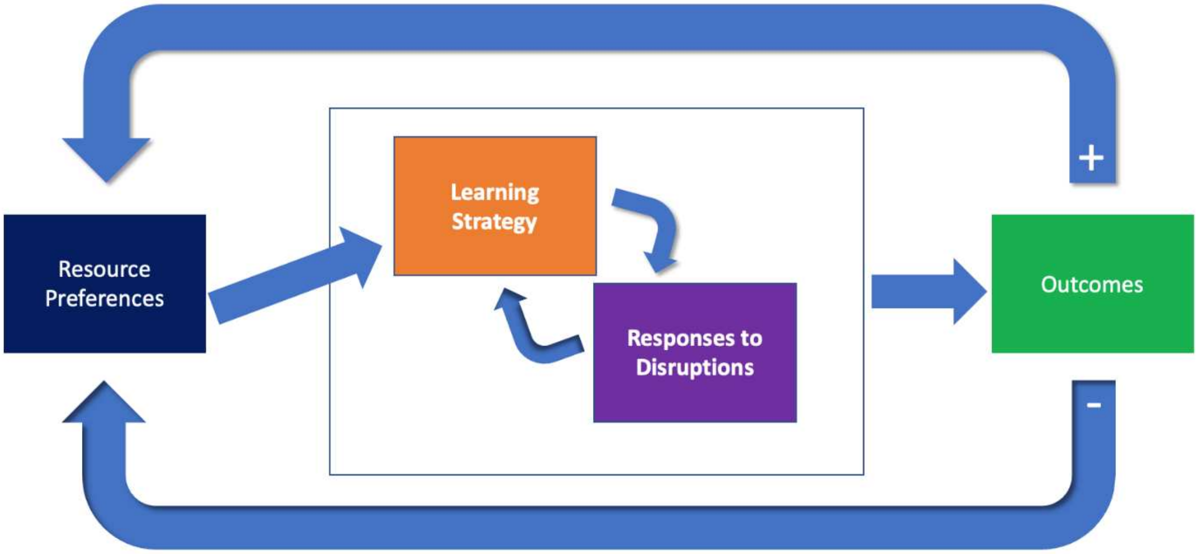


FIGURE 5.1. Initial Framework Following Inductive Thematic Analysis

Initial framework following inductive thematic analysis.

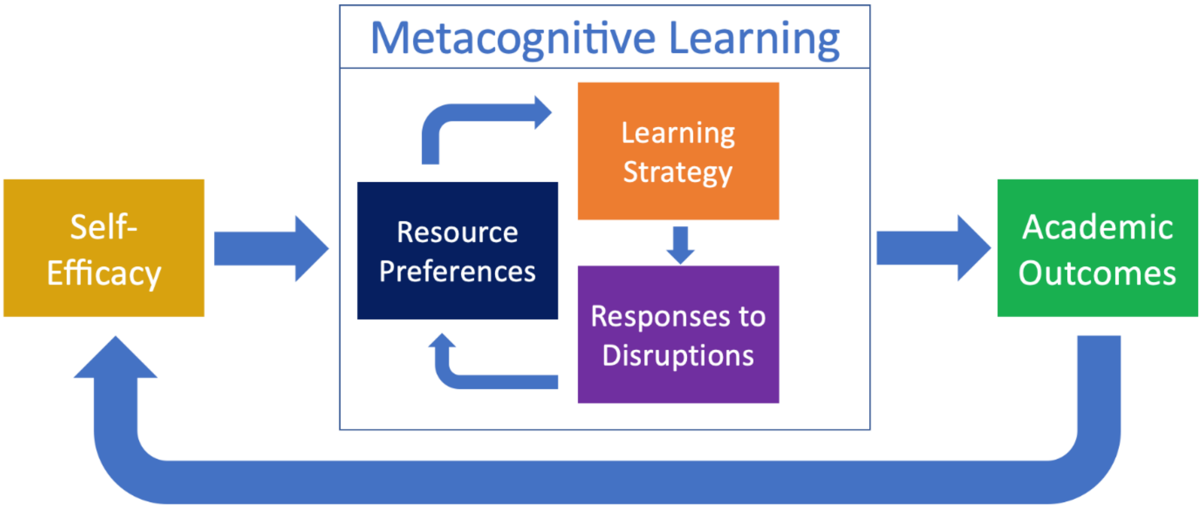


FIGURE 5.2. Revised Framework Following Deductive Thematic Analysis

FIGURE 5.3. Example Image of an Articulated Emu Pelvis in Left-Lateral position.

CHAPTER VI – CONCLUSIONS, RECOMMENDATIONS, AND FUTURE RESEARCH

6.1. DISSERTATION SUMMARY

My dissertation uses a variety of methodological approaches to identify relationships between gross anatomy learning resources and performance on summative assessments. In the first chapter, I examine how gross anatomy learning resources have been presented in existing literature. An opportunistic study was made available when a global pandemic necessitated a shift from laboratory-based instruction to online virtual laboratory-based instruction (Chapter II). Two categories of learning resources, content mastery, and learning ecology navigation, are described. To evaluate concerns of insuficient instructional duration, a retrospective study of prosection and dissection-based animal anatomy examination scores is included in this dissertation (Chapter III). A correlation between prior examination experience and dissection examination scores is described. Anatomy education literature suggests that gross anatomy examination performance variation is associated with visuospatial ability. To examine this, a randomized experimental study compared two measures of spatial ability, Mental Rotation Task and Landmark Position with Map, with anatomy assessment performance following single- and multi-resource instruction (Chapter IV). Mental Rotation Task scores correlated positively with multi-resource course performance but inconsistently correlated following single resources study. Landmark Position with Map was negatively correlated with virtual reality instructions. A mixed-method analysis (Chapter V) connected student interviews and surveys with course examination performance. By performing a hybrid inductive-deductive thematic analysis, I generated an experimental framework that describes how learning resource selection influences the relationship between self-efficacy and prosection-based examination performance. By leveraging this framework, different learning resources were identified for low- and high-self-efficacy learners. Low self-efficacy learners benefit from increased time attending in-person didactic lectures and teaching assistant mediated open laboratories. Conversely, high self-efficacy learners benefited from independent exploration of prosected cadavers and prepared specimens provided during open-laboratory periods. Interestingly, when instruction was limited to a single-resource, preference was not correlated with medium-term recall of novel anatomy..

6.2. RECOMMENDATIONS

As an anatomy educator, these chapters represent a data-driven approach to instructional design. While a single list ranking learning resources from most to least effective would simplify curriculum development and offer strong guidelines for academic intervention strategies, such a solution remains unavailable. That said, some recommendations for helping students learn can be derived from the chapters of this dissertation.

Students value resources independent of their total time spent using them. Anatomy resources are often expensive and time-consuming to maintain. While it may be tempting to reduce or remove the less frequently used resources, educators must consider that the resources may be masking a deficiency within the learning ecology. I encourage a closer inspection of how resources are used by those who use them before removing access.

A recurring point in anatomy education literature is that instructional time is at historic lows. While this dissertation does report that examination scores within an undergraduate prosecution-based anatomy course improve when students double their instructional time by enrolling concurrently in a dissection-based gross anatomy course, there are two points to consider. First, students who do not receive additional instruction are academically successful.

Second, differences in dissection-based examination scores were correlated with experience completing an examination with similar formatting and not correlated with previous exposure to the material.

Of the resources investigated within this dissertation, cadaveric access provided the greatest learning equality across both measures of spatial ability. Organizations concerned about learning equality should maintain a cadaveric laboratory. If that objective is not a priority or such resources are unavailable, spatial ability assessments can provide insight for selecting cadaveric alternatives. Learners with high Mental Rotation Task scores benefit from access to an anatomical atlas or rotatable 3D models on a computer screen. Learners with low Landmark Position on Map scores benefit from virtual reality instruction.

Finally, anatomy educators are encouraged to consider assessing student self-efficacy. Learners with low self-efficacy benefit from increased direction. Consider changes that encourage attending in-person lectures and providing lower student-to-staff laboratory experiences. If self-efficacy assessments are difficult to implement, develop a strategy for documenting resource utilization. Over time, high self-efficacy learners are more likely to optimize their study behavior by reducing the number of resources they access and the time spent using them. By contrast, low self-efficacy learners may change which resources they use but do not optimize their study.

6.3. FUTURE RESEARCH

The primary objective of this dissertation was to address learning resource selection in anatomy education. Unfortunately, comparisons of resources used in isolation do not represent how students report using resources within prosection- and dissection-based animal anatomy courses. It was reported in Chapter I that students value multiple learning resources during laboratory-based instruction. What other variables influence how attention is distributed among valued resources remains to be seen. Possible explanations include, but are not limited to, prior learning history, undocumented disabilities, or access problems. Interactions between such variables may produce additional demands that require a more dynamic selection of learning resources.