Title

Impact of gross anatomy laboratory on student written examination performance: A 3-year study of a large-enrollment undergraduate anatomy course

Authors

Anderson, Hana
Weil, Jennifer A
Tucker, Richard P
et al.

Publication Date

2023-08-21

DOI

10.1002/ase.2327

Peer reviewed

Received: 3 March 2023 | Revised: 25 June 2023 | Accepted: 28 July 2023

DOI: 10.1002/ase.2327

RESEARCH REPORT

Anatomical Sciences Education WILEY

Impact of gross anatomy laboratory on student written examination performance: A 3-year study of a large-enrollment undergraduate anatomy course

Hana Anderson[1,2] | Jennifer A. Weil[1,3] | Richard P. Tucker[1] | Douglas S. Gross[1]

[1]Department of Cell Biology and Human Anatomy, University of California, Davis, Davis, California, USA

[2]Department of Internal Medicine, University of California, Davis, Davis, California, USA

[3]School of Health Professions, Eastern Virginia Medical School, Norfolk, Virginia, USA

Correspondence
Dr. Hana Anderson, Department of Cell Biology and Human Anatomy, University of California, Davis, 1 Shields Avenue, Davis, CA 95616, USA.
Email: hnanderson@ucdavis.edu

Abstract

The efficacy of the various pedagogies that are used in human anatomy laboratories has been extensively debated. Nevertheless, an important question remains relatively unexamined—how the learning experience in the anatomy laboratory impacts students' mastery and application of anatomical knowledge beyond the laboratory setting. In this study, the effect of a prosection-based anatomy laboratory on overall comprehension and mastery of anatomical knowledge was evaluated in an upper division undergraduate anatomy curriculum that consists of a mandatory lecture course and an optional laboratory course. This flexible curricular structure permitted assessing the merit of laboratory learning on the written examination performance of the lecture course. In 2019 and 2022, the anatomy laboratory was taught in-person using prosections, while in 2021 due to the Covid-19 pandemic related regulations, it was taught remotely with live-streaming of prosections using document cameras. In both in-person and remote instructive formats, written examination scores of the lecture course were compared between two cohorts of students: Those enrolled in lecture only and those enrolled in both lecture and laboratory. Results showed that the cohort enrolled in both lecture and laboratory courses consistently outperformed the lecture-only cohort by one full letter grade. Furthermore, when the degrees of improvement on written examination scores were compared between the two instructive formats, in-person laboratory had a greater increase compared to remote laboratory. Altogether this study demonstrates that the prosection-based anatomy laboratory enhances students' mastery of anatomical knowledge beyond the laboratory setting by promoting comprehension of spatial relationships of anatomical structures.

KEYWORDS
anatomy laboratory, gross anatomy education, in-person instruction, remote instruction, student performance, undergraduate education

INTRODUCTION

Learning anatomy is hard—this is a common sentiment expressed by first-time learners of human anatomy. Challenges inherent to learning anatomy, such as unfamiliar anatomical terminologies, three-dimensional comprehension of anatomical structures, and the sheer volume of complex information, have led to the perception of anatomy as a tough, even daunting, subject (Smith & Mathias, 2010; Cheung et al., 2021). Human anatomy is a key foundational subject in the education of future health professionals. In the baccalaureate level nursing, allied health, and pre-health sciences programs, the challenging nature of anatomy courses can create stumbling blocks for the students to advance in their curriculum (Griff, 2016; Schutte, 2016; Royse et al., 2020). In addition, it has been reported that 95% of physician assistant programs in the United States require anatomy as a prerequisite (Snyder et al., 2022). Considering its significant weight in the education of future health professionals, designing human anatomy curricula that effectively support learning and performance of undergraduate students is an important mission for anatomy educators.

Literature suggests that instruction of anatomy across undergraduate and graduate level programs is most effectively implemented through hands-on learning experiences and many anatomy curricula use a human cadaver-based gross anatomy laboratory to provide such experiences (Estai & Bunt, 2016; Wilson et al., 2018). Originating as an instruction based on dissection (Ghosh, 2017), anatomy laboratories today embrace a wide variety of modalities to teach the structural intricacy of the human body (Estai & Bunt, 2016; Wilson et al., 2018). There have been active discussions regarding the effectiveness of various types of anatomy laboratory teaching modalities, for example, dissection versus prosection (Lackey-Cornelison et al., 2020; Pather, 2020; Thompson & Marshall, 2020; Huynh et al., 2021; McWatt et al., 2021; Koh et al., 2023), cadaver-based versus virtual/imaging technology-based (Griksaitis et al., 2012; Peterson & Mlynarczyk, 2016; Henssen et al., 2020; Narnaware & Neumeier, 2021; Zibis et al., 2021), and cadaver-based versus plastic models (Mitrousias et al., 2020; Zibis et al., 2021). Despite the varying views on the efficacy of the specific types of teaching modalities, it has been shown that what matters most to the novice anatomy learners is the learning experience in the laboratory itself, rather than a type of modality, yet few studies have investigated the efficacy of the laboratory experience, regardless of the modalities used (Codd & Choudhury, 2011; Lombardi et al., 2014; Anderton et al., 2016; Wilson et al., 2018). In addition to the laboratory teaching modalities described above, the format of delivering anatomy laboratory learning experience, that is, in-person versus remote, has also attracted active discussions in recent years, further stimulated by the instructional restrictions triggered by the pandemic (Longhurst et al., 2020; Harrell et al., 2021; Thom et al., 2021; Abualadas & Xu, 2023). Considering the trend of shortening anatomy laboratory hours that anatomy educators at medical schools are facing (Gillingwater, 2008; McBride & Drake, 2018), it is becoming increasingly critical to clearly demonstrate the value and efficacy of the anatomy laboratory learning experience on broader learning outcomes in the discipline of human anatomy.

The objective of this study was to evaluate the efficacy of a prosection-based undergraduate (baccalaureate) gross anatomy laboratory on the mastery of anatomical knowledge in the broad sense. For this, we used written examination scores as a measure of effectiveness of anatomy laboratory to support students' knowledge acquisition beyond the laboratory setting. To achieve the objective, we conducted the following analyses: (1) assessing the effect of prosection-based gross anatomy laboratory learning experience on written knowledge test performance, (2) comparing the effect of in-person and remote anatomy laboratories on written knowledge test performance to see whether the two delivery formats differ in their effectiveness, and (3) if the anatomy laboratory experience enhanced learner's anatomy comprehension, defining the types of questions that were most significantly associated with its benefit.

Examining the efficacy of the anatomy laboratory in a large-enrollment course as conducted in this study will aid in determining the importance and relevance of hands-on learning experience in anatomy education beyond the laboratory, and therefore assist the designing of an effective anatomy course in the time of declining anatomy curricular hours.

MATERIALS AND METHODS

At the University of California, Davis, the undergraduate anatomy curriculum consists of a large-enrollment lecture course and a concurrently run anatomy laboratory course. Although the majority of students take both concurrently, they have the option to enroll just in the lecture course and if they desire, can take the laboratory course separately afterward. This flexible organization, combined with our experience of delivering both courses remotely during the reduced in-person operations (as mandated during the height of the Covid-19 pandemic), presented a unique opportunity for us to conduct the analyses described in this study.

Human subjects research exemption was granted by the Internal Review Board (IRB) at the University of California, Davis (IRB ID: 1944066-1). Student identifiers such as names and identification numbers were removed from the examination scores and GPAs analyzed in this study to maintain anonymity.

Organization of the gross anatomy lecture course (CHA 101) and gross anatomy laboratory course (CHA 101L), and Covid-19-related modifications

CHA 101 is a 10-week long comprehensive regional gross anatomy course delivered through didactic lectures including clinical cases/correlates, and hands-on draw-together activities using a document camera. The course content includes an introduction to the peripheral and central nervous systems, upper limb, pelvis and perineum, vertebral column and back, thorax, abdomen, skull, cranial nerves, oral and nasal cavities, neck, pharynx and face, larynx, orbit, ear, and lower limb. During the 10-week period, four 50-min lectures are given Monday through Thursday weekly (total of 36 lecture hours). The course is targeted at upper level undergraduate students and includes introductory biology as a prerequisite. Two different instructors teach essentially the same content, one in Winter Quarter (WQ) and the other in Spring Quarter (SQ). The two instructors have taught gross anatomy together for almost 30 years and review each other's examination questions to maintain equivalent difficulty levels and types of questions between the two quarters. In 2019, audio recordings of the lectures were provided to the enrolled students, but neither the video recordings of lectures nor lecture slides were made available. In 2021, due to Covid-19-related state and local public health policies, all lectures were prerecorded and uploaded to the course website in advance for students to view on their own. In 2022, in-person lectures were reinstated with modifications to support the transition; namely, all in-person lectures were recorded and uploaded to the course website following the completion of lectures. Due to a high local Covid-19 infection rate, the first 4 weeks of CHA 101 were taught remotely in WQ 2022, and in-person instruction was subsequently resumed. In addition to the instructor, there are two to three Lecture Aides for CHA 101 each quarter. Each Lecture Aide holds a weekly 2-h review session on lecture content taught in the previous week, and as a team they offer pre-midterm and final exam review sessions. These reviews focus on lecture-specific material and are open to all students enrolled in CHA 101. Typical attendance at the weekly lecture reviews is 5%–10% of enrolled students (10–50 students), with attendance at the preexamination review sessions reaching 30%–40% of enrolled students (100–150 students). Students enrolled in CHA 101 also had access to learning support materials such as a comprehensive course syllabus providing a summary of learning materials prepared jointly by the two instructors and the Visible Body Human Anatomy Atlas application accessible from the course website.

CHA 101L is a 9-week comprehensive regional anatomy laboratory course using human prosections, skeletons and disarticulated bones, and anatomical models. The course is offered in parallel with CHA 101 in WQ and SQ, and the anatomical regions covered in the laboratory course follow the same sequence as the lecture course. Although students can enroll in the lecture course without concurrent enrollment in the laboratory course, students cannot enroll in CHA 101L without taking CHA 101 either concurrently or previously. The students enrolled in CHA 101L were subdivided into multiple sections (10 sections in WQ and eight sections in SQ) to limit the number of enrollments in each section to 44 students. Each CHA 101L section was led by a fully trained teaching assistant (TA) recruited from a pool of medical and graduate students, with guided hands-on instruction facilitated by laboratory aides (LAs). LAs are former CHA 101L students who successfully completed the course and were chosen through a competitive selection process. Students enrolled in CHA 101L participated in two 3-h laboratory sessions per week, which amounted to 17 laboratories and 51 h of directed instruction by TAs and LAs. During the regular laboratory sessions, students identified the anatomical structures on prosections and models that are specified in the laboratory instruction manual under the guidance of their TA and LAs. To supplement the students' hands-on prosection-based education, each TA held two office-hours per week, and 10–12 h of open laboratory time (Open Lab) were offered weekly as well. During Open Labs, students did independent or small-group self-directed study on the prosections with guidance provided by LAs.

In 2021, due to Covid-related restrictions, CHA 101L was converted entirely to a remote teaching format. Regularly scheduled weekly anatomy laboratory sessions were delivered remotely as synchronous online laboratories using Zoom (Zoom Video Communications, Inc., San Jose, CA). Special permission was granted by the UC Davis Body Donation Program to allow teaching from the prosections using document cameras with appropriate safeguards to maintain donor privacy. As such, to maintain donors' anonymity, head and neck anatomy was taught using images from anatomy atlases and publicly available cadaveric image databases (e.g., University of Michigan BlueLink). All other regions were taught using livestreaming of prosections from the laboratory. Open Labs were also converted to a synchronous online format, with LAs teaching from prosections using document cameras in the laboratory and students logged in to Zoom breakout rooms. Breakout rooms were divided according to anatomical regions, and students could navigate among breakout rooms according to their specific questions and needs. Students could ask questions either verbally by unmuting or using the chat feature. In 2022, due to a high local Covid infection rate, the first week of WQ CHA 101L was taught remotely, and in-person instruction was resumed thereafter. Learning materials for the laboratory included the laboratory instruction manual (mandatory), Visible Body Human Anatomy Atlas, and any major anatomy atlas (optional).

In this study, the data from the year 2020 were excluded because of the following irregularities in the course operations. CHA 101 and CHA 101L were mostly offered in-person during WQ 2020, though the final examinations needed to be offered remotely. Due to state and local public health requirements, CHA 101 and CHA 101L were offered remotely during SQ 2020. The laboratory course was taught using atlases and models over Zoom, and the lecture course was impacted by both the pandemic restrictions and civil unrest, which led to the extraordinary decision to make the final examination optional.

Student performance assessment

The lecture course (CHA 101) used two midterm and one final multiple choice question examinations to assess students' knowledge acquisition. The first midterm examination covered the content of the first 12 lectures, the second midterm covered the following 12 lectures, and the final examination was cumulative with additional emphasis on materials covered after the first 24 lectures. Midterm and final examinations contained 40 and 60 questions, respectively—each with five answer choices. On average, breakdown of questions based on cognitive levels are: 14.2% (±7.7% SD) for primary, 60.6% (±6.1% SD) for secondary, and 25.2% (±3.8% SD) for tertiary questions. In 2021, to maintain the integrity of examinations taken remotely, the examinations were built in Canvas and the exam software LockDown Browser® and Respondus Monitor® (Respondus, Inc.) was used to deliver the examinations, which utilizes fully automated proctoring for online examinations. Specifically, students were required to install and utilize the LockDown Browser to take all CHA 101 examinations, which prevented them from accessing any files on their computer and websites other than Canvas for purposes of completing the examination. The Respondus Monitor component of the software required students to submit a video of their testing environment to show they were in a private location with no unauthorized materials (e.g., notes, textbooks) present on or around their desk. To confirm student identity, students had to submit a photo of themselves in their testing environment along with a photo of their student identification card. The entire testing session was recorded using the student's webcam, and the Respondus Monitor software utilized predictive algorithms to flag suspicious behaviors potentially indicative of cheating. Instructors received copies of all student testing sessions, including notes from the predictive algorithm regarding suspicious behaviors noted during the examination period. In this test-taking environment, the time limitation equal to that of in-person examination was applied. To assess internal consistency reliability of all examinations, Kuder–Richardson formula (KR-20) was calculated. KR-20 scores of the CHA 101 examinations analyzed in this report ranged between 0.83 and 0.94 (a summary of KR-20 values for all CHA 101 examinations is presented as Table S1).

The laboratory course (CHA 101L) used one midterm practical examination at the half-way point and a cumulative final practical examination at the conclusion of the course to assess students' learning of laboratory materials. Both examinations contained 60 multiple choice questions with five answer choices, and students were given 70 min to complete the examination (1 min per question, with an additional 10 min to revisit any questions). In the midterm practical examination, questions identifying tagged structures accounted for half of the questions, and the rest were secondary or tertiary questions. In the final practical examination, questions identifying tagged structures accounted for two thirds, and the rest were secondary or tertiary questions. Specimens tagged on the 2019 and 2022 midterm and final practical examinations included human cadaver prosections (80%–90% of specimens on each examination) and a combination of bones, skeletons, and/or models (10%–20% of specimens on each examination). In 2021, practical examinations were built in Canvas and conducted remotely using the same exam software LockDown Browser® and Respondus Monitor® as was utilized for all CHA 101 examinations, with identical specifications as noted above. Because of the remote nature of these midterm and final practical examinations, photographs (including of prosections in the CHA 101L teaching collection [50%–75% of images on each examination] and publicly available photographs from University of Michigan BlueLink [10%–20% of images on each examination]) and anatomical drawings/images (5%–10% of images on each examination) were digitally tagged using colored arrows and dots, and these images were uploaded into Canvas and paired with their corresponding question. While the students did have access to publicly available cadaveric photographs, such as those provided through University of Michigan BlueLink, and to anatomical drawings/images, such as those provided in the Netter Atlas of Human Anatomy and Grant's Atlas of Anatomy, students did not have access to any of the photographs taken of the prosections in the CHA 101L teaching collection, which comprised the bulk of each practical examination, due to limitations placed by the UC Davis Body Donation Program on distribution of donor anatomical images. Students did receive real-time exposure to the prosections in the CHA 101L teaching collection through their regular synchronous Zoom-based laboratories, office hours, and Open Labs. In this remote test-taking environment, although the time limitation equal to that of the in-person practical examinations was applied (i.e., total examination time of 70 min), the 1 min per question time limit was not enforced, as implementing a question-specific time limit within Canvas prevented students from revisiting previous question. KR-20 scores of the CHA 101L practical examinations analyzed in this report ranged between 0.84 and 0.90 (a summary of KR-20 values for all CHA 101L practical examinations is presented as Table S1).

Analysis of examination scores of CHA 101 and CHA 101L courses

The written examination scores of CHA 101 and the practical examination scores of CHA 101L were collected from 2019, 2021, 2022 WQ and SQ grade records. Records of overall GPA of the students whose examination scores were analyzed in this study were obtained from the office of UC Davis registrar.

For statistical analysis, the Shapiro–Wilk test was applied first to assess the normality of data distributions. When data fit a normal distribution, the *t*-test and one-way ANOVA were used, and when they did not, the Mann–Whitney *U* test and Kruskal–Wallis *H* test (one-way ANOVA on ranks) were used to compare either two or multiple datasets. When statistically significant difference was detected by the Kruskal–Wallis *H* test, a pairwise post hoc comparison using the Nemenyi test was performed subsequently to identify pairs of datasets demonstrating a statistical difference.

To evaluate the effect of taking CHA 101L laboratory course on knowledge gain tested by written examinations in CHA 101 lecture course, students in the CHA 101 course were divided into two cohorts: Those who took both CHA 101 and CHA 101L (termed as “both”), and those who took CHA 101 only (termed as “lecture-only”). Scores of each examination (i.e., midterms and final) as well as the total combined scores of all examinations were compared between these two cohorts. The numbers of students in each cohort whose examination scores were analyzed are summarized in Table 1. Scores of students who completed all regularly scheduled examinations were included in the analysis. Similarly, students in the CHA 101L course were divided into two cohorts, those who took both CHA 101L and CHA 101 and those who took CHA 101L only

TABLE 1 Number of students whose examination scores were analyzed.

	CHA 101			CHA 101L		
Year and quarter	CHA101/101L[b]	CHA101 only	Total	CHA101L/101[b]	CHA101L only	Total
2019 (in-person)						
Winter	375	36	411	380	5	385
Spring	158	33	191	159	19	178
Combined	533	69	602	539	24	563
2021 (remote)						
Winter	354	71	425	350	14	364
Spring	241	57	298	236	29	265
Combined	595	128	723	586	43	629
2022 (in-person)						
Winter[a]	394	48	442	393	7	400
Spring	248	45	293	254	17	271
Combined	642	93	735	647	24	671

[a]The first 4 weeks of 2022 Winter Quarter were remotely delivered due to a high local Covid infection rate.

[b]The numbers of the two columns in each quarter are not identical because students who missed any one of the regularly scheduled examinations in either course were excluded from the analysis. Those numbers were different between the two courses.

TABLE 2 Comparisons of overall GPA between the two cohorts.

	Cohort-based CHA 101 students' GPA				Cohort-based CHA 101L students' GPA			
Year and quarter	Cohort	Mean ± SD	Median	*p* Value	Cohort	Mean ± SD	Median	*p* Value
2019[a]								
Winter	CHA101/101L	3.24 ± 0.40	3.28	0.045*	CHA101L/101	3.24 ± 0.40	3.27	0.047*
	CHA101 only	3.10 ± 0.40	3.02		CHA101L only	2.82 ± 0.37	2.88	
	Whole class	3.23 ± 0.40	3.26	–	Whole class	3.23 ± 0.40	3.27	–
Spring	CHA101/101L	3.15 ± 0.47	3.19	0.86	CHA101L/101	3.16 ± 0.44	3.20	0.28
	CHA101 only	3.15 ± 0.54	3.20		CHA101L only	3.06 ± 0.35	2.97	
	Whole class	3.15 ± 0.47	3.18	–	Whole class	3.15 ± 0.43	3.18	–
2021								
Winter	CHA101/101L	3.35 ± 0.38	3.42	0.50	CHA101L/101	3.35 ± 0.38	3.42	0.042*
	CHA101 only	3.36 ± 0.47	3.46		CHA101L only	3.10 ± 0.47	3.10	
	Whole class	3.35 ± 0.40	3.42	–	Whole class	3.35 ± 0.39	3.42	–
Spring	CHA101/101L	3.31 ± 0.39	3.33	0.41	CHA101L/101	3.31 ± 0.39	3.33	0.69
	CHA101 only	3.25 ± 0.41	3.31		CHA101L only	3.36 ± 0.41	3.35	
	Whole class	3.30 ± 0.39	3.33	–	Whole class	3.31 ± 0.39	3.34	–
2022								
Winter	CHA101/101L	3.43 ± 0.40	3.47	0.12	CHA101L/101	3.42 ± 0.39	3.47	5.1E−04*
	CHA101 only	3.34 ± 0.45	3.38		CHA101L only	2.89 ± 0.24	2.96	
	Whole class	3.42 ± 0.40	3.44	–	Whole class	3.42 ± 0.40	3.44	–
Spring	CHA101/101L	3.31 ± 0.44	3.38	0.16	CHA101L/101	3.31 ± 0.45	3.38	0.74
	CHA101 only	3.23 ± 0.37	3.24		CHA101L only	3.35 ± 0.45	3.31	
	Whole class	3.30 ± 0.43	3.35	–	Whole class	3.31 ± 0.44	3.38	–

Note: Overall GPAs of the two cohorts in CHA 101 and CHA 101L were compared using Mann–Whitney *U* test.

Abbreviation: SD, standard deviation.

[a]No statistical significance between the two cohorts was found when the data from Winter and Spring quarters were combined ($p = 0.16$).

*$p < 0.05$.

(termed as "laboratory-only"), and practical examination scores were compared between these two cohorts. None of the examination scores of CHA 101 as well as CHA 101L fit a normal distribution, therefore, the nonparametric statistical tests described above were used for analysis to calculate *p*-value and effect size *r*. Effect size *r* was also converted to Cohen's *d* using the formula described by Fritz et al. (2012). As a guideline, Cohen (1988) provides the following interpretation for different effect sizes; $d = 0.2–0.4$ (small), $d = 0.5–0.7$ (medium), and $d \geq 0.8$ (large).

Detailed results of inter-quarter and inter-academic year comparisons of CHA 101 and CHA 101L examination scores are presented as Tables S2 and S3, respectively. These analyses did not identify systematic statistically significant differences, that is, no consistent WQ versus SQ difference was detected. In addition, comparisons of practical examination performance among sections in the CHA 101L course were performed using 2022 WQ and SQ as representative (Table S4). Of these, scores of two sections (2022 SQ) were statistically lower than the rest, however, it was not significant enough to cause skewing in the analysis reported in this study.

Overall GPAs were compared between the two cohorts whose examination scores were analyzed for CHA 101 and CHA 101L to assess the comparability of the two student cohorts. Statistical analysis was conducted as described above and summarized as Table 2.

As we observed that the in-person laboratory was more effective than the remote laboratory in improving written examination scores of the CHA 101 lecture course, we examined whether this difference was intrinsic to the content delivery format. To do so, the following analyses were performed for both CHA 101 and CHA 101L. First, to see if there was a distinctive trend in the change in examination scores based on the delivery format, the Kruskal–Wallis *H* test and the Nemenyi test were applied to identify pairs of examination scores exhibiting a statistical difference within each examination category. Second, distributions of total combined scores of 2019 and 2022 (in-person) and 2021 (remote) were compared for each of CHA 101 and CHA 101L using the Mann–Whitney *U* test to assess if these two different formats in a broad sense resulted in a statistically significant difference in student examination performance.

Question item analysis

To characterize how taking the CHA 101L laboratory course benefited students in anatomical knowledge gain tested in the CHA 101 lecture course, we evaluated what types of questions (i.e., anatomical regions, levels of cognition, requirement of spatial relationships, and the presence of images) were more difficult for the lecture-only cohort compared to the cohort who concurrently took the laboratory course. These questions were the focus of our analysis since they would highlight the benefit of the laboratory learning experience compared to questions with comparable performance. Questions more difficult for the lecture-only cohort were defined as those on which they performed at or below 80% compared to the cohort who took both courses. To identify these questions, first we compared the correct answer rates of all final examination questions between the two cohorts to calculate relative correct answer rates ("lecture-only" cohort to "both" cohort). In this calculation, the smaller the ratio is, the poorer the lecture-only cohort performed on the question. A total of 107 questions met this criterion, and were further analyzed based on: (1) anatomical regions, (2) levels of cognition (i.e., lower-level [primary] or higher-order [secondary and tertiary]), referred to Bloom's taxonomy (Krathwohl, 2002) and the Blooming Anatomy Tool (BAT) (Thompson & O'Loughlin, 2015), (3) spatial comprehensions, and (4) the presence or absence of images. Within each classification, a ratio of each type of questions (e.g., upper limb questions, primary questions) in the 107 questions to that in all final questions was calculated as a relative ratio of representation.

Regarding the cognition levels, primary questions are "what is the name of" or "identify the structure," which correspond to the lower order "Knowledge" type questions (Thompson & O'Loughlin, 2015). Questions considered as secondary are those that require connecting the anatomical structure with its function or understanding of substructures of the anatomical structure, which correspond to the lower order "Comprehension" type questions. Tertiary questions are those requiring the integration of knowledge such as the identity of an anatomical structure with its function and blood/nerve supply in order to infer the consequence of an injury to the structure, corresponding to the higher order "Application" type of questions. Regarding the requirement of spatial comprehensions, questions were classified into two types: (1) comprehension of "spatial relationships" as an understanding of anatomical structures as they physically relate to one another and how they are oriented spatially within the body (e.g., the gastroduodenal artery is posterior to the first part of the duodenum), and (2) comprehension of "3-dimensional (3D) relationships" as an understanding of the movement or manipulation of anatomical structures and/or relationships in space (e.g., how a given muscle moves a limbs). Coding of questions was performed through a discussion between HA and JAW to achieve the consensus. Select examples of the secondary and tertiary questions, and questions requiring the comprehension of spatial and 3D relationships, are provided as Supplementary Material.

RESULTS

The anatomy laboratory course (CHA 101L) has a positive effect on student performance in the anatomy lecture course (CHA 101)

To evaluate the effect of concurrently taken CHA 101L on performance in CHA 101, written examination scores of CHA 101 were compared between the cohort who took both (CHA 101/CHA 101L) and the lecture-only cohort (CHA 101 only). In addition, overall GPAs of the two cohorts were compared to verify that their academic aptitude was comparable. As shown in Table 2, GPAs of the two cohorts in CHA 101 were statistically comparable with one

TABLE 3 Comparison of written examination performance between the two cohorts across three years.

	Midterm 1 (full score = 40)		Midterm 2 (full score = 40)		Final (full score = 60)		Total (full score = 140)	
	Mean ± SD	Median	Mean ± SD	Median	Mean ± SD	Median	Mean ± SD	Median
2019 (in-person)								
CHA101/101L	33.2 ± 4.64	34	34.2 ± 4.64	36	49.9 ± 8.01	52	117.3 ± 15.7	121
CHA101 only	30.3 ± 5.91	31	31.0 ± 6.22	32	42.7 ± 10.4	42	103.9 ± 20.3	106
p Value	5.60E−05		6.81E−06		1.01E−08		2.98E−08	
Effect size *r* (Cohen's *d*)	0.157 (0.319)		0.177 (0.360)		0.229 (0.470)		0.221 (0.453)	
2021 (remote)								
CHA101/101L	33.0 ± 5.62	34	34.6 ± 5.64	37	49.8 ± 9.40	53	117.4 ± 18.9	123
CHA101 only	30.2 ± 7.03	32	31.6 ± 7.17	33	43.8 ± 12.9	47	105.6 ± 25.2	112
p Value	1.52E−05		2.77E−06		2.16E−07		6.45E−08	
Effect size *r* (Cohen's *d*)	0.155 (0.314)		0.169 (0.343)		0.188 (0.383)		0.196 (0.401)	
2022 (in-person)								
CHA101/101L	32.7 ± 5.78	34	34.5 ± 4.97	36	50.2 ± 7.93	52	117.4 ± 17.2	121
CHA101 only	28.4 ± 6.88	29	29.9 ± 7.30	31	41.1 ± 11.1	44	99.3 ± 23.5	106
p Value	1.19E−09		4.26E−10		3.17E−17		2.46E−14	
Effect size *r* (Cohen's *d*)	0.220 (0.452)		0.226 (0.465)		0.297 (0.622)		0.278 (0.579)	

Note: Analysis was conducted using the Mann–Whitney *U* test to compare the examination scores between the cohort who took concurrently CHA 101 and CHA 101L (CHA101/101L) and the cohort who took only CHA 101. Effect size *r* was converted to Cohen's *d* according to the formula described by Fritz et al. (2012). A summary of score comparisons in which data of Winter and Spring quarters were processed separately is provided as Table S5.

Abbreviation: SD, standard deviation.

TABLE 4 CHA 101 examination score comparisons between the two subgroups within the CHA 101-only cohort.

CHA 101-only subgroup (student number)	Midterm 1 (full score = 40)		Midterm 2 (full score = 40)		Final (full score = 60)		Total (full score = 140)	
	Mean ± SD	Median	Mean ± SD	Median	Mean ± SD	Median	Mean ± SD	Median
No CHA 101L (224)	29.6 ± 6.8	30	30.6 ± 7.1	32	42.4 ± 12.0	44	102.6 ± 23.9	107
CHA 101 then CHA 101L (66)	29.8 ± 6.7	31	31.8 ± 6.7	33.5	43.6 ± 11.4	46	105.2 ± 23.0	110.5
p Value	0.413		0.109		0.253		0.215	

Note: Students in each subgroup were combined across the 3 years (2019, 2021, and 2022) and their examination scores were compared between the two subgroups using Mann–Whitney *U* test.

Abbreviation: SD, standard deviation.

exception (2019 WQ) showing a minor inter-cohort difference, which was unlikely to cause a significant difference in examination performance between the two cohorts. As summarized in Table 3, the cohort who concurrently took both courses consistently outperformed the lecture-only cohort on all examinations with a significant statistical difference. Based on the effect size, score improvement was most substantial on final examinations. When the point difference in score was converted to a grading scale, across all 3 years of analysis the cohort who took both outperformed their lecture-only counterparts by at least one full letter grade. In 2019 and 2021, the cohort who took both earned an overall average score equivalent to a B grade, while the lecture-only cohort earned an overall average score equivalent to a C grade. This difference was even more pronounced in 2022, as while the lecture-only cohort earned an overall average score equivalent to a C-, the cohort who took both earned an overall average score equivalent to a B. These results clearly demonstrate that the concurrent anatomy laboratory learning experience significantly boosted examination performance in the lecture course.

Next, we compared the effect of in-person and remote anatomy laboratories on CHA 101 examination performance. When effect sizes were compared among 2019, 2021, and 2022 within each examination category, the in-person years (2019 and 2022) were consistently associated with greater effect sizes compared to the remote year (2021) (Table 3). The difference in effect size was most significant on final examinations: in-person years (2019 and 2022), $r = 0.229$ (Cohen's $d = 0.470$) and $r = 0.297$ (Cohen's $d = 0.622$); and remote year (2021), $r = 0.188$ (Cohen's $d = 0.383$).

Some of the lecture-only cohort took CHA 101L after they completed CHA 101. There was a possibility that these students might have performed better than those who never took anatomy laboratory course afterward. To evaluate this possibility, we compared CHA 101 examination scores between these two subgroups (Table 4). To gain statistical power, data of each subgroup were combined across 3 years. Although there is a trend that the subgroup who subsequently took the laboratory course performed slightly better than the other subgroup, the difference was not statistically significant.

Practical examination performance in CHA 101L is not dependent on when the CHA 101 lecture course was taken

As mentioned above, some students elect to take CHA 101 first and then take CHA 101L in the subsequent quarter. Here, we performed similar analysis as described above by comparing practical examination scores of the cohort who took both CHA 101L and CHA 101 concurrently and the cohort who took the laboratory course only. As summarized in Table 5, no statistically significant difference was found between the scores of the two cohorts. Though GPAs of the laboratory-only cohort in three Winter Quarters were statistically lower compared to the cohort took both (Table 2), this difference was likely due to the small sample size of this cohort causing skewing in the data (Table 1). Therefore, in contrast to the positive impact of the concurrently taken anatomy laboratory course on written examination performance in the lecture course, the CHA 101 lecture course—taken concurrently or later—did not significantly affect students' practical examination performance.

Different instruction formats, in-person and remote, do not affect student examination performance within each course

As indicated by the larger effect size, in-person laboratory was more effective in improving written examination scores in the CHA 101 lecture course (Table 3). Here, we examined whether this effect is attributable to the in-person laboratory learning experience or to a broader effect of the delivery format, in-person or remote. To do so, we first compared examination scores of CHA 101 and CHA 101L within each examination category (i.e., midterm or final) as described in the Materials and Methods section. This analysis identified 11 combinations (nine in CHA 101 and two in CHA 101L, Tables S2 and S3, respectively) associated with a statistically significant difference ($p < 0.05$); however, these differences did not constitute a discernable pattern attributable to the two delivery formats and appeared stochastic, indicating that the two delivery formats did not have a systematic impact on student examination performance in either course. Second, when total combined examination scores of WQ and SQ in the years of 2019, 2021, and 2022 were compared between years and quarters, no statistical difference was found for either CHA 101 or CHA 101L ($p = 0.069$ and $p = 0.307$, Tables S2 and S3, respectively). Therefore, in-person and remote delivery formats were, when assessment was aligned with the instructional method, equally effective in supporting mastery of anatomical knowledge.

To further confirm this observation, we compared the distributions of total combined scores between the in-person years (2019 and 2022) and the remote year (2021). No statistically significant difference was found for either CHA 101 ($p = 0.077$) or CHA 101L ($p = 0.28$). Figures 1 and 2 summarize the score distributions of in-person years (2019 and 2022) and the remote year (2021) for CHA 101 and CHA 101L, respectively. Taken together, these

TABLE 5 Comparison of practical examination performance between the two cohorts across 3 years.

	Midterm (full score = 60)		Final (full score = 60)		Total (full score = 120)	
	Mean ± SD	Median	Mean ± SD	Median	Mean ± SD	Median
2019 (in-person)						
CHA101L & 101	49.1 ± 6.83	50	48.7 ± 7.48	50	97.8 ± 13.5	101
CHA101L only	50.0 ± 5.93	51	48.3 ± 5.58	48	98.4 ± 10.4	97.5
p Value	0.283		0.201		0.420	
2021 (remote)						
CHA101L and 101	49.0 ± 8.47	51	48.6 ± 8.51	51	97.6 ± 16.0	101
CHA101L only	48.6 ± 7.90	51	47.9 ± 7.38	49	96.5 ± 14.1	97
p Value	0.300		0.147		0.198	
2022 (in-person)						
CHA101L and 101	48.0 ± 7.80	50	49.8 ± 7.77	51	97.8 ± 14.7	101
CHA101L only	48.5 ± 8.77	51.5	47.0 ± 9.17	48.5	95.5 ± 16.5	99
p Value	0.264		0.068		0.262	

Note: Analysis was conducted using the Mann–Whitney *U* test to compare the practical examination scores between the cohort enrolled in both CHA 101L and CHA 101 and the cohort enrolled only in CHA 101L.

Abbreviation: SD, standard deviation.

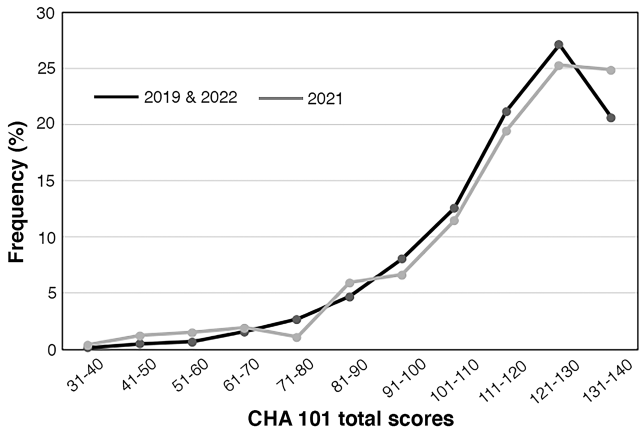

FIGURE 1 Comparison of the total written examination scores of CHA 101 between in-person and remote instructions. Distributions of the total combined examination scores of CHA 101 from 2019 to 2022 (in-person, black line) and those from 2021 (remote, gray line) are shown. The *x* axis indicates bins of scores (the full score = 140) and the *y* axis indicates % frequency of students whose scores corresponded to each bin. The pooled scores of 2019 and 2022 ($N = 1337$) were compared with those of 2021 ($N = 723$) using the Mann–Whitney *U* test. Median scores and mean scores ± standard deviations for 2019 and 2022 (in-person) were 120 and 115.4 ± 18.1, and for 2021 (remote) 121 and 115.3 ± 20.7. No statistically significant difference was found between the in-person (2019 and 2022) and remote (2021) cohorts ($p = 0.077$).

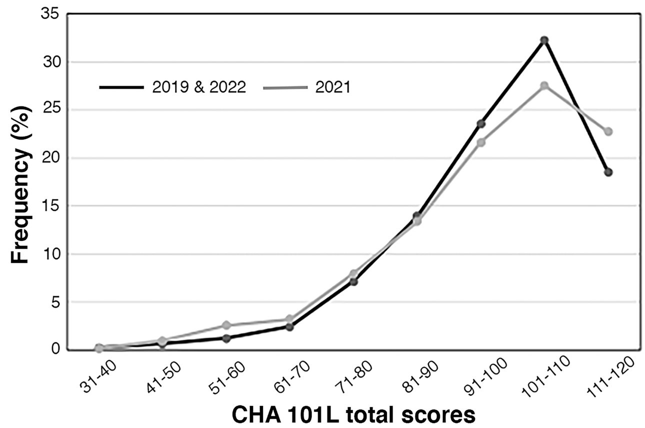

FIGURE 2 Comparison of the total practical examination scores of CHA 101L between in-person and remote laboratories. Distributions of the total combined practical examination scores of CHA 101L from 2019 and 2022 (in-person, black line) and those from 2021 (remote, gray line) are shown. The *x* axis indicates bins of scores (the full score = 120) and the *y* axis indicates % frequency of students whose scores corresponded to each bin. The pooled scores of 2019 and 2022 ($N = 1234$) were compared with those of 2021 ($N = 629$) using the Mann–Whitney *U* test. Median scores and mean scores ± standard deviations for 2019 and 2022 (in-person) were 101 and 97.8 ± 14.1, and for 2021 (remote), 101 and 97.5 ± 15.9. No statistically significant difference was found between the two cohorts ($p = 0.28$).

observations show that in-person and remote instruction were equally effective to support students learning anatomy; however, in-person laboratory learning experience provided additional support for students to perform better on their written examinations in the lecture course.

Prosection-based anatomy laboratory learning experience supports the development of spatial comprehension for anatomical structures

Using the criterion described in the Materials and Methods section, we identified 107 final examination questions on which the lecture-only cohort underperformed significantly compared to the cohort who took the lecture and laboratory courses simultaneously. The differential in performance on these 107 questions are summarized in Table S6. To explore how laboratory learning experience benefited the students, we analyzed the characteristics of these 107 questions by anatomical regions, comprehension of spatial and 3D relationships, levels of cognition, and questions with or without images.

As summarized in Table 6, among the anatomical regions examined, questions on the pelvis, abdomen, and thorax (31 in total) were overrepresented in the 107 questions by 70%, 20%, and 20% (i.e., 1.7, 1.2, and 1.2 in relative ratio of representation), respectively. Of these 31 questions, 16 questions required knowing the locations of neurovasculature and ducts, five questions required knowing attachments of muscles, and six questions required knowing relationships of anatomical regions within a structure. Interestingly, although head and neck questions were highly represented in these questions (32.7%), they were not overrepresented when compared to all final examination questions (i.e., 0.9 in relative ratio of representation). This may reflect the common observation that the head and neck anatomy is considered challenging by learners in general (Hall et al., 2018), and not by a specific group of learners.

Based on the high representation of questions requiring the knowledge on relative locations and spatial relationships of anatomical structures, we postulated that the anatomy laboratory learning experience most effectively aids students in developing comprehension of spatial and 3D relationships between them. We were also interested in examining if reasoning at higher cognitive levels was enhanced by anatomy laboratory learning. To explore these two points, we calculated relative representation of each type of questions (i.e., spatial and 3D comprehension; cognitive levels) in the 107 questions. In addition, we examined if the presence of images affected the lecture-only cohort's performance. Table 7 summarizes these analyses. In accordance with our postulation, questions requiring spatial and 3D comprehension were overrepresented, but not those without that requirement (i.e., relative ratios of representation for spatial relationship and 3D vs. no spatial relationship: 1.25 and 1.16 vs. 0.6). Consistent with this observation, when relative correct answer rates (lecture-only/both) of these three types of questions were compared, those requiring spatial relationship (SR) and three-dimensional (3D) relationship showed significantly lower relative correct answer rates compared to those requiring no spatial relationship (NO)—corresponding relative correct answer

TABLE 6 Analysis of questions lecture-only cohort underperformed by anatomical regions.

Regions/systems	Number of questions lecture-only cohort performed at or below 80%	% Representation	Number of questions in all final exams[a]	% Representation	Relative ratio of representation
Head and neck	35	32.7	124	36.6	0.9
Lower limb	22	20.6	81	23.9	0.9
Upper limb	17	15.9	52	15.3	1.0
Abdomen	12	11.2	31	9.1	1.2
Pelvis	10	9.3	19	5.6	1.7
Thorax	9	8.4	24	7.1	1.2
Nervous system	2	1.9	8	2.4	0.8
Total	107	100	339	100	

Note: Lecture-only cohort performed at or below 80% compared to the lecture/laboratory (both) cohort on 107 final examination questions. These questions were categorized based on their corresponding anatomical regions and systems, and their relative representations compared to all final examination questions were calculated.

[a]Because questions on the Back were consistently well performed by the lecture-only cohort and were not selected into the 107 questions, they were excluded from this analysis. Questions dropped or double-keyed were also excluded.

TABLE 7 Analysis of questions lecture-only cohort underperformed by spatial comprehension requirement and cognitive levels.

Types of questions	% Representation in the questions lecture-only cohort underperformed	% Representation in all final examination questions	Relative ratio of representation
No spatial relationships required	21.5	35.7	0.60
Spatial relationships required	52.3	41.7	1.25
3D comprehension required	26.2	22.6	1.16
Primary	15.9	14.5	1.10
Secondary	62.6	59.7	1.05
Tertiary	21.3	25.8	0.83
Image based	19.6	20.9	0.94

Note: Analysis was performed on the 107 final examination questions lecture-only cohort performed at or below 80% compared to lecture/laboratory (both) cohort. These questions were categorized based on: (1) required comprehension of spatial relationships between anatomical structures, and 3D movement and interactions of anatomical structures in space, (2) cognitive levels, and (3) use of an image. A relative representation of each question category was calculated by comparing to their representation in all final examination questions.

rates: SR: 0.84; 3D: 0.85; NO: 0.88 (p values, 2×10^{-4} for 0.84 vs. 0.88; 8×10^{-3} for 0.85 vs. 0.88).

Regarding the cognitive levels, while primary and secondary questions were slightly overrepresented (i.e., relative ratios of representation, 1.10 and 1.05, respectively), tertiary level questions were not (i.e., relative ratio of representation, 0.83) (Table 7), This observation may reflect the general nature of tertiary questions, which is more heavily reliant on logical reasoning than comprehension of spatial relationships of anatomical structures. Among the primary questions, one half were on the pelvic, abdominal, and thoracic anatomies, reiterating the difficulty of these anatomical region for the lecture-only cohort. Similarly, 78% of the secondary questions were classified as requiring the comprehension of spatial and 3D relationships. Therefore, not the levels of cognition, but the difficult anatomical region and the required comprehension of spatial and 3D relationships to get correct answers appear to have resulted in a slight overrepresentation of the primary and secondary questions.

Image-based questions did not present significant challenges to the lecture-only cohort (i.e., relative ratio of representation, 0.94). Also, no significant difference in the relative correct answer rate (lecture-only/both) was found between the questions with and without images ($p=0.42$, 0.84 ± 0.15 vs. 0.85 ± 0.12, mean $\pm$ SD, respectively). Since the images used in examinations showed two-dimensional (2D) relationships of anatomical structures that were sufficient to answer the questions, laboratory learning experience would not have provided an additional advantage. This result is consistent with the supposed benefit of prosection-based anatomy laboratory, which is to solidify understanding of spatial and 3D relationships of anatomical structures that are difficult to attain from 2D atlas images used in didactic anatomy instruction.

DISCUSSION

This study examined the effect of prosection-based gross anatomy laboratory on acquisition of anatomical knowledge in the broad sense. The curricular structure of undergraduate human gross anatomy at UC Davis allowed us to conduct an in-depth analysis of the effect of anatomy laboratory on students' knowledge acquisition beyond the laboratory setting. We observed that the concurrently taken prosection-based anatomy laboratory significantly improved students' written examination performance in the anatomy lecture course, likely because the anatomy laboratory learning experience promoted the development of spatial comprehension between anatomical structures. Moreover, in-person anatomy laboratory was superior to remote laboratory in terms of the betterment of written examination performance.

The efficacy of a cadaver-based human gross anatomy laboratory on learning human anatomy has been extensively debated for the anatomy curricula of health professions schools (Korf et al., 2008; Pizzimenti et al., 2016; Ghosh, 2017; Kinirons et al., 2019; Lackey-Cornelison et al., 2020; Mitrousias et al., 2020; Thompson & Marshall, 2020; Koh et al., 2023), and to a lesser extent for the undergraduate level anatomy curricula (Gonsalvez et al., 2015; McWatt et al., 2021). These reports focused mainly on a direct impact measured by laboratory practical examination performance, and a broader effect on anatomical knowledge acquisition measured by written examination performance has been less frequently reported (Gonsalvez et al., 2015; Pizzimenti et al., 2016; Thompson & Marshall, 2020; McWatt et al., 2021; Tucker & Anderson, 2021).

We observed that the cohort of students who concurrently took both human anatomy lecture and laboratory courses outperformed the lecture-only cohort by at least one full letter grade on the anatomy lecture written examinations (Table 3). Notably, the in-person anatomy laboratory was superior to the remotely delivered laboratory based on an increase in the examination scores. Both the in-person and remote laboratories were taught using prosections. In the case of in-person laboratory course, however, students could handle the prosections directly, whereas in the synchronously delivered remote laboratory course, students observed passively the pertinent anatomical structures in prosections presented by their TAs or LAs. The fact that even those students who learned through passive observation outperformed the lecture-only cohort is a testament of the value of cadaver-based laboratory in learning anatomy. The visual learning experience gained from examination of an actual human body, albeit indirectly, significantly improved the comprehension of spatial and three-dimensional relationships of anatomical structures. That said, as stated above, we observed that the extent of written examination score improvement was more significant among the students who learned from prosections directly, most notably on their final examinations.

A positive correlation between hands-on cadaver-based learning and the performance in an undergraduate anatomy laboratory course has been reported from other institutions. For example, Dennis and Creamer (2022) showed that practical examination performance was positively impacted the more the students engaged with direct handling of prosections. Another report also described a positive correlation between the duration of time spent on hands-on learning with dissection or prosection/models and students' performance on practical and written (theoretical) examinations (Gonsalvez et al., 2015). Taken together, for novice anatomy learners, direct hands-on learning experience is highly valuable in promoting comprehension of spatial relationships of anatomical structures.

When we compared the performance on laboratory practical examinations between in-person and remote laboratories, the students' scores with the two delivery formats were equivalent (Figure 2). This result likely suggests that a well-implemented remotely delivered anatomy laboratory course can provide an effective learning experience for students to become proficient at identifying and interpreting relationships of anatomical structures. For a large enrollment undergraduate anatomy course, the equal efficacy of online anatomy laboratory to in-person laboratory has also been reported (Attardi & Rogers, 2015; Attardi et al., 2018). In their reports, they described a carefully designed online anatomy course consisting of lectures and interactive laboratory instructions using 3D computer models. The cohort utilizing the remote computer-assisted instruction performed equally to the in-person cohort. It should be noted that alignment of the format in instruction and assessment (i.e., how the course content was delivered and how the students' knowledge acquisition was tested) is critical to accurately evaluate the efficacy of learning activities. In the study by Attardi et al. (2018), the in-person laboratory cohort studied from prosections while the online laboratory cohort studied from 3D computer models, however, both cohorts took the same practical examinations based on images of cadavers and 3D models. Because of this disparity, on the first practical examination the in-person cohort performed poorly compared to the online cohort, although their performance caught up on the subsequent examinations and eventually on final scores.

In our remote anatomy laboratory course, practical examinations were based on photographic images of the prosections and models from which students learned during their online laboratory sessions, aligning the format of learning activities and assessments. The observation that practical examination performance of the remote cohort was equivalent to that of the in-person cohort confirms that the alignment was successful. In March of 2020, upon initiation of the Covid-19 instructional policies, we had to make an emergency transition of both CHA 101 and CHA 101L into an online format with little preparation time. This placed us in the situation resulting in numerous ad hoc and continual adjustments, an experience shared by numerous anatomy educators (e.g., Brassett et al., 2020; Cuschieri & Calleja Agius, 2020; Evans et al., 2020; Pather et al., 2020; Harmon et al., 2021; Iwanaga et al., 2021; McWatt, 2021; Yun et al., 2022). The following year (2021), online implementation of both CHA 101 and CHA 101L was more smoothly executed based on our experience gained from the previous year; both the laboratory sessions and practical examinations were delivered in an orderly manner with students' performance equivalent to the in-person years (Figures 1 and 2). It would be noteworthy to further investigate how well the cohort of remote laboratory learners would perform if they were tested on real prosections in the laboratory setting. It has been shown that learners who studied with multimedia anatomy simulation were significantly disadvantaged on cadaver-based practical examination (Saltarelli et al., 2014). More recently, Wainman et al. (2020) illustrated that the experience of stereoscopic learning is the crucial element for comprehending anatomical structures in a cadaveric specimen. Applicability of anatomical knowledge learned from non-cadaveric materials to human patients in the clinical setting is, therefore, an important subject warranting careful discussion given that anatomy education is shifting to incorporating more computer-based learning modalities (Aziz et al., 2002).

Analysis of the questions the lecture-only cohort found more difficult provided us an insight into the benefits of prosection-based anatomy laboratory for novice learners. Since these questions were selected based on the underperformance by the lecture-only cohort, they highlighted the disadvantage experienced by the students who did not learn anatomy on prosections and models. Regarding anatomical regions, the pelvis, abdomen, and thorax were overrepresented in the more difficult questions to the lecture-only cohort (Table 6). That the pelvic anatomy was found more difficult by the lecture-only student cohort supports previous observations describing this region challenging to medical students (Kramer & Soley, 2002; Hall et al., 2018). In comparison, the thoracic and abdominal anatomies have not previously been reported as problematic regions for learners. This observation was informative for us to extrapolate the benefits of studying anatomy with prosections. These questions required the in-depth knowledge of spatial and 3D relationships involving neurovasculature, ducts, and muscles. Not exposed to the spatial relationships of these anatomical structures in the laboratory environment, therefore, likely put the lecture-only cohort at a disadvantage. These suppositions agree with the finding that the lecture-only students significantly underperformed on questions requiring the comprehension of spatial and 3D relationships. In addition, we assessed whether the presence of 2D images in questions would alter the performance of the lecture-only cohort and found no significant effect. This observation is consistent with a report describing a neutral effect of image inclusion in examinations of undergraduate histology (Holland et al., 2015) and anatomy (Notebaert, 2017; Bahlmann, 2018). In a more detailed analysis of image inclusion into questions, Vorstenbosch et al. (2013) compared the use of organ images accompanied with lettered arrows and a text list for answer options and demonstrated that the effect of images on item difficulty and discrimination was context dependent. Sagoo et al. (2021) examined the effect of anatomical and radiological images on an online examination and showed that medical students performed better on questions with the images. In our written examinations, the images used were mostly selected from the lecture slides and the student guide. Therefore, to answer these questions correctly, 3D comprehension acquired from laboratory learning would not have been necessary. Accordingly, the lecture-only cohort performed similarly on questions with or without images.

Limitations of the study

The analysis of this study was performed by comparing the two cohorts of students to evaluate the effect of the anatomy laboratory learning experience on written examination performance. There are possible behavioral variations among the enrolled students that could have influenced examination performance. First, an argument can be made that the cohort of students who took both the lecture and laboratory courses simply studied longer hours, and therefore obtained better grades. One observation that could counter this reasoning is the reciprocal situation in the laboratory course. In this case, the students who took both the laboratory and lecture courses and those who solely took the laboratory course performed equally on their practical examinations. Therefore, the time may not be the sole reason that the students who took both the lecture and laboratory courses did better on their written examination; rather, comprehension of spatial and 3D relationships of anatomical structures gained by the laboratory learning experience gave them an advantage on answering questions requiring these comprehensions. Also, this cohort likely benefitted from reinforcement and spaced repetition of the content learned in lectures. A second possibility is that the lecture-only cohort might have been less motivated to study compared to the cohort who took both. We showed that though the students who subsequently took laboratory did slightly better compared to those who never took laboratory, the difference in their performance was not statistically significant. A third possibility is that students who took in-person laboratory and lecture courses might have studied in a group more frequently than those who studied remotely, which could have created a more supportive environment for in-person cohort's learning. Finally, the nature of anatomical knowledge evaluated by practical examinations in the in-person and remote laboratories may likely not be identical, especially in terms of the ability to identify anatomical structures on real cadaveric specimens. Even with this caveat, the remote anatomy laboratory was nevertheless effective in improving written examination scores of the learners.

CONCLUSIONS

This study described a significantly positive effect of prosection-based human gross anatomy laboratory learning experience on written examination performance in a human gross anatomy lecture course. The cohort of students who took both courses concurrently outperformed the lecture-only cohort by at least one full letter grade. Although both in-person and remote laboratories were effective, the in-person laboratory was superior in terms of the degree of score improvement. We observed that enhancement of the spatial and 3D comprehension of the anatomical structures was likely the benefit for the students who took laboratory. The anatomy laboratory learning experience has, therefore, an irreplaceable position in gross anatomy education.

ACKNOWLEDGMENTS

The authors thank the teaching assistants and laboratory aides for their hard work in successfully delivering the courses, especially during the difficult time presented by the pandemic. We are also grateful to the staff of the UC Davis Body Donation Program for their support for the operation of anatomy laboratories and express gratitude to the individuals who donated the bodies and their families for the advancement of anatomy education and education research. We also thank

Dr. Bridget O'Brien for her critical review of the manuscript. The authors have no conflicts of interest to report. An abstract summarizing this study was accepted as an oral presentation for the Educational Research Platform at American Association for Anatomy's Anatomy Connected occurring March 25–27, 2023 in Washington, DC.

ORCID

Hana Anderson https://orcid.org/0000-0001-9159-1472
Jennifer A. Weil https://orcid.org/0000-0002-5718-5064
Richard P. Tucker https://orcid.org/0000-0001-8552-5401

REFERENCES

Abualadas HM, Xu L. Achievement of learning outcomes in non-traditional (online) versus traditional (face-to-face) anatomy teaching in medical schools: a mixed method systematic review. Clin Anat. 2023;36:50–76.

Anderton RS, Chiu LS, Aulfrey S. Student perception to teaching undergraduate anatomy in health sciences. Int J High Educ. 2016;5:201–16.

Attardi SM, Barbeau ML, Rogers KA. Improving online interactions: lessons from an online anatomy course with a laboratory for undergraduate students. Anat Sci Educ. 2018;11:592–604.

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

Aziz MA, Mckenzie JC, Wilson JS, Cowie RJ, Ayeni SA, Dunn BK. The human cadaver in the age of biomedical informatics. Anat Rec. 2002;269:20–32.

Bahlmann O. Illustrated versus non-illustrated anatomical test items in anatomy course tests and German Medical Licensing examinations (M1). GMS J Med Educ. 2018;35:Doc 25.

Brassett C, Cosker T, Davies DC, Dockery P, Gillingwater TH, Lee TC, et al. COVID-19 and anatomy: stimulus and initial response. J Anat. 2020;237:393–403.

Cheung CC, Bridges SM, Tipoe GL. Why is anatomy difficult to learn? The implications for undergraduate medical curricula. Anat Sci Educ. 2021;14:752–63.

Codd AM, Choudhury B. Virtual reality anatomy: is it comparable with traditional methods in the teaching of human forearm musculoskeletal anatomy? Anat Sci Educ. 2011;4:119–25.

Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers; 1988.

Cuschieri S, Calleja Agius J. Spotlight on the shift to remote anatomical teaching during Covid-19 pandemic: perspectives and experiences from the University of Malta. Anat Sci Educ. 2020;13:671–9.

Dennis JF, Creamer BA. To touch or not to touch: evaluating student laboratory outcomes of hands-on versus visual examination of prosected cadavers. Educ Sci. 2022;12:519.

Estai M, Bunt S. Best teaching practices in anatomy education: a critical review. Ann Anat. 2016;208:151–7.

Evans DJR, Bay BH, Wilson TD, Smith CF, Lachman N, Pawlina W. Going virtual to support anatomy education: a STOPGAP in the midst of the Covid-19 pandemic. Anat Sci Educ. 2020;13:279–83.

Fritz CO, Morris PE, Richler JJ. Effect size estimates: current use, calculations, and interpretation. J Exp Psychol Gen. 2012;141:2–18.

Ghosh SK. Cadaveric dissection as an educational tool for anatomical sciences in the 21st century. Anat Sci Educ. 2017;10:286–99.

Gillingwater TH. The importance of exposure to human material in anatomical education: a philosophical perspective. Anat Sci Educ. 2008;1:264–6.

Gonsalvez DG, Ovens M, Ivanusic J. Does attendance at anatomy practical classes correlate with assessment outcome? A retrospective study of a large cohort of undergraduate anatomy students. BMC Med Educ. 2015;15:231.

Griff ER. Changing undergraduate human anatomy and physiology laboratories: perspectives from a large-enrollment course. Adv Physiol Edu. 2016;40:388–92.

Griksaitis MJ, Sawdon MA, Finn GM. Ultrasound and cadaveric prosections as methods for teaching cardiac anatomy: a comparative study. Anat Sci Educ. 2012;5:20–6.

Hall S, Stephens J, Parton W, Myers M, Harrison C, Elmansouri A, et al. Identifying medical student perceptions on the difficulty of learning different topics of the undergraduate anatomy curriculum. Med Sci Educ. 2018;28:469–72.

Harmon DJ, Attardi SM, Barremkala M, Bentley DC, Brown KM, Dennis JF, et al. An analysis of anatomy education before and during Covid-19: May–August 2020. Anat Sci Educ. 2021;14:132–47.

Harrell KM, McGinn MJ, Edwards CD, Warren Foster K, Meredith MA. Crashing from cadaver to computer: Covid-driven crisis-mode pedagogy spawns active online substitute for teaching gross anatomy. Anat Sci Educ. 2021;14:536–51.

Henssen DJHA, van den Heuvel L, De Jong G, Vorstenbosch MATM, van Cappellen van Walsum A-M, Van den Hurk MM, et al. Neuroanatomy learning: augmented reality vs. cross-sections. Anat Sci Educ. 2020;13:353–65.

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

Huynh N, Burgess A, Wing L, Mellis C. Anatomy by whole body dissection as an elective: student outcomes. J Surg Educ. 2021;78:492–501.

Iwanaga J, Loukas M, Dumont AS, Tubbs RS. A review of anatomy education during and after the COVID-19 pandemic: revisiting traditional and modern methods to achieve future innovation. Clin Anat. 2021;34:108–14.

Kinirons SA, Reddin VM, Maguffin J. Effects of alternating dissection with peer teaching and faculty prosected cadaver demonstrations in a physical therapy and occupational therapy gross anatomy course. Anat Sci Educ. 2019;12:468–77.

Koh ZJ, Yeow M, Srinivasan DK, Ng YK, Ponnamperuma GG, Chong CS. A randomized trial comparing cadaveric dissection and examination of prosections as applied surgical anatomy teaching pedagogies. Anat Sci Educ. 2023;16:57–70.

Korf H-W, Wicht H, Snipes RL, Timmermans J-P, Paulsen F, Rune G, et al. The dissection course—necessary and indispensable for teaching anatomy to medical students. Ann Anat. 2008;190:16–22.

Kramer B, Soley JT. Medical students perception of problem topics in anatomy. East Afr Med J. 2002;79:408–14.

Krathwohl DR. A revision of Bloom's taxonomy: an overview. Theory Pract. 2002;41:212–8.

Lackey-Cornelison WL, Bauler LD, Smith J. A comparison of the effectiveness of dissection and prosection on short-term anatomic knowledge retention in a reciprocal peer-teaching program. Adv Physiol Educ. 2020;44:239–46.

Lombardi SA, Hicks RE, Thompson KV, Marbach-Ad G. Are all hands-on activities equally effective? Effect of using plastic models, organ dissections, and virtual dissections on student learning and perceptions. Adv Physiol Educ. 2014;38:80–6.

Longhurst GJ, Stone DM, Dulohery K, Scully D, Campbell T, Smith CF. Strength, weakness, opportunity, threat (SWOT) analysis of the adaptations to anatomical education in the United Kingdom and Republic of Ireland in response to the Covid-19 pandemic. Anat Sci Educ. 2020;13:301–11.

McBride JM, Drake RL. National survey on anatomical sciences in medical education. Anat Sci Educ. 2018;11:7–14.

McWatt SC. Responding to Covid-19: a thematic analysis of students' perspectives on modified learning activities during an emergency transition to remote human anatomy education. Anat Sci Educ. 2021;14:721–38.

McWatt SC, Newton GS, Umphrey GJ, Jadeski LC. Dissection versus prosection: a comparative assessment of the course experiences, approaches to learning, and academic performance of non-medical undergraduate students in human anatomy. Anat Sci Educ. 2021;14:184–200.

Mitrousias V, Karachalios TS, Varitimidis SE, Natsis K, Arvanitis DL, Zibis AH. Anatomy learning from prosected cadaveric specimens versus plastic models: a comparative study of upper limb anatomy. Anat Sci Educ. 2020;13:436–44.

Narnaware Y(R), Neumeier M. Use of a virtual human cadaver to improve knowledge of human anatomy in nursing students: research article. Teach Learn Nurs. 2021;16:309–14.

Notebaert AJ. The effect of images on item statistics in multiple choice anatomy examinations. Anat Sci Educ. 2017;10:68–78.

Pather N. Prosection and dissection laboratory sessions: design and implement to enhance learning. In: Chan LK, Pawlina W, editors. Teaching anatomy: a practical guide. 2nd ed. Cham: Springer International Publishing; 2020. p. 259–69.

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;13:284–300.

Peterson DC, Mlynarczyk GSA. Analysis of traditional versus three-dimensional augmented curriculum on anatomical learning outcome measures. Anat Sci Educ. 2016;9:529–36.

Pizzimenti MA, Pantazis N, Sandra A, Hoffmann DS, Lenoch S, Ferguson KJ. Dissection and dissection-associated required experiences improve student performance in gross anatomy: differences among quartiles. Anat Sci Educ. 2016;9:238–46.

Royse EA, Sutton E, Peffer ME, Holt EA. The anatomy of persistence: remediation and science identity perceptions in undergraduate anatomy and physiology. Int J High Educ. 2020;9:283–99.

Sagoo MG, Vorstenbosch MATM, Bazira PJ, Ellis H, Kambouri M, Owen C. Online assessment of applied anatomy knowledge: the effect of images on medical students' performance. Anat Sci Educ. 2021;14:342–51.

Saltarelli AJ, Roseth CJ, Saltarelli WA. Human cadavers vs. multimedia simulation: a study of student learning in anatomy. Anat Sci Educ. 2014;7:331–9.

Schutte AF. Who is repeating anatomy? Trends in an undergraduate anatomy course. Anat Sci Educ. 2016;9:171–8.

Smith CF, Mathias HS. Medical students' approaches to learning anatomy: students' experiences and relations to the learning environment. Clin Anat. 2010;23:106–14.

Snyder J, Brown D, Quincy B. Course prerequisites: here, there, and everywhere. Time to standardize? J Physician Assist Educ. 2022;33:164–70.

Thom ML, Kimble BA, Qua K, Wish-Baratz S. Is remote near-peer anatomy teaching an effective teaching strategy? Lessons learned from the transition to online learning during the Covid-19 pandemic. Anat Sci Educ. 2021;14:552–61.

Thompson AR, Marshall AM. Participation in dissection affects student performance on gross anatomy practical and written examinations: results of a four-year comparative study. Anat Sci Educ. 2020;13:30–6.

Thompson AR, O'Loughlin VD. The Blooming Anatomy Tool (BAT): a discipline-specific rubric for utilizing Bloom's taxonomy in the design and evaluation of assessments in the anatomical sciences. Anat Sci Educ. 2015;8:493–501.

Tucker RP, Anderson H. Dissection experience and performance on a human gross anatomy written examination: lessons learned during the Covid-19 pandemic. Anat Sci Educ. 2021;14:169–70.

Vorstenbosch MATM, Klaassen TPFM, Kooloos JGM, Bolhuis SM, Laan RFJM. Do images influence assessment in anatomy? Exploring the effect of images on item difficulty and item discrimination. Anat Sci Educ. 2013;6:29–41.

Wainman B, Pukas G, Wolak L, Mohanraj S, Lamb J, Norman GR. The critical role of stereopsis in virtual and mixed reality learning environments. Anat Sci Educ. 2020;13:401–12.

Wilson AB, Miller CH, Klein BA, Taylor MA, Goodwin M, Boyle EK, et al. A meta-analysis of anatomy laboratory pedagogies. Clin Anat. 2018;31:122–33.

Yun YH, Jo DH, Jeon SK, Kwon HY, Jeon YM, Shin DH, et al. The impact of the modified schedules of anatomy education on students' performance and satisfaction: responding to COVID-19 pandemic in South Korea. PLoS One. 2022;17:e0266426.

Zibis A, Mitrousias V, Varitimidis S, Raoulis V, Fyllos A, Arvanitis D. Musculoskeletal anatomy: evaluation and comparison of common teaching and learning modalities. Sci Rep. 2021;11:1517.

AUTHOR BIOGRAPHIES

Hana Anderson, Ph.D., is an associate professor in the Department of Cell Biology and Human Anatomy and the Department of Internal Medicine, University of California, Davis, School of Medicine. She teaches anatomy to first-year medical students and undergraduate students. Her research interest is the impact of pedagogies on student learning.

Jennifer A. Weil, M.S., is academic coordinator for undergraduate human anatomy programs in the Department of Cell Biology and Human Anatomy at the UC Davis School of Medicine. She teaches anatomy to undergraduate and first-year medical students. She is pursuing a Ph.D. in Medical and Health Professions Education at EVMS.

Richard P. Tucker, Ph.D., is a professor in the Department of Cell Biology and Human Anatomy, University of California, Davis, School of Medicine. He teaches gross anatomy to first-year medical students and undergraduate students. His research interests are in developmental and evolutionary biology, as well as in anatomy education.

Douglas S. Gross, M.D., Ph.D., is a senior lecturer in the Department of Cell Biology and Human Anatomy, University of California, Davis, School of Medicine. He teaches anatomy to first-year medical students and undergraduate students, and clinical neuroanatomy to undergraduate students. His research interests are global health, disaster medicine, and medical education.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Anderson H, Weil JA, Tucker RP, Gross DS. Impact of gross anatomy laboratory on student written examination performance: A 3-year study of a large-enrollment undergraduate anatomy course. Anat Sci Educ. 2023;00:1–14. https://doi.org/10.1002/ase.2327