HAPS
HUMAN ANATOMY & PHYSIOLOGY SOCIETY

Volume 29, Issue 2 | August, 2025

The EDUCATOR

Journal of the Human Anatomy and Physiology Society

Established in 1989 by Human Anatomy & Physiology Teachers

- Consistent Leader for Gross Anatomy
- Student-Created MCQs for Physiology
- Anatomical Dissection and Student Outcomes
- P-V Curve of the Respiratory System
- Classroom Activity for Renal Physiology
- Low-Stakes Online Quizzes and A&P Outcomes
- Teaching Acid-Base Disorders using AI
- Smart Tablet Technology for Anatomy Teaching

HAPS EDUCATOR JOURNAL

SUMMER 2025

Cover art:
Image from
Luis D. Rosado's article which begins on page 92

TABLE OF CONTENTS

EDUCATIONAL RESEARCH

CURRENT TOPICS

PERSPECTIVES ON TEACHING

HAPS Educator Journal of the Human Anatomy and Physiology Society

The *HAPS-Educator, The Journal of the Human Anatomy and Physiology Society,* aims to foster teaching excellence and pedagogical research in anatomy and physiology education. The journal publishes articles under three categories. Educational Research articles discuss pedagogical research projects supported by robust data. Perspectives on Teaching articles discuss a teaching philosophy or modality but do not require supporting data. Current Topics articles provide a state-of-the-art summary of a trending topic area relevant to anatomy and physiology educators. All submitted articles undergo peer-review. Educational Research articles will additionally be reviewed for the quality of the supporting data. All issues of the HAPS Educator are freely available, and individual articles are uploaded to the Life Science Teaching Resource Community (and link to https://www.lifescitrc.org/) and available in the Education Resources Information Center (ERIC).

The HAPS Educator is published in April, August and December. The deadlines for submission are March 1, July 1 and November 1.

Submission Guidelines for Authors

Information for authors on the terms of submission, the submission procedure, formatting the manuscript, formatting the references, the submission of illustrations, and the peer review process, is available HERE.

Submission Link

When ready, please follow this **link**
to submit your manuscript to the *HAPS Educator.*

You do not need to be a member of the Human Anatomy and Physiology Society (HAPS) to publish in the *HAPS Educator*. For more information see the complete submission guidelines using the link above.

Human and animal research subjects

Research that includes dissection and manipulation of animal tissues and organs must adhere to the Human Anatomy and Physiology Society (HAPS) *Position Statement on Animal Use*, which states that the use of biological specimens must be in strict compliance with federal legislation and the guidelines of the National Institutes of Health and the United States Department of Agriculture. The use of humans or animals in research must fulfill clearly defined educational objectives.

Experimental animals must be handled in accordance with the author's institutional guidelines and informed consent must be obtained for studies on humans. It is the responsibility of the author(s) to secure IRB approval for research on humans.

Plagiarism

Authors must obtain permission to reproduce any copyright material and the source of this material must be acknowledged in their manuscript.

Disclaimer

Responsibility for (1) the accuracy of facts, (2) the expression of opinion and (3) the authenticity of any supporting material presented by the author rests solely with the author. The *HAPS-Educator*, its publishers, editors, reviewers and staff, take no responsibility for these things.

CONTACT THE *HAPS-Educator* Editor if you have additional questions or concerns.

HAPS EDUCATOR
Journal of the Human Anatomy and Physiology Society
EDUCATIONAL RESEARCH
Volume 29, Issue 2 | August, 2025
https://doi.org/10.21692/haps.2025.008

Examining the Effectiveness of a Consistent Leader on Gross Anatomy Performance

Timothy M. Dasinger, PhD, Yuefeng Lu, MD, PhD, Erin Norcross, PhD, Nathan Tullos, PhD, and Gongchao Yang, MD

University of Mississippi Medical Center, Jackson, MS, USA

Corresponding author: tdasinger@umc.edu

Abstract

Gross anatomy is a challenging pre-clinical course in medical school. A lab combined with lecture adds extra pressure, especially if students do not have prior anatomy exposure. One possible strategy to alleviate academic stress within anatomy is to designate table leaders. Some universities implement a rotating leader over the course of the semester as compared to our consistent leader system. Before the semester, all gross anatomy faculty distributed the students to 34 dissection tables based on demographic information and previous anatomy learning, ensuring that each table had a member with some anatomy experience; the students chose their own leaders. The objective was to examine if volunteering or being selected as a leader resulted in better anatomy laboratory performance compared to students who were not leaders. Independent t-tests were run to examine differences in grades between groups, as well as comparing table leaders and the rest of the table. An ANOVA compared differences among tables. Finally, a repeated measures factorial ANOVA was run to test the interaction of table number, group, and table leader on performance. The t-tests revealed no difference between groups. Similarly, the ANOVA comparing tables showed that no difference existed. The independent t-test examining table leaders and the other members was statistically significant. These results suggest that allowing students to choose a leader may make the students take more accountability and result in higher scores for the table leader. https://doi.org/10.21692/haps.2025.008

Key words: leadership, pedagogy, gross anatomy, teaching method

Introduction

Gross anatomy is consistently considered one of the most challenging pre-clinical disciplines in medical school. Students often struggle with the amount of information they are expected to learn, interpreting the spatial organization and relationships of the body, and understanding the clinical relevance of gross anatomy (Cheung et al., 2021; Roach et al., 2021). From the faculty perspective, the decrease in the number of contact hours for both the laboratory and lecture components of anatomy and limited number of faculty present during the laboratory portion of the course also provide challenges for student learning. One possible strategy to alleviate the reduced number of faculty present during gross anatomy lab sessions could be the implementation of table leaders in the lab. In our lab, table leaders served as supervisors of dissection in the lab by assigning roles for the members of the group for each session, ensuring that the dissection is of high quality and completed efficiently, and communicating any issues or concerns their group may have directly with faculty.

Previous research by Teli and Kate (2020) determined that specifying table leaders and having clear roles in gross anatomy positively influenced group dynamics and helped students learn better and develop leadership and communication skills. Similarly, Kyrch and colleagues (2005) found that cooperative learning during gross anatomy labs helped to increase content knowledge and communication skills. Finally, Pawlina et al. (2006) reported that gross anatomy table leaders who delegated roles and heavily participated in gross anatomy dissection had higher exam scores and more integrity and professionalism. While some medical schools have implemented a rotating leader who would change either with each body region or each examination (Wheeler et al., 2015), this study is based on a single leader at each dissection table throughout the course. It is important to note that all of the groups retained their leader for the entire duration of the class. Also, the roles and expectations of the leaders only applied to the time spent in the lab and they were given no guidance to extend leader responsibilities outside of laboratory sessions.

continued on next page

Methods

Participants

In this study, the gross anatomy lab consisted of 167 first-year medical students who were divided into 17 dissection tables. Then, the 17 tables were further divided into two groups (A and B) who alternated dissections so that there was a total of only four or five students at each dissection table during each class session. Dissection groups were determined by the gross anatomy faculty prior to the start of the semester. Before the semester began, the gross anatomy faculty allocated the 167 students to the 34 total table groups based on demographic information including sex, race/ethnicity, and previous gross anatomy experience so that the tables were balanced and had at least one member with some prior exposure to anatomy. Prior gross anatomy exposure included taking a gross anatomy course with a cadaveric component.

On the first day of gross anatomy lab, each table was tasked with determining a leader for their table who would remain the leader for the entirety of the course. The faculty were not involved in this process; the table groups were given full autonomy to choose their leader. The purpose of this study was to examine if volunteering or being selected as a table leader resulted in better gross anatomy laboratory performance compared to students who were not table leaders.

Measures

Exam scores
Scores on the two gross anatomy practical exams were analyzed; the exams consisted of 50 questions each and were worth a total of 100 points. The exams were held in the gross anatomy lab and consisted of 40 cadaveric questions, five bone questions, and five surface anatomy questions. The students had one minute to answer each question, and they rotated from question to question with rest stations after every five stations. The first practical exam covered the back and upper limb regions, while the second practical exam included the thorax, abdomen, pelvis, lower limb, neck, and head.

Demographic information
The following demographic information was recorded: gross anatomy group (A or B), table number (1-17), and whether or not the student was a table leader.

Procedure
This project was approved by the Institutional Review Board of the University of Mississippi Medical Center. Since this was a retrospective study, a waiver of consent was applied. After the gross anatomy course was over, the researchers examined the demographic information and the grades on the two gross anatomy practical exams.

Statistical Analysis
Means and standard deviations were calculated for each student, table, table leader, and group for each exam. Then, independent t-tests were used to examine differences between groups A and B on exam scores as well as comparing table leaders and the rest of the table members. A one-way analysis of variance (ANOVA) was then used to compare differences among tables on the gross anatomy practical exams. After this, bivariate correlations were applied to compare relationships between the two exam scores and a subsequent regression analysis to predict the score on exam two based on exam one performance. Finally, a repeated factorial ANOVA was used to test the interaction of table number, group, and table leader on gross anatomy performance.

Results

The summary of gross anatomy practical exam scores based on demographic information is displayed in Table 1. The independent t-test comparing the means of groups A and B revealed no significant difference on either exam one (t = -0.54, p = 0.59) or exam two (t = -1.07, p = 0.29). Similarly, the ANOVA comparing means among gross anatomy dissection tables showed no significant difference on either examination, (F = 0.89, p = 0.58; F = 1.16, p = 0.31, respectively). Regarding the table leaders' outcomes on the practical exams, an independent t-test comparing their average to those who were not table leaders resulted in a significant difference with a moderate effect on both examinations [(exam one, t_{165} = 3.66; p = 0.002, Cohen's d = 0.62) (exam two, t_{165} = 2.71, p = 0.007, Cohen's d = 0.53)].

	Exam 1 Score (mean +/- SD)	Exam 2 Score (mean +/- SD)
Entire Class	81.8 +/- 10.69	80.95 +/- 12.47
Group A	80.89 +/- 11.01	82.69 +/- 12.40
Group B	81.01 +/- 10.35	80.92 +/- 12.63
Table Leader	86.48 +/- 8.93	87.12 +/-10.61
Non-Table Leader	79.56 +/- 10.70	80.44 +/- 12.57

Table 1. *Comparisons of exam scores.*

A bivariate correlation examining the relationship between exam one and exam two revealed that there was a strong, positive correlation between the two exam scores (r = 0.701; p < 0.001) with almost half of the performance on the first practical exam accounting for the performance on the second exam. The regression analysis predicting the score on the second exam from exam one was also statistically significant (R^2 = 0.492; F= 159.658; p < 0.001). Finally, the only significant value from the repeated measures factorial ANOVA analysis was the influence from the table leader (F = 7.78; p = 0.007). A summary of the output of this ANOVA is displayed in Table 2.

continued on next page

	F	*p*
Table number	1.11	0.36
Group	1	0.32
Table leader	7.78	0.01*
Table number * Group	0.46	0.96
Table number * Table leader	1.06	0.40
Group * Table leader	0.70	0.41
Table number * Group * Table leader	0.46	0.96

Table 2. *Repeated measures factorial ANOVA comparing outcomes for Exams 1 and 2.*

Discussion

The aim of this study was to examine the impact of being a table leader throughout an entire gross anatomy lab course on exam performance. Based on the results, volunteering or being selected as a leader on day one of the course contributed to a significantly higher performance on gross anatomy practical exams compared to the rest of the table members. This is consistent with previous research (Teli & Kate, 2020; Wheeler et al., 2015). These results demonstrate that allowing students to choose a group leader may increase the academic performance for the leader. One possible explanation for this effect may be an increased sense of accountability on the leader's part which might translate into higher scores. This study utilized one table leader through the entire gross anatomy lab as compared to other studies that switched leaders during the course (Darling et al., 2015, Pawlina et al., 2006; Wheeler et al., 2015). Being a table leader for the entire semester resulted in significantly higher scores as compared to Darling et al. (2015) in which a rotating leader only outscored the other members in the group by 1.25% points on examinations.

Previous studies have demonstrated a relationship between being a team leader and higher academic performance on gross anatomy assessments (Teli & Kate, 2020; Wheeler et al., 2015). However, compared with previous studies evaluating a rotating leader, possible advantages of maintaining one leader may include consistency of expectations and roles in student groups and improved communication between students and faculty due to having a stable liaison. Conversely, this means that the other students in the group are at a disadvantage by not choosing to be the table leader. Potential ways to lessen the gap in performance between table leaders and the other members of the group may include training the table leaders in leadership techniques including being more intentional in sharing knowledge with others. Another important aspect to consider is the cohesion of the table and the impact of the leader in uniting the group despite the lack of research indicating a significant relationship between team cohesion and academic performance in gross anatomy (Holman et al., 2016).

Limitations of this study include a convenience sample of students and not assessing prior anatomical exposure. Although prior anatomical exposure was not measured in this study, the students were told that this was not to be the only factor to consider in choosing their table leader and not every table leader had prior gross anatomy experience. Future research should focus on students' perceptions of leaders and their impact on performance. Additionally, using a theoretical framework could be advantageous to better investigate the impact of table leaderships. One possible framework would be tenets of autonomy-supportive environments (Liu et al., 2016) from Ryan and Deci's Self-Determination Theory. Also, exploring the other students' perceptions of their table leader via reflective surveys could provide insight into the group dynamics and students' thoughts regarding the effectiveness of their table leader.

About the Authors

All authors are faculty at the University of Mississippi Medical Center in the Department of Advanced Biomedical Education. Tim Dasinger, PhD, is an assistant professor and was responsible for data analysis and writing of the manuscript. Yuefeng Lu, PhD, MD, is a professor and was responsible for conceptualization and methodology. Erin Norcross, PhD, and Gongchao Yang, MD are professors and Nathan Tullos, PhD is an associate professor; all three helped with methodology and writing/editing of the manuscript.

Literature Cited

Cheung, C. C., Bridges, S. M., & Tipoe, G. L. (2021). Why is anatomy so difficult to learn? The implications for undergraduate medical curricula. *Anatomical Sciences Education, 14*(6), 752-763. https://doi.org/10.1002/ase.2071

Liu, W.C., Wang, J. C. K., & Ryan, R. M. (Eds.). (2016). *Building autonomous learners: Perspectives from research and practice using self-determination theory.* Springer.

Darling, R., Notebaert, A., Lu, Y., Conway, M., Maklad, A., & Sinning, A. (2015, March 28-April 1). *Assessing organizational changes to a first year medical gross anatomy dissection course* [Conference presentation]. Federation of American Societies for Experimental Biology, Boston, MA, USA. https://doi.org/10.1096/fasebj.29.1_supplement.551.4

Holman, M. A., Porter, S. G., Pawlina, W., Juskewitch, J. E., & Lachman, N. (2016). Does emotional intelligence change during medical school gross anatomy course? Correlations with students' performance and team cohesion. *Anatomical Sciences Education, 9*(2), 143-149. https://doi.org/10.1002/ase.1541

continued on next page

Kyrch, A. J., March, C. N., Bryan, R. E., Peake, B. J., Pawlina, W., & Carmichael, S. W. (2005). Reciprocal peer teaching: Students teaching students in the gross anatomy laboratory. *Clinical Anatomy, 18*(4), 296-301. https://doi.org/10.1002/ca.20090

Pawlina, W., Hromanik, M. J., Milanese, T. R., Dierkhising, R., Viggiano, T. R., & Carmichael, S. W. (2006). Leadership and professionalism curriculum in the gross anatomy course. *Annals of the Academy of Medicine, Singapore, 35,* 609-614.

Roach, V. A., Mi, M., Mussell, J., Van Nuland, S. E., Lufler, R. S., DeVeau, K. M., et al. (2021). Correlating spatial ability with anatomy assessment performance: A meta-analysis. *Anatomical Sciences Education, 14*(3), 317-329. https://doi.org/10.1002/ase.2029

Teli, C., & Kate, N. (2020). Effect of group dynamics on performance of first year medical students. *Indian Journal of Clinical Anatomy and Physiology, 7*(1), 36-41. https://doi.org/10.18231/j.ijcap.2020.009

Wheeler, K., Martin, C., & Edmonson, A. (2015). Evaluation of the impact of peer teaching within the gross anatomy laboratory on student academic performance. *Federation of American Societies for Experimental Biology, 29*(S1). https://doi.org/10.1096/fasebj.29.1_supplement.690.14

Back to TOC

EDUCATIONAL RESEARCH

HAPS EDUCATOR
Journal of the Human Anatomy and Physiology Society

Volume 29, Issue 2 | August, 2025
https://doi.org/10.21692/haps.2025.009

Student Engagement and Assessment Satisfaction with Student-Created Multiple-Choice Question Exams in Human Physiology

Laura S. Kabiri, PT, DPT, PhD[1], Thomas M. McCabe, PhD[2], Catherine R. Barber, PhD[2], Augusto X. Rodriguez, PhD[1]

[1]Department of Kinesiology, Rice University, Houston, Texas, USA

[2]Center for Teaching Excellence, Rice University, Houston, Texas, USA

Corresponding author: laura.kabiri@rice.edu

Abstract

Student-centered teaching techniques and equitable assessments are crucial in high-quality education. Asking students to write multiple-choice questions (MCQs) over course content and including them on exams may address both concurrently. Much of the work surveying students' responses to this practice involved graduate level education in specialized programs (e.g., medical or pharmacology school) and generated inconsistent or conflicting findings. Furthermore, minimal research has investigated whether students within majority and minority identity groups perceive this practice similarly. To explore undergraduate students' perceptions of student-created MCQ exams, this study analyzed open-ended and closed-ended survey responses from 40 students who wrote MCQs in an undergraduate human physiology course. Closed-ended responses were positive, indicating increased engagement, assessment satisfaction, and self-efficacy. The survey results indicated that there were no differences in perception between minority and majority groups. Open-ended responses were also largely positive, reflecting improved metacognition, content knowledge, and socioemotional state. Concerns included question wording resulting in difficulty variations and perceptions that the practice is best applied to questions requiring basic knowledge or recall. Student-created MCQ exams were largely well-received and appear to be an engaging student learning activity. Future studies should focus on improving student question writing techniques and expand into larger populations and more diverse content. https://doi.org/10.21692/haps.2025.009

Key words: assessment, engagement, equity, undergraduate

Introduction

Instructors who teach anatomy and physiology courses have increasingly recognized the importance of student-centered learning, in which students take an active role in the learning process rather than being passive recipients of transmitted knowledge (Goodman et al., 2018). Using student-centered teaching strategies in the classroom can improve outcomes of students from historically marginalized groups, including females, underrepresented minorities, first generation students, and students who identify as low income, without harming students who are not members of these groups (Arif et al., 2021).

One aspect of student-centered teaching and learning that is particularly important is equitable assessment (Meyer & Cui, 2019). An equitable assessment method must elicit comparable results for members of various identity groups and be perceived as equally fair by those groups. Equitable assessment also favors more authentic assessment over simple recall of facts and formulas, which allows students to capitalize on their individual preferences and strengths (Ralph et al., 2022). Qualters (2016) has identified student-created test questions as one such example of an inclusive assessment strategy. However, minimal research has explicitly studied whether student-created test questions produce equitable outcomes or whether students from various identity groups perceive these questions as fair.

In a previous paper (Kabiri et al., 2024), we reported on students' superior performance on student-created multiple-choice question (MCQ) exams, compared to instructor-

continued on next page

created MCQ exams, in a human physiology course. The performance advantage of student-created MCQs was seen for all sub-groups; however, the method did not close a performance gap for first generation students or students identifying as low income, suggesting that additional work is needed to understand the benefits and limitations of student-created MCQ exams, particularly viewed through an equity lens.

In the current paper, we examined students' perceptions of student-created MCQ exams, focusing on two research questions:

- ***How did students in a human physiology course perceive the process of developing and assessing their learning with student-created exams, compared to writing instructor-created exams?***
- ***Did students from various underrepresented identity groups view the student-created exams differently from their counterparts?***

Related Literature

Facilitating students' creation of MCQs is a common student-centered teaching strategy. Creating MCQs entails numerous cognitive and meta-cognitive processes, including cognitive elaboration of course concepts and self-assessment of one's understanding of the course material (Yu & Wu, 2016). Previous research has shown that students can create MCQs similar to instructor-created MCQs in terms of item difficulty (Shah et al., 2019) and discrimination (Kabiri et al., 2024; Pham et al., 2023), both of which are measures of item quality (Towns, 2014). In contrast, student-created MCQs tend to require lower levels of cognitive skill than instructor-created MCQs (Khashaba, 2020, Pham et al., 2023), suggesting that students may need guidance to create MCQs that tap into such skills as application and evaluation (Grainger et al., 2018; Hancock et al., 2018). However, the goal of student-created MCQs is not simply to replicate instructor-created MCQs but to provide a more student-centered, equitable approach to assessment.

To achieve equity, student-created MCQs must not only be rigorous, but also fair and effective in eliciting comparable outcomes for various demographic groups.

Although equity has not been addressed in previous research on student-created MCQs, numerous studies support student-created MCQs as both a formative and a summative assessment method. For example, Kay et al. (2020) conducted a multi-year, multi-institution study of PeerWise, an online tool that facilitates student-created MCQs for formative assessment (PeerWise - Documentation). In their sample of more than 3000 students in STEM courses, Kay et al. (2020) found that creating and responding to MCQs in PeerWise during the semester was positively correlated with final exam performance, even after controlling for prior ability. In the summative assessment domain, we found large differences in student performance favoring student-created MCQ exams versus instructor-created MCQ exams in human physiology, although some of the differences may have been attributable to variable levels of item difficulty and complexity (Kabiri et al., 2024).

A successful assessment strategy should not only produce expected academic results but also be well-received by students. Overall, students' experiences of student-created MCQs are positive, with students describing both cognitive and emotional benefits associated with authoring MCQs. A review by Touissi et al. (2022) found that creating MCQs was generally satisfactory to students, though it was also perceived as time-consuming because of task complexity. Similarly, Khashaba's 2020 review of PeerWise in physiology courses found that students responded favorably to creating MCQs, with positive impacts on interest in the material and perceived learning competence.

Most studies indicate that students who create MCQs as part of a formative assessment practice believe that the process deepens their understanding of and confidence in the course material (Kurtz et al., 2019; Pittinger & Lounsbery, 2011; Schullo-Feulner, 2014). Another perceived advantage of authoring MCQs is that it provides practice with generating and answering such questions, which may benefit students in future testing situations (Aflalo, 2021; Kurtz et al., 2019; Shah, 2019). Furthermore, students may experience student-created MCQs as a less stressful but equally valid measure of their learning. For example, Saucier et al. (2022) replaced traditional exams with student-created exams during the COVID-19 pandemic. Although the questions were not MCQ format, the process was similar to other research on student-created exams in that students answered questions they had written. Students found the exams to be valid measures of their learning, enjoyed writing and answering the exam questions, and experienced reduced test anxiety with this format (Saucier et al., 2022).

Concerns about student-created MCQs in the context of summative assessment include the potential ambiguity of peers' questions (Green, 1997) and the perceived challenge of writing questions (Guilding et al., 2021; Papinczak et al., 2012). A common theme related to these concerns is a desire expressed by students for instruction on how to write good MCQ questions (Doyle et al., 2019). Indeed, in Guilding and coworkers' study (2021), some students had concerns about their own question writing ability, suggesting that instructor guidance would be beneficial to improving students' confidence in and satisfaction with the activity.

Several studies have shown that students endorse the use of student-created MCQs in future courses (Bottomley & Denny, 2011; Schullo-Feulner et al, 2014). However, other research has found more negative or mixed perceptions about whether creating MCQs is a valuable activity that should be continued. For example, Grainger et al. (2018) found that

continued on next page

medical students did not see the task of creating MCQs as valuable for their learning; they were dissatisfied with the activity and did not believe that it should continue in the future. Similarly, only about one-third of Papinczak et al.'s (2012) sample indicated that writing MCQ exam questions supported their learning or their ability to problem-solve. However, the majority of students found the bank of questions helpful and supported continuing the activity with future students. Students' perceptions of the pros and cons of authoring MCQs, whether in a formative or a summative context, appear to be complex. Furthermore, previous studies have not examined whether students' perceptions of creating MCQs vary according to demographic variables.

The current study fills a gap in the existing literature by exploring students' perceptions of student-created MCQ exams compared to instructor-created exams within undergraduate human physiology.

Methods

Students enrolled in a three-credit-hour undergraduate human physiology lecture course in fall of 2023 were invited to participate in this study. This project was approved by the Institutional Review Board at Rice University, and informed consent was obtained from all participants. During the initial class meeting, participants were informed about the research study, and additional information was available in the syllabus. Except for one voluntary demographic question, all research study assessment practices aligned with standard course procedures. Participation in the research study was not tied to course performance or assessment, and students could decline participation at any time through a written right of refusal to the provided email address. Opportunities for in-person questions and discussion occurred during the initial class session and throughout the semester.

During the semester, students created and completed four sets of non-cumulative exams, each comprising one exam version with 25 student-created MCQs and one version with 25 instructor-created MCQs. Students were provided with a handout containing MCQ examples and best practices, which was reviewed by the instructor with them along with guided practice and peer review during the first week of class. At the end of each class period, students continued to write and peer review 2-3 MCQs throughout the semester. Subtopics from the lecture were assigned to each group to cover all course material resulting in a total of 75-150 MCQs per topic from which 5 questions were chosen for the exam.

It's important to note that it was not possible for students to easily or ethically preview the questions created by their peers. The campus where this study was conducted requires students to pledge themselves to an honor code, and sharing student-generated questions would be a direct violation of the instructor's explicit instructions and represent a willful violation of institutional policy. While this is not a foolproof solution for academic dishonesty, it was expected to be a major discouragement to inappropriate sharing of the questions. Furthermore, a student wanting to preview the student-generate MCQs would not only have to acquire them by knowingly engaging in academically dishonest behavior but also memorize dozens of questions from which a small subset would appear on any given assessment.

Exam questions were randomly selected for both versions and were excluded if they assessed a topic already tested by another MCQ on the exam. Only minor grammatical, formatting, or punctuation edits were made by the instructor to any MCQs included on the exam. Version author (student or instructor) was randomly assigned using a coin flip for each exam series and students were given the option to take either version first. Students were also blinded to the exam version author during assessment. Students took both versions of the exam, with the higher score retained as their grade of record. After grades were reviewed and posted by the instructor, students were informed which version was instructor-created and student-created. Question quality was assessed using multiple methods including the percentage of correct responses, a discriminatory coefficient, and cognitive complexity score. More detailed information regarding exam creation, performance, and quality has been described elsewhere (Kabiri et al., 2024).

Students voluntarily completed a researcher-created survey assessing their engagement and satisfaction with student-created exams at the end of the course. This survey contained both closed-ended and open-ended questions to generate quantitative and qualitative data, respectively. An optional demographic question asked students to select any group with which they identified to determine differences in engagement and satisfaction among underrepresented populations in STEM. Students could select any number of the following: first generation students (first in their family to attend college), low-income college students, underrepresented racial/ethnic/cultural minority, female/feminine/woman, or none of the above.

The next seven survey questions, which assessed students' responses to the exam formats, were informed by input from experts in teaching and learning and used a seven-point Likert scale (strongly disagree to strongly agree). Compared to instructor-created exams, students were asked to rate their overall preference for student-created exams as well as the exams' ability to increase the following constructs: critical thinking about course material, how likely students felt they would succeed on the exam, ability to learn course material, ability to retain course material, engagement in the testing process, and satisfaction with the testing process.

continued on next page

The final three questions were open-ended. Students described other ways in which the use of student-created exams affected their learning experience in the course and how the process affected their testing preparation and/or course performance. Finally, students provided their perspectives on whether student-created exams would be desirable in other courses and were encouraged to explain their answers.

Students received the link to a Google form containing the survey questions during class at the end of the course prior to final grade release. Responses were optional and anonymous, and no identifiers were collected. A sample survey is provided in Appendix 1.

Statistical Analysis

Sample demographics were described as frequencies and percentages. Descriptive statistics were calculated for the seven Likert scale items, including mean, median, mode, and standard deviation. Exploratory factor analysis assessed the possible underlying structure of the seven researcher-written survey items. Subsequently, Cronbach's *alpha* values for the identified factors were calculated to examine internal consistency as a measure of reliability. *T* tests (parametric) or Mann-Whitney U tests (non-parametric for violations of the assumption of normality) were used to compare responses between participants who identified as members of an underrepresented group and those who did not. All statistical tests were run using *Jamovi* (version 2.5.3), with an alpha value of ≤ 0.05 to indicate significance. Using the factors identified by the exploratory factor analysis, a composite score for each factor was calculated for each student by averaging their scores for the questions ascribed to each factor. These composite scores were plotted as a histogram (Figure 1) to visualize trends across the factors.

Qualitative Analysis

For the open-ended questions, a codebook was developed to categorize student responses to questions targeting how student-created exams affected their learning, test preparation and performance, as well as interest in using student-created exams beyond the current course. Two researchers (CB and TM) independently coded participant responses to all three open-ended questions using an inductive approach. One round of interrater checks was performed to identify codes, followed by a second round of coding. Finally, coders resolved any discrepancies by consensus. The summarized qualitative data is presented using frequencies and representative quotations.

Further qualitative analysis explored whether students from different identity groups had different levels of agreement with one of the free response questions compared to their counterparts. The two coders rated the valence of each student's response to the question, "Having created and taken peer-created exams this semester, would you want to have this type of exam option in other classes? Please explain why or why not." The responses were then sorted to isolate the responses of students from each of the four demographic categories tested in this study and compared against the spread of responses in the group of students who did not identify with being first generation, low income, underrepresented racial or ethnic minority, nor female, respectively.

Results

Executive Summary

In all, 40 of 46 students enrolled in this human physiology course completed the survey in full for a response rate of 87%. A majority (73%) of the responding students identified as female/feminine/woman, and 44% of identified as underrepresented racial, ethnic, or cultural minority. Detailed demographics of the sample are found in Table 1, while descriptive statistics for the seven Likert-scale items are found in Table 2. Exploratory factor analysis identified two latent factors and the composite scores for these factors were calculated by averaging for each student. These composite scores trended positive for both factors with most students responding from slightly to strongly agree for both factors (Figure 1).

Students expressed positive affective responses towards the technique, with 29.3% of responding students rating this construct at slightly agree or above. They indicated that, compared to instructor-created exams, they preferred the student-created exams, they felt more likely to succeed on student-created exams, and were more satisfied with the assessment process after creating and taking student-created exams. The cognitive factor ranked toward the positive for most students, with 78% of students with composite cognitive scores rated at slightly agree or above. This trend appears driven by students' agreement with the statement that they were more engaged with the assessment process. Students reported generally positive, but lesser agreement with the impact of student-created exams on their ability to learn and retain the course material. These positive sentiments seen in the quantitative results were mirrored in students' qualitative responses to open ended questions. Most student statements (77/112 coded statements) focused on positive effects of creating and being assessed with student-created exams, highlighting both cognitive and social-emotional consequences.

continued on next page

Identity	Frequency (%)
First Generation College Student	8 (20%)
Low Income College Student	9 (22.5%)
Underrepresented Racial, Ethnic, or Cultural Minority	18 (45%)
Female/Feminine/Woman	30 (75%)
None of These	6 (15%)

Table 1. *Demographic identities of the sample (n = 40). Students were allowed to select as many of the listed identities as applied. Students not identifying with any of the above answered "None of These." All but one participant provided demographic information.*

Compared to an instructor-created exam…	Mean (± SD)	Median	Mode	Factor
peer-created exams made me think more critically about the material.	3.63 (± 1.50)	3	3	Cognitive
peer-created exams improved my ability to retain the material.	4.63 (± 1.34)	5	4	Cognitive
peer-created exams improved my ability to learn the material	4.80 (± 1.42)	5	5	Cognitive
peer-created exams increased my engagement in the testing process.	5.15 (± 1.61)	5	7	Cognitive
I prefer peer-created exams.	5.73 (± 1.40)	6	7	Affective
I felt more likely to succeed on peer-created exams.	5.88 (± 1.45)	6	7	Affective
peer-created exams increased my satisfaction with the testing process.	5.63 (± 1.64)	6	7	Affective

Table 2. *Aggregate Likert-scale survey item descriptive statistics (n = 41). Students responded to seven researcher-written Likert-scale items by rating their agreement with each statement on a scale from "Strongly Disagree" (1), "Disagree" (2), "Slightly Disagree" (3), "Neutral" (4), "Slightly Agree" (5), "Agree" (6), to "Strongly Agree" (7).*

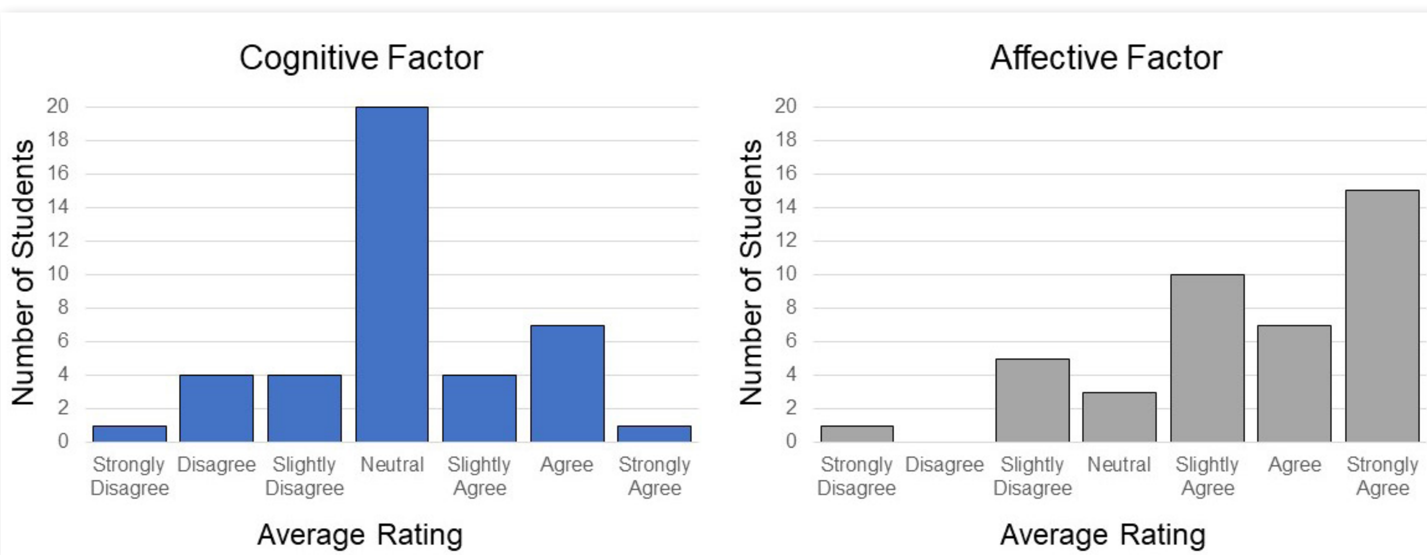


Figure 1. *Histograms of average student scores for cognitive and affective factors (n = 41) displaying the number of students with a composite score corresponding to each level on the seven-point Likert scale.*

continued on next page

Exploratory Factor Analysis of Researcher-Written Survey Items

Exploratory factor analysis was conducted using a principal axis factoring extraction method. Given the sample size and inter-factor correlation, oblimin rotation was selected. The analysis indicated a two-factor solution with an inter-factor correlation of 0.39 and Eigenvalues of 3.39 and 1.30, cumulatively explaining 75% of the variance. There were no violations of the test assumptions, permitting interpretation of the results. The KMO Measure of Sampling Adequacy for the overall model was 0.76, and Barlett's Test of Sphericity returned a p value < 0.001, rejecting the null hypothesis that the correlation matrix is an identity matrix.

Factor loadings, communality values, and factor correlations are found in Table 3. The first factor included four items ($\alpha = 0.88$) describing a subscore related to effects on students' cognition as a result of taking student-created exams. In particular, the items related to increases or improvement in ability to learn, engagement in the assessment process, critical thinking, and ability to retail course material. The second factor included the remaining three items ($\alpha = 0.94$) describing a subscore related to students' affective responses to the student-created exams. Specifically, the items asked students if they felt more likely to succeed, had increased satisfaction with the assessment process, and their preference for peer- or instructor- created exams.

Statistical Analysis of Groupwise Differences in Survey Responses

Independent samples *t* tests for the cognitive subscores and Mann-Whitney U tests for the affective subscores were calculated for differences in each of four demographic groups: first generation students (first in their family to attend college), low income college students, underrepresented racial/ethnic/cultural minority, and female/feminine/woman. No significant differences in mean score for any comparison occurred, showing that students identifying with any of the demographic subgroups did not differ significantly from their counterparts in their responses (Appendix 2).

Qualitative Analysis

Participants returned a total of 85 coded utterances. Blank responses were not counted or coded in this analysis. A single utterance could contain multiple codes and all but one code was identified in response to all three questions asked. Inductive coding by two researchers (CB and TM) returned four major codes and nine subcodes (Table 4). Student answers were coded as cognitive (C_) or social and emotional (SE_) responses to creating and assessing learning with student-created exams. Participants also identified limitations or drawbacks (LD_) to student-created exams, and instances where they did not experience a difference between the two exam types were labeled neutral (NEU). A full table of the utterances and their codes can be found in Appendix 3.

Factor ID	Survey Item	Factor 1	Factor 2	Communality
Cognitive	Improved Ability to Learn Material	**0.91**	0.05	0.87
	Increased Engagement in Assessment Process	**0.79**	-0.002	0.62
	Improved Critical Thinking	**0.78**	-0.14	0.55
	Improved Ability to Retain Material	**0.77**	0.08	0.65
Affective	Felt More Likely to Succeed	-.002	**0.97**	0.92
	Increased Satisfaction with Assessment Process	0.02	**0.94**	0.90
	Prefer Peer-Created	0.01	**0.86**	0.74

Table 3. *Communalities and factor loadings for exploratory factor analysis of Likert-scale survey items (cognitive and affective) to determine whether and then how the seven researcher-written survey items grouped together. Factor loadings represent the correlation between the individual items and a theoretical factor (Factors 1 & 2). Positive factor loadings closer to 1 indicate a stronger correlation and negative values indicate a negative correlation. The communality scores provide information about the proportion of variance for that item explained by the factors identified by the analysis. (n = 41)*

continued on next page

Main Code	Subcode Name	Short Subcode	Frequency
Cognitive	Metacognition	C_MC	29
	Content Relevance	C_CR	21
Social and Emotional	Emotion	SE_EM	17
	Peer Relationships	SE_PR	10
Limitations or Drawbacks	Instructor Language Difficult	LD_DIFF	7
	Wording of Questions	LD_WORD	8
	Not Appropriate for Application	LD_APP	6
	Negative Critiques	LD_NEG	6
Neutral	Neutral or No Change	NEU	8

Table 4. *Qualitative codes and subcodes with frequencies (n = 33). These qualitative codes were developed using an inductive approach by analyzing the total pool of answers from students to all three open-ended questions. 84 unique responses were analyzed; however, not all received a code and some received multiple codes. The frequencies represent the number of utterances that received the respective subcode.*

When discussing the cognitive aspects of student-created exams, participants described how the process of creating and assessing learning using student-created exams affected their metacognition (C_MC). These responses explained how writing exam items positively affected their study process and acted as a review of the material. Greater attention to the material was often mediated by the process of developing exam questions. Sample student responses demonstrating this include:

> "I think this exam format helps students review material while learning it, increasing long term retention."

> "...[this exam option] definitely opens up more opportunities to prove my knowledge of the content, beyond instructor-generated questions."

The second subcode under cognition connected the student-created exam process to content relevance (C_CR). Here, participants described how the process affected their participation in class, such as seeing a direct connection with the course materials provided. Participants also recognized how this process related to their experience of the content discussed during the face-to-face portions of the course. Related responses included:

> "They were based on activities we did in class - so I could relate peer-questions to experiences from class."

> "...I felt the peer exams were a better reflection of what I was actually learning rather than what I was 'expected' to figure out."

The social and emotional (SE) code incorporated two subcodes. The first was a range of responses where participants described emotional (SE_EM) responses to the process of creating and taking student-created exams. All utterances identified with the SE_EM code were instances where participants described positive effects of the student-created exams on their engagement, testing anxiety, confidence in test-taking, ability to learn, security during their study time, and reductions in stress towards test preparation. The following comments illustrate these subcodes:

> "...I think I perform better with peer-created exams and it facilitates my desire to study and learn the material."

> "The peer-created exams made me feel more confident in continuing to learn the material as I knew I had a greater chance of understanding the questions."

In addition to these more affective responses, participants identified positive interactions with their peers as a result of participating in the student-created MCQ-exam process. The peer relations (SE_PR) code described responses where students shared how the process allowed them to learn alongside other peers, showed similar thoughts as their peers, and was reflective of students' course experiences as presented below:

> "I also think [student-created exams] decreases the amount of testing anxiety one may have because you know the people who made the exam are your peers who are also learning the material."

continued on next page

> "I knew if it was a question made by students, I should definitely know it as well, so I knew I could do well."

Participants also returned several insights into the limitations and drawbacks (LD) surrounding both student- and instructor-created exams. For some, the student-created exams were perceived as easier in comparison to the instructor-created exams (LD_DIFF). For example, one student wrote:

> "I think the instructor-made exams usually had questions that better represented the level at which we were to be learning the material while the peer-made exams were more general questions that emphasized the important/main parts of a given topic."

The second subcode addressed the wording (LD_WORD) of the questions. Responses coded as LD_WORD often identified disparities in the wording of questions on the student-created exams or praised the consistency of the language in the instructor-created exams. Related comments included:

> "I feel like the questions varied in difficulty, not because of the content, but because of the awkward wording of some questions. The exams didn't feel consistent."

> "I noticed that peer questions are sometimes hyper specific and other times way too general, thus, the difficulty level of the average peer questions is all over the place as compared to instructor questions, which are usually consistent in their difficulty."

Third, a number of participants provided generally negative feedback about the student-created exams, labeled as LD_NEG. Occasionally, this code appeared with the previous LD_WORD where participants expressed a negative view towards the student-created exams in part because of perceived differences between student and instructor language. Some participants also expressed no benefit to their studying or understanding of the material as a result of developing and taking student-created exams. However, there was not always an explanation as to why they felt this way. Responses included:

> "I think we can get lazy with making questions and it doesn't really help with learning the material. In addition, without someone actually reviewing the questions, I feel like the questions themselves and answer choices can be misleading."

The last subcode was the only code identified in reference to a specific survey question. Responding when asked if they would want to have student-created exams in other classes, several participants identified that the student-created exam process may not be appropriate for courses that focused on application (LD_APP). Rather, these questions were better suited for courses with an emphasis on knowledge acquisition by memorization. Two responses indicating this concern are:

> "Yes, but only in classes in which material is not application heavy, I believe these tests are optimal for memorization heavy material since questions are made in a short amount of time, encouraging less application based questions and promoting simple memorization questions."

> "Yes, but probably not in classes that are very heavy in detail and calculations/problems (i.e., physics, organic chemistry); these classes are less memorization based, so the questions are a lot less straightforward. The questions are also easy to convolute and need to be worded very carefully. I felt that sometimes the peer-created exams in this class were even worded less well/clearly."

The final major code was all-inclusive for utterances where students expressed that they did not experience a difference between the two exam types, labeled neutral (NEU). This code was used for a variety of reasons; two utterances were reported "N/A" when asked whether they detected a difference between the exam types. For others, the source of the exam items did not affect how they approached the course. Lastly, some expressed ambivalence because they did not see any difference in their performance on either exam type.

Valence Analysis of the Question regarding Student-Created Exams in Other Classes

The valence analysis conducted on the open response question regarding students' interest in having the student-created exam technique used in other classes revealed some trends in how students identifying from different demographic groups viewed the technique (Table 5). Among students who identified as first generation, a lower relative proportion (25%) affirmed that they would want to use the technique in other classes compared to students who did not identify as first generation (60%). The larger proportion (75%) of first-generation students did not want to use the technique in other classes or had a mixed response. Students provided a range of reasons explaining why they may not want to use the technique such as wanting more preparation in designing questions, quality control of questions used on assessments, and commentary about how instructors should also be contributing questions, "because [instructors] know what topics and phrases will best test [students'] knowledge of the material."

continued on next page

Identity Group	Positive	Neutral	Negative
First Generation (n = 8)	25%	50%	25%
Not First Generation (n = 32)	60%	28%	12%
Low Income (n = 9)	33%	45%	22%
Not Low Income (n = 31)	58%	29%	13%
URM (n = 18)	44%	31%	25%
Not URM (n = 22)	59%	35%	6%
Female (n = 30)	65%	26%	9%
Not Female (n = 10)	25%	50%	25%

Table 5. *Results of the valence analysis of student responses to the question "Having created and taken peer-created exams this semester, would you want to have this type of exam option in other classes? Please explain why or why not." These numbers represent relative proportions of students identifying with and not identifying with underrepresented identity groups responding in the affirmative (Positive), mixed valence (Neutral), and negative (Negative).*

A similar trend was seen with students identifying as low-income college students responding in the affirmative less frequently compared to students who did not identify as having a low income. However, a majority (67%) of the students identifying as having a low income also identified as first generation students, partially explaining this similar trend.

Students identifying from underrepresented racial or ethnic minorities (URM) had numbers of neutral opinions (31% URM v. 35% non-URM) about the technique that were similar to those not identifying as an underrepresented minority. However, there were more negative responses from underrepresented students (25% URM v. 6% non-URM), also reflected in a lower percentage of URM students reporting positively (44% URM v. 59% non-URM).

Finally, students identifying as female returned a similar number of neutral or negative responses to the question as their non-female counterparts, but they represented most (15/17) of the fully positive responses. Female students made up the majority of the population (73%), but on average were more likely to want the technique used in different courses. Their reasoning for why they would support the use of the technique spanned the qualitative code book described above, representing positive outcomes on their metacognition, a better understanding of the relevance of the material they were learning, and a variety of positive social and emotional consequences.

Discussion

This study attempted to better understand how students perceived the process of designing and taking student-created compared to instructor-created MCQ exams in a human physiology course. Meyer and Cui (2019) provided a framework for understanding how an assessment can be deemed equitable by assessing the equity of the performance of various learners and that all learners perceive the process to be equitable. A previously published companion piece to the current study showed that all groups tested had better performance on student-created MCQ exams compared to instructor-created MCQ exams, with a large effect size. However, a performance gap persisted for students from some underrepresented identity groups with first-generation and low-income students scoring significantly lower than their counterparts on both student- and instructor-created MCQ exams (Kabiri et al., 2024). Closing this gap is the focus of future work, but the data in this study help to elucidate whether students responded similarly to the technique, regardless of their identity group. Survey data and open response answers provide evidence that students had generally positive attitudes towards the technique, though there were some differences associated with students' identities. The results point to suggested improvements for future use of the technique in anatomy and physiology classrooms.

A researcher-designed survey of Likert items was used to investigate students' preference for student-created MCQ exams and their perceptions of how participating in the student-created exam process contributed to critical thinking about course material, likelihood of succeeding on the exam, ability to learn course material, ability to retain course

continued on next page

material, engagement in the testing process, and satisfaction with the testing process. A factor analysis of the survey items revealed two latent factors, one associated with effects on students' cognitive processes, and another with students' affective responses to the technique. The high levels of agreement with statements related to both latent factors in this study match well with other studies of student-created MCQ exams, indicating that the technique is widely liked by a variety of students (e.g., Khasaba, 2020; Touissi, 2022).

This generally positive sentiment was further expressed in the types of benefits that students reported through open-ended questions. These observations were similar to other reports in the literature. For example, the act of writing new questions and being assessed using them enhanced some students' approach to studying for the class and provided an opportunity for students to review material in a different mode than first presented, as also described by Aflalo (2021) and Kurtz and colleagues (2019). In addition, students in this study population noted that they better understood the material, explaining that the process of writing questions was useful to their study process and that they were able to review the material more than they would normally. Others have also noted improvement to students' understanding of the material and an increase in students' perception of their learning ability after the process of developing MCQs (Khasaba, 2020; Kurtz et al., 2019; Pittinger & Lounsbery, 2011). Using this technique enhanced students' understanding of the relevance of particular information to the larger course structure, helping them to connect between different activities and the provided course materials.

Students in this study also reported a range of social and emotional consequences after using the technique. The open-ended questions were not written to address any specific aspect of students' social and emotional experience of the course, but student responses matched similar consequences reported elsewhere: increases to their engagement (Pittinger & Lounsbery, 2011), reduction of anxiety and stress towards assessment (Saucier et al. 2022), and increased confidence in their ability to learn and perform on assessments (Kurtz et al. 2019; Schullo-Feulner et al. 2014).

Novel to this study, students reported benefits connected to how they related to other students in the same class. There were several responses indicating that using the technique allowed them to see how their learning matched their peers.

> "...I feel the peer exam was more representative of what is known as a student learning the material for the first time than an instructor who has been learning the material and teaching the class for years."

Students reported that they identified having similar ideas to their peers and ultimately that the assessments were reflective of the lived experiences of students in the course. This camaraderie is notable with regard to the goal of creating more equitable assessments. It indicates that students participating in the larger project of assessment development may feel more connected to a local group of scholars, which theoretically could impact their engagement, the course climate, and their performance in a course (Gopalan & Brady 2019; Herman & Hilton, 2017). Focusing on how students interact with each other through the process of developing MCQs may also address difficulties mentioned by this and other study populations such as the complexity and challenge of writing high quality, clear questions or writing questions at an appropriate level of difficulty.

In the present study, students were given basic information and training about writing MCQs at the beginning of the semester. Throughout the rest of the semester, students wrote new questions and received peer feedback at the end of most class sessions. The peer feedback was designed to help lighten instructor workload by reducing the number of individual instructor-student interactions and hopefully subvert the need for extensive training or a robust peer mentoring system. However, given concerns raised by the students about the difficulty of the questions, the introduction of a peer-instructor/teaching assistant (e.g., a student who had taken a course with student-created exams) who could provide experience and guidance toward the development of better assessment items may be a worthwhile addition to this type of student-created exam approach. Also, adjustments to the nature of peer interactions in the writing process could correct an apparent lack of training of students (Doyle et al. 2019; Guilding et al. 2021).

While students, on average, reported positive responses to the technique, they also provided feedback about how the process was less beneficial. Part of this feedback included open-ended responses addressing perceived differences in the quality of the question language. Some students found comfort in the consistency of the instructor-created questions in terms of the wording and perceived question difficulty. Others found student-created questions easier to understand, but this came with inconsistent levels of difficulty across the student-created questions. Other researchers have documented similar findings, highlighting student concerns that more training or guidance was warranted to improve question quality (e.g., Green 1997; Guilding et al. 2019; Papinczak et al. 2012). Presumably, a better understanding of how to generate effective assessment items would also address question consistency issues and critiques of some students in other studies regarding how long the process took them to complete (Papinczak et al. 2012). Participants in the present study did not mention this last concern. But this is likely connected with students reporting that they used the writing process as another means of engaging with course materials.

continued on next page

A final indicator of students' acceptance of student-created MCQ exams was their level of endorsement of the technique outside of the physiology course where they encountered it. Students generally agreed that they preferred the student-created MCQs over the instructor-created MCQs and approximately half of participants agreed that they would want to see the technique in other courses. Additionally, most participants answering with a more neutral response said that they would want to see the technique used in other courses, but with some adjustment to the process as they experienced it in the present course. Interestingly, a consistent comment from students who had a more negative view of the technique was that they felt that the questions were less complex and only tested students on knowledge acquisition. Other studies have also observed a similar trend for students to write questions requiring lower cognitive skills (Kabiri et al., 2024; Pham et al., 2023). Additionally, several students responding with a more negative view said they did not think the technique would be successful in courses like physics and organic chemistry where the exams required the application of information or more complex cognitive skills. Fortunately, other researchers have provided some insight into the development of more cognitively challenging MCQs (e.g., Crowe et al., 2008; Tractenberg et al., 2013) and have demonstrated that subject matter non-experts can accurately rate MCQs for their level of cognitive complexity. By extension, it may be possible to train non-experts (e.g., students) to write MCQs at varying levels of cognitive complexity.

With regards to the second research question addressing acceptability amongst groups of students, the data did not provide evidence indicating that underrepresented students viewed the student-created exams differently from their counterparts. Regardless of their identity(ies), students reported similar levels of agreement in terms of the technique's impact on both their cognitive and social-emotional experiences of the course. This is encouraging as, at least for this sample, the process of designing and taking student-created exams appeared similarly acceptable. As indicated above, students generally expressed positive perceptions of student-created exams, encouraging further research into the utility of this assessment design.

Comparing the open-ended responses of students from underrepresented groups to their non-identifying peers, there were varying levels of agreement as to the use of this technique in other courses. A majority of students identifying as first generation and low-income college students did not want the technique in other courses. Those identifying as underrepresented racial or ethnic minorities largely agreed with their counterparts. Sample size and the nature of the demographic questions did not allow us to explore whether other factors may explain variations in students' responses to the technique. Finally, most of the positive responses to this question came from students identifying as female with only two non-female students affirming their interest in using the technique in other classes. The scope of this valence analysis is limited, but it provides some leads in monitoring how students from various intersectional identities understand and appreciate student-created exams. However, another potential conclusion is that no gross differences in how students appreciated the technique occurred and that some level of equity was demonstrated. Taking the quantitative and qualitative results together, most students, regardless of identity, saw benefits to using the technique.

Practical Implications

In an academic environment fraught with issues of inequity and anxiety, the practice of writing and assessing students with MCQs written by their peers is a promising tool for addressing both. The largely positive response from students on both open- and closed-ended questions regarding student-created MCQ exams indicated minimal to no difference in preference among underrepresented minority groups. The increased self-efficacy and perceived ability to succeed as well as direct quotes from the open-ended survey questions also clearly reported decreased test anxiety. While further improvements are possible in the system, this initial study shows promise for applying this technique in undergraduate human physiology courses.

One additional benefit noted by students and the professor alike includes the sense of community and belonging that accompanied the practice. Throughout the term, students worked with their peers to receive feedback on their MCQs. These rounds of feedback appeared regularly in the course schedule as students wrote questions for all content sections of the course. Students were encouraged to meet with new peers each time and this, in turn, allowed students to identify similar struggles and ideas during the question writing process.

> "Yes, having peer created exams has been helpful and a more creative way of test taking. Because we all have similar ways of understanding the knowledge and thought processes the questions will be questions we will want to engage with."

> "I enjoyed taking peer-created exams as I was able to learn alongside with my peers and test each other on the information that we learned collectively. I also found the questions more relatable to the content we were learning…"

Recognizing these commonalities can contribute to a sense of community and build networking skills, as noted in open-ended student survey responses. Additionally, contributing questions for exam material can foster a sense of belonging among students. Knowing that questions crafted from their personal viewpoint held value and contributed to the group test bank lent weight and import to the practice. Rather than seeing it as "busy work," students seemed to feel they were making meaningful contributions to the group learning and assessment process.

continued on next page

From an instructor's viewpoint, the increased engagement with the material and collaboration seen during question writing and peer review was evident. Any classroom or assessment technique that increases these skills as well as increases satisfaction with the assessment process as indicated by student surveys is worthwhile. Anecdotally, complaints and challenges regarding exam scoring were also noticeably less or absent altogether. While the exact reason or rate of decrease was not a focus of this research, fewer students requesting regrades or private exam review was a perceived benefit to the instructor (Kabiri et al., 2024). Additionally, the practice of requiring students to write MCQs creates a living test bank with numerous questions addressing each course content topic added weekly. With only around 5% of questions from each test bank included in the actual exam, there was still an abundance of unused questions to potentially incorporate into practice quizzes, future exams, or even exams for other course sections or sessions. The sheer volume of questions created per topic using this technique serves as an additional benefit noted by the instructor.

Limitations and Opportunities

The scope of this study was narrow, looking at a small sample of predominantly female students in a physiology course at a private doctoral university with very high research activity in the southern US. However, all four underrepresented groups in STEM were represented, with almost half of the sample identifying as an underrepresented racial, ethnic, or cultural minority. Nevertheless, future studies should survey larger samples to strengthen these findings. In addition, this study did not rigorously analyze the connection between a student's identity and the quality of their responses. Further limitations include the use of self-reported surveys with the potential for self-selection bias. However, the 87% response rate is strong and helped minimize this potential bias. The satisfaction survey responses may have been affected by other forms of bias. For example, students may have responded positively because they understood which questions were student- versus instructor-created by the time they completed the satisfaction survey and expressed a preference towards the questions they created. It is also possible that students were able to infer that, on average, scores for the student-created MCQs were higher than instructor-created or they may have noticed a difference in the cognitive level of the questions (Kabiri et al., 2024). However, they would have had to come to these conclusions independently, as these research findings were not available during the time that students were taking the course. Future research will continue to increase our understanding of student reception of this technique and elucidate ways of improving the design and implementation process.

Conclusion

Results from this study further support the use of student-created MCQs and exams among undergraduates in human physiology. Student responses to Likert scale questions were largely positive and cited benefits in learning and retaining course material as a result of writing MCQs. Moreover, as an assessment process, students reported feeling not only improved engagement and satisfaction with the assessment process, but they also reported feeling more likely to succeed, indicating higher self-efficacy with this practice. As minimal differences were seen between underrepresented groups and their non-underrepresented peers, our study also suggests the practice may be a valuable component of equitable assessment.

In response to open-ended questions, student feedback was also generally positive. More specifically, most students stated positive effects on their metacognition, understanding of course content, and socioemotional state with respect to the course. Student comments on limitations and drawbacks of student-created MCQ exams revealed concerns about wording inconsistencies between student- and instructor-created questions. Some students thought their peers worded questions in a way that made them easier to understand, while others said they were more difficult to interpret. An additional concern with student-created MCQ exams included the perception that the practice was more suited to less complex cognitive skills like basic recall and that using the same practice for application-level responses might be more difficult to write.

Continued use and investigation of student-created MCQs and exams as an assessment tool in the tested populations, expansion of the tool into other content areas, and recommendations on how to improve student training in designing assessments to improve the quality of assessment items and student preparation for the assessment process are warranted. Future research will adjust how students are trained to write questions in future courses and provide additional training on selecting questions for exams that require application.

About the Authors

Laura S. Kabiri, PT, DPT, PhD, is an assistant teaching professor in the Department of Kinesiology at Rice University. Her main teaching responsibilities include human anatomy, human physiology, and medical terminology. Her research interests include educational experiences and assessment in undergraduate STEM courses, particularly among underrepresented minorities. Catherine R. Barber, PhD, is an assistant teaching professor and associate director of the Center for Teaching Excellence at Rice University. Her teaching and research interests include exploring the intersection of cognitive, social, and emotional aspects of learning, using evidence-based strategies for developing

continued on next page

effective and engaging online courses and workshops, and implementing action research to improve teaching and learning. Thomas M. McCabe, PhD, is an assistant teaching professor and associate director of the Center for Teaching Excellence at Rice University. His research and teaching interests include post-secondary biology education research, professional development, and curricular design. Augusto X. Rodriguez, PhD, is an associate teaching professor and chair of the Department of Kinesiology at Rice University. His areas of expertise include research design and statistical analysis.

Literature Cited

Aflalo, E. (2021). Students generating questions as a way of learning. *Active Learning in Higher Education*, *22*(1), 63–75. https://doi.org/10.1177/1469787418769120

Arif, S., Massey, M. D. B., Klinard, N., Charbonneau, J., Jabre, L., Martins, A. B., et al. (2021). Ten simple rules for supporting historically underrepresented students in science. *PLOS Computational Biology*, *17*(9), Article e1009313. https://doi.org/10.1371/journal.pcbi.1009313

Bottomley, S. & Denny, P. (2011). A participatory learning approach to biochemistry using student authored and evaluated multiple-choice questions. *Biochemistry and Molecular Biology Education, 39*(5), 352-361. https://doi.org/10.1002/bmb.20526

Crowe, A., Dirks, C., & Wenderoth, M. P. (2008). Biology in Bloom: Implementing Bloom's taxonomy to enhance student learning in biology. *CBE—Life Sciences Education*, *7*(4), 368–381. https://doi.org/10.1187/cbe.08-05-0024

Doyle, E., Buckley, P., & Whelan, J. (2019). Assessment co-creation: An exploratory analysis of opportunities and challenges based on student and instructor perspectives. *Teaching in Higher Education*, *24*(6), 739–754. https://doi.org/10.1080/13562517.2018.1498077

Goodman, B. E., Barker, M. K., & Cooke, J. E. (2018). Best practices in active and student-centered learning in physiology classes. *Advances in Physiology Education*, *42*(3), 417–423. https://doi.org/10.1152/advan.00064.2018

Gopalan, M. & Brady, S. T. (2020). College students' sense of belonging: A national perspective. *Educational Researcher*, *49*(2), 134–137. https://doi.org/10.3102/0013189X19897622

Grainger, R., Dai, W., Osborne, E., & Kenwright, D. (2018). Medical students create multiple-choice questions for learning in pathology education: A pilot study. *BMC Medical Education*, *18*(1), Article e201. https://doi.org/10.1186/s12909-018-1312-1

Green, D. H. (1997). Student-generated exams: Testing and learning. *Journal of Marketing Education*, *19*(2), 43–53. https://doi.org/10.1177/027347539701900205

Guilding, C., Pye, R. E., Butler, S., Atkinson, M., & Field, E. (2021). Answering questions in a co-created formative exam question bank improves summative exam performance, while students perceive benefits from answering, authoring, and peer discussion: A mixed methods analysis of PeerWise. *Pharmacology Research & Perspectives*, *9*(4), Article e00833. https://doi.org/10.1002/prp2.833

Hancock, D., Hare, N., Denny, P., & Denyer, G. (2018). Improving large class performance and engagement through student-generated question banks. *Biochemistry and Molecular Biology Education*, *46*(4), 306–317. https://doi.org/10.1002/bmb.21119

Herman, J., & Hilton, M. (Eds.). (2017). *Supporting students' college success: The role of assessment of intrapersonal and interpersonal competencies*. The National Academies Press. https://doi.org/10.17226/24697

Kabiri, L. S., Barber, C. R., McCabe, T. M., & Rodriguez, A. X. (2024). Student performance and exam quality in student-versus-instructor-created exams in human physiology. *HAPS Educator, 28*(2), 35-46. https://doi.org/10.21692/haps.2024.011

Kay, A. E., Hardy, J., & Galloway, R. K. (2020). Student use of PeerWise: A multi-institutional, multidisciplinary evaluation. *British Journal of Educational Technology*, *51*(1), 23–35. https://doi.org/10.1111/bjet.12754

Khashaba, A. S. (2020). Evaluation of the effectiveness of online peer-based formative assessments (PeerWise) to enhance student learning in physiology: A systematic review using PRISMA guidelines. *International Journal of Research in Education and Science*, *6*(4), 613–628.

Kurtz, J. B., Lourie, M. A., Holman, E. E., Grob, K. L., & Monrad, S. U. (2019). Creating assessments as an active learning strategy: What are students' perceptions? A mixed methods study. *Medical Education Online*, *24*(1), Article e1630239. https://doi.org/10.1080/10872981.2019.1630239

Meyer, E. R. & Cui, D. (2019). Diversity and inclusion in anatomy and physiology education, degree programs, and professional societies. *HAPS Educator*, *23*(2), 396–419. https://doi.org/10.21692/haps.2019.012

Papinczak, T., Peterson, R., Babri, A. S., Ward, K., Kippers, V., & Wilkinson, D. (2012). Using student-generated questions for student-centred assessment. *Assessment & Evaluation in Higher Education*, *37*(4), 439–452. https://doi.org/10.1080/02602938.2010.538666

Pham, H., Court-Kowalski, S., Chan, H., & Devitt, P. (2023). Writing multiple choice questions—Has the student become the master? *Teaching and Learning in Medicine*, *35*(3), 356–367. https://doi.org/10.1080/10401334.2022.2050240

continued on next page

Pittenger, A. L., & Lounsbery, J. L. (2011). Student-generated questions to assess learning in an online orientation to pharmacy course. *American Journal of Pharmaceutical Education*, *75*(5), Article e94. https://doi.org/10.5688/ajpe75594

Qualters, D. M. (2016). Inclusive assessment: Equal or equitable? In M. Bart (Ed.), *Diversity and inclusion in the college classroom* (pp. 22-23). Faculty Focus.

Ralph, V. R., Scharlott, L. J., Schafer, A. G. L., Deshaye, M. Y., Becker, N. M., & Stowe, R. L. (2022). Advancing equity in STEM: The impact assessment design has on who succeeds in undergraduate introductory chemistry. *JACS Au*, *2*(8), 1869–1880. https://doi.org/10.1021/jacsau.2c00221

Saucier, D. A., Schiffer, A. A., & Jones, T. L. (2022). "Exams by you": Having students write and complete their own exams during the COVID-19 pandemic. *Teaching of Psychology*, *51*(2), 158-165. https://doi.org/10.1177/00986283221097617

Schullo-Feulner, A., Janke, K. K., Chapman, S. A., Stanke, L., Undeberg, M., Taylor, C., et al. (2014). Student-generated, faculty-vetted multiple-choice questions: Value, participant satisfaction, and workload. *Currents in Pharmacy Teaching and Learning*, *6*(1), 15–21. https://doi.org/10.1016/j.cptl.2013.09.019

Shah, M. P., Lin, B. R., Lee, M., Kahn, D., & Hernandez, E. (2019). Student-written multiple-choice questions – A practical and educational approach. *Medical Science Educator*, *29*(1), 41–43. https://doi.org/10.1007/s40670-018-00646-5

Touissi, Y., Hjiej, G., Hajjioui, A., Ibrahimi, A., & Fourtassi, M. (2022). Does developing multiple-choice questions improve medical students' learning? A systematic review. *Medical Education Online*, *27*(1), Article e2005505. https://doi.org/10.1080/10872981.2021.2005505

Towns, M. H. (2014). Guide to developing high-quality, reliable, and valid multiple-choice assessments. *Journal of Chemical Education, 91*(9), 1426-1431. https://doi.org/10.1021/ed500076x

Tractenberg, R. E., Gushta, M. M., Mulroney, S. E., & Weissinger, P. A. (2013). Multiple choice questions can be designed or revised to challenge learners' critical thinking. *Advances in Health Sciences Education*, *18*(5), 945–961. https://doi.org/10.1007/s10459-012-9434-4

Yu, F.-Y., & Wu, C.-P. (2016). The effects of an online student-constructed test strategy on knowledge construction. *Computers & Education*, *94*, 89–101. https://doi.org/10.1016/j.compedu.2015.11.005

continued on next page

Appendix 1. End of Course Response to Peer-Created Exams Survey

1) Identify as (please check all that apply):
    a) First generation college student (I am the first in my family to attend college)
    b) Low-income college student
    c) Underrepresented racial, ethnic, or cultural minority
    d) Female/feminine/woman
    e) None of the above
2) Compared to an instructor-created exam, I prefer peer-created exams.
    a) Strongly agree
    b) Agree
    c) Slightly agree
    d) Neutral
    e) Slightly disagree
    f) Disagree
    g) Strongly disagree
3) Compared to an instructor-created exam, peer-created exams made me think more critically about the material.
    a) Strongly agree
    b) Agree
    c) Slightly agree
    d) Neutral
    e) Slightly disagree
    f) Disagree
    g) Strongly disagree
4) Compared to an instructor-created exam, I felt more likely to succeed on peer-created exams.
    a) Strongly agree
    b) Agree
    c) Slightly agree
    d) Neutral
    e) Slightly disagree
    f) Disagree
    g) Strongly disagree
5) Compared to an instructor-created exam, peer-created exams improved my ability to learn the material.
    a) Strongly agree
    b) Agree
    c) Slightly agree
    d) Neutral
    e) Slightly disagree
    f) Disagree
    g) Strongly disagree

continued on next page

6) Compared to an instructor-created exam, peer-created exams improved my ability to retain the material.
    a) Strongly agree
    b) Agree
    c) Slightly agree
    d) Neutral
    e) Slightly disagree
    f) Disagree
    g) Strongly disagree
7) Compared to an instructor-created exam, peer-created exams increased my engagement in the testing process.
    a) Strongly agree
    b) Agree
    c) Slightly agree
    d) Neutral
    e) Slightly disagree
    f) Disagree
    g) Strongly disagree
8) Compared to an instructor-created exam, peer-created exams increased my satisfaction with the testing process.
    a) Strongly agree
    b) Agree
    c) Slightly agree
    d) Neutral
    e) Slightly disagree
    f) Disagree
    g) Strongly disagree

9) Describe any other way(s) in which the use of peer-created exams affected your LEARNING experience in this course.

10) Describe any other way(s) in which the use of peer-created exams affected your TESTING preparation and/or performance in this course.

11) Having created and taken peer-created exams this semester, would you want to have this type of exam option in other classes? Please explain why or why not.

continued on next page

Appendix 2. Contingency Table of Group Wise Comparisons for the Cognitive and Affective Factors

COGNITIVE Factor (*t* tests)

Group Identity	In Group Mean	In Group *SD*	Out Group Mean	Out Group *SD*	*t(38)*	*p*	Cohen's *d*
URM	4.63	11.32	4.58	1.21	0.11	0.91	0.036
Low Income	4.44	1.28	4.65	1.25	0.42	0.68	0.16
First Gen	4.63	1.16	4.59	1.28	-0.06	0.95	-0.025
Female/ Feminine	4.65	1.22	4.45	1.22	-0.44	0.67	-0.16

AFFECTIVE FACTOR (Mann-Whitney U tests)

Group Identity	In Group Mean	In Group *SD*	Out Group Mean	Out Group *SD*	*z(38)*	*p*	Cohen's *d*
URM	5.43	1.78	5.98	1.05	174	0.50	0.12
Low Income	5.07	1.92	5.92	1.23	99.5	0.19	0.29
First Gen	5.38	2.06	5.82	1.26	121	0.81	0.059
Female/ Feminine	6.00	1.04	5.50	2.11	112	0.23	0.25

continued on next page

Appendix 3. Table of Student Responses to Open-Ended Questions with Codes

Describe any other way(s) in which the use of peer-created exams affected your LEARNING experience in this course.	
N/A	NEU
Encouraged me to practice review before exam time	C_MC
Peer created exams helped me stay more engaged in the material and forced me to look back through my notes and review what I had written often.	C_MC, SE_EM
learning felt the same	NEU
I enjoyed taking peer-created exams as I was able to learn alongside with my peers and test each other on the information that we learned collectively. I also found the questions more relatable to the content we were learning, rather than very difficult and abstruse questions from instructor created exams.	C_CR, SE_EM, SE_PR
I knew if it was a question made by students, I should definitely know it as well, so I knew I could do well	SE_PR, SE_EM
The peer-created exam made me hone in on the fine details of the class material	C_CR, C_MC
The peer created exams test material over the content that was discussed in the videos which make you want to pay attention to the video and take good notes so that when it's time to study you will know the common terms and majority of the content that will be tested	C_CR, C_MC
They were based on activities we did in class - so I could relate peer-questions to experiences from class.	C_CR
I felt that peer-created exams were more reflective of what we focused on in activities in class, as well as the lectures.	C_CR
Forced me to think about the content and try to understand it more when I was writing the questions	C_MC
I would say I felt more engaged with the material I was learning by creating peer exams	SE_EM
I think that creating the exam questions everyday in class does help me learn the exam material better because I am able to go back to look at the content I have taken notes on to come up with questions.	C_MC
It made me more relaxed while learning.	SE_EM
forced me to review content to make questions	C_MC
Making the questions for peer created exams made it easier to understand what kind of material would be on the exams.	C_CR
I feel that some (not all) of the questions were made in a way that's easier for students to better rule out and remember answers.	C_MC, LD_DIFF
The peer created exams made me feel more confident in continuing to learn the material as I knew I had a greater chance of understanding the questions	C_MC, SE_EM
I knew that most students only watched the lecture videos and took the facts as is straight from the lecture videos, so I felt like my process of study primarily involved just memorizing the legitimate script of the videos. So, I felt as if I didn't really learn the material that well or understand the mechanisms behind why things were the way they were. Also, I felt as if the lecture videos kind of put together in a piece-wise fashion, making it difficult to understand the full picture idea.	C_CR, C_MC, SE_PR, LD_NEG
Peer created exams were helping to review my knowledge since being able to write questions myself was a good way to test my own knowledge.	C_MC
I feel like the questions varied in difficulty, not because of the content, but because of the awkward wording of some questions. The exams didn't feel consistent.	LD_NEG, LD_WORD
Using peer-created exams really helped with my confidence towards exams so I felt more keen on learning the content to not only make better questions but to make sure I knew the material going forward	C_MC, SE_EM
I think the instructor-made exams usually had questions that better represented the level at which we were to be learning the material while the peer-made exams were more general questions that emphasized the important/ main parts of a given topic.	C_CR, LD_DIFF
It helped me learn the material by also creating the questions for the exams, so that I could engage with the material in a different way rather than just reading the information passively.	C_MC
Creating the questions for the exams helped me retain information for the exam.	NEU

continued on next page

Describe any other way(s) in which the use of peer-created exams affected your TESTING preparation and/or performance in this course.	
N/A	NEU
helped me to review by looking at the questions I made throughout the unit	C_MC
The questions in the peer-created exams were way more relevant to what we actually went over in class and in the videos, so it was a lot easier and less stressful to prepare for the exams.	C_CR, SE_EM
Improved performance	
Allowed me to think about how I would formulate a test question while watching the videos so I was able to prepare well for the exams by thinking about how I would ask something.	C_MC
I knew students would make easier questions so I didn't have to try as hard :(	LD_DIFF
Peer-created exams helped me with my note taking and review for exams as I knew I did not have to review a lot of textbook jargon as most students would write questions pertaining to the content taught in lecture.	C_MC, C_CR
It did not affect my studying, I would have studied the same for both exams.	C_MC, NEU
The peer created exam helped me reinforce the skills of rote memorization to prepare for the test	C_MC
Peer created exams are worded better in my opinion compared to the instructor created exams. It doesn't really try to trick except for the occasional 1 or 2. For the most part it's a content based question from the video which assess how well you have grasped the knowledge from the videos.	C_CR, LD_DIFF
The instructor created exams seemed to come from the textbook generally while the peer exams came from the videos/class activities generally, so the peer exams made more sense because I never touched the textbook.	C_CR
Peer created questions were easier to understand and I felt they touched on the content that was discussed more. Sometimes the instructor ones asked questions I felt we didn't go over	C_CR, LD_DIFF
I think the peer created exams created less testing anxiety because it was made by my peers in the class	SE_EM
Using peer created exams impacted testing preparation because I was able to practice making exam questions in order to prepare as a strategy.	C_MC
It provided me with an additional sense of security during the studying process.	SE_EM
peers made it easier	LD_DIFF
I noticed that peer questions are sometimes hyper specific and other times way too general, thus, the difficulty level of the average peer question is all over the place as compared to instructor questions, which are usually consistent in their difficulty.	LD_WORD
I knew that my peers would use material provided in the lecture videos for their exam questions, so that made it easier to know what to study before the exams.	C_CR
I do not feel as if I did much better on any particular type of exam so I feel ambivalent towards the peer exam	NEU
I feel that with the peer created exams I was more worried about what was on my notes rather than studying material outside of class.	C_CR
I sometimes chose the right answer by thinking about what answer was most close to the information as it was spoken as in the video.	C_CR
I feel like I did much better on the instructor created exams vs the peer created.	LD_NEG
As mentioned above, it helped with testing confidence going into exams and I felt the questions were more likely to hit content at the level we had learned it with a few curveballs here and there.	C_CR, SE_EM
It made reviewing for the test easier because I felt that I would focus on the same or simulate things as my peers. Often on the instructor exams, quest sometimes pooped up on topics I felt weren't highlighted much in the lecture, thus I was more likely to miss points on those.	C_MC, C_CR, SE_PR
I think I prepared just as much for both exams every time. I felt at more ease knowing I had two chances though.	SE_EM, NEU
I was able to use the questions I created as somewhat of a study-guide to remember important topics.	C_MC

continued on next page

Having created and taken peer-created exams this semester, would you want to have this type of exam option in other classes? Please explain why or why not.	
Yes, writing questions for myself helped me to be able to study much more effectively	C_MC
Yes, I felt the peer exams were a better reflection of what I was actually learning rather than what I was "expected" to figure out.	C_CR
Yes I would, as I think creating exams also helps with the process of learning the material.	C_MC
Yes, because it allows me to study sooner than I would and feel more confident going in to the exam	C_MC, SE_EM
Yes I would because the peer created exams would go over what most people actually studied.	SE_PR
Definitely	
Yes, I think I perform better with peer created exams and it facilitates my desire to study and learn the material.	SE_EM
in stem classes yes	
Yes, I would love to have this type of exam option in other courses. However, I would appreciate if we could see all the peer created questions that were created to review the more difficult questions.	
Yes, I feel like it's fair and it makes me feel more comfortable when taking an exam.	SE_EM
It depends on the contact of the class. In classes that are more application based, I would advise no. However, if there is a situation where I need to memorize material on the fly a peer created exam is more than adequate	LD_APP
Yes, having peer created exams has been helpful and a more creative way of test taking. Because we all have similar ways of understanding the knowledge and thought processes the questions will be questions we will want to engage with.	SE_PR
Yes, this way you know what peers have retained or learned during the course and if you don't know the material on a peer exam you can't blame the instructor.	SE_PR
Yes, but probably not in classes that are very heavy in detail and calculations / problems (i.e. physics, organic chemistry); these classes are less memorization based, so the questions are a lot less straightforward. The questions are also easy to convolute and need to be worded very carefully. I felt that sometimes the peer-created exams in this class were even worded less well/clearly.	LD_APP, LD_WORD
Yes absolutely! I think this exam format helps students review material while learning it, increasing long term retention. I also think it decreases the amount of testing anxiety one may have because you know the people who made the exam are your peers who are also learning the material.	C_MC, SE_EM, SE_PR
Yes, I feel the peer exam was more representative of what is known as a student learning the material for the first time than an instructor who has been learning the material and teaching the class for years.	SE_PR
I would, as it definitely opens up more opportunities to prove my knowledge of the content, beyond just instructor-generated questions.	C_CR, C_MC
I think so, for anything that is more content and memorization based.	LD_APP
Yes, but only in classes in which material is not application heavy. I believe these tests are optimal for memorization heavy material since questions are made in a short amount of time, encouraging less application based questions and promoting simple memorization questions.	LD_APP
yes and no; it is more work for us but it is a little bit easier	NEU
I would be interested in piloting this option in other classes, because I feel that while this model may succeed in some of the more memorization-based classes, it would be a complete nightmare in classes that are mainly application-based.	LD_APP
I would not want to have this option in other classes because I feel like there is a reason why instructors create their own exam questions, because they know what topics and phrases will best test our knowledge of the material.	LD_NEG
I don't think it'd be something I'd desire unless I feel the teacher wasn't creating exams that test the material well or test what is important. In some classes I feel like that is the case but for most classes that would offer this as an option I don't think the testing would even be of concern as these types of teachers are more caring and receptive	LD_NEG

continued on next page

Overall yes, I think the peer-created exams were better for keeping my engagement and interest in the class material. I would only hope for questions to be better looked over for mistakes	SE_EM, LD_WORD
No. I think we can get lazy with making questions and it doesn't really help with learning the material. In addition, without someone actually reviewing the questions, I feel like the questions themselves and answer choices can be misleading.	LD_NEG, LD_WORD
Yes I would. I think peer created exams give more perspective into what the students as a whole feel like they are learning. Sometimes professor made exams can make students feel like they were missing parts of the information regardless of how much they studied, so I think student made exams more accurately reflect the knowledge of students as a whole.	SE_PR
No; In my opinion, the instructor created exams are more consistent in the wording, content, and the way the content is tested.	LD_WORD
Yes I'd love to, but maybe some courses do require instructor questions here and there. Classes like orgo and synthesis units or many comp classes might not have students prepared to create exam questions robust enough for other students!	LD_DIFF
I think it depends on the class. With class that are memorization based and have multiple choice questions, I think this is pretty great. However if I'm in a class where long answer questions are more common, or a question can have multiple answers, I rather instructor creator exams. In general I'm also more used to instructor exams and do prefer the fact that the formatting is usually more standardized. However I feel like the student ones appeal to my logic more.	LD_APP, LD_WORD
Sure, but only if instructor-made is also an option like it was here. I often found that Dr. (BLIND)'s exams had more complex questions--> but they couldn't be misunderstood. They were straight to the point and made it clear what they were wanting from you. This was a cool research. Thank you for doing it and letting us be a part of it!	LD_WORD
Yes, I think it better reflected the information that we learned in class and the depth we learned about it, and it also created a way for us to actively learn the material as we also created the questions.	C_CR, C_MC
I think so but I would personally want more guidance about how to create questions	LD_NEG
I would not because sometimes the details on there do not seem to improve my retention as a learner and instead seem to just be there with not as much of a purpose.	LD_NEG, C_MC

Back to TOC

HAPS EDUCATOR
Journal of the Human Anatomy and Physiology Society

EDUCATIONAL RESEARCH

Volume 29, Issue 2 | August, 2025
https://doi.org/10.21692/haps.2025.011

Anatomical Dissection of Whole-Body Donors Improves Anatomy Lecture Performance in Undergraduate Students

Katherine Baker, BS[1], Jim Thomas, MS[1], Wei Fang, PhD[2], Miriam Leary, PhD[1]

[1]Division of Exercise Physiology, School of Medicine, West Virginia University, Morgantown, WV, USA

[2]West Virginia Clinical and Translational Science Institute, WVU Health Sciences Center Erma Byrd Biomedical Research Center, Morgantown, WV, USA

Corresponding author: Miriam.Leary@hsc.wvu.edu

Abstract

Proficiency in anatomy is pivotal in the education of future healthcare workers and ensures safe medical practice. Traditional anatomy lectures can be supplemented by a laboratory component, such as human whole-body donor dissection. The primary purpose of this study was to determine if students who take a dissection-based laboratory concurrently with lecture improve lecture grades, compared with lecture alone. Eight consecutive semesters of ordinal course grades for the lecture and laboratory sections, as well as students' GPA, were used. Additionally, student responses to randomly selected primary and tertiary questions were recorded, and the percentage correct for each group (Lecture versus Lecture + Lab) was calculated. There was a significant difference in the likelihood of higher grades between the two groups ($p<0.0001$). The odds of higher grades in Lecture + Lab group were 4.013 times the odds of higher grades in the Lecture group (adjusted odds ratio (AOR): 4.013, 95% CI: 2.8356, 5.6794). In addition, GPA was significantly associated with lecture grades ($p<0.0001$). Students that took lab concurrently were more likely to answer primary (Lecture + Lab: 85% ± 20% vs. Lecture: 72% ± 23%; $p<0.0001$) and tertiary (Lecture + Lab: 71% ± 28% vs. Lecture: 55% ± 30%; $p<0.0001$) questions correctly than students that took lecture alone. Anatomical dissection of whole-body donors improves performance in didactic anatomy examinations and should be recommended to ensure proficiency in anatomical education. https://doi.org/10.21692/haps.2025.011

Key words: anatomy education, anatomical dissection, higher-order thinking, critical thinking

Introduction

Anatomy, the intricate study of the human body, sets the foundation of undergraduate education for future health care professionals. It is regarded as a core discipline of health science (McCumber et al., 2021) and is a prerequisite for many health-related programs (Asante et al., 2021). Proficiency in anatomy is essential for developing clinical skills in healthcare, particularly in surgical fields (Abdellatif, 2020). Concerns that insufficient anatomical knowledge may lead to increased morbidity and mortality has led to the inclusion of a dissection experience in some surgical residencies (Memon, 2018). Beyond surgery, support for dissection in the development of practical skills was supported across a surveyed group of medical professionals at all levels (Ahmed et al., 2010). Surveyed physical therapists indicated that they felt anatomy was important for safe clinical practice (Rompolski et al., 2023). Moreover, human gross anatomy provides an opportunity for bioethical and humanistic reflection, through which instructors strive to instill respect, compassion, and appreciation within enrolled students – important traits in a future health care professional (Weeks et al., 1995).

To achieve this proficiency, anatomy courses are typically instructed as a didactic lecture, with or without a practical laboratory component. The type of practical laboratory component varies between institutions and their respective programs. Laboratory types commonly include, but are not limited to, whole-body donor dissection, whole-body donor prosection, surface anatomy and palpation, osteology, virtual dissection/3-D software, plastinated models, body

continued on next page

painting, clay modeling, and virtual reality (Rompolski, 2023; Rompolski et al., 2023). Whole-body donor dissection, specifically, is a form of experiential learning conducted on a human body donated to science for purposes of education and research (Korf et al., 2008). The dissection experience consists of student groups exploring anatomical regions of the body (i.e., back, upper extremity, etc.) to reveal and study muscles, nerves, and blood vessels, among other tissues. Compared with a non-dissection laboratory, dissection has several benefits, as it provides improved spatial understanding of the body (Abdellatif, 2020) including the compartmentalization of the body through fascia and connective tissues (Korf et al., 2008), an appreciation for human variation (Abdellatif, 2020; Huynh et al., 2021), increases motivation and spatial relationships (Nwachukwu et al., 2015; Thompson & Marshall, 2020; Asante et al., 2021; Huynh et al., 2021), and promotes professionalism (Nwachukwu et al., 2015).

Modality has been an area of debate over the last several years. One study found that medical students, trainees, and specialists agreed in a traditional Likert survey statement that "dissection is better than prosection," and was the most recommended learning tool across the groups (Ahmed et al., 2010). Similarly, a 2021 qualitative study reported that medical students found dissection to be interesting, the better way to learn, assists in retention, and should not be replaced by other forms of learning (Asante et al., 2021). Other studies have shown similar support for dissection (Abdellatif, 2020; Huynh et al., 2021; Korf et al., 2008; Nwachukwu et al., 2015; Pizzimenti et al., 2016; Phitayakorn et al., 2024; Thompson & Marshall, 2020). In contrast, recent literature has shown support for alternate forms of learning anatomy. A 2015 study argued that dissection is valuable but not essential (Patel et al., 2015). Additionally, one report argued that students retained benefits of human body donor laboratory - clinical reasoning, teamwork, and empathy, in a course that operates without the use of whole-body donors (Rompolski, 2023). Other studies have shown support for the use of alternate methods in anatomy instruction (Lombardi et al., 2014; Memon, 2018; Wilson et al., 2018).

While modality has been a debate, the use of practical laboratory component has steadily declined (McBride & Drake, 2018). Dissection, in particular, has declined in recent years due to limited availability of whole-body donors, decreased time allocated for lab, improvements in preservation, increased use of plastinated specimens, and technology (Abdellatif, 2020), as well as expense of donor body acquisition and preservation, time and skill required to dissect, health risks related to preservation agents, the COVID-19 pandemic, and increasing scientific information to cover in medical education (Phitayakorn et al., 2024). As of 2019, just over half of US medical schools incorporate donor body dissection in their curriculum (Schroeder et al., 2020).

Amid decline in use, qualitative views and quantitative measures on the benefits of anatomy lab remains largely positive (Ahmed et al., 2010). Increased time spent in dissection anatomy laboratory experiences increased laboratory and written examination scores for first year medical students, with the most significant effect for students in the lower-quartile (Pizzimenti et al., 2016). Similarly, undergraduate students who participated in prosection-based anatomy laboratory out performed students in lecture alone on written anatomical exams (Anderson et al., 2024). These studies, as well as several others in the literature support the inclusion of some form of laboratory component in anatomy courses given the positive impact on student performance. However, to our knowledge, no studies have previously specifically examined the difference in lecture examination performance between students enrolled in both a dissection-based laboratory and lecture, compared to students who take lecture alone.

Performance in anatomy is primarily assessed through written examination. Written exams conducted in anatomy lecture or lab vary in style but are commonly in multiple choice question (MCQ) format. In one study, MCQs in an anatomy course were categorized as primary, secondary, or tertiary, and sorted as such based on the level of comprehension of spatial and 3-D relationships required to answer the question (Anderson et al., 2024). Based on these constraints, primary questions were denoted lower-order knowledge type, such as *"what is the name"* or *"identify the structure."* Secondary questions required students to connect an anatomical structure with its function or blood/nerve supply, while tertiary questions required connection to both in order to determine outcome in the case of injury (Anderson et al., 2024). An example of such would be, "*A 32-year-old cyclist falls and sustains a midshaft fracture of the humerus. Following the injury, he is unable to extend his wrist and fingers, and experiences numbness over the dorsum of his hand. What nerve is likely injured, and what muscle group is primarily affected?"* Answering higher-order (secondary and tertiary) questions correctly suggests deeper comprehension, and was observed in students taking a prosection lab with anatomy lecture compared to the lecture alone (Anderson et al., 2024).

The purpose of the present study was to determine if undergraduate students taking a dissection-based anatomy laboratory concurrently with lecture earn higher final lecture grades and answer lecture-based higher order questions correctly more often than those taking lecture alone. We hypothesize that students participating in dissection will achieve higher overall course grades and answer more higher order questions correctly compared with students taking anatomy lecture alone.

continued on next page

Methods

Human subjects research exemption was granted by the Internal Review Board (IRB) at West Virginia University (WVU) (IRB ID: 2401898785), an R1 research, land-grant institution in Appalachia. This research was conducted in the Exercise Physiology program, a large (700+ student) undergraduate pre-health program. Within this program, Anatomy for Exercise Physiology, a three-credit required human anatomy lecture course, provides an in-depth, integrative understanding of human anatomy with applications toward exercise in both healthy and diseased populations. The course runs concurrently with an optional Gross Anatomy Laboratory, a whole-body donor dissection laboratory section offered as an elective.

Course Design

The lecture course is taught in 80-minute traditional didactic lectures twice a week for 16-weeks. The course format is a regional approach to anatomy, including regions of the back, upper extremities, lower extremities, thorax, abdomen, as well as the head and neck. Within each region, the identification of human anatomical structures of the musculoskeletal, nervous, and circulatory systems are central course objectives. Upon completion of unit material, information is reinforced by a standard 52 question multiple choice (MC) unit examination (exam grades are calculated as number correctly answered/50 questions). MC questions are pulled from a question bank and include questions ranging in difficulty from primary to tertiary questions. Additional learning resources for the lecture include an optional textbook, optional 3-D human anatomy application, homework assignments, quizzes, and the option to enroll in the laboratory section as an elective. Any student that does not take the laboratory section concurrently is invited to attend the lab section for observation, but not dissection.

The laboratory section is a dissection-based anatomy lab, also taught 80-minutes twice a week for 16 weeks, and the material covered in lab runs concurrently with the material covered in lecture. Students work in groups of four on their respective whole-body donor. The members of each group work together to meticulously dissect the same anatomical regions covered in lecture, in the same sequence.

Modifications of the course design were made for Spring 2020 semester cohort amid the COVID-19 pandemic. All students enrolled in anatomy at that time took the first eight weeks of the course in person as normal, learning the back and upper extremity units. West Virginia University then moved all classes online for the remaining 8 weeks of the semester, and thus the lower extremity, abdomen/thorax/pelvis, and head/brain/brain units were instructed online using voice over PowerPoints and online exams. For the lab component, there were a few videos of dissection available and online exams required students to identify structures in still shot images.

The lecture had been taught for several years before the whole-body donor lab section was offered. In the first few years of the dissection laboratory, there were not enough seats to permit all students in the program to enroll in the dissection laboratory, so placement was prioritized by GPA. Students that had less than a 3.0 GPA were not eligible for the dissection laboratory. This restriction was removed for the Spring 2023 cohort when more sections were offered so all students that were interested in enrolling in the dissection laboratory were guaranteed a seat. Therefore, statistical analysis controlled for GPA.

Outcomes

The primary outcomes of this study were overall lecture grades, unit exam grades and performance on types of questions (primary vs. tertiary). Ordinal course grades, with A > B > C > D, for the lecture and laboratory sections, together with students' GPA, were pulled from the WVU Registrar's Office for eight consecutive semesters (Fall 2019-Spring 2023). Student identification was anonymous as data was shared with an ID number instead of names. To determine whether the participation in the dissection laboratory improved students' ability to answer higher order questions correctly, unit exam questions and student responses were pulled from Fall 2023 records. Two to five primary and tertiary questions from each unit exam were randomly selected for comparisons between groups.

Statistical Analysis

The percent of students achieving each ordinal grade was determined and separated by students' GPA, as displayed in Table 1. Because GPA was a prerequisite requirement for lab seven of eight semesters, most data points in the 2.5 – 3.0 GPA group were Lecture only. Therefore, for comparisons, any data points that existed for Lecture + Lab were matched based on GPA to Lecture only data points. A regression analysis using a proportional odds model was conducted to determine if there was a significant difference in the likelihood of higher lecture grades between the two groups (Lecture vs. Lecture + Lab) controlling for GPA. Student responses to selected primary and tertiary questions were recorded in an Excel spreadsheet as "Yes" or "No," depending on whether the student answered the randomly-selection question correctly. The percent correct ± SD for each group were reported and compared using Chi-squared tests. All statistical analyses were performed using R 4.4.1 (R Core Team, Vienna Austria) and the "rms" (Version 6.8-1) package.

continued on next page

Results

The original data collected from the registrar's office included 713 students, including 348 that took Lecture alone and 365 that enrolled in Lecture + Lab. The percent of students achieving each ordinal grade are presented in Table 1.

Lecture Only			Lecture + Lab	
N = 348			N = 365	
GPA	3.302 ± 0.371		GPA	3.582 ± 0.304
A	28% (96)		A	69% (252)
B	36% (127)		B	24% (86)
C	19% (67)		C	6% (22)
DFW	17% (58)		DFW	1% (5)
3.5 – 4.0 GPA				
N = 103			N = 229	
GPA	3.758 ± 0.158		GPA	3.777 ± 0.154
A	63% (65)		A	81% (185)
B	24% (25)		B	16% (37)
C	9% (9)		C	2% (4)
DFW	4% (4)		DFW	1% (3)
3.0 – 3.5 GPA				
N = 169			N = 122	
GPA	3.249 ± 0.137		GPA	3.30 ± 0.137
A	14% (24)		A	52% (64)
B	50% (84)		B	35% (43)
C	19% (32)		C	11% (13)
DFW	17% (29)		DFW	2% (2)
2.5 – 3.0 GPA				
N = 14			N = 14	
GPA	2.858 ± 0.126		GPA	2.856 ± 0.130
A	0% (0)		A	21% (3)
B	36% (5)		B	43% (6)
C	50% (7)		C	36% (5)
DFW	14% (2)		DFW	0% (0)

Because GPA was a prerequisite requirement for lab seven of eight semesters, most data points in the 2.5 – 3.0 GPA group were lecture only. Therefore, for comparisons, any data points that existed for Lecture + Lab were matched based on GPA to Lecture only data points.

Table 1. *Ordinal grade achievement by GPA category*

There was a significant difference in the likelihood of higher grades between the two groups ($p<0.0001$). In particular, the odds of higher grades in Lecture + Lab group were 4.013 times the odds of higher grades in the Lecture group (adjusted odds ratio (AOR): 4.013, 95% CI: 2.8356, 5.6794). In addition, GPA was significantly associated with the lecture grades ($p<0.0001$). It is of note that GPA was non-linearly associated with lecture grades, as shown in Figure 1, when GPA was greater than 3.75, the likelihood of higher lecture grades increased more rapidly than when GPA was less than 3.75.

continued on next page

Question data analysis was conducted on 113 students enrolled in anatomy in fall 2023. Students that took lab concurrently were more likely to answer primary (Lecture + Lab: 85% ± 20% vs. Lecture: 72% ± 23%; p<0.0001) and tertiary (Lecture + Lab: 71% ± 28% vs. Lecture: 55% ± 30%; p<0.0001) questions correctly than students that took lecture alone. These results are demonstrated in Figure 2.

Discussion

The purpose of this study was to determine if students taking a dissection-based human gross anatomy laboratory course concurrently with lecture earn higher final grades compared to those taking lecture alone. The results support the hypothesis that those participating in dissection achieve higher overall course grades than those students taking lecture without lab. Further, students taking lecture and laboratory sections concurrently were more likely to answer primary and tertiary questions correctly. These findings support that whole-body donor dissection improves performance on anatomy lecture written exams and adds to the body of literature in support for inclusion of a practical anatomy laboratory component when learning anatomy.

In this study, students earned higher grades when concurrently enrolled in anatomy lecture and lab which is consistent with former research examining the impact of concurrent lab and lecture, despite different course design (Pizzimenti et al., 2016; Anderson et al., 2024). For instance, Anderson et al. (2024) found that students enrolled in a prosection laboratory outperformed their non-lab counterparts by a full letter grade (Anderson et al., 2024). Pizzimenti et al. (2016) incorporated a rotating laboratory experience, meaning not every laboratory student was actively dissecting each class, but instead some students dissected while others completed quizzes and other activities or watched the dissection for "peer-based learning". Pizzimenti et al. (2016) found a significant increase in examination performance with greater time spent dissecting within the anatomy laboratory (Pizzimenti et al., 2016). Indeed, many scholars argue that time spent in the laboratory is the most important factor toward success in an anatomy course. Both greater time spent in the laboratory and better performance on the act of dissection itself was found to be associated with better overall

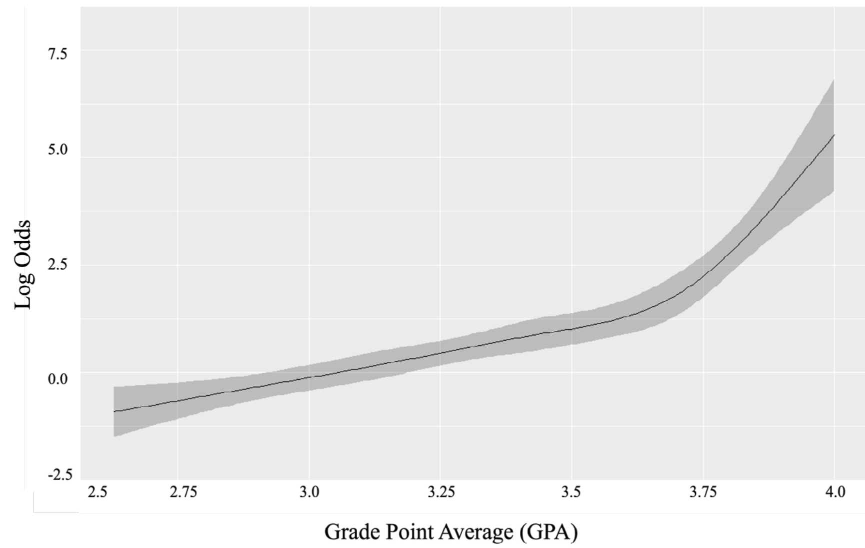


Figure 1. *Association between lecture grades and GPA is demonstrated with a smoothing curve and 95% confidence interval in grey shading, based on the Loess Smoothing Method.*

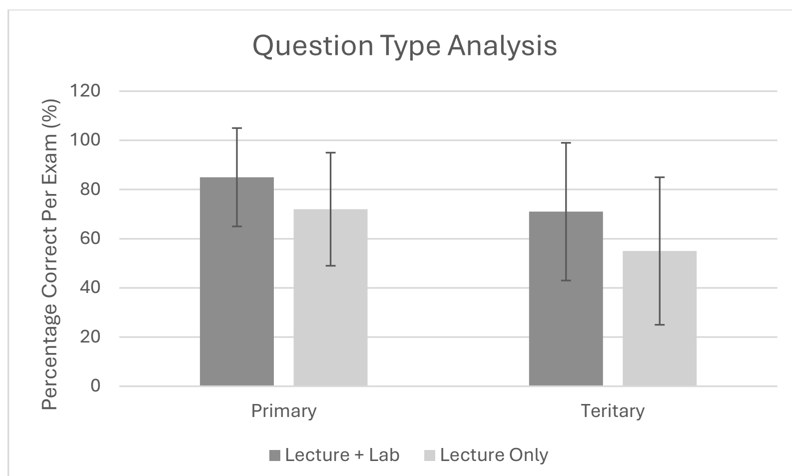


Figure 2. *Performance on primary vs tertiary question types across lab and non-lab students.*

final assessment performance in previous reports (Nwachukwu et al., 2015; Abdellatif, 2020). In the present study, the material covered in lecture and lab ran concurrently; thus, those students enrolled in both lecture and laboratory sections received twice the amount of time on the same information compared with students taking lecture alone. The current literature, in agreement with the results of the present study, support that anatomy laboratory courses are beneficial for mastering anatomical understanding and that the time spent in lab, regardless of laboratory type (e.g., dissection

continued on next page

vs. prosection) improves overall performance (Lombardi et al., 2014; Pizzimenti et al., 2016; Wilson et al., 2018; Anderson et al., 2024). Other researchers argue that what matters most when learning anatomy is the experience in the lab itself (Wilson et al., 2018; Anderson et al., 2024).

The present study found students that took lecture and lab concurrently performed better on primary and tertiary questions, in alignment with previous findings (Anderson et al., 2024). Tertiary questions are more complex and require a deeper level of understanding, suggesting that students in anatomy laboratories develop higher order critical thinking and application skills. Higher order critical thinking and application skills are crucial for pre-health students as they enable students to analyze complex medical situations, make informed decisions, and apply theoretical knowledge to real-world scenarios. Critical thinking helps students evaluate evidence, recognize patterns, and consider multiple perspectives, while application skills ensure that students can translate classroom learning into practical skills, which is vital in clinical settings where theoretical knowledge must be applied to patient care. The development of these skills in undergraduate courses can benefit long term academic performance for pre-health students. Previously, Thomas et al. (2021) found enrollment in an undergraduate anatomy course was associated with better performance on anatomy-related assessments in medical school (Thomas et al., 2021).

Previous work showed greater time spent participating in dissection increased laboratory practical examination scores for many students, with the most significant effect for students in the lower-quartile (Pizzimenti et al., 2016). This study suggests there is correlation between study time and academic achievement that might preferentially benefit lower achieving students. The present study somewhat contradicts this finding as GPA was non-linearly associated with lecture grade: the likelihood of higher lecture grades increased more rapidly than when GPA > 3.75. Unfortunately, seven of eight semesters used in the present dataset had a GPA restriction set for the whole-body donor dissection elective, which likely limits the analyses. Therefore, future research should more fully explore the relationship between dissection and anatomy academic performance among lower achieving students.

The results support the hypothesis that participation in dissection can achieve higher overall course grades when compared to students taking lecture without a laboratory component. These data confirm the benefit of undergraduate students enrolling in anatomy lecture and laboratory concurrently. These findings may influence program development and future policies and may positively impact undergraduate and graduate participation in dissection internationally. As noted via the tertiary question analysis, students in dissection develop and use more critical thinking and application skills. Building capacity in these areas equips students with the ability to handle challenging undergraduate coursework as well as the rigors of professional and graduate programs in their future. Finally, participation in dissection allows for a richer learning experience, enabling students to obtain hands-on practice in collaboration with their peers. This study thus provides rationale arguing against the previously mentioned trend in decline of whole-body donor dissection (McBride & Drake, 2018; Abdellatif, 2020; Schroeder et al., 2020; Phitayakorn et al., 2024).

Despite these strengths, this study is not without limitations. Conducted at a single university in Appalachia, the sample included was relatively in terms of race/ethnicity, socioeconomic status, and geographic background, with most participants identifying as White/Caucasian. Future studies should aim to recruit more diverse samples, considering additional factors such as socioeconomic background, racial/ethnic composition, and rural versus urban residence. The question type analysis was limited to a single semester, and thus further examination on how laboratory components can promote higher order thinking skill development is warranted in future research.

Additionally, post-GPA requirement data is limited to one semester. Further the benefit of dissection on anatomy lecture examination performance in lower-quartile students could be elucidated in future studies by comparing student quartile performance, similar to previous work (Pizzimenti et al., 2016). Future research should also aim to include student perceptions of participation in laboratory, as former qualitative research has reported positive findings (Edmund Atta Asante et al., 2021; Huynh et al., 2021). Notably, data was included from Spring 2020, which as previously described was modified amid the COVID-19 pandemic and did not impact data significantly, despite lack of an in-person laboratory experience for half of the semester. As previously mentioned, there are a wide variety of resources available to students (textbooks, applications, etc.), and there is no way to control for which resources students elected to utilize. Finally, the primary limitation is that the Lecture + Lab cohort receives twice the amount of time on the same information compared with students taking lecture alone, serving as a large confounder. Given that both former studies utilizing a different form of laboratory component similar positive effects on grades to dissection (Lombardi et al., 2014; Patel et al., 2015; K. Rompolski, 2023; Wilson et al., 2018) and the lack of comparison of another type of laboratory activity in the present study and its impact on student performance, earning higher grades in this study may be better attributed to greater exposure of the course material including a hands-on experience rather than directly to dissection itself. This does however add to the body of knowledge on the benefits of a laboratory component in anatomy instruction.

Despite these limitations, findings of the study were overall positive regarding the dissection of whole-body donors and its impact on performance in anatomy, and adds to the body of knowledge on the benefits of a laboratory component in anatomy instruction.

continued on next page

Conclusion

Anatomy is a foundational subject in undergraduate pre-health education and a critical prerequisite for future health care providers as proficiency in the subject ensures safe practice. Compared with students that take lecture alone, this study demonstrates that undergraduate students who take a dissection-based anatomy laboratory concurrently with lecture are four times more likely to earn higher grades in lecture. Additionally, these students are more likely to answer tertiary questions correctly, suggesting better understanding and development of higher order thinking skills. Anatomical dissection of whole-body donors improves academic performance and mastery of anatomical understanding and should be recommended to ensure proficiency in anatomical education.

About the Authors

Katherine Baker graduated with a BS in Exercise Physiology from WVU in 2024 and is currently pursuing an MD from WVU. Jim Thomas, MS, is an associate professor within the Division of Exercise Physiology at WVU with expertise in teaching human anatomy lecture and dissection laboratory courses. Wei Fang has a PhD in statistics and works as a biostatistician at the West Virginia Clinical and Translational Science Institute (WVCTSI). Miriam Leary, PhD, is an associate professor and the assistant chair in the Division of Exercise Physiology at WVU. Her research focuses on evidence-based teaching and learning practices with an emphasis on student success and retention of at-risk populations.

Acknowledgments

Research reported in this publication was supported by the National Institute of General Medical Sciences of the National Institutes of Health under Award Number 5U54GM104942-08. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

The authors thank Dr. Krista Rompolski, Dr. David Rasicci, Dr. H. Wayne Lambert, and Dr. Ashley Petrone for their valuable contributions to the manuscript.

Literature Cited

Abdellatif, H. (2020). Time Spent in Practicing Dissection Correlated with Improvement in Anatomical Knowledge of Students: Experimental Study in an Integrated Learning Program. Cureus. https://doi.org/10.7759/cureus.7558

Ahmed, K., Rowland, S., Patel, V., Khan, R. S., Ashrafian, H., Davies, D. C., et al. (2010). Is the structure of anatomy curriculum adequate for safe medical practice? The Surgeon, 8(6), 318–324. https://doi.org/10.1016/j.surge.2010.06.005

Anderson, H., Weil, J. A., Tucker, R. P., & Gross, D. S. (2024). Impact of gross anatomy laboratory on student written examination performance: A 3-year study of a large-enrollment undergraduate anatomy course. Anatomical Sciences Education, 17(1), 114–127. https://doi.org/10.1002/ase.2327

Asante, E. A., Maalman, R. S., Ali, M. A., Donkor, Y. A., & Korpisah, J. K. (2021). Perception and Attitude of Medical Students towards Cadaveric Dissection in Anatomical Science Education. Ethiopian Journal of Health Sciences, 31(4). https://doi.org/10.4314/ejhs.v31i4.22

Huynh, N., Burgess, A., Wing, L., & Mellis, C. (2021). Anatomy by Whole Body Dissection as an Elective: Student Outcomes. Journal of Surgical Education, 78(2), 492–501. https://doi.org/10.1016/j.jsurg.2020.07.041

Korf, H-W., Wicht, H., Snipes, R. L., Timmermans, J-P., Paulsen, F., Rune, G., et al. (2008). The dissection course – necessary and indispensable for teaching anatomy to medical students. Annals of Anatomy - Anatomischer Anzeiger, 190(1), 16–22. https://doi.org/10.1016/j.aanat.2007.10.001

Lombardi, S. A., Hicks, R. E., Thompson, K. V., & Marbach-Ad, G. (2014). Are all hands-on activities equally effective? Effect of using plastic models, organ dissections, and virtual dissections on student learning and perceptions. Advances in Physiology Education, 38(1), 80–86. https://doi.org/10.1152/advan.00154.2012

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

McCumber, T. L., Latacha, K. S., & Lomneth, C. S. (2021). The state of anatomical donation programs amidst the SARS-CoV-2 (Covid-19) pandemic. Clinical Anatomy (New York, N.Y.), 34(6), 961–965. https://doi.org/10.1002/ca.23760

Memon, I. (2018). Cadaver Dissection Is Obsolete in Medical Training! A Misinterpreted Notion. Medical Principles and Practice, 27(3), 201–210. https://doi.org/10.1159/000488320

Nwachukwu, C., Lachman, N., & Pawlina, W. (2015). Evaluating dissection in the gross anatomy course: Correlation between quality of laboratory dissection and students outcomes. Anatomical Sciences Education, 8(1), 45–52. https://doi.org/10.1002/ase.1458

Patel, S. B., Mauro, D., Fenn, J., Sharkey, D. R., & Jones, C. (2015). Is dissection the only way to learn anatomy? Thoughts from students at a non-dissecting based medical school. Perspectives on Medical Education, 4(5), 259–260. https://doi.org/10.1007/S40037-015-0206-8

Phitayakorn, R., Stearns, D. & Esther K. (2024). Is Cadaver Dissection Still Necessary in Surgical Education? American College of Surgeons. https://www.facs.org/for-medical-professionals/news-publications/journals/rise/articles/cadaver-dissection-still-necessary/

continued on next page

Pizzimenti, M. A., Pantazis, N., Sandra, A., Hoffmann, D. S., Lenoch, S., & Ferguson, K. J. (2016). Dissection and dissection-associated required experiences improve student performance in gross anatomy: Differences among quartiles. Anatomical Sciences Education, 9(3), 238–246. https://doi.org/10.1002/ase.1574

Rompolski, K. (2023). Confessions of a converted anatomist: Teaching without anatomical donors. Advances in Physiology Education, 47(3), 484–490. https://doi.org/10.1152/advan.00004.2023

Rompolski, K. L., Fojas, C. L., & Taylor, M. A. (2023). How do practicing physical therapists perceive anatomy education? Anatomical Sciences Education, 16(5), 979–988. https://doi.org/10.1002/ase.2277

Schroeder, T., Elkheir, S., Farrokhyar, F., Allard-Coutu, A., & Kahnamoui, K. (2020). Does exposure to anatomy education in medical school affect surgical residency applications? An analysis of Canadian residency match data. Canadian Journal of Surgery, 63(2), E129–E134. https://doi.org/10.1503/cjs.019218

Thomas, R., Yancey, T., Skidmore, C., Ferrin, N., Zapata, I., Williams, J., et al. (2021). The Effects of Pre-medical Anatomy and Clinical Experiences on Medical School Anatomy-Related Academic Performance. Medical Science Educator, 31(6), 1839–1849. https://doi.org/10.1007/s40670-021-01372-1

Thompson, A. R., & Marshall, A. M. (2020). Participation in Dissection Affects Student Performance on Gross Anatomy Practical and Written Examinations: Results of a Four-Year Comparative Study. Anatomical Sciences Education, 13(1), 30–36. https://doi.org/10.1002/ase.1859

Weeks, S. E., Harris, E. E., & Kinzey, W. G. (1995). Human gross anatomy: A crucial time to encourage respect and compassion in students. Clinical Anatomy, 8(1), 69–79. https://doi.org/10.1002/ca.980080113

Wilson, A. B., Miller, C. H., Klein, B. A., Taylor, M. A., Goodwin, M., Boyle, E. K., et al. (2018). A meta-analysis of anatomy laboratory pedagogies. Clinical Anatomy, 31(1), 122–133. https://doi.org/10.1002/ca.22934

Back to TOC

HAPS EDUCATOR
Journal of the Human Anatomy and Physiology Society
CURRENT TOPICS
Volume 29, Issue 2 | August, 2025
https://doi.org/10.21692/haps.2025.015

A Learning Tool for the P-V Curve of the Respiratory System

Vander Baptista, PhD

Department of Physiological Sciences, Centre of Biological Sciences, Federal University of Santa Catarina, Florianópolis, SC, Brazil

Corresponding author: vander.baptista@ufsc.br

Abstract

The respiratory system's pressure-volume (P-V) relationship is crucial for understanding respiratory mechanics and invaluable for teaching ventilation principles. The Rahn diagram (Rahn et al., 1946) is particularly significant, as it illustrates the distinct P-V curves for the lungs and thoracic cage, demonstrating how their integration creates the overall curve of the respiratory system. A thorough mastery of this diagram is essential for a comprehensive understanding of respiratory mechanics. The three intersecting sigmoid curves shown in the Rahn diagram can pose various interpretive challenges, hindering its effectiveness as a teaching tool. A didactic model replicating the Rahn diagram has been proposed to overcome these challenges. The model is supported by a theoretical background that aids in explaining the associated concepts. The first component of this model uses a simplified linear approach based on spring mechanics, elucidating the relationship between two elastic elements and paving the way for the development of sigmoidal P-V curves. The second component comprises a mechanical model specifically designed to simulate data acquisition, thus allowing the accurate reconstruction of the Rahn diagram. This approach was developed as a pedagogical tool for classroom slide presentations, dispensing with the need for a physical model, while providing explanatory reading material that students are expected to engage with as part of the learning process. It seeks to enhance the teaching and learning process regarding the respiratory system's P-V curves while highlighting the Rahn diagram's educational value. https://doi.org/10.21692/haps.2025.015

Key words: respiratory mechanics, Rahn diagram, didactic model

Introduction

The respiratory system's pressure-volume (P-V) relationship is rooted in the elastic mechanical behavior of the lungs and the chest wall. This relationship has been utilized as a physiological tool for research, diagnostic applications, and the monitoring of respiratory mechanics across a range of clinical conditions, as well as for guiding safe ventilation pressures in mechanical ventilators (Brochard, 2006; Maggiori & Brochard, 2001; Venegas et al., 1998). Additionally, P-V curves serve as invaluable pedagogical resources for elucidating the mechanics of ventilation, including the elastic properties of the lungs and chest wall. Particularly noteworthy is the P-V relationship presented by Rahn et al. (1946), which, in diagram form, illustrates three relaxation pressure curves: one representing the entire respiratory system and the other two depicting its components: the lungs and the combined elasticity of the rib cage, diaphragm, and other elastic structures. When correctly understood, this diagram holds significant educational value. As Mitzner (2011) noted, it is frequently featured in respiratory physiology textbooks, underscoring its pedagogical relevance.

Despite its considerable instructional value, many undergraduate courses overlook exploring the Rahn diagram, typically reserving it for more advanced studies. This omission by instructors may arise from two principal concerns regarding the students. First, practical graph interpretation is essential and requires cognitive engagement in analytical thinking, proportional reasoning, and visuospatial skills, areas where undergraduate students may not consistently excel, thus limiting their ability to extract meaningful insights (Angra & Gardner, 2018; Glazer, 2011). This challenge is particularly pertinent to Rahn's P-V diagram, which features three intersecting exponential

continued on next page

and sigmoidal curves that illustrate complex physiological concepts, often rendering them difficult and intimidating for some students. Moreover, understanding alveolar ventilation is also challenging, as it fundamentally involves mechanical processes that demand understanding of basic physical principles (West, 2008), which students may not always have. Without a clear grasp of the mechanical interactions between the lungs and the chest wall, interpreting the diagram can be daunting, potentially complicating instead of enhancing comprehension.

It is essential to acknowledge that the diagram is not designed to initiate studies on respiratory mechanics but rather to summarize and conclude existing explanations. In this sense, analyzing and interpreting the diagram challenges, tests, and complements the knowledge already acquired regarding alveolar ventilation. By adopting this pedagogical approach, the Rahn diagram acts as a tool for synthesizing and reinforcing fundamental concepts, guiding learners to integrate theoretical principles with visual representations of respiratory system behavior. This methodology effectively bridges the gap between isolated physiological concepts and the overall mechanical behavior of the respiratory system, thereby enhancing the understanding of the dynamics of alveolar ventilation. I believe that those who grasp the diagram also attain a true mastery of respiratory mechanics.

Acknowledging the Rahn diagram's difficulties while recognizing its significant pedagogical role, this work aims to provide alternative reference materials for undergraduate and postgraduate courses in respiratory physiology, particularly in the study of respiratory mechanics. The goal is to foster a deeper understanding of the diagram's underlying principles, emphasizing its importance in illustrating the complex interrelations of the elastic structures of the respiratory system in the mechanics of breathing.

The content is organized into two sections. The first section introduces a simplified model that illustrates the linear mechanical behavior of two elastic elements, individually and in combination. This foundation sets the stage for the second section, which delves into the sigmoidal P-V relationship.

In both sections, a theoretical framework contextualizes the subject matter, combining the key theoretical elements underpinning the elastic behavior of the structures involved. The framework facilitates a comprehensive understanding of the concepts discussed, serving as a teaching-learning guide. All compliance curves presented are constructed from simulated data acquisition in abstracted experiments using didactic models. Thirteen figures throughout the text illustrate the models, providing a visual structure that guides students in simulating data acquisition and constructing curves, ultimately leading to structuring the Rahn diagram.

This approach was designed as a pedagogical tool for classroom slide presentations and reading assignments. The figures can be displayed sequentially in class, encouraging active student engagement, fostering discussion and analysis, and promoting a deeper understanding of the Rahn diagram. It is important to note that this didactic model serves as a conceptual representation supported by explanatory figures to facilitate comprehension; it is not a physical model.

Section 1 - Strain-Force Relationship

The mechanics of ventilation comprise both active contractions performed by the respiratory muscles and passive movements governed by the elastic recoil properties of pulmonary structures and the thoracic cage. While these two components are mechanically interdependent, it is beneficial to examine their elastic characteristics separately.

Lung Elastic Components

The lungs are not mere hollow, balloon-like structures. Instead, their parenchyma comprises numerous microscopic polygonal air sacs known as alveoli, which share walls with neighboring sacs, creating a honeycomb-like arrangement when viewed two-dimensionally. The primary constituents of the extracellular matrix in the interalveolar septa are elastin and collagen, which endow the alveoli with elasticity. Consequently, we can infer that each alveolus operates mechanically like a micro balloon, engaging in isotropic expansion and contraction (Toshima et al., 2004). The extracellular fibroelastic matrix is continuous throughout the lung, extending inward along the alveolar septa from the hilum and visceral pleura, creating an interconnected network of alveolar walls. This design fosters mechanical interdependence among the alveoli, allowing mechanical stress to be distributed throughout the lung parenchyma and promoting uniform expansion of the airspaces (Mead et al., 1970).

Another elastic component of the lungs is the surface tension at the air-liquid interface within the alveolar space, which tends to cause the alveoli to collapse. However, the presence of surfactant reduces surface tension as the alveoli shrink, while a decrease in its concentration during alveolar expansion leads to increased surface tension. This mechanism helps equalize the alveoli's pressure and prevents collapse, thereby ensuring stability (Andreassen et al., 1985).

Considering these concepts and employing a didactic simplification, we can posit that the mechanical behavior of the lung arises from the linear aggregation of the individual responses of all alveoli or, to use our analogy, all micro balloons. Consequently, an excised lung can be conceptualized as a single large balloon, as depicted in Figure 1A, where inflation and deflation processes parallel those of inspiration and expiration. By applying further pedagogical simplification, if we focus solely on the linear portion of the P-V relationship of the balloon, we can refine the lung's behavior from that of a balloon to that of a simple

continued on next page

unidirectional spring (Figure 1B). In this context, the stretching and relaxing of the polymer molecules in the balloon wall can be likened to the expansion and contraction of a spring, which we will refer to as the lung spring (L-spring) in our model.

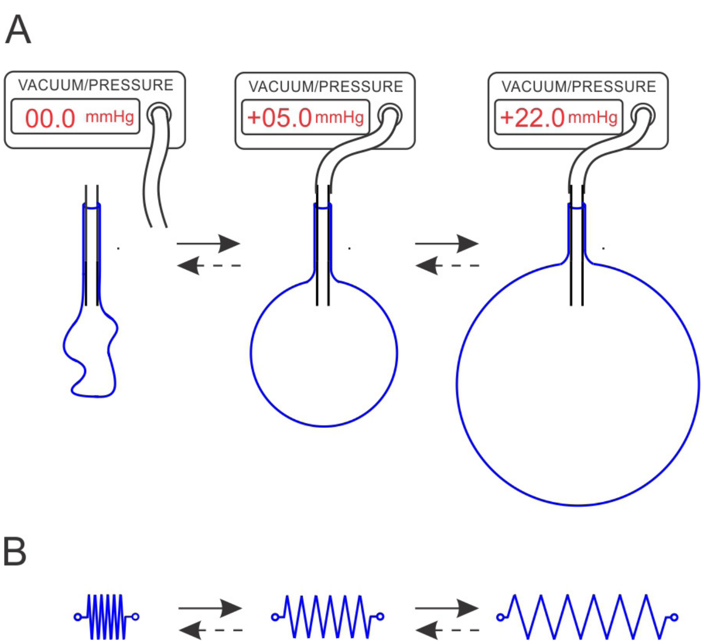


Figure 1. *Lung models. A. The lung is modeled as a balloon with inflation and deflation of the balloon being akin to inspiration and expiration. B. The lung is modeled as a spring (L-spring). Solid arrows represent active movement, while dashed arrows indicate passive movement.*

Let us explore the elastic behavior of the L-spring, which Hooke's law can describe in the following form:

$$S = \frac{F_L}{k} \quad (equation\ 1)$$

Where *S* (strain) represents the change in length relative to the original length, F_L is the applied force (in the L-spring), and *k* is the elastic constant of the L-spring. In this context, F_L is the independent variable, meaning that a given force (F_L) is applied, and the resulting displacement, S, is measured.

In the theoretical simulation of the experiment shown in Figure 2A, two different forces (weights) are applied, and the final strain (static equilibrium condition) of the L-spring is measured. The theoretically simulated data are plotted on a graph as strain versus force (*S-F* relationship), and equation 1 fits them as shown in Figure 2B.

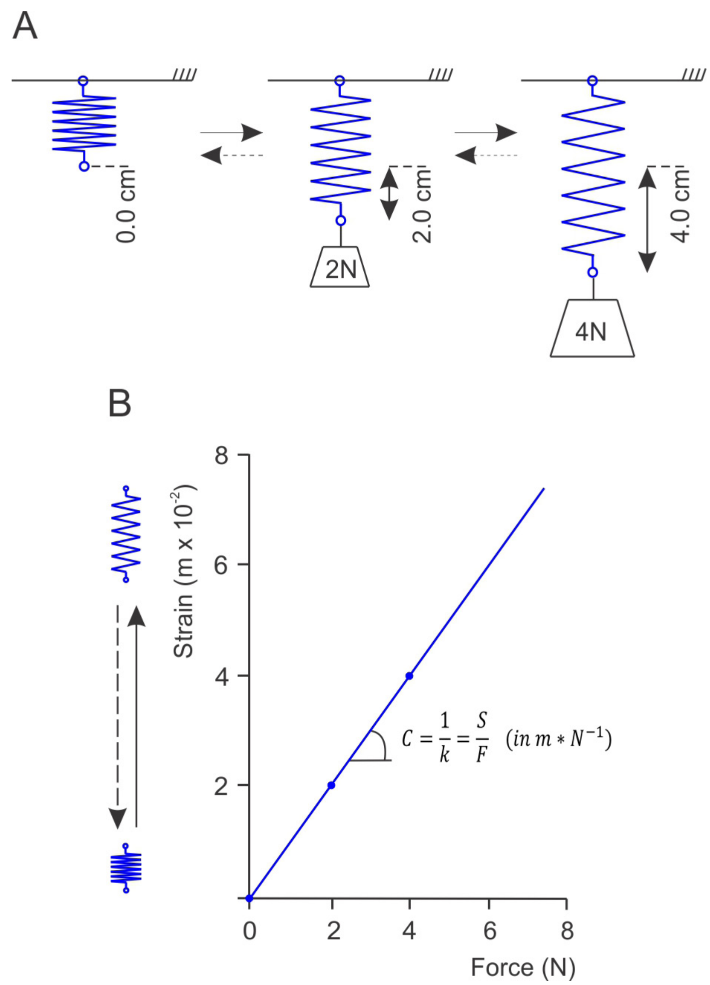


Figure 2. *Spring elastic behavior. A. Theoretical simulated experiment with the expansion of the L-spring. B. The linear strain-force relationship of the L-spring. Solid arrows represent active movement, while dashed arrows indicate passive movement.*

From the graph, three key observations can be made: i) The slope (*1/k*), measured in meters per Newton, represents the compliance of the spring. This value indicates how readily the spring extends in response to an applied force. ii) The values on the x-axis represent the force exerted on the L-spring. These values can also be interpreted in a static scenario as the elastic recoil force stored within the spring. Consequently, a positive force (positive values on the x-axis) indicates the spring's tendency to return to its resting position. iii) Since the L-spring can only elongate from its equilibrium position, its *S-F* line is entirely located in the first quadrant.

continued on next page

Chest wall elastic elements

Figure 3 presents a model of the chest that aids in visualizing fundamental aspects of breathing mechanics. Panel 3A features a lateral view sketch of a human thorax, illustrating how the ribs articulate with the vertebrae and the sternum. Panel 3C simplifies this by depicting a single rib and its rotational movements at the costovertebral joint, with downward movement shown in Figure 3B and upward movement in Figure 3D, both affecting the position of the sternum to the vertebral column. Panel 3F further abstracts the representation, illustrating the rib, vertebral column, and sternum as three articulated rods stabilized by elastic ligaments. These ligaments represent the costovertebral and sternocostal ligaments, which possess elastic properties.

The sternocostal cartilage (shown in red or blue in Figures 3B, 3C, and 3D) and the diaphragm (not depicted) also exhibit elasticity. At the equilibrium point illustrated in Figure 3F, none of the ligaments are under tension (or if they are, they offset each other, maintaining the entire system in an elastic equilibrium). The ribs can move in both directions from their elastic resting points, thereby altering the volume of the chest wall. However, an active force is necessary to disrupt the equilibrium position. For instance, if an external force is applied to lower (Figure 3E) or raise (Figure 3G) a rib, specific ligaments are stretched (shown as red lines), and the cartilage flexes (in red in Figures 3B and 3D), allowing for the storage of elastic energy. Once the external force is removed, the system naturally reverts to its equilibrium state, meaning the elastic recoil force pulls the system back into balance.

Let us consider a scenario in which only the *S-F* curve's linear region is analyzed for this system. In this context, we can simplify further by consolidating all the elastic elements into a single entity represented by a single bidirectional spring (Figure 3I), which we will refer to as the chest wall spring (CW-spring). This spring, akin to the ribs, can be compressed (Figure 3H) or extended (Figure 3J) by an external force, thereby storing elastic potential energy. Once the force is removed, the spring returns to its equilibrium position (Figure 3I).

Assuming that the CW-spring at the equilibrium position (neither compressed nor extended) is 5 cm longer (see Figure 4B) than the resting length of the L-spring (Figure 2A), we can conduct another theoretical simulation in an experiment similar to that with the L-spring. Using an appropriate device, a 5N weight fully compresses the CW-spring (Figure 4A), reducing it to the same length as the natural size (equilibrium position) of the L-spring (Figure 2A). Conversely, a 2N weight applied in the opposite direction extends the CW-spring (Figure 4C). To fit the theoretically simulated data, Hooke's law is expressed as follows:

$$S = \frac{F_{CW}}{k} + 5 \ \ (equation\ 2)$$

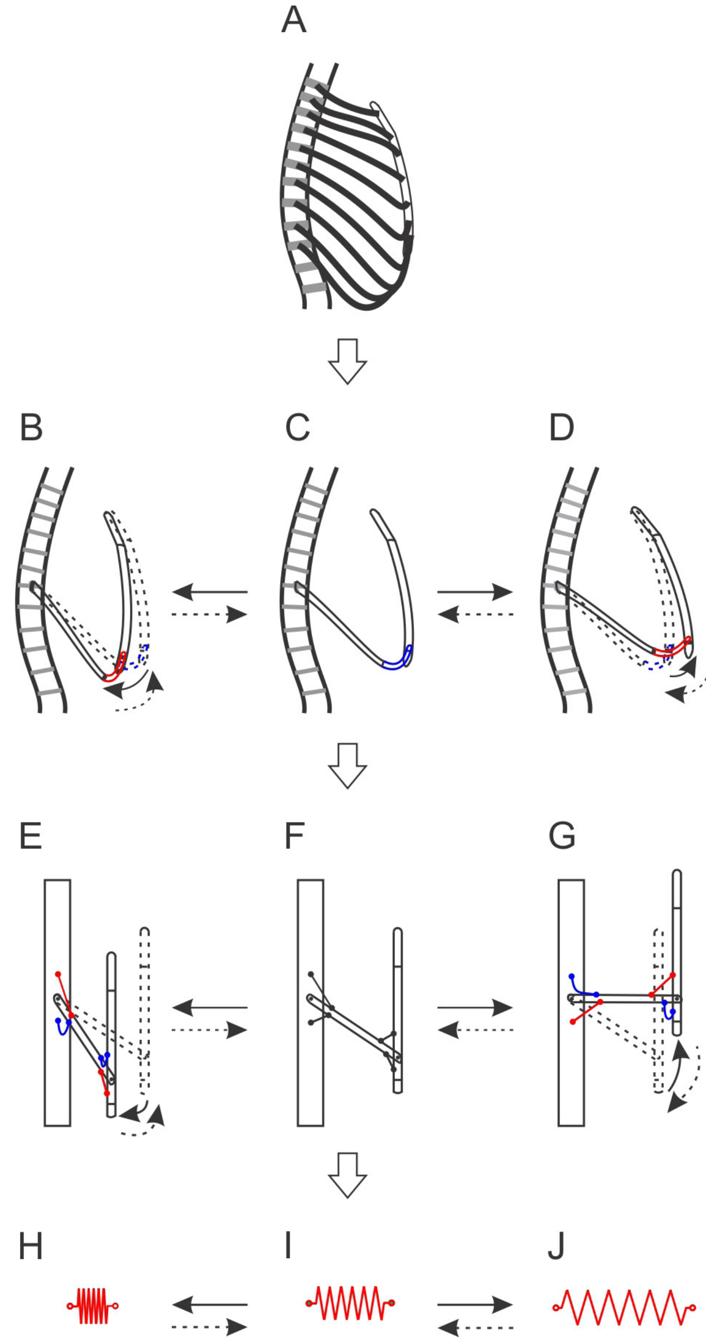


Figure 3. *Chest wall models. A. Sketch of the lateral view of the human thorax. B, C, D. A single rib and rotational movements at the costovertebral joint. E, F, G. The rib, vertebral column, and sternum are represented as three articulated rods held in a balanced position by elastic ligaments. H, I, J. The chest wall is modeled as a spring (CW-spring). Solid arrows represent active movement, while dashed arrows indicate passive movement.*

continued on next page

Where F_{CW} is the force applied to the CW-spring. This equation indicates that the fitted line is displaced five units upwards to the L-spring. Furthermore, for simplicity and educational clarity, we assign the CW-spring the exact value of *k* as that of the L-spring, resulting in a fitted line parallel to that of the L-spring (Figure 4D).

Three key observations emerge from this experiment: i) The CW-spring can elongate and contract from its elastic equilibrium point. ii) The *S-F* line spans the first two quadrants, intersecting the y-axis at its elastic resting point ($F = 0$) at 5 cm. iii) While the positive segment of the x-axis indicates the values of the elastic retraction force, the negative segment represents the values of the elastic expansion force.

Elastic behavior of the association of the L- and CW-springs

Like the respiratory system, where the lungs and chest wall are mechanically coupled by subatmospheric pressure in the pleural space (see Baptista, 2010, for this coupling mechanism), the L-spring and CW-spring will be interconnected. A preliminary examination of Figure 13 indicates that the P-V curve for the entire respiratory system exhibits a less steep incline than that of the lung and chest wall when considered separately, implying a lower compliance. To capture the stiffening behavior associated with the respiratory system, components in the spring model, the L- and CW-springs, will be arranged in parallel. This configuration leads to a system that is less compliant than each of the springs individually, thus accurately modeling the physiological interaction between the lungs and the chest wall (see Figure 5). With these foundational assumptions in place, we can proceed to conduct a theoretical simulation as illustrated in Figure 5. During this simulation, under static conditions (with no movement of the springs involved), the data points obtained will be displayed in Figure 5G. This figure also incorporates the individual line fits for the *S-F* (*Strain-Force*) relationships of the L- and CW-springs, detailed separately in Figures 2B and 4D.

To establish the parallel association, first, a weight of 5 N is applied to the L-spring, which undergoes an elongation of 5 cm, reaching the length of the CW-spring when at elastic rest (see Figures 5A and 5B). Following this, we connect the two springs in parallel, culminating in a system designated as the respiratory system spring (RS-spring) (Figure 5C). To reach the steady state of the system, the 5N weight is released; the system RS then retracts due to the elastic potential energy accumulated in the L-spring. During retraction of the RS-spring, the CW-spring is gradually compressed until it accumulates potential expansion energy that equals that of retraction of the L-spring, establishing a new steady state for the RS system. This scenario is depicted in Figure 5D, where the RS-spring reaches its elastic equilibrium position, characterized by a total force of zero. From this point, the RS-spring can either compress or extend, which can only occur under the influence of an external force. Over

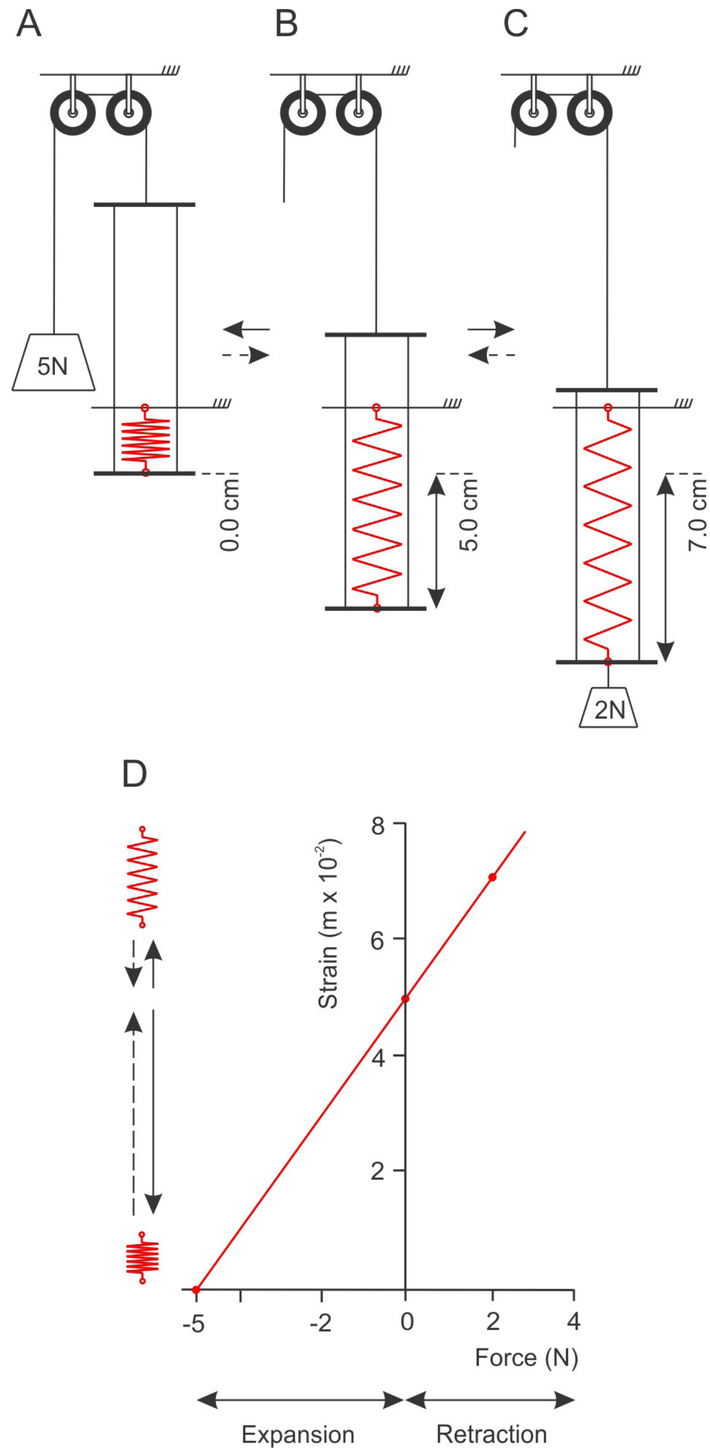


Figure 4. *Spring elastic behavior. A. Theoretical simulated experiment with expansion and compression of the CW-spring. B. The linear strain-force relationship of the CW-spring. Solid arrows represent active movement, while dashed arrows indicate passive movement.*

continued on next page

the complete excursion range of the RS-spring from 0 to approximately 8 cm in our experimental setup, the strain at the equilibrium position is documented as 2.5 cm (Figure 5D). At this juncture, the total force balances out, with the retraction force of the L-spring counteracting the expansion force of the CW-spring. This position, set at 2.5 cm stretch, becomes the foundational data point for the construction of the *S-F* relationship for the RS-spring system (notated as point 'a' in Figure 5G). Subsequently, further contraction is accomplished by progressively adding weights to the end of the line traversing through the pulley, allowing the RS-spring to achieve full compression under a cumulative force of 5N (Figure 5E, point 'b' in Figure 5G). At this point, the CW-spring component stores elastic energy for expansion while the spring L remains in an elastic resting state. Upon the release of the RS-spring from the 5N weight, it passively returns to its elastic equilibrium position (illustrated in Figure 5D and corresponding to point 'a' in Figure 5G).

A series of progressively heavier weights is applied to the bar connecting the lower ends of the two springs, extending the RS-spring system from its equilibrium position. Figures 5C and 5F illustrate two of these weights. With a force of 5N, the RS-spring stretches to an extension of 5 cm (as shown in Figure 5C and marked as point 'c' in Figure 5G), while a weight of 9N further extends it to 7 cm (illustrated in Figure 5F and marked as point 'd' in Figure 5G).

We can now derive the equation to fit the data obtained from the RS-spring experiment. In a parallel configuration, the deformation of both springs in the system is equal, while the forces acting on them combine to produce the total force. Consequently, the total force on the RS-spring (F_{RS}) can be expressed as the sum of the forces exerted by the L-spring (F_L) and the CW-spring (F_{CW}):

$$F_{RS} = F_L + F_{CW} \quad (equation\ 3)$$

From Equation 1, we have:

$$F_L = k * S$$

From Equation 2, we obtain:

$$F_{CW} = k * S - 5k$$

Substituting these into Equation 3 yields:

$$F_{RS} = k * S + k * S - 5k$$

$$F_{RS} = 2k * S - 5k$$

Expressing *S* in terms of *F*, we get:

$$S = \frac{F_{RS} + 5k}{2k} \quad (equation\ 4)$$

This equation defines the *S-F* relationship for the RS-spring based on the experiment depicted in Figures 5C-5F, resulting in the green line in the graph shown in Figure 5G.

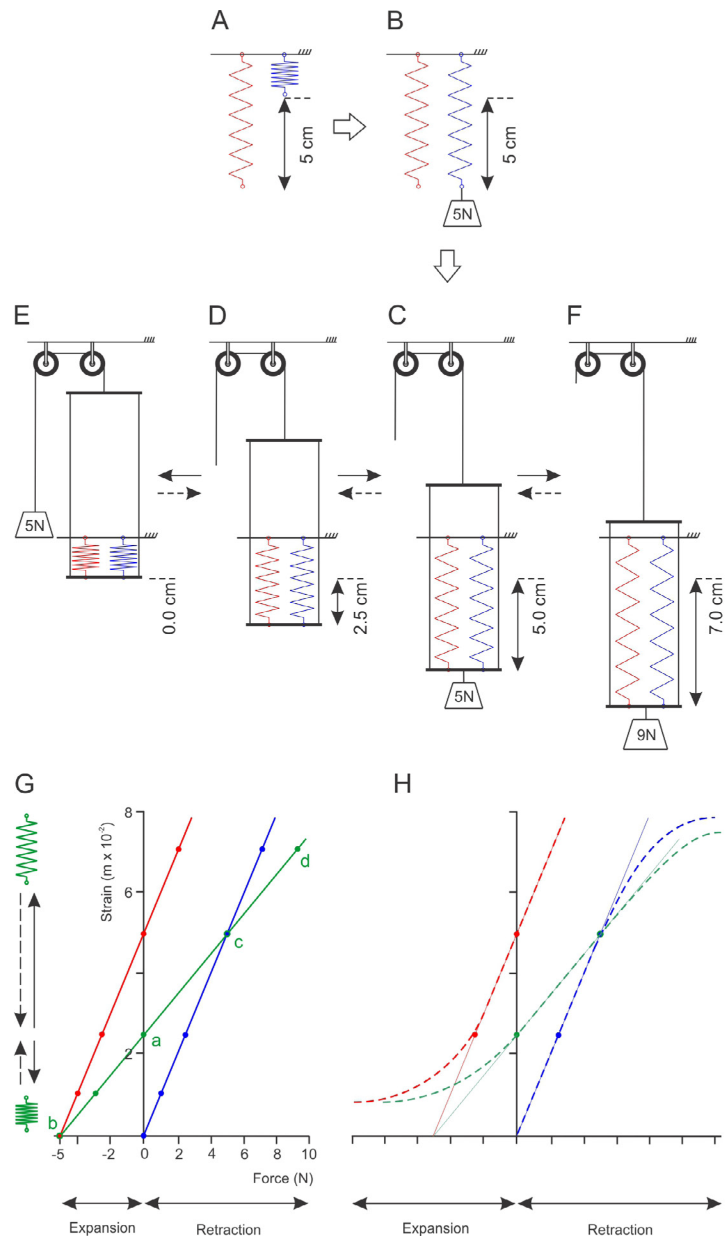


Figure 5. *Elastic behavior of two springs in association, forming the RS-spring. A, B, C. Sequence to place the L-spring (blue) and CW-spring (red) in parallel association. D. Elastic equilibrium point of the RS-spring. E. Full compression of the RS-spring. F. Full expansion of the RS-spring. G. The strain-force relationship of the L-spring (blue) and CW-spring (red) separately and in association (RS-spring in green). H. The dotted lines superimposed on the S-F relationship of the RS-spring provide a preliminary visualization of the expected P-V relationship of the respiratory system. Horizontal solid arrows represent active movement, while dashed arrows indicate passive movement.*

continued on next page

From the analysis of the presented graph, eight key considerations emerge:

i. A notable observation is that the compliance of the RS-spring is significantly less than that of either the L-spring or the CW-spring when considered individually. This characteristic is typical of springs arranged in parallel.

ii. The *S-F* line of the RS spring represents the sum of the *S-F* lines of the L-spring and the CW-spring at any given elongation. Thus, as indicated by equation 3, the force needed to extend the RS-spring equals the sum of the forces required for each spring when stretched independently, showcasing the increased stiffness of the overall system.

iii. The *S-F* line of RS-spring spans the first two quadrants, intersecting the y-axis at its elastic resting point ($F = 0$) at 2.5 cm.

iv. At a length of 2.5 cm, the RS-spring achieves elastic equilibrium; the total elastic forces exert zero net force. This situation signifies that the CW-spring's elastic expansion force matches the L-spring's elastic retractive force. It is essential to highlight that the RS-spring remains in static equilibrium, not thermodynamic equilibrium. Adjustments from this state, whether compression or elongation, necessitate the application of an external force.

v. When the RS-spring is elongated to 5 cm, only the L-spring accumulates elastic energy, prompting a tendency to recoil. Contrarily, the CW-spring maintains its elastic equilibrium at this position.

vi. If an external force compresses the RS-spring to a dimension less than 2.5 cm, it retains potential elastic energy for expansion. Thus, upon removal of the external force, the RS-spring will revert to its elastic equilibrium at 2.5 cm (identified as point 'a' in Figure 5G). On the other hand, if stretched beyond 2.5 cm, it stores potential elastic energy for retraction, enabling it to return spontaneously to equilibrium.

vii. For lengths less than 5 cm, the two components of the RS-spring system, the L-spring and CW-spring, exert forces in opposing directions: the CW-spring seeks expansion while the L-spring aims for retraction. In contrast, for lengths exceeding 5 cm, both components act in concert, collectively directing a retraction.

viii. Ultimately, it is evident through simple observation that the L- and CW-springs are both physically and functionally integrated. The emerging properties of the RS-spring represent a fascinating balance of the characteristics inherent to each spring.

In summary, this section sought to establish a fundamental understanding of the linear mechanical behavior of two elastic elements, individually and in combination. This foundation will support our exploration in the next section, where we will delve into the sigmoidal P–V relationship. Although the *S–F* plot depicted in Figure 5G is linear, its three straight lines correspond with the three sigmoidal curves of the P–V plot of the respiratory system (refer to Figure 13). A preliminary comparison reveals similar visual characteristics, such as slope and intercepts. Moreover, the observations made regarding the L- and CW-springs, as well as the eight points described for the RS-spring system, can be qualitatively extrapolated to the three related curves in the Rahn diagram constructed in Figure 13. Finally, the dotted lines superimposed on the *S-F* relationship of the spring system (Figure 5H) provide an initial visualization of the anticipated P-V relationship of the respiratory system (see Figure 13). This illustration highlights the mechano-elastic similarities between the two systems.

Section 2 - Pressure-Volume relationship

Like the *S-F* relationship illustrated in Figure 5G, the P-V curve is also constructed under static conditions. Thus, it eliminates the effects of resistance and impedance on pressure, allowing only compliance to be accessed (Harris, 2005). Each point on the P-V curve is obtained by temporarily halting tidal breathing at various lung volumes, ranging from residual volume to total lung capacity, and measuring the corresponding pressures. The resulting data is plotted on a graph with volume on the y-axis and pressure on the x-axis, although it is commonly called the P-V relationship. Volume can be directly measured using spirometry, whereas pressure necessitates indirect measurement methods, which can sometimes confuse students when interpreting the x-axis of the P-V curves. Establishing precise terminology for the different pressures within the respiratory system is essential to mitigate this potential confusion. Three key pressures to consider are: i) Atmospheric pressure (P_{AT}), which serves as the reference point. ii) Intrapleural pressure (P_{IP}) represents the pressure within the space between the parietal and visceral pleura. iii) Alveolar pressure (P_A) refers to the alveoli's pressure.

continued on next page

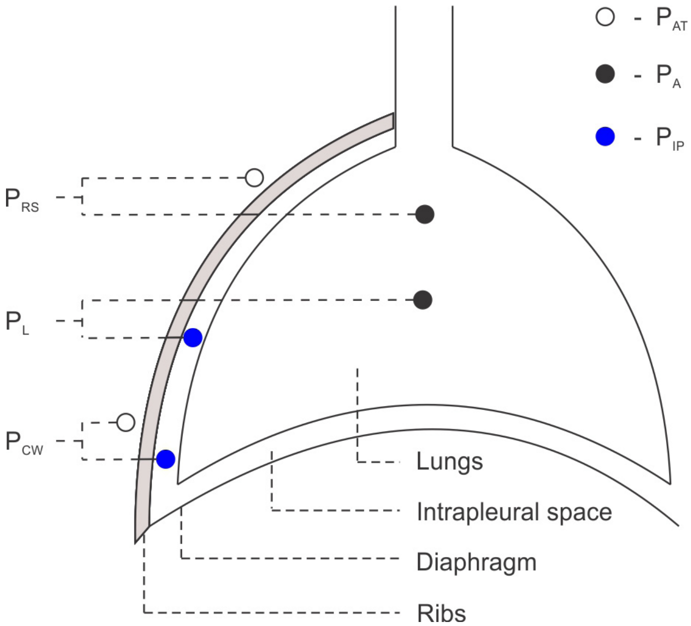


Figure 6. *Conventional terminology for the different respiratory system pressures. The black complete cycle represents alveolar pressure (P_A), the blue entire cycle represents intrapleural pressure (P_{IP}), and the empty cycle represents atmospheric pressure (P_{AT}). P_{RS} is the total pressure of the respiratory system, P_L is pulmonary pressure, and P_{CW} is chest wall pressure.*

From these pressures, we can derive three distinct pressure drops (refer to Figure 6). The pressure drop across the respiratory system (P_{RS}) is the difference between alveolar and atmospheric pressure, expressed as $P_{RS} = P_A - P_{AT}$. The intrapleural pressure (P_{IP}) acts as a pressure divider, separating P_{RS} into two components: the pressure drop across the chest wall ($P_{CW} = P_{IP} - P_{AT}$), and the pressure drop across the lungs ($P_L = P_A - P_{IP}$). These pressure drops yield the following relationship, as described by Rahn et al. (1946):

$$P_{RS} = P_L + P_{CW} \quad (equation\ 5)$$

Knowing the value of intrapleural pressure is essential to determine each pressure drop. While intrapleural pressure can be measured directly with a needle, which is often impractical in clinical settings, a less invasive method involves approximating it through esophageal pressure, given that the esophagus runs through the intrapleural space. Esophageal pressure is typically assessed using a differential transducer placed in the distal third of the esophagus, aligned with the middle of the lungs (Gibson, 2020). Direct measurement of alveolar pressure is also not practical. Instead, it is approximated by measuring proximal airway pressure, which can be assessed using the multiple occlusion technique on a mechanical ventilator. During short occlusions, with the lungs inflated to various volumes, pressure equilibrates throughout the system, enabling proximal airway pressure to estimate alveolar pressure (Dostal & Dostalova, 2023).

A didactic model for constructing P-V curves.

With a focus on educational objectives and on mitigating the technical complexities associated with obtaining intrapleural (or esophageal) and alveolar (or proximal airway) pressures, we implemented a didactic model (Figure 7) to simulate an equivalent pressure measurement process. This model is intended to enhance the visualization of data acquisition for constructing respiratory P-V curves. Furthermore, it aims to aid in data interpretation, allowing students to concentrate more effectively on understanding the interrelationships of respiratory pressures that govern breathing mechanics.

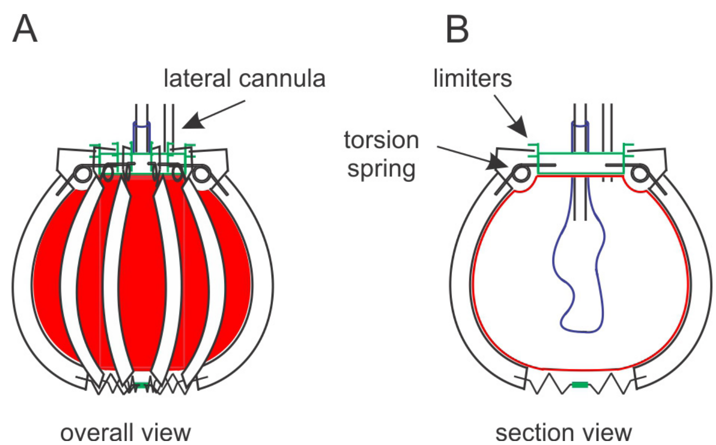


Figure 7. *Respiratory system model. A. Full view of the barrel structure. The articulated barrel boards represent the ribs, and the parietal pleura is shown in red. B, C, D. Different views of the model, as indicated.*

The model depicts the chest wall as a barrel-shaped structure containing a balloon representing the lungs. As illustrated in Figure 7, the barrel boards, symbolizing the ribs, are articulated to a central framework by connecting to torsion springs, enabling expansion and retraction movements. These joints emulate the costovertebral joints, and the springs their elastic components. At the opposing end, the boards are linked by compression and extension springs, representing other elastic components of the chest wall, including the diaphragm and skin. The inner surface of the barrel is lined with an impermeable material, representing the parietal pleura (depicted in red in Figure 7). This configuration creates a sealed cavity between the balloon and the inner surface of the barrel, representing the pleural space. From this point forward, we will refer to the balloon as the modeled lung (M-lung) and the barrel as the modeled chest wall (M-chest wall). Collectively, the M-lung and

continued on next page

M-chest wall, which simulate the respiratory system, will be referred to as the modeled respiratory system (M-respiratory system), and the space between them will be known as the modeled pleural space (M-pleural space).

Figures 7A and 7B display the model's full and sectional views for enhanced visualization. Note that a lateral cannula allows access to the M-pleural space. Furthermore, the M-chest wall's movement limiters establish structural constraints on its expansion and retraction.

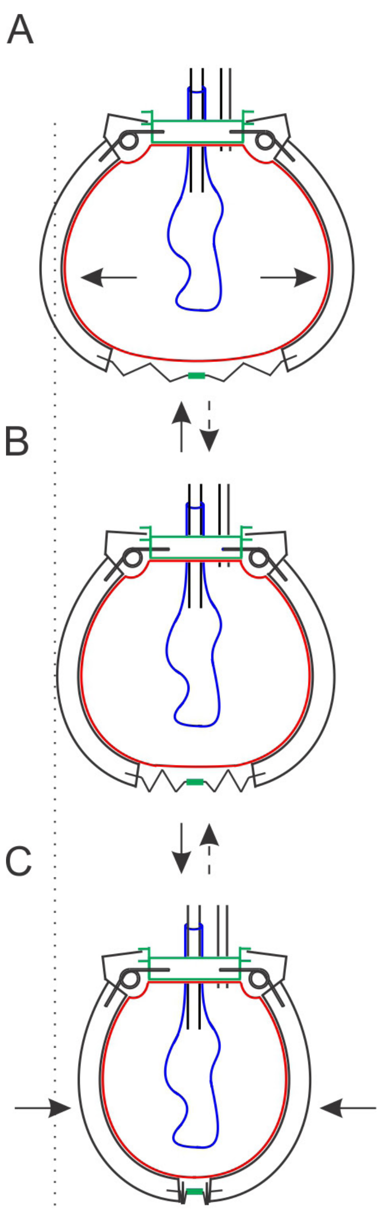


Figure 8. *Full range of movement excursion of the model's chest wall component (M-chest wall). A. Maximal expansion. B. Elastic equilibrium point. C. Maximal contraction. The horizontal arrows indicate external force. Vertical solid arrows represent active movement, while dashed arrows indicate passive movement.*

Figure 8 highlights the functional aspect of the M-chest wall. Starting from its elastic equilibrium point—when none of the springs are under tension (or, if they are, they cancel each other out) (Figure 8B): an external force can induce the full range of motion of the M-chest wall, with maximum expansion and retraction restricted by structural constraints, as illustrated in Figures 8A and 8C, respectively. Assuming that the M-lung and the M-chest wall possess elastic and mechanical properties akin to those of natural lungs and the chest wall, respectively, and that they function together as the entire respiratory system, we will simulate data acquisition to replicate the three P-V relationship curves illustrated in the Rahn diagram.

P-V curve for isolated lung

Data for this curve can be obtained through *ex vivo* experiments using cadaver lungs outside the thoracic cavity, where PL = $P_A - P_{AT}$, or through *in vivo* procedures, where $P_L = P_A - P_{IP}$. Utilizing our didactic model, we can simulate an *ex vivo* experiment with the M-lung, miming an excised lung. This approach appears to be the most effective way for students to visualize the P-V curve of isolated lungs, as the lungs are genuinely isolated in this excised state. Figure 9A illustrates the experimental protocol. In this simulation, the M-lung is inflated incrementally from its relaxed volume until it reaches a maximum, which, while simulating physiological conditions, is approximately 6.3 liters (close to dV/dP = 0). The vacuum-pressure machine measures the internal pressure of the M-lung (relative to P_{AT}) and the volume under static conditions at each step. This scenario simulates the glottis being closed and the respiratory muscles being relaxed, meaning there is no airflow. The data points acquired in this manner are plotted on a graph as volume versus pressure (Figure 9B). While the literature provides equations that adequately fit actual data regarding the P-V relationship of the lungs, for the sake of simplicity, the exponential line connecting the simulated theoretical points has been manually adjusted. This graph is straightforward to interpret, especially considering the following observations:

i. A typical toy balloon's P-V relationship exhibits a sigmoidal shape characterized by two inflection points that divide the curve into three distinct segments. Initially, inflating the balloon requires substantial effort due to its low compliance. However, as the balloon expands, compliance increases, making inflation noticeably easier. Eventually, the inflation pressure rises again (Stein, 1958). This behavior stems from the interaction of geometric factors and the coiled arrangement of flexible polymer chains within the balloon wall, which collectively contribute to the sigmoidal nature of the P-V curve.

ii. Despite the toy balloon's characteristics, it is assumed that the M-lung in our experiment shares the same elastic properties as an excised lung, which

continued on next page

does not display the curve section before the lower inflection point. In a fully collapsed excised lung, the more distal airways (lacking cartilaginous support) and the alveoli collapse due to the lung's elastic recoil pressure and decreased internal pressure. This results in an abrupt cessation of volume change before reaching any potential lower inflection point. To inflate a collapsed excised lung, a critical opening pressure must be achieved to re-expand these structures, prompting a sudden change in the slope of the P-V curve (Ranieri & Slutsky, 1999; Terragni et al., 2003). In our simulated experiment, this corresponds to point 'a' in Figure 9B.

The P-V curve of the excised lung (Figure 9B) demonstrates that as pressure increases, the volume initially expands almost linearly, followed by an exponential rise at higher pressures. This behavior indicates hyperinflation, with overdistension nearing the breaking point. Like the extension spring used in the analogy of Figures 1 and 2, the empty excised lung can only expand; thus, the entire P-V relationship curve remains in the first quadrant. Consequently, the positive segment of the x-axis in the lung P-V curve signifies elastic recoil pressures, reflecting the lungs' natural tendency to return to their resting volume.

P-V curve for isolated chest wall

Considering that $P_{CW} = P_{IP} - P_{AT}$, the data for this curve can be derived from esophageal pressure measurements since this serves as a surrogate for P_{IP}. Alternatively, the curve can be calculated. Given that $P_{RS} = P_{CW} + P_L$, it follows that P_{CW} equals the difference between P_{RS} and P_L. Assuming the M-chest wall exhibits similar elastic properties to the actual chest wall, the static P_{CW} can be measured more directly using the didactic mode. This approach offers students a straightforward method for visualization. The initial step involves assessing the volumetric capacity of the chest wall at its elastic equilibrium position, following the protocol outlined below. With the lateral cannula open (M-pleural space connected to the atmosphere and the M-lung not mechanically engaged with the M-chest wall, as in a pneumothorax), the M-lung is inflated until it fills the entire space of the relaxed M-chest wall (see Figures 10A and 10B). The volume is then recorded using the vacuum-pressure machine. The physiological value is approximately 72% of vital capacity (Rahn et al., 1946), corresponding to about 4.8 liters in our simulation; this represents the first point of the M-chest wall P-V curve (point 'a' in Figure 10F).

To obtain additional points on the curve, the M-lung is withdrawn from the M-chest wall, and the lateral cannula is sealed (Figure 10C), effectively isolating the M-chest wall. In this initial state, no elastic component is being stretched; thus, the M-chest wall is in its elastic equilibrium position ($P_{CW} = 0$ and a volume of approximately 4.8 liters). From here, it can expand or contract by applying positive or negative pressure steps using the vacuum-pressure machine. As

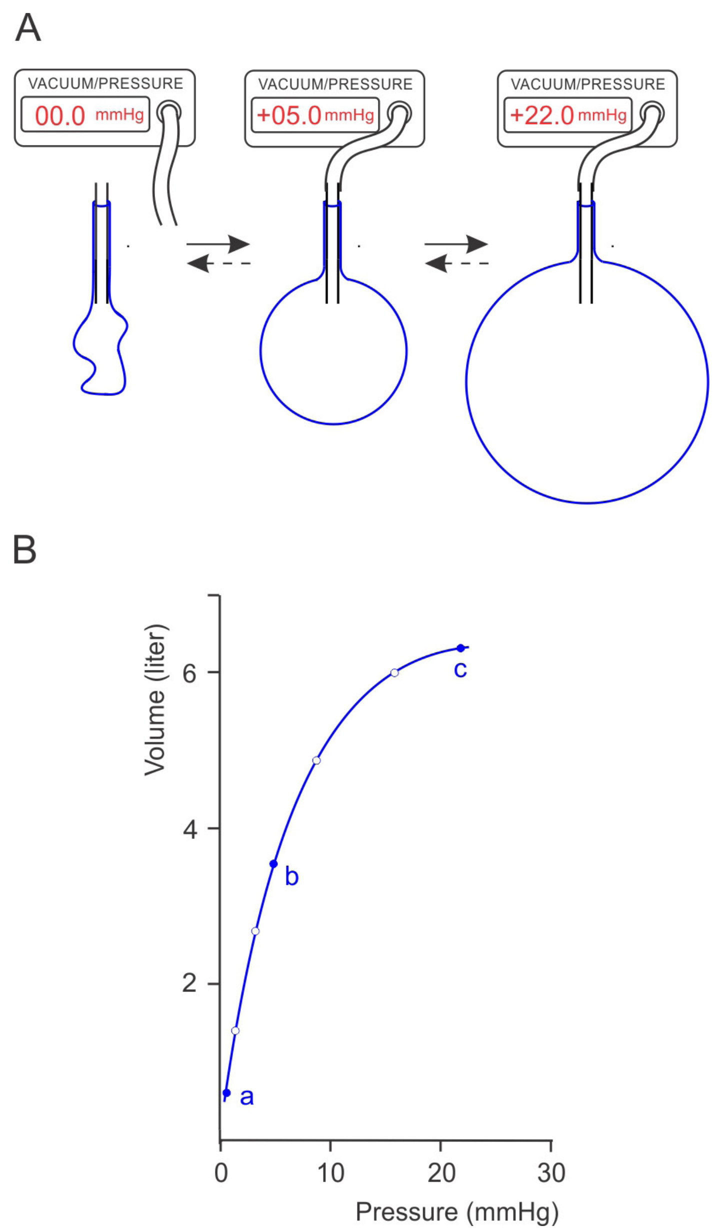


Figure 9. *The balloon lung model (M-lung). A. The Inflation and deflation of the balloon are analogous to inspiration and expiration. B. The M-lung P-V relationship. Solid arrows represent active movement, while dashed arrows indicate passive movement.*

continued on next page

illustrated in Figure 10D, the application of positive pressure causes the M-chest wall to be pushed outward, reaching the maximum expansion of 6 liters, as limited by structural factors, at +2.5 mmHg (point 'c' in Figure 10F). Physiologically, this point reflects total lung capacity. Conversely, when applying negative pressure steps, the M-chest wall is drawn inward. This occurs because the internal pressure of the M-chest wall falls below the external pressure, allowing atmospheric pressure to compress its wall, resulting in a maximum contraction of 1.1 liters, restricted by structural limiters. The maximum contraction, physiologically corresponding to the residual volume, occurs at a negative pressure of -23 mmHg (Figure 10E and point 'b' in Figure 10F).

In Figure 10F, a fit line has been manually adjusted to align with the data points. Notably, from the equilibrium position (as shown in Figure 10C and point 'a' in Figure 10F), the M-chest wall expands linearly until it reaches its maximum capacity, which occurs suddenly and without warning. In this curve segment, the M-chest wall compliance is relatively constant, abruptly dropping to zero. Physiologically, this sudden limitation is attributed to the structure of the joints and the overall architecture of the chest wall, which restrict further expansion. This scenario is akin to the elbow joint, which facilitates smooth movement of the forearm but imposes a sharp limit on extension when the angle between the upper arm and forearm reaches 180 degrees. It is important to note that the maximum volume of M-chest wall is achieved before the lungs reach their maximum volume; in essence, the constraints imposed by the chest wall act as a protective mechanism, safeguarding the lungs from overstretching and potential rupture of the lung parenchyma.

The compliance of the M-chest wall remains relatively constant from maximum volumetric capacity to smaller retraction volumes, after which it begins to decline sharply. This indicates that significant muscular effort is required to achieve progressively smaller respiratory capacities, eventually reaching the residual volume of approximately 1.1 liters. The increased chest wall stiffness at diminished volumes reflects the heightened rigidity of the costosternal cartilages, the distensibility limits of the costovertebral

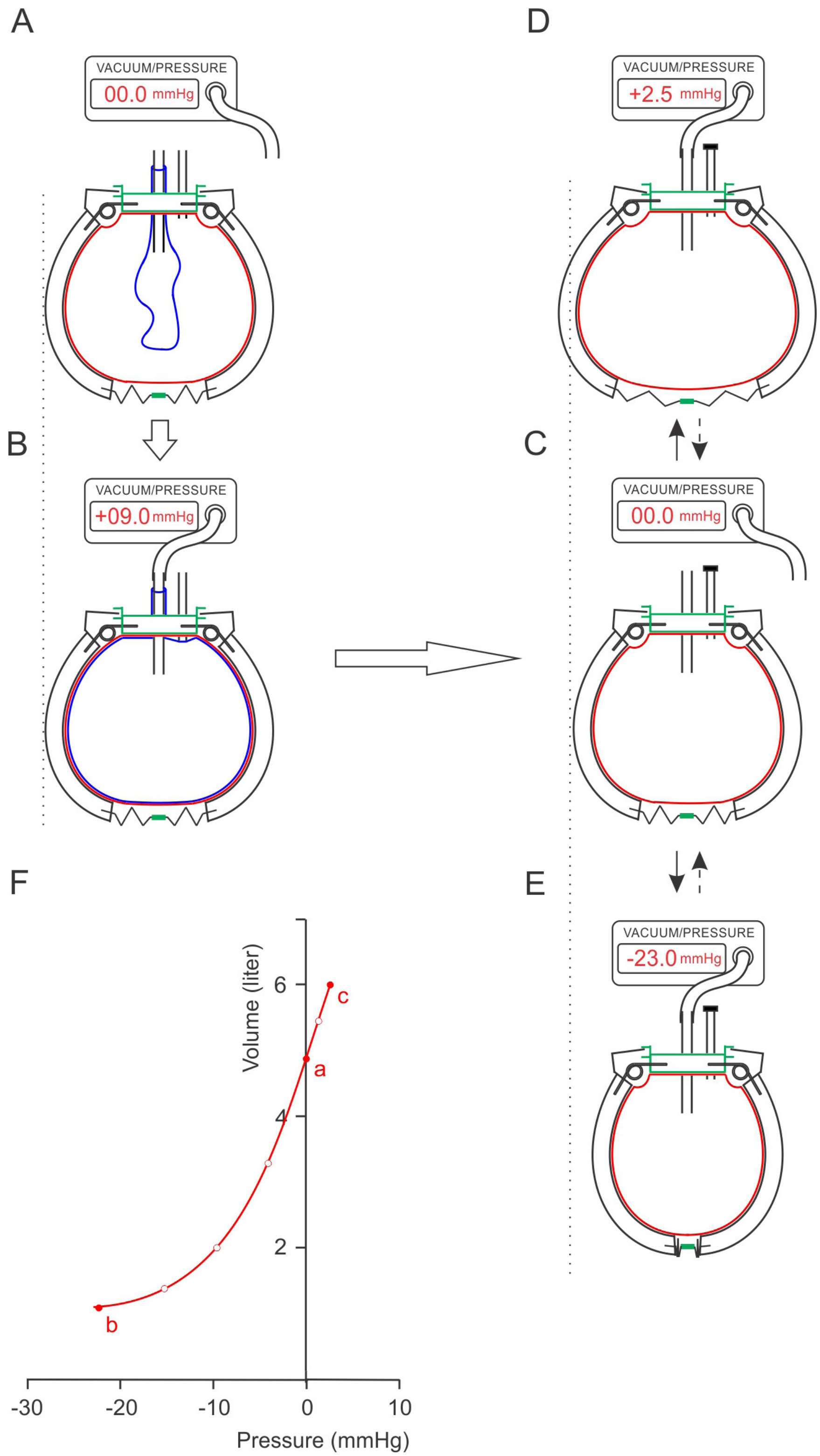


Figure 10. *The barrel chest wall model (M-chest wall) of the M-respiratory system. A, B. Procedure to measure the volumetric capacity of the M-chest wall. C. Elastic equilibrium point of the M-chest wall. D, E. Maximal expansion and contraction of the M-chest wall. F. The P-V relationship of the M-chest wall. Solid arrows represent active movement, while dashed arrows indicate passive movement.*

continued on next page

ligaments, and the movement restrictions imposed by the multiple joints of the bony structure. Just as the upper expansion limit of the rib cage protects the lungs from overstretching, the lower limit imposed by the chest wall structure safeguards the lungs against collapse. Physiologically, at this lower volumetric limit of the chest wall, the lungs retain the residual volume and cannot retract further, as they are held in an expanded state by the negative pressure of the pleural space. As a final observation, it is noteworthy that, similar to the spring illustrated in Figure 4D (CW-spring), the M-chest wall can move in both directions, either expanding or contracting from its elastic equilibrium position. Consequently, its P-V relationship curve occupies the first two quadrants, intersecting the y-axis at the elastic equilibrium point of the M-chest wall (point 'a' in Figure 10F). Thus, while the positive segment of the x-axis in the P-V graph represents elastic recoil pressure, the negative segment corresponds to elastic expansion pressure, reflecting the chest wall's tendency to return to its resting volume.

P-V curve for the entire respiratory system

Considering that $P_{RS} = P_A - P_{AT}$, the data for this curve can be obtained from airway pressure measurements, as it serves as a substitute for P_A. Alternatively, since $P_{RS} = P_{CW} + P_L$, it can be constructed indirectly by summing P_{CW} and P_L.

By utilizing M-respiratory system, it becomes possible to simulate the measurement of static P_{RS} directly, providing an educational tool for constructing the P-V curve of the respiratory system. To achieve this and simulate physiological conditions, it is essential to mechanically couple the M-chest wall to the M-lung and generate negative pressure in the M-pleural space. Figure 11 illustrates the strategy employed in this didactic model. Initially, approximately 20 mL of water, representing pleural fluid, is introduced into the M-pleural space through the lateral cannula (Figure 11A). Next, with the lateral cannula open, the M-lung is inflated until it reaches the elastic resting volume of the M-chest wall, approximately 4.8 liters (Figure 11B). This process spreads the pleural fluid uniformly across all surfaces, creating a water film interface along the entire contact area between the inner surface of the M-chest wall and the outer surface of the M-lung, mimicking physiological conditions. Then, the lateral cannula is sealed, creating an airtight space between the M-lung and the M-chest wall, referred to as the M-pleural space. Following this, the M-respiratory system is disconnected from the vacuum-pressure machine, exposing the M-lung to atmospheric pressure (Figure 11C). In this state, the M-lung passively deflates toward its elastic equilibrium position—its minimal volume. However, the P-V relationship of the M-respiratory system does not align with that of the M-lung, as the M-lung is no longer isolated; it is coupled to the M-chest wall.

As the M-lung deflates, it tends to separate from the M-chest wall, leading to an increase in the virtual volume of the M-pleural space. According to Boyle's law, this results in a growing negative intrapleural pressure, which progressively pulls the M-chest wall toward retraction, following the movement of the M-lung. This effect occurs because intrapleural pressure decreases relative to the surrounding atmospheric pressure. In this scenario, P_{AT} exerts pressure on the M-chest wall toward the M-lung. The cohesive force of the pleural fluid, which 'binds' the two surfaces together, further enhances this mechanical coupling. It becomes clear that as the M-lung retracts toward its elastic equilibrium, the M-chest wall moves away from its equilibrium point while the negativity of the M-pleural space increases in magnitude. Eventually, the three forces reach a state of balance. The elastic tension of both structures becomes equivalent, though directed oppositely, establishing the system's point of stationary or dynamic elastic equilibrium, where $P_{RS} = 0$ (see Figure 11C). The M-chest wall does not expand toward its elastic resting point, nor does the M-lung retract completely towards deflation, as the negative pressure of the M-pleural space maintains the connection between the two structures. Physiologically, this elastic resting point corresponds to the conclusion of a resting expiratory process, which is defined as the residual functional capacity (RFC), approximately 3.3 liters in our experiment. Since the M-respiratory system is in equilibrium, only applying an external force can displace it from this state, which physiologically occurs when the respiratory muscles contract. It is also worth noting that, in this position, alveolar pressure equals atmospheric pressure, and there is no airflow, which is compatible with the end of a resting expiratory process.

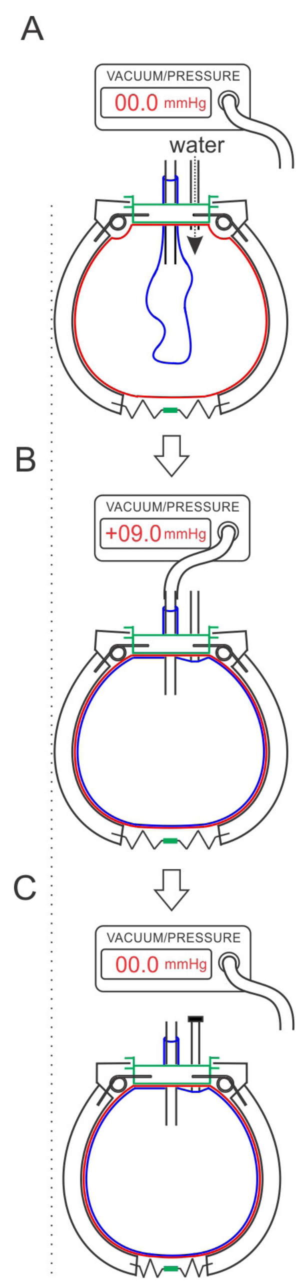


Figure 11. *M-lung and M-chest wall association. A, B, C. The sequence for mechanically coupling the M-lung to the M-chest wall.*

continued on next page

The M-respiratory system configuration depicted in Figure 11C, replicated in Figure 12B, serves as a starting point for a theoretical simulation of an experiment to gather data points to construct a P-V relationship similar to that of the entire respiratory system, as illustrated in the Rahn diagram. The values of volume (~3.3 liters) and pressure (0 mmHg) of this M-respiratory system position mark the initial data point for developing the P-V curve (designated as point 'a' in Figure 13). Utilizing the didactic model, the pressure within the M-lung can be adjusted to contract or expand the M-respiratory system across the full range of vital capacity, similar to the operation of a mechanical ventilator. Throughout the experiment, the acquired data points will be plotted in Figure 13 to form the P-V curve of the M-respiratory system. As shown in Figure 12, the M-lung is initially connected to a pressure-vacuum machine, and a series of progressively more negative pressure steps is applied, resulting in a gradual decrease in M-lung volume and, consequently, the overall volume of the M-respiratory system.

The following line of reasoning can conceptualize this process. The internal pressure of the M-lung and the intrapleural pressure act jointly to expand the M-lung. In contrast, the elastic resistance of the M-lung wall opposes these forces and strives to retract the M-lung. With each step of negative pressure applied by the pump, the expanding forces diminish, ultimately becoming weaker than the retraction force, leading to the M-lung's retraction. The elastic recoil of the M-lung, which seeks to expand the M-pleural space, makes the intrapleural pressure more negative, thereby pulling the M-chest wall inward in a movement that follows the M-lung's retraction. It is important to note that a new dynamic elastic equilibrium is established and actively sustained by the vacuum-pressure machine at each step under static conditions.

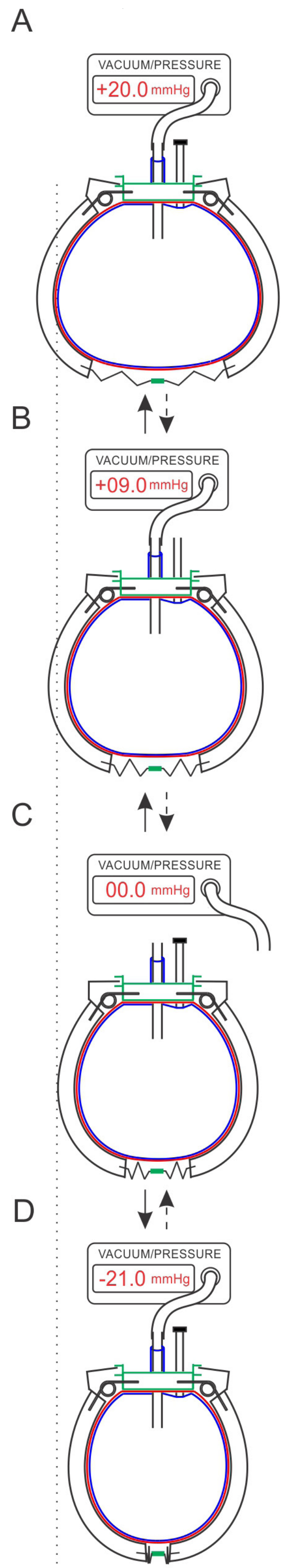


Figure 12. *The whole movement excursion of the M-respiratory system model. A to D. The sequence from maximal expansion (A) to maximal retraction (D). The elastic equilibrium point is at B.*

The negative pressure application procedure continues until a residual volume of approximately 1.1 liters is reached. This lower limit is imposed by structural factors, preventing further contractions, as seen physiologically. The volume change is recorded at each pressure step and under static conditions, resulting in a new data point on the P-V curve, as illustrated in Figure 13. An external force from the pressure-vacuum machine is necessary to maintain the M-respiratory system at residual volume—or any volume distinct from the residual functional capacity. In physiological conditions, this force corresponds to the action of the expiratory muscles during contraction.

To illustrate the active retraction of the M-respiratory system from its elastic resting point and to facilitate the construction of the P-V curve, we can decouple the pressure vacuum machine. This action allows the M-respiratory system to relax elastically and passively expand to its elastic equilibrium point (as represented in Figure 12C and point 'a' in Figure 13). Following the established protocol, we restart from this point and couple the pressure-vacuum machine, initiating a sequence of positive pressure steps. At each step, the M-lung undergoes expansion, which prompts the M-chest wall to expand concurrently. This dynamic can be understood by considering the two forces that maintain the M-chest wall in a static position at resting functional capacity: the intrapleural pressure, which draws it inward, and the elastic force stored within its elastic components, which counters intrapleural pressure. In the executed protocol, the expansion of the M-lung compresses the pleural space, resulting in a P_{IP} that becomes smaller than the elastic expansion force of the M-chest wall. Consequently, the M-chest wall expands in synchrony with the M-lung's expansion.

Under static conditions, the intrapleural pressure becomes progressively less negative with each positive pressure step applied to the M-lung until it reaches zero or equilibrates with atmospheric pressure (P_{AT}). At this juncture, the elastic pressure of the M-chest wall also becomes zero, indicating that it neither expands nor

continued on next page

compresses the M-pleural space. In this state, the M-chest wall is at its elastic equilibrium. At this pressure and volume level, the P-V relationship of the M-respiratory system coincides with that of the isolated M-lung, as the contribution of the M-chest wall to pressure generation is effectively null (see Figure 12B and point 'c' in Figure 13). As further positive pressure steps are applied, the expansion of the M-lung compresses the M-pleural space, which then pushes the M-chest wall outward, prompting it to expand actively. Additional positive pressure steps continue until maximal pulmonary capacity is reached. The structural constraints of the M-chest wall determine this point (refer to Figure 12A and point 'd' in Figure 13). With each pressure increment applied to the M-lung, P_{IP} becomes more positive because the M-chest wall attempts to retract toward its elastic equilibrium point, thereby compressing the M-pleural space against the outer surface of the M-lung. The vacuum-pressure pump work (or the inspiratory muscle contraction) maintains the system expanded.

The five points derived from the simulated theoretical experiment depicted in Figure 12 and shown in Figure 13 enable us to illustrate the M-respiratory system P-V relationship, which aligns with the P-V curves found in the Rahn diagram. The line fitting of the data, illustrated in green in Figure 13, was conducted manually. Its sigmoidal shape reflects the combined characteristics of the M-lung and M-chest wall curves, underscoring the mechanical connectivity between these two structures. For quantifying and characterizing respiratory P-V curves, the literature offers sigmoidal equations that fit the data with remarkable precision, producing robust and physiologically relevant parameters (Ferreira et al., 2011; Venegas et al., 1998). It is essential to highlight that the M-respiratory system facilitates the attainment of a static condition (no flow) with each increment in applied pressure, thereby eliminating pressure changes caused by resistive elements. This static condition is challenging to achieve under physiological circumstances, resulting in what is referred to as a 'quasi-static' P-V curve (Harris, 2005).

Examining the compliance and mechanical properties of the M-respiratory system reveals essential insights into its functional dynamics, mirroring observations from the *S-F* curves. The following points encapsulate the key findings related to the P-V relationships of M-respiratory system, M-lung, and M-chest wall:

i. The compliance of the M-respiratory system is notably less than that of the M-lung or M-chest wall when assessed individually.

ii. The P-V curve representing the M-respiratory system reflects the sum of the P-V curves of both the M-lung and M-chest wall at each static volume. This implies that the pressure necessary for the movement of the M-respiratory system corresponds to the sum of the pressures required for the M-lung and M-chest wall, demonstrating increased stiffness in the system.

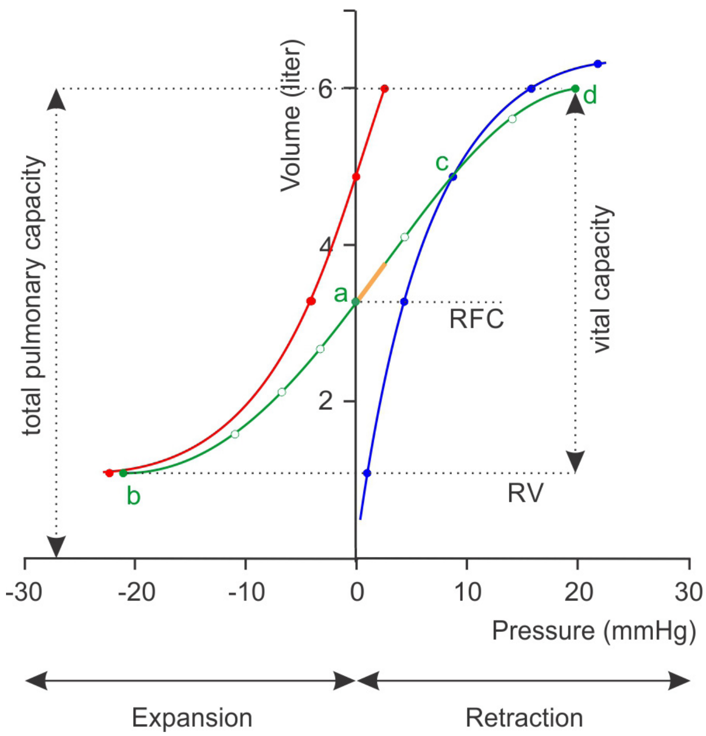


Figure 13. *Reproduction of Rahn's diagram. The sigmoidal P-V curve of the M-respiratory system (green) is derived from the P-V curves of the M-lung (blue) and M-chest wall (red). RFC = residual functional capacity. RV = residual volume. The small orange segment represents the tidal volume (~0.5 L), illustrating that only a small fraction of vital capacity is utilized during resting breathing.*

iii. At the resting functional capacity, the M-respiratory system attains a state of elastic equilibrium, where the elastic pressures of its components are balanced, leading to a zero net pressure. This equilibrium occurs because the M-chest wall's elastic expansion pressure matches the M-lung's elastic retraction pressure. Consequently, external force—physiologically with respiratory muscles—is required to alter the M-respiratory system state through compression or expansion.

continued on next page

iv. At a volume of about 4.8 liters, it is observed that only the M-lung is capable of storing elastic energy while attempting to recoil. In contrast, the M-chest wall remains in a balanced state of elastic equilibrium.

v. When the M-respiratory system is compressed below the residual functional capacity, either mechanically, through the pressure-vacuum apparatus, or physiologically through expiratory muscle activity, it accumulates elastic energy for subsequent expansion. Thus, upon disconnection from the pressure-vacuum apparatus, the M-respiratory system reverts to elastic equilibrium, expanding back to the resting functional capacity. Conversely, if the M-respiratory system is broadened beyond the residual functional capacity, it also stores elastic energy, promoting a spontaneous retraction towards its equilibrium position.

vi. For volumes below 4.8 liters, the M-lung and M-chest wall function antagonistically, with the M-chest wall promoting expansion while the M-lung facilitates retraction. However, at volumes exceeding 4.8 liters, both components act synergistically in the same direction, contributing to retraction.

vii. To emphasize the scale of lung capacity in the diagram, the tidal volume, approximately 0.5 liters during resting breathing, is depicted as a small orange segment. This visual cue underscores that only a minor fraction of the vital capacity is utilized at rest, reinforcing the notion that most of the lung's elastic reserve remains quiescent under basal conditions.

Discussion and Conclusions

The elastic characteristics of the lungs and chest wall, which are pivotal in governing respiratory mechanics, are deeply intertwined with the respiratory system's P-V relationship. A key educational resource in this domain is the diagram formulated by Rahn et al. (1946). This diagram illustrates three distinct curves that encapsulate the elastic behavior of the lungs and chest wall, elucidating the complex interplay between these anatomical structures. To enhance the educational value of the Rahn diagram, we propose a didactic model to visualize the construction of these curves and their interrelationships. The model was designed as a pedagogical representation illustrated by a set of 13 figures intended for sequential presentation and discussion in the classroom. The primary objective is to facilitate the interpretation of the entire diagram and its intricate physiological concepts, thereby enhancing comprehension of ventilatory mechanics. This approach is structured around three fundamental points: First, a theoretical framework that encompasses the essential elements of the respiratory system's elastic behavior is introduced and gradually developed throughout the text, in parallel with the presentation of the figures. This content is designed to fill knowledge gaps and mitigate misunderstandings. Second, the subject matter is presented in two sections, commencing with more straightforward and familiar concepts before progressing to more complex physiological principles. The first session focuses on linear elastic behavior in springs, establishing a foundation for the second session, which addresses the sigmoid relationship inherent in the elastic structures of the respiratory system. Each session is complemented by mechanical models and abstract experiments that empower students to construct compliance curves. Third, including thirteen figures alongside textual explanations demystifies complex concepts, making them more accessible and engaging for students.

The approach is designed to progressively build students' understanding of P-V relationships in the respiratory system. Analyzing and interpreting the diagram challenges students to apply their knowledge of alveolar ventilation, compliance, and respiratory pressures. The sequence of didactic models and abstract experiments guides students through the step-by-step construction of compliance curves, illustrating how individual components (lungs and chest wall) contribute to the total respiratory system behavior. By following this structured approach, students gradually reengage with key concepts of respiratory mechanics, allowing them to visually and conceptually connect these principles to the Rahn diagram, making its interpretation and application more intuitive.

continued on next page

While a formal analytical assessment of the practical effects of this didactic approach in the classroom is lacking, my observations from teaching experience indicate that it holds significant educational value. Over the years, my perception from classroom experience on applying this educational strategy has been that:

i. The Rahn diagram serves as a final step in the study of breathing mechanics, effectively synthesizing, consolidating, and complementing the knowledge previously acquired.

ii. Students are able to visualize and relate the models to the elastic structures of the respiratory system. They tend to respond positively to this approach, actively engaging in the theoretical experiments and drawing conclusions based on the simulated data, thereby deepening their understanding of compliance curves.

iii. The sequential presentation of the figures—detailing the models and experimental protocols—promotes student engagement in the constructive execution of simulated data acquisition.

iv. The step-by-step construction of each curve through theoretical simulations highlights the underlying mechanisms and their interactions, facilitating the analysis and interpretation of the Rahn diagram.

v. This approach enables me to communicate ideas more clearly and effectively, ultimately enhancing the teaching–learning process.

I hope this didactic strategy may also prove useful to others.

Acknowledgments

I thank Clariswaldo Baptista and M. Marly R. Baptista for their unwavering support, guidance, and encouragement.

About the Author

Vander Baptista is an associate professor in the Department of Physiology at the Federal University of Santa Catarina (UFSC), where he teaches courses in physiology and biophysics.

Literature Cited

Andreassen, S., Steimle, K.L., Mogensen, M.L., Bernardino de la Serna, J., Rees, S. E., Karbing, D.S., (1985). The effect of tissue elastic properties and surfactant on alveolar stability. *Journal of Applied Physiology, 109*(5), 1369-1377. https://doi.org/10.1152/japplphysiol.00844.2009

Angra, A. & Gardner, S.M. (2018). The graph rubric: Development of a teaching, learning, and research tool. *CBE Life Sciences Education*. *17*(4), 1-18. https://doi.org/10.1187/cbe.18-01-0007

Baptista, V. (2010). A qualitative analogy for respiratory mechanics. *Advances in Physiology Education. 34*(4), 239-243. https://doi.org/10.1152/advan.00014.2010

Brochard, L. (2006). What is a pressure-volume curve? *Critical Care 10*, Article e156. https://doi.org/10.1186/cc5002

Dostal, P. & Dostalova, V. (2023). Practical aspects of esophageal pressure monitoring in patients with acute respiratory distress syndrome. *Journal of Personalized Medicine. 13*(1), Article e136. https://doi.org/10.3390/jpm13010136

Ferreira, J.C., Benseñor, F.E.M., Rocha, M.J.J., Salge, J.M., Harris, R.S., Malhotra, A., et al. (2011). A sigmoidal fit for pressure-volume curves of idiopathic pulmonary fibrosis patients on mechanical ventilation: clinical implications. *Clinics*. *66*(7), 1157-1163. https://doi.org/10.1590/S1807-59322011000700006

Gibson, G.B. (2020). Lung and chest wall elasticity. In R. L. Maynard, S. J. Pearce, B. Nemery, P. D. Wagner, & B. G Cooper (Eds.) *Cotes' Lung Function* (pp. 187-201). John Wiley & Sons Ltd. https://doi.org/10.1002/9781118597309

Glazer, N. (2011). Challenges with graph interpretation: A review of the literature. *Studies in Science Education, 47*(2), 183–210. https://doi.org/10.1080/03057267.2011.605307

Harris, R.S. (2005). Pressure-volume curves of the respiratory system. *Respiratory Care, 50*(1), 78–99. PMID: 15636647

Maggiore, S.M., & Brochard, L. (2001). Pressure-volume curve: Methods and meaning. *Minerva Anestesiologica, 67*(4), 228–237. PMID: 11376515.

Mead, J., Takishima, T., & Leith, D. (1970). Stress distribution in lungs: A model of pulmonary elasticity. *Journal of Applied Physiology, 28*(5), 596–608. https://doi.org/10.1152/jappl.1970.28.5.596

Mitzner, W. (2011). Mechanics of the lung in the 20th century. *Comprehensive Physiology, 1*(4), 2009–2027. https://doi.org/10.1002/cphy.c100067

continued on next page

Rahn, H., Otis, A.B., Chadwick, L.E., & Fenn, W.O. (1946). The pressure-volume diagram of the thorax and lung. *American Journal of Physiology, 146*(2), 161-178. https://doi.org/10.1152/ajplegacy.1946.146.2.161

Ranieri, V. M., & Slutsky, A. S. (1999). Respiratory physiology and acute lung injury: The miracle of Lazarus. *Intensive Care Medicine, 25*, 1040–1043. https://doi.org/10.1007/s001340051010

Stein, R.S. (1958). On the inflating of balloons. *Journal of Chemical Education, 35*(4), 203–204. https://doi.org/10.1021/ed035p203

Terragni, P.P., Rosboch, G.L., Lisi, A., Viale, A.G., & Ranieri, V.M. (2003). How respiratory system mechanics may help in minimizing ventilator-induced lung injury in ARDS patients. *European Respiratory Journal, 22*(42 suppl), 15s–21s. https://doi.org/10.1183/09031936.03.00420303

Toshima, M., Ohtani, Y., & Ohtani, O. (2004). Three-dimensional architecture of elastin and collagen fibers networks in the human and rat lung. *Archives of Histology and Cytology, 67*(1), 31–40. https://doi.org/10.1679/aohc.67.31

Venegas, J.G., Harris, R.S., & Simon, B.A. (1998). A comprehensive equation for the pulmonary pressure-volume curve. *Journal of Applied Physiology, 84*(1), 389–395. https://doi.org/10.1152/jappl.1998.84.1.389

West, J.B. (2008). Challenges in teaching the mechanics of breathing to medical and graduate students. *Advances in Physiology Education, 32*(3), 177–184. https://doi.org/10.1152/advan.90146.2008

Back to TOC

PERSPECTIVES ON TEACHING

HAPS EDUCATOR
Journal of the Human Anatomy and Physiology Society

Volume 29, Issue 2 | August, 2025
https://doi.org/10.21692/haps.2025.010

Renal Physiology: Classroom Activity to Review Basic Functions

Lindsey A. Pleasant, BA [1], Josie G. Ayers, BS[2], Catharine C. Whiting, PhD[2], and Mary D. Kinkel, PhD[1]

[1]Department of Biology, Appalachian State University, Boone, NC, USA

[2]Department of Biology, University of North Georgia, Oakwood, GA, USA

Corresponding author: kinkelmd@appstate.edu

Abstract

This paper describes a renal physiology review activity that can be completed during one class session. The activity is a tactile simulation in which colored beads represent various components of the blood, glomerular filtrate, tubular fluid, and interstitial fluid. Students simulate major functions of the nephron by moving beads between compartments represented by simple diagrams. The diagrams include the glomerulus and Bowman's capsule, proximal tubule with peritubular capillaries, and the loop of Henle with vasa recta. Students transport the beads along the nephron and, in each region, make decisions regarding what is transported, beginning with filtration and reabsorption in early cortical segments. In the medulla, students consider the functions of the descending versus ascending limbs in both the loop of Henle and the vasa recta. Here, the focus is on mechanisms of water and salt transport that set up and maintain the osmotic gradient in the medullary interstitium as well as concentrate and dilute the urine. As students proceed along the nephron they answer review questions. They are also given a short clinical scenario and asked to apply their knowledge of transport across membranes to predict the effects of loop diuretics on urine formation. Finally, a set of additional review questions is included that students can complete outside of class. This active learning approach to teaching renal physiology provides students the opportunity to develop higher-order cognitive skills as they move beyond rote memorization and basic understanding to application and evaluation. https://doi.org/10.21692/haps.2025.010

Key words: filtration, reabsorption, countercurrent mechanism, juxtamedullary nephron, active learning

Introduction

Background

While renal physiology is a challenging topic to master, its study provides opportunities for students to apply their knowledge of basic physiological processes that are shared among organ systems. Before encountering the renal system, students will have ideally completed units covering basic principles (e.g., flow down gradients, membrane transport) and the cardiovascular system (e.g., capillary structure and function, Starling forces, blood composition). General physiology textbooks (e.g., Fox & Rompolski, 2022; Sherwood, 2016; Silverthorn, 2019; Widmaier et al., 2023) present almost all other organ systems before the renal system. Thus, by following a typical topic order, students can build a foundation that supports mastering the fundamental mechanisms of the renal system. Indeed, studying the renal system provides an opportunity for integrating knowledge across the multiple organs systems that interact with the kidneys, particularly for maintaining fluid and electrolyte balance.

However, despite being familiar with relevant mechanisms and core concepts, students can have difficulty transferring knowledge from one system to another, as discussed in Goodman et al. (2018) and Michael et al. (2017). For example, it is relatively easy to understand the Starling forces that influence filtration and reabsorption in systemic capillary beds. However, the glomerulus and peritubular capillaries are not quite like systemic capillary beds. The renal vessels often seem like strange new territory, and it is not necessarily straightforward to transfer knowledge of systemic capillary beds to understand glomerular filtration

continued on next page

or peritubular capillary reabsorption. Similarly, students with a solid foundation in transport across membranes, including osmosis and cotransport, may still find themselves at a loss when first faced with the countercurrent mechanism and the vertical osmotic gradient in the renal medulla. Hands-on, active learning approaches may be particularly useful for overcoming these difficulties.

Despite the benefits of active learning in STEM (Freeman et al., 2014), there are few published classroom activities for renal physiology suitable for the undergraduate classroom. Of note, Hull (2016) developed a hands-on simulation using a game board and beads for modeling renal clearance, Dirks-Naylor (2016) described an activity in which students draw and label the nephron for learning the locations and basic functions of renal transporters, and Giffen and Carvalho (2015) described using paper diagrams of specific renal cell types, transporters, channels, and plasma components to move material at the basolateral or luminal cell surfaces. Along these lines, we have developed an in-class activity that guides students through a hands-on simulation of several basic nephron functions including filtration, reabsorption and urine concentration and dilution. The complete instructions and diagrams are included in the appendices. The activity uses beads and diagrams to simulate renal physiology and was inspired by activities from an undergraduate anatomy and physiology lab manual (Whiting, 2019). The overall objective of the activity described here is to reinforce understanding of basic mechanisms essential to the main function of the kidneys, which is to maintain, on a continuous basis, the extracellular fluid volume and the solute composition of that fluid (Koeppen & Stanton, 2019; Mulroney & Myers, 2025).

Description of the Activity

This activity was designed for a junior-level physiology course for biology majors, as a review of several key concepts for the renal system. In this course, the activity is scheduled for the last class of a unit on the renal system and corresponds to the last day of class for the semester. Students have already encountered all material in the activity except for a final set of questions in which they are asked to use their existing knowledge to predict the effect of loop diuretics. Thus, the activity is used as a self-test and review before the final exam.

The basic strategy of the activity is to use prompts that guide students through a simulation of renal filtration, reabsorption, and concentration and dilution of the urine. Brief summaries of mechanisms are included throughout, to provide context. Four simple diagrams represent specific regions of the juxtamedullary nephron. Two dishes of beads represent fluid compartments: the blood and, variously, the filtrate or tubular fluid. Different bead colors and sizes represent specific components of the fluids, including blood cells, plasma proteins, water, sodium, and organic molecules.

As the students work through the guided steps and the diagrams, the beads are transported into or out of the fluid compartments to visualize specific renal functions. For example, students decide which beads to filter into Bowman's space and which to reabsorb from the proximal tubule into the peritubular capillaries. In the renal medulla, they decide which limb of the loop of Henle has aquaporins versus Na-K-2Cl (NKCC2) transporters. Beads representing water or ions are subsequently moved out of the tubular fluid to build the vertical osmotic gradient in the medulla. Students also move beads (water and ions) between the medullary interstitium and the vasa recta. Thus, the review activity provides a tactile simulation of filtration, reabsorption, countercurrent multiplication, and countercurrent exchange.

After students simulate these mechanisms, they answer written review questions. Most questions probe basic recall from the lecture material while others require application or interpretation. These questions can be answered with a partner during or after class, as time permits. Alternatively, a group of students could use the think-pair-share method during class. The activity requires ~40 minutes, including distributing the materials and orienting the students, and can be incorporated into either a lecture or lab class.

The activity consists of three main parts:

1. *Filtration and reabsorption*
2. *Urine concentration in the juxtamedullary nephron*
3. *Additional review questions*

In Part 1, *Filtration and reabsorption*, students use a diagram showing the glomerulus, Bowman's space, and the proximal tubule. A dish, holding 59 beads representing the blood, is placed in the glomerulus. An empty dish is placed in Bowman's space to hold the filtrate. The students answer basic questions including *Which blood components are freely filtered into Bowman's space?* and *What is normally not filtered?* As they answer these questions, they are prompted to move the beads representing filtered components from the "blood" to the "filtrate" (Figure 1A). Next, they're prompted to move the "filtrate" dish to a diagram representing the proximal tubule while the "blood" dish is moved to the adjacent peritubular capillaries. In the proximal tubule, students move beads from the "tubular fluid" back to the "blood" to simulate reabsorption (Figure 1B).

continued on next page

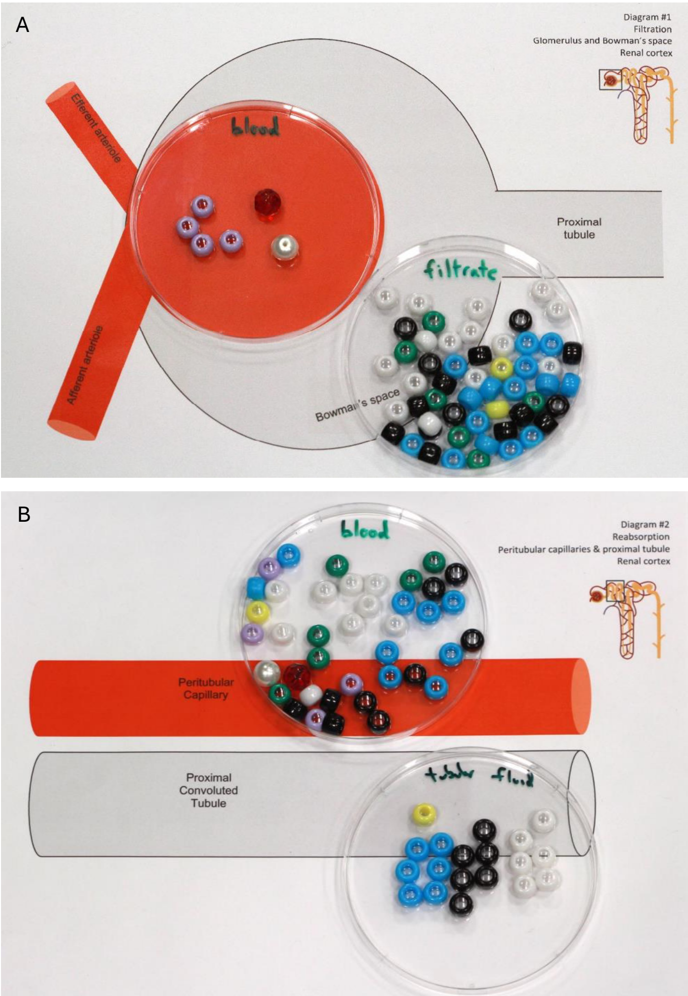


Figure 1. *Bulk flow in the renal cortex. Bead colors indicate water (blue), sodium (ivory), other ions (black), plasma proteins (purple), urea (yellow), red blood cells (red, large), white blood cells (white, large).* ***A)*** *Filtration (Bowman's space and the proximal tubule). In Bowman's space, all blood components that can be filtered are moved to the "filtrate" dish: water, nutrients, urea, sodium, and other ions. In the glomerulus, the "blood" dish contains the non-filtered components: red blood cells, white blood cells, and plasma proteins.* ***B)*** *Reabsorption representing the peritubular capillaries and proximal tubule. The "blood" dish holds the components of the filtrate that are reabsorbed. To simulate reabsorption, all beads representing nutrients have been transferred from the "tubular fluid" to the "blood" along with a proportion of the ions, water, and urea.*

In Part 2 of the activity, *Urine concentration in the juxtamedullary nephron,* students are guided to build the vertical osmotic gradient in the renal medulla. This uses a diagram showing the long loop of Henle and the interstitium of the renal cortex and medulla. Building the osmotic gradient is done by first labeling the diagram, then working with the bead dishes to transport water and Na+ (Figure 2). Using the diagram, students start by considering which limb of the loop of Henle transports Na+ versus water and labeling the appropriate limb with NKCC2 transporters or aquaporins. Next, they are guided through a series of steps to label the vertical osmotic gradient in the renal medulla, and then they label the loop of Henle to show changes in tubular fluid osmolarity as water and Na+ are transported. A labeled diagram is shown in Figure 2A.

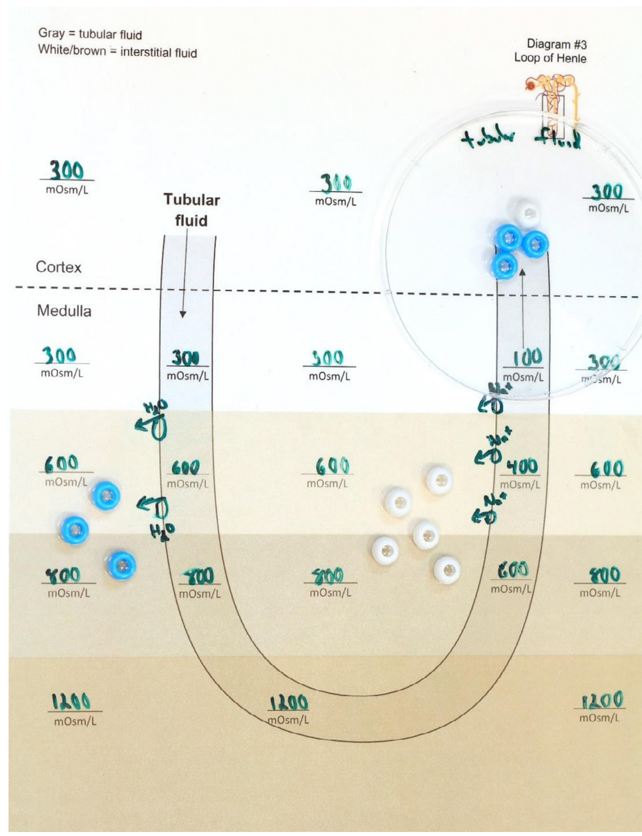


Figure 2. *Reduction of urine volume by the loop of Henle. For simplicity, the beads representing other ions and urea have been removed. Bead colors indicate water (blue), sodium (ivory). Students first label the diagram to show Na^+ transporters and aquaporins on the appropriate limbs. Then, they label the osmolarity of the medullary interstitium and the tubular fluid. Beads are moved to simulate water transport in the descending limb and sodium transport in the ascending limb. The "tubular fluid" dish shown above represents the last step in this part of the activity, after the tubular fluid has been concentrated and then diluted.*

continued on next page

To concentrate the urine, the "tubular fluid" dish is moved down the descending limb of the loop of Henle. Until this point, water and Na+ have moved proportionally by bulk flow. Here in the medulla, the students simulate the effects of transporting water in the absence of Na+ transport. The "tubular fluid" dish, as it enters the medulla, contains 6 "water" beads and 6 "Na+" beads. The students are prompted to transfer half of the "water" beads from the "tubular fluid" to the surrounding medullary interstitial fluid. As the students transport the water, they visualize that the osmolarity of the "tubular fluid" has increased. At the bottom of the loop of Henle, the "tubular fluid" dish contains 3 "water" beads but 6 "Na+" beads. This provides a visual demonstration that the presence of aquaporins in the absence of NKCC2 transporters results in a tubular fluid osmolarity matching that of the surrounding interstitial fluid.

Next, the students dilute the urine by moving the dish up the ascending limb of the nephron loop, to return to the cortex. Along the way, they are prompted to transfer 5 of the 6 "Na+" beads from the dish into the medullary interstitial fluid. As the dish enters the cortex it now has only 3 "water" beads and 1 "Na+" bead (Figure 2B). Because the dish began with 12 beads (6 "water" and 6 "Na+") before traveling along the loop, the students can visualize that the tubular fluid has been both concentrated and diluted by the loop of Henle.

To simulate countercurrent exchange, the students overlay a transparent diagram of the vasa recta so that the blood vessels are adjacent to the loop of Henle (Figure 3). They then work through prompts to simulate reabsorption of water and Na+ from the interstitial fluid. By the end of this series of prompts, students can visualize how osmolarity of the plasma is restored to 300 mOsm/L, as blood flows through the vasa recta from the bottom of the medulla into the cortex. This part of the activity concludes with "clinical correlations" in which students are asked to consider the effects of loop diuretics on the medullary interstitium. Finally, Part 3 *Additional review questions*, provides a set of 12 questions that students can complete outside of class.

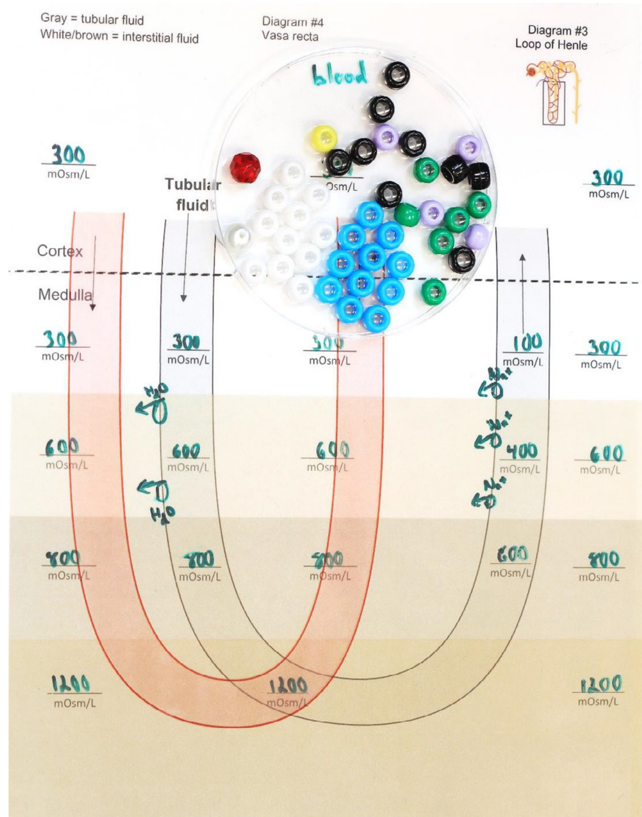


Figure 3. *Countercurrent exchange (vasa recta). Bead colors indicate water (blue), sodium (ivory), other ions (black), plasma proteins (purple), urea (yellow), red blood cells (red, large), white blood cells (white, large). As students move the "blood" dish through the loop, they transfer beads to simulate changes in osmolarity. When the dish is returned to the renal cortex via the ascending vasa recta, its proportion of water and sodium beads represents restoration of the blood to 300 mOsm/L.*

The purpose of this activity is to review renal filtration, reabsorption, and urine concentration using simple diagrams and models of the blood and other fluid compartments. The strategy is to use a tactile, active learning approach combined with review-type questions to reinforce key concepts and provide practice answering basic questions related to renal physiology. With regard to learning objectives, by the end of the activity, students should be able to:

- Apply the principles of bulk flow to explain which components of the blood are filtered
- Explain which components of the filtrate are reabsorbed from the proximal tubule
- Distinguish between the functions of the descending and ascending limbs of the loop of Henle
- Describe how the countercurrent mechanism establishes and maintains the medullary vertical osmotic gradient
- Predict the effects of loop diuretics on urine formation

continued on next page

Methods

Instructions/Review Questions Packet

There should be enough sets of instructions so that each student gets a packet (Appendix 1). The packet is designed to be self-explanatory and could be used for independent work. However, students are asked to work in pairs (or groups of three if materials are limited) and discuss the answers with their partners. The students keep the instruction packet to help review the material outside of class. Some orientation from the instructor is useful before beginning. An instructor answer key with notes shows example answers for the review questions (Appendix 6).

Diagrams

Each pair of students uses one of each diagram. There are four 8.5 x 11″ diagrams (Appendices 2-5) of segments of the nephron: Diagram 1, *Filtration* (glomerulus and Bowman's space, renal cortex); Diagram 2, *Reabsorption* (peritubular capillaries and proximal tubule, renal cortex); Diagram 3, *Loop of Henle*; Diagram 4, *Vasa recta*. Diagrams 1-3 are laminated or put into sheet protectors. Diagram 4, *Vasa recta* is a color transparency. The transparency is overlaid on Diagram 3, *Loop of Henle* to demonstrate countercurrent exchange. The diagrams are collected afterwards to be reused.

Bead Dishes

Each pair of students uses a 100 mm Petri dish containing 59 beads representing various fluids. Each dish contains 1 large red bead (representing red blood cells) and 1 large white bead (representing white blood cells). The remaining smaller beads include 2 yellow (urea), 4 purple (plasma proteins), 6 green (organic nutrients), 15 blue (water molecules), 15 ivory (Na+), 15 black (other ions). Students are asked to confirm that they have these beads in their dish before beginning the activity. Extra beads are on hand for backup. The beads were obtained from a craft store.

Additional Supplies

Each pair of students uses an erasable marker and a 50 mL plastic beaker. The marker is used to label the diagrams and Petri dishes. The beaker is used to hold excess beads in part 2 of the review activity.

Discussion

This activity is intended to be a focused review of basic functions of the nephron. To design the activity, deliberate decisions were made regarding the scope, level of detail, and terminology. In its first iteration, the scope was broader, to include functions of the distal tubule and collecting duct, acid-base balance, and aspects of the renin-angiotensin-aldosterone system (RAAS). This broader version was beta-tested in 2019 by at least 28 students who provided detailed written feedback. The students generally indicated that the most useful modules were the ones included here: Part 1. *Filtration and reabsorption*, and Part 2. *Urine concentration in the juxtamedullary nephron*. Students reported that they had most difficulty conceptually with those early events and that it was eye-opening to physically move the beads and observe the effects. Further, students generally indicated that once the topics in the early segments were mastered, it was relatively easy to understand the functions of the later segments. Some students indicated that while renal functions related to, for example, the RAAS and acid-base balance are challenging, they found that a bead-based activity was less useful than other study resources. Based on this feedback, the review activity was revised to focus on the renal functions presented here.

To decide on the level of detail, the review was initially drafted to be consistent with the course textbook (Sherwood, 2016), lecture material, and the HAPS learning objectives (Human Anatomy and Physiology Society, 2019). The lecture material was informed by multiple sources including the primary literature. By necessity, textbooks omit many details and provide relatively superficial descriptions of complex mechanisms. Moreover, textbooks tend to not point out gaps in knowledge, areas of ongoing research, or that some mechanisms are presented as established knowledge rather than as models that are expected to be refined as we learn more. Walter B. Cannon noted that the "… textbook is a selection of facts that the author has regarded as important and is not the whole story and, further, that the summary set forth in a textbook may neglect or smooth out discrepancies" (Cannon, 1945/1968, p 83). As a review activity in an undergraduate course, the materials presented here likewise omit many details and provide superficial descriptions. However, the materials may be easily modified to change the scope, details, and review questions that are included.

The most challenging aspects of designing the activity included deciding how best to describe the vertical osmotic gradient in the renal medulla and how to describe and draw the vasa recta. Textbooks often show the renal medulla with a gradient of 300 mOsm/L closest to the cortex and 1200 or 1400 mOsm/L at the deepest point (Hall & Hall, 2012; Mulroney & Myers, 2025; Sherwood, 2016; Silverthorn, 2019; Widmaier et al., 2023). For this activity, we chose values that

continued on next page

are consistent with the course textbook (Sherwood, 2016), to avoid potential confusion and with the idea that the instructor can emphasize that these values are examples and are not constant. To support in-depth discussion on this topic we have found it useful to consult some of the vast literature on urine concentration (e.g., Danztler et al., 2014, Jamison, 2014; Kuang, 2023; Nawata & Pannabecker, 2018; Sands & Layton 2013; Sands et al., 2020).

We also considered whether (and how) to include the contribution of urea transport and cycling to the osmotic gradient. A complete consideration would require adding a diagram of the collecting duct. Based on feedback of the first iteration of the activity, which included the collecting duct, we predicted that students would not need a bead-based activity to simulate urea transport. Further, as the role of urea in establishing or maintaining the vertical osmotic gradient is either not mentioned or not covered in detail in any of the undergraduate physiology textbooks that we consulted, we chose to omit this topic from the review materials (Fox & Rompolski, 2022; Sherwood, 2016; Silverthorn, 2019; Widmaier et al., 2023). Therefore, we confined urea handling to filtration in Bowman's capsule and reabsorption in the proximal tubule.

We also found it somewhat challenging to develop a simplified diagram of the vasa recta. It is difficult to represent these extensive vessels in a realistic manner without obscuring the loop of Henle. Most textbooks represent the vasa recta associated with one nephron as a simple loop with few to many anastomoses. The vessels are often colored entirely red or entirely blue but figures sometimes use red for the descending vasa recta, blue for the ascending vasa recta, and a red/blue/purple mix for the anastomoses between the straight vessels.

To contrast with textbook figures, lecture material may supplement those figures by using images such as found in Kriz & Kaiisling (2013), Pallone and Cao (2013), or Navar et al. (2020). For modeling the vasa recta here, we considered various possibilities and ultimately followed the simple red loop as shown in Figure 31.5 in Whiting (2019). Using a simple loop for the vasa recta on a transparency allows a clear view of the underlying loop of Henle and the surrounding interstitial space. Further, we anticipate that before encountering the review activity materials, students will have completed a unit on renal physiology and, as a result, they will have already seen multiple images of the kidney and nephron including scanning electron micrographs. Therefore, students should be generally prepared to interpret the highly simplified diagrams used here.

Another difficulty with modeling the vasa recta lies with inconsistencies between textbooks and the primary literature regarding what type of vessels these are. In textbooks, the vasa recta are often referred to as capillaries (e.g., Widmaier et al., 2023) or peritubular capillaries (Hall & Hall, 2021; Sherwood, 2016; Silverthorn, 2019). However, the ultrastructure of the vasa recta is distinct from that of the cortical peritubular capillaries. These vessels share characteristics of capillaries but can be interpreted as structurally unique, as reviewed by multiple groups (Kriz & Kaissling, 2013; Navar et al., 2020; Pallone & Cao, 2013). For this review activity, we have avoided referring to these vessels as capillaries and have simply called them the vasa recta.

In general, we have used traditional terminology to describe the nephron, despite the HAPS anatomy and physiology learning outcomes referencing *Terminologia Anatomica* 2nd Edition to encourage removing eponyms (Human Anatomy and Physiology Society, 2019). Terms such as "Bowman's capsule" and "loop of Henle" are firmly established in the literature and remain in common use among researchers and in comprehensive nephrology texts such as Verlander & Clapp (2020). It may be safe to assume that students who learn the eponymous terms will have little trouble recognizing "glomerular capsule" and "nephron loop", and instructors can be careful to use the alternative terms interchangeably. Moreover, the traditional eponymous terms may lead interested students to discover the fascinating historical literature such as Bowman's original descriptions (Bowman, 1842) or the more recent (and more accessible) reviews of the loop of Henle by Koulouridis & Koulouridis (2014), Marcoux et al. (2022), or Morel (1999).

The materials presented here have been used for several years for an in-class review session and students are consistently highly engaged. Because of positive student feedback and improved performance on exams, this renal physiology activity has become a regular feature of the course. The following are representative student comments from an in-class survey:

> *"I thought [the review questions] were useful for making connections and placing the information in the context of the big picture."*
>
> *"To do the activity and 'see' it happening it really helps my understanding."*
>
> *"I found it fun moving the beads around and simulating how water and sodium move within the loop, very informative and detailed."*
>
> *"I like that it makes you think about it independently instead of telling you what to do. It makes you actually process what's happening …"*

continued on next page

In sum, this in-class activity provides a review of the most basic functions of the kidney that relate to the kidney's role in maintaining an appropriate volume and solute composition of the extracellular fluid. Reviewing filtration, reabsorption, and the concentration of urine is intended to provide a firm foundation for understanding more complex renal processes.

Acknowledgments

Sarah Arrington kindly directed one of us (MK) to the published renal physiology laboratory activity that led us to design our in-class activity. The nephron insets for Diagrams 1-3 were created in BioRender: Pleasant, L. (2024) https://BioRender.com/b69e509 (Diagram 1); https://BioRender.com/u96a010 (Diagram 2); https://biorender.com/u64b812 (Diagram 3).
We thank the physiology students at Appalachian State University for evaluating an earlier version of the kidney review materials and providing substantive feedback that helped us make improvements. We also thank the anonymous reviewers for suggestions that greatly improved the manuscript.

About the Authors

Lindsey Pleasant is a graduate of Appalachian State University with a BA in biology. Josie Ayers is a summa cum laude graduate of the University of North Georgia with a BS in biology (chemistry minor), and will be attending graduate school at Augusta University. Catharine Whiting is a professor of Biology at the University of North Georgia. Mary Kinkel is an associate professor of Biology at Appalachian State University.

Literature Cited

Bowman, W. (1842). On the structure and use of the Malpighian bodies of the kidney, with observations on the circulation through that gland. *Philosophical Transactions of the Royal Society of London*, *132*(1842), 57–80. http://www.jstor.org/stable/108143

Cannon, W. B. (1968). *The way of an investigator. A scientist's experiences in medical research.* Hafner Publishing Company. (Original work published in 1945 by W. W. Norton & Company Inc.).

Dantzler, W. H., Layton, A. T., Layton, H. E., & Pannabecker, T. L. (2014). Urine-concentrating mechanism in the inner medulla: Function of the thin limbs of the loops of Henle. *Clinical Journal of the American Society of Nephrology*, *9*(10), 1781–1789. https://doi.org/10.2215/CJN.08750812

Dirks-Naylor A. J. (2016). An active learning exercise to facilitate understanding of nephron function: Anatomy and physiology of renal transporters. *Advances in Physiology Education*, *40*(4), 469–471. https://doi.org/10.1152/advan.00111.2016

Fox, S. I., & Rompolski, K. (2022). *Human physiology, 16th edition*. McGraw Hill LLC.

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, *111*(23), 8410–8415. https://doi.org/10.1073/pnas.1319030111

Giffen, Z. C., & Carvalho, H. (2015). Development of a manipulative for nephron physiology education. *Advances in Physiology Education*, *39*(1), 39–41. https://doi.org/10.1152/advan.00087.2013

Goodman, B. E., Barker, M. K., & Cooke, J. E. (2018). Best practices in active and student-centered learning in physiology classes. *Advances in Physiology Education*, *42*(3), 417-423. https://doi.org/10.1152/advan.00064.2018

Hall, J. E., & Hall, M. E. (2021). *Guyton and Hall textbook of medical physiology, 14th edition*. Elsevier, Inc.

Hull K. (2016). Renal clearance: Using an interactive activity to visualize a tricky concept. *Advances in Physiology Education*, *40*(4), 458–461. https://doi.org/10.1152/advan.00059.2016

Human Anatomy and Physiology Society. (2019). *HAPS A&P learning outcomes*. https://www.hapsweb.org/haps-learning-outcomes/haps-ap-learning-outcomes-los/

Jamison R. L. (2014). Resolving an 80-yr-old controversy: The beginning of the modern era of renal physiology. *Advances in Physiology Education*, *38*(4), 286-295. https://doi.org/10.1152/advan.00105.2014

Koeppen, B. M., & Stanton, B. A. (2019). *Renal physiology, 6th edition.* Elsevier Inc.

Koulouridis, E., & Koulouridis, I. (2014). The loop of Henle as the milestone of mammalian kidney concentrating ability: A historical review. *Acta Medico-Historica Adriatica*, *12*(2), 413–428.

Kriz, W., & Kaissling, B. (2013) Structural organization of the mammalian kidney. In R. J. Alpern, M. J. Caplan, O. W. Moe (Eds.), *Seldin and Giebisch's the kidney: Physiology and pathophysiology*, *5th edition*. (pp. 595-691), Academic Press.

Kuang S. Y. (2023). A better explanation of countercurrent multiplication in the formation of the corticopapillary osmotic gradient in the outer medulla. *Advances in Physiology Education*, *47*(3), 665-671. https://doi.org/10.1152/advan.00227.2022

Marcoux, A. A., Tremblay, L. E., Slimani, S., Fiola, M. J., Mac-Way, F., Haydock, L., et al. (2021). Anatomophysiology of the Henle's loop: Emphasis on the thick ascending limb. *Comprehensive Physiology*, *12*(1), 3119–3139. https://doi.org/10.1002/cphy.c210021

continued on next page

Michael, J., Cliff, W., McFarland, J., Modell, H., & Wright, A. (2017). *The core concepts of physiology: A new paradigm for teaching physiology*. Springer.

Morel, F. (1999). The loop of Henle, a turning-point in the history of kidney physiology, *Nephrology Dialysis Transplantation, 14*(10), 2510–2515. https://doi.org/10.1093/ndt/14.10.2510

Mulroney, S. E., & Myers, A. K. (2025). *Netter's essential physiology, 3rd edition*. Elsevier, Inc.

Navar, L. G., Maddox, D. A., & Munger, K. A. (2020). The renal circulations and glomerular filtration. In A. S. L. Yu, G. M. Chertow, V. A. Luyckx, P. A. Marsden, K. Skorecki, & M. W. Taal (Eds.), *Brenner & Rector's the kidney, 11th edition*. (pp. 80-114). Elsevier Inc.

Nawata, C. M., & Pannabecker, T. L. (2018). Mammalian urine concentration: A review of renal medullary architecture and membrane transporters. *Journal of Comparative Physiology. B*, *188*(6), 899-918. https://doi.org/10.1007/s00360-018-1164-3

Pallone, T. L., & Cao, C. (2013). Renal cortical and medullary microcirculations: Structure and function. In R. J. Alpern, M. J. Caplan, O. W. Moe (Eds.), *Seldin and Giebisch's the kidney: Physiology and pathophysiology*, *5th edition*. (pp. 803-857), Academic Press.

Sands, J. M., & Layton, H. E. (2013). The urine concentrating mechanism and urea transporters. In R. J. Alpern, M. J. Caplan, O. W. Moe (Eds.), *Seldin and Giebisch's the kidney: Physiology and pathophysiology*, *5th edition*. (pp. 1463-1510). Academic Press.

Sands, J. M., Layton, H. E., & Fenton, R. A. (2020). Urine concentration and dilution. In A. S. L. Yu, G. M. Chertow, V. Luyckx, P. A. Marsden, K. Skorecki, & M. W. Taal (Eds.), *Brenner and Rector's the kidney, 11th edition*. (pp. 274-302). Elsevier.

Sherwood, L. (2016). *Human physiology: From cells to systems, 9th edition*. Cengage Learning.

Silverthorn, D. U. (2019). *Human physiology: An integrated approach, 8th edition*. Pearson Education, Inc.

Verlander, J. W., & Clapp, W. L. (2020). Anatomy of the kidney. In A. S. L. Yu, G. M. Chertow, V. Luyckx, P. A. Marsden, K. Skorecki, & M. W. Taal (Eds.), *Brenner and Rector's the kidney, 11th edition*. (pp. 38-79). Elsevier Inc.

Whiting, C. C. (2019). *Human anatomy & physiology laboratory manual, cat version*. Pearson Education, Inc.

Widmaier, E. P., Raff, H., & Strang, K. T. (2023). *Vander's human physiology: The mechanisms of body function, 16th edition*. McGraw Hill LLC.

continued on next page

Appendix 1. Renal Physiology Review

Purpose: Review urine formation using simple diagrams and models of the nephron and its blood supply.

Objectives:

By the end of the activity, you should be able to:

- Apply the principles of bulk flow to explain which components of the blood are filtered
- Explain which components of the filtrate are reabsorbed from the proximal tubule
- Distinguish between the functions of the descending and ascending limbs of the loop of Henle
- Describe how the countercurrent mechanism establishes and maintains the medullary vertical osmotic gradient
- Predict the effects of loop diuretics on urine formation

Task: Follow the step-by-step instructions and answer each question before proceeding to the next step.

Answer short sets of additional review questions.

1. Filtration and reabsorption

Introduction

This activity simulates filtration and reabsorption using beads to represent components of the blood. In an adult, the average glomerular filtrate rate (GFR) is ~115 mL/min (females) or ~125 mL/min (males), and ~20% of the plasma is filtered each minute. For this simulation, assume that the systemic blood pressure and GFR are normal.

Materials

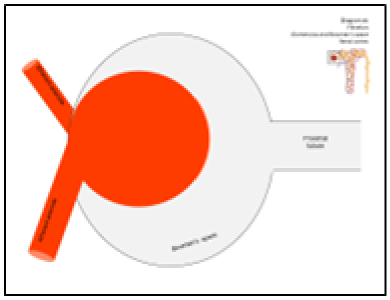
Petri dish with beads

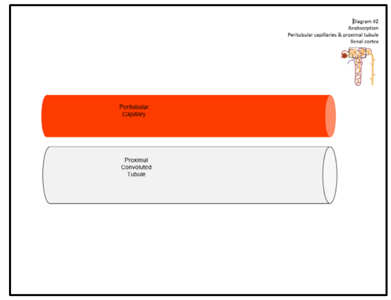
Diagram #1 (glomerulus and Bowman's space)

Diagram #2 (peritubular capillary and proximal convoluted tubule)

Erasable marker for labeling the Petri dish

continued on next page





Setup

1. Make two fluid compartments:

Label the Petri dish bottom "blood".

Label the Petri dish lid "filtrate".

2. The components of the "blood" dish are listed below. Make sure all of these components are in your dish.

1 red blood cell (RBC), large red bead

1 white blood cell (WBC), large white bead

2 urea molecules (organic waste), yellow beads

4 plasma proteins, purple beads

6 organic nutrients (amino acids, glucose), green beads

15 water molecules, blue beads

15 Na+, ivory beads

15 other ions, black beads

3. On diagram #1, move the "blood" dish through the afferent arteriole and into the glomerulus. Put the "filtrate" dish into Bowman's space.

Filtration simulation

Filtration is determined by the permeability of the glomerular capillaries and by bulk flow (Starling forces).

Complete the following steps to simulate filtration.

1. Which blood components are freely filtered into Bowman's space? List these below.

2. Move all of the beads representing filtered material into the dish lid (the "filtrate").

(For simplicity, this exercise will not consider the GFR. All blood components that can be filtered should be moved, instead of 20% of them.)

continued on next page

3. What is not normally filtered? List these below.

4. Keep the non-filtered material in the dish bottom (the "blood").

5. Move the "filtrate" dish to the proximal tubule (diagram #2). Re-label this dish "tubular fluid". Move the "blood" dish to the peritubular capillary, adjacent to the proximal tubule.

Reabsorption simulation (proximal tubule)

Of the filtered blood components, a proportion of each is reabsorbed by the proximal convoluted tubule and returned to the blood. Recall that the mechanism for reabsorption is transepithelial transport from the tubular fluid to the plasma.

6. First, check your work: The "blood" dish should contain 1 RBC (large red bead), 1 WBC (large white bead), and 4 plasma proteins (purple beads).

7. Which substances are 100% reabsorbed from the proximal convoluted tubule? List these below.

8. Which substances are 50-80% reabsorbed from the proximal convoluted tubule? List these below.

9. Move all of the 100% reabsorbed materials into the peritubular capillary ("blood" dish). Move 1 yellow bead (urea) to represent that 50% is reabsorbed here. For the other reabsorbed materials, move 9 beads of each color into the "blood" dish to represent that a proportion has been reabsorbed.

10. Check your work: In the tubular fluid, you should have 6 ivory beads (Na+), 6 black beads (other ions), 6 blue beads (water), and 1 yellow bead (urea).

2. Urine concentration in the juxtamedullary nephron

Introduction

The loop of Henle is the countercurrent multiplier. This loop establishes the vertical osmotic gradient in the renal medulla. The key features of the countercurrent mechanism are that the two limbs of the loop of Henle run in opposite directions, and they have different permeabilities.

Having a vertical osmotic gradient in the interstitial fluid means that sodium is progressively more concentrated as we move from the outer medulla to the innermost region. As tubular fluid travels through the loop, the surrounding osmotic gradient influences the movement of water out of the tubular fluid and into the interstitial fluid. The fluid can then be reabsorbed by the adjacent vasa recta. Ultimately, this movement of water results in conserving body water by concentrating the urine.

continued on next page

Materials

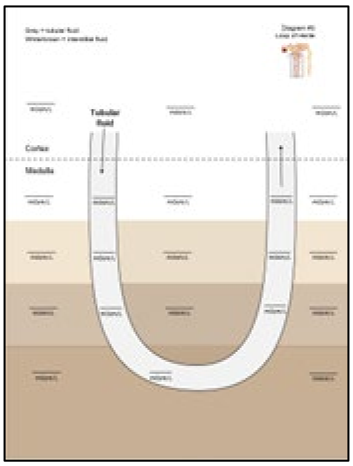
Diagram #3 (loop of Henle)

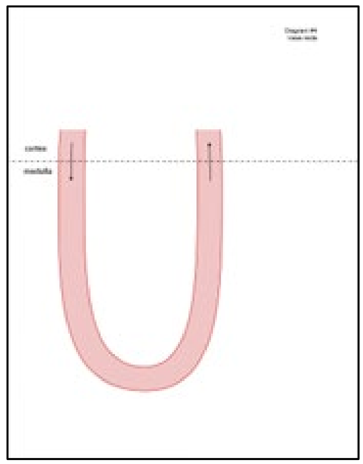
Diagram #4 (vasa recta)

Erasable marker for labeling the diagrams

Petri dishes with beads carried over from earlier work

Beaker for transferring extra beads





Part A: The loop of Henle

Build the vertical osmotic gradient using the countercurrent multiplier

1. Start with diagram #3. Consider which limb of the loop of Henle is permeable to Na+ versus water. On diagram #3, draw Na+ transporters on the appropriate limb of the loop. Make sure that the transporters are oriented to move Na+ out of the tubule.

2. Draw aquaporin (AQP) channels on the other limb. Draw arrows to indicate that the direction of water movement is out of the tubule.

3. <u>Interstitial fluid in the renal cortex</u>. The cortex is the white space above the dashed line. Fill in all 3 labels in the cortex to show that the interstitial fluid in the renal cortex is 300 mOsm/L.

continued on next page

4. Interstitial fluid in the renal medulla. The renal medulla is the area below the dashed line, surrounding the loop of Henle. To show that the interstitial fluid in the renal medulla has a vertical osmotic gradient, fill in all 11 labels in the medulla following these steps:
   a) At the first level below the cortex, fill in all 3 labels to indicate that the interstitial fluid is 300 mOsm/L at the top of the gradient.
   b) At the next level down, fill in the next 3 labels. These labels should indicate that the interstitial fluid at this level is 600 mOsm/L.
   c) At the next level down, fill in the next 3 labels. These labels should indicate that the interstitial fluid at this level is 900 mOsm/L.
   d) At the bottom of the figure, fill in the remaining 2 labels. These labels should indicate that the interstitial fluid at the bottom of the gradient is 1200 mOsm/L.
5. Tubular fluid in the descending limb of the loop of Henle. Label the diagram of the loop to show the osmolarity of the tubular fluid (gray):
   a) At the bottom of the loop, indicate 1200 mOsm/L.
   b) In the descending limb, the osmolarity of the tubular fluid matches the interstitial fluid gradient, so use increments of 300 mOsm/L to label the remaining values.
6. Tubular fluid in the ascending limb of the loop of Henle. In the ascending limb, Na+ transport (secondary active transport) decreases the osmolarity of the tubular fluid. Na+ transport sets up a 200 mOsm/L gradient between the tubular fluid and the interstitial fluid.

   Note: In the ascending limb, the gradient between the tubular fluid and the interstitial fluid is *not vertical*, it is *across,* at each horizontal level. For example, where the interstitial fluid is 1200 mOsm/L, the tubular fluid is 1000 mOsm/L. This is because Na+ transporters move Na+ from the tubular fluid to the interstitial fluid.

   Label the diagram of the loop to show the osmotic gradient of the tubular fluid in the ascending limb:
   a) Start by labeling the level that is 700 mOsm/L.
   b) Continue labeling, making sure that each label is 200 mOsm/L lower than the surrounding interstitial fluid at the same horizontal level.

continued on next page

Concentrate the urine

7. Set up the Petri dishes. For simplicity, you can remove all of the black beads ("other ions") from the tubular fluid. (Transfer these beads to the beaker.) You may also want to remove the yellow urea bead from the tubular fluid. This will allow you to focus on the movement of water (blue) and sodium (ivory).
8. Move the "tubular fluid" dish to the start of the descending limb of the loop of Henle. Note that the tubular fluid is 300 mOsm/L.
9. Move the dish down the descending limb. As the tubular fluid descends through the osmotic gradient, water will be pulled into the interstitial fluid by osmosis. Transfer 3 water molecules (blue) to the interstitial fluid.
10. Move the "tubular fluid" dish to the bottom of the loop of Henle. Notice that the dish has 3 water molecules (blue) and 6 Na+ (ivory). The osmolarity of the tubular fluid has increased because water has moved out.

Dilute the urine

11. Move the "tubular fluid" dish beyond the hairpin turn, to the start of the ascending limb of the loop of Henle. The tubular fluid is initially hyperosmotic relative to the interstitial fluid. As the fluid travels up the ascending limb, Na+ is transported out.
12. Move 5 Na+ (ivory) from the tubular fluid to the interstitial fluid.
13. Check your work: At the top of the ascending limb of the loop of Henle, your tubular fluid should be hypoosmotic to the interstitial fluid. It should have 3 water molecules (blue) and 1 Na+ (ivory).

Part B: Vasa recta

Show how the countercurrent exchange of the vasa recta maintains the vertical osmotic gradient

The vasa recta essentially make a loop through the medulla, similar to the nephron's loop of Henle. Capillary blood flow begins in the cortex and then enters the vasa recta and travels _down_ through the vertical osmotic gradient of the medulla. At the bottom of the vasa recta, the vessels make hairpin turns so that blood travels _back up_ through the gradient, heading back towards the cortex.

continued on next page

The vasa recta function as a countercurrent exchanger with the interstitial fluid. As the blood flows through the vertical osmotic gradient, it picks up salt and water from the interstitial fluid. For the vasa recta, the movement of Na+ and water is by passive transport down their respective concentration gradients (passive diffusion of Na+ and osmosis of water).

14. Overlay diagram #4 (vasa recta) on top of diagram #3 (loop of Henle).

15. The vasa recta have properties of capillaries. Are these vessels most like continuous or fenestrated capillaries?

16. Are the vasa recta highly permeable to both water and Na+?

17. Descending vasa recta. For water and Na+, consider the direction of movement (into or out of the blood vessels) at a given horizontal level. The direction of movement is influenced by concentration gradients across the interstitial fluid and plasma.

    The blood coming from the cortex is 300 mOsm/L. On diagram #4, label the osmolarity of the descending blood as it enters the medulla.

    Transfer the "blood" dish to the start of the vasa recta.

18. Move the "300 mOsm/L blood" dish down the path of blood flow. Notice that as the blood descends, the blood is lower in osmolarity than the surrounding fluid.

19. For the descending vasa recta, what is the direction of Na+ movement? (Is Na+ moving out of or into the blood vessels?)

20. Note that the diffusion of water and Na+ is in opposite directions here. Transfer 4 Na+ (ivory) into the blood. Transfer 2 water molecules (blue) out of the blood. Note that the blood is picking up solute from the salty interstitial fluid.

21. At the bottom of the vasa recta loop, the osmolarity of the blood is 1200 mOsm/L. Label this on diagram #4.

22. Ascending vasa recta. At the bottom of the vasa recta loop, the blood is hyperosmotic to the interstitial fluid. As the plasma travels back up through the medulla, it will take in water and lose salt. This will restore the osmolarity of the blood to 300 mOsm/L as it leaves the medulla.

    Move the "blood" dish up the ascending vasa recta.

    Notice that because the blood is now hyperosmotic compared to the surrounding interstitial fluid, water will move into the blood by osmosis.

continued on next page

23. Transfer 4 water molecules (blue) into the "blood" dish. Transfer 2 Na+ (ivory) out of the blood. Note that the blood is picking up water from the interstitial fluid.

24. Check your work: The blood should have an equal number of water molecules (blue) and Na+ (ivory). This demonstrates that by the time the blood flow exits the medulla, the osmolarity of the plasma will have been restored to 300 mOsm/L.

Part C: Clinical correlations

Consider the effects of disruptions to the interstitial concentration gradient in the renal medulla

Loop diuretics such as furosemide and bumetanide are used to treat hypertension and edema. Diuretics cause more water to be excreted in the urine. Loop diuretics function by inhibiting Na+ transport in the ascending limb of the loop of Henle.

25. If Na+ transport is disrupted, what would happen to the vertical osmotic gradient of the medulla?

26. Would the solute load of the tubular fluid be increased or decreased compared to the normal load?

27. Would there be more water than normal in the tubular fluid? Why?

3. Additional review questions

<u>Filtration</u>:

1. Consider that not all substances are filtered by the glomerular capillaries. What factors determine what is or is not filtered?

<u>Reabsorption</u>:

2. Glucose is 100% reabsorbed from the proximal tubule unless the glucose concentration exceeds the transport maximum. How is glucose transported here? How does the transport mechanism cause a limitation for how much glucose can be transported per unit time?

3. Beyond the proximal tubule, where else along the nephron is Na+ reabsorbed without the action of hormones?

4. Beyond the proximal tubule, where else along the nephron is K+ reabsorbed?

<u>Secretion</u>:

5. Urea is filtered, reabsorbed, and secreted by the nephron. Explain the physiological basis of urea handling.

6. In what context would you expect K+ to be secreted?

continued on next page

7. Which nephron segments can secrete H+?

8. In what context would you expect H+ to be secreted?

Countercurrent mechanism:

9. In the renal medulla, what structure is the countercurrent multiplier?

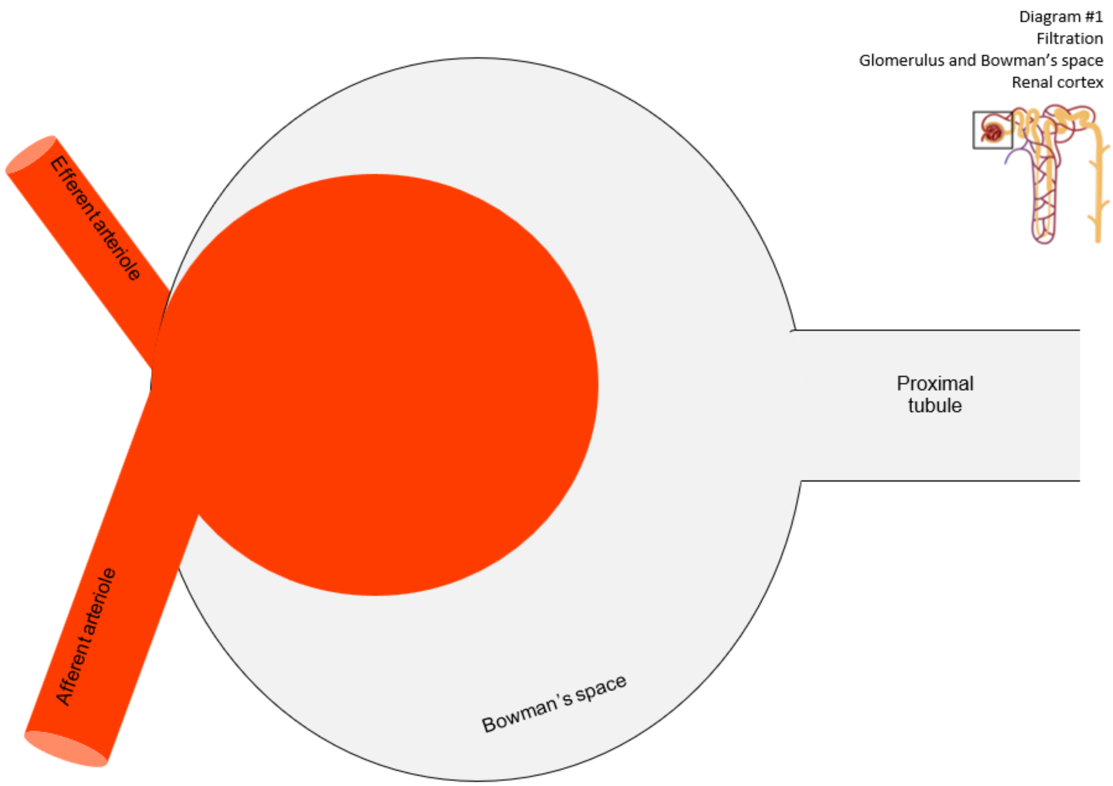
10. Define countercurrent multiplication.

11. In the renal medulla, what structure is the countercurrent exchanger?

12. What is the effect of countercurrent exchange for the renal medulla?

continued on next page

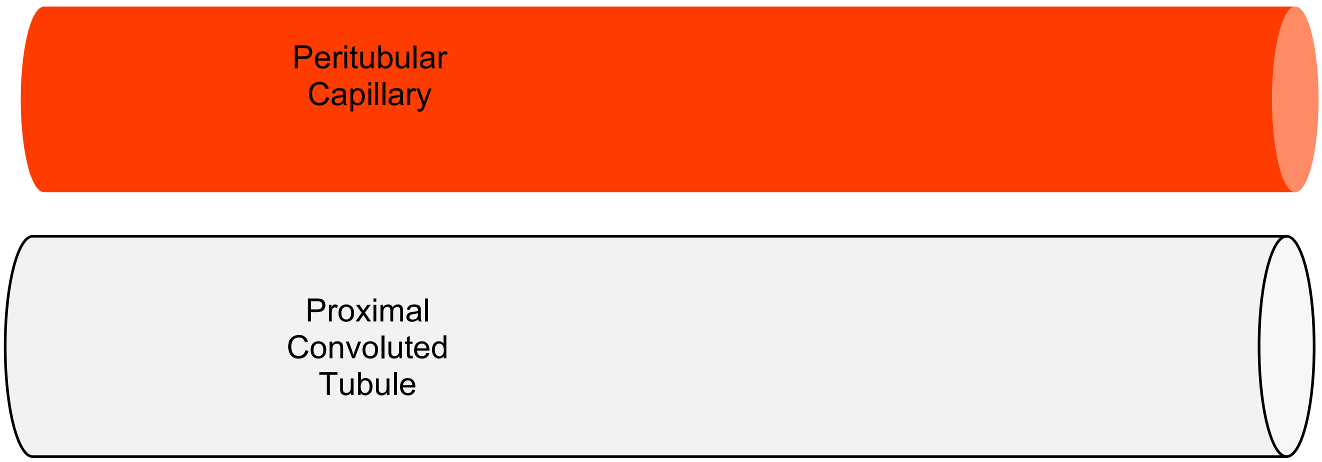
Appendix 2. Diagram #1



continued on next page

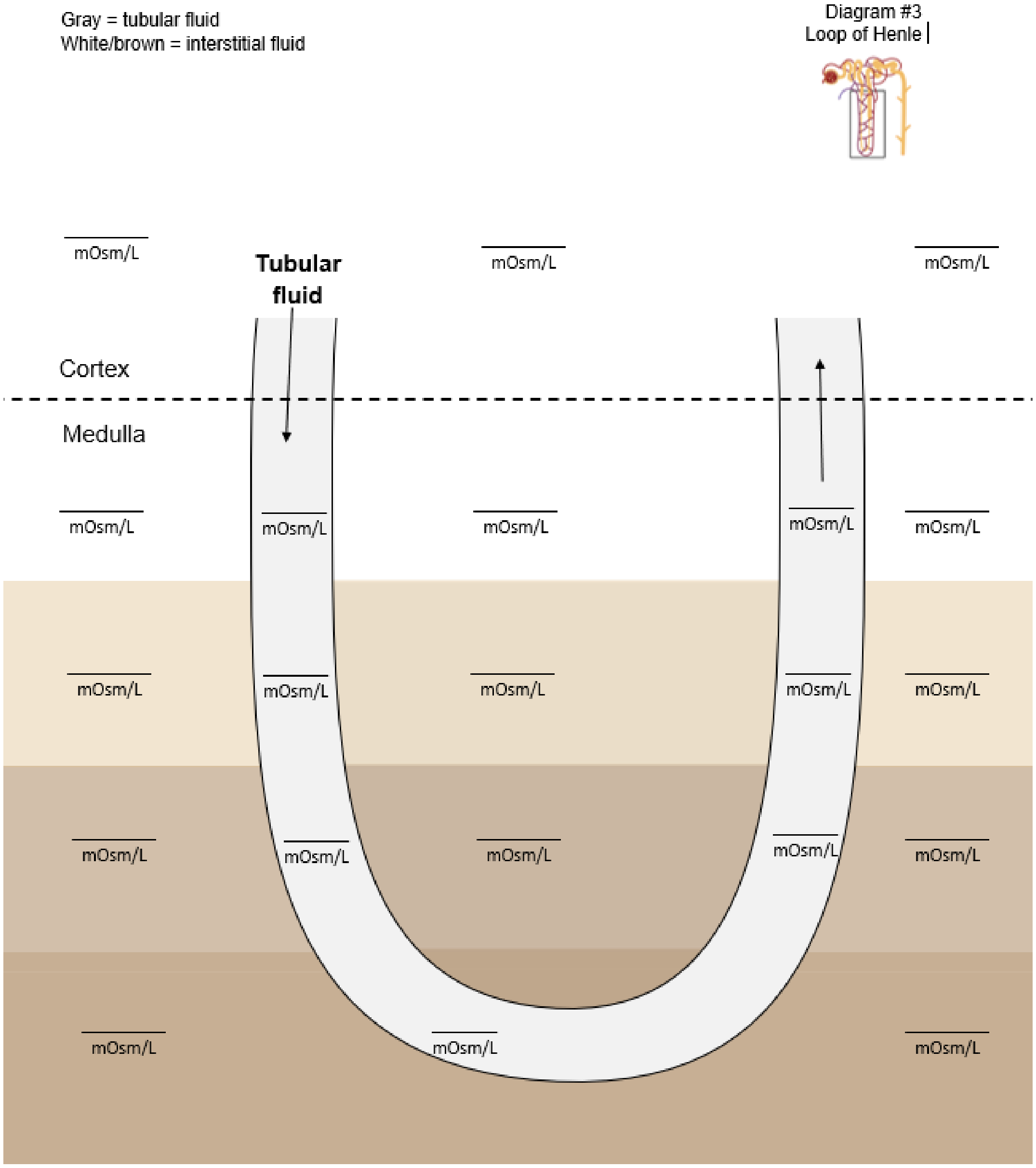
Appendix 3. Diagram #2

Diagram #2
Reabsorption
Peritubular capillaries & proximal tubule
Renal cortex



continued on next page

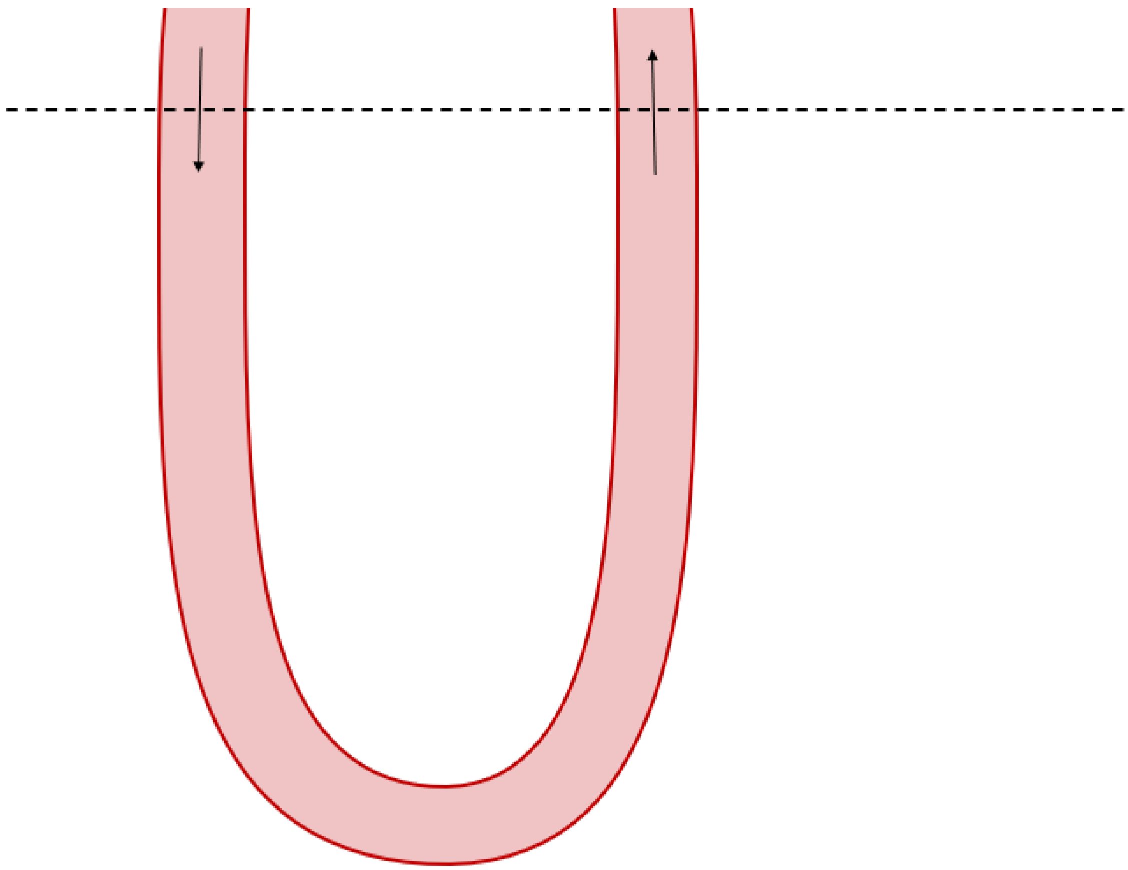
Appendix 4. Diagram #3



continued on next page

Appendix 5. Diagram #4

Diagram #4
Vasa recta



continued on next page

Appendix 6. Answer Key with Notes

1. Filtration and reabsorption

1. Which blood components are freely filtered into Bowman's space? List these below.

Answer: organic molecules (nutrients and waste products), water, ions

Nutrients include amino acids and glucose.

Waste products include creatinine and urea.

3. What is not normally filtered? List these below.

Answer: cells, platelets, plasma proteins

7. What substances are 100% reabsorbed by the proximal convoluted tubule? List these below.

Answer: amino acids, glucose

8. What substances are 50-80% reabsorbed by the proximal convoluted tubule? List these below.

Answer: water, ions, urea

2. Urine concentration in the juxtamedullary nephron

Part A: The loop of Henle

1. Start with diagram #3. Consider which limb of the loop of Henle is permeable to Na^+ versus water. On diagram #3, draw Na+ transporters on the appropriate limb of the loop. Make sure that the transporters are oriented to move Na+ out of the tubule.

Answer: The transporters should be drawn on the **(thick) ascending limb**.

Note: The transporter uses secondary active transport. The transporter is the NKCC2 symporter, carrying Na+, Cl, and K+. (There are also Na+/Cl^- cotransporters.) The descending limb is not permeable to NaCl.

2. Draw aquaporin (AQP) channels on the other limb. Draw arrows to indicate that the direction of water movement is out of the tubule.

Answer: The aquaporins should be drawn on the **(thin) descending limb**.

Note: The descending limb is permeable to water. The ascending limb is not permeable to water.

Part B: Vasa recta

15. The vasa recta have properties of capillaries. Are these vessels most like continuous or fenestrated capillaries?

continued on next page

Answer: Fenestrated

Note: The answers may vary. The descending vasa recta (DVR) and ascending vasa recta (AVR) show some structural differences, and there is some disagreement in the literature regarding their histology. The DVR are generally described as having a continuous epithelium in proximal regions that becomes fenestrated distally, while the AVR are generally described as highly fenestrated.

16. Are the vasa recta highly permeable to both water and NaCl?

Answer: Yes

19. For the descending vasa recta, what is the direction of Na+ movement? (Is Na+ moving out of or into the blood vessels?)

Answer: Na+ moves into the descending vasa recta.

Part C: Clinical correlations

25. If Na+ transport is disrupted, what would happen to the vertical osmotic gradient of the medulla?

Answer: The gradient would be disrupted. The interstitial fluid would be isosmotic.

26. Would the solute load of the tubular fluid be increased or decreased compared to the normal load?

Answer: The solute load would be increased.

27. Would there be more water than normal in the tubular fluid? Why?

Answer: Yes. The increased solute load of the tubular fluid would draw water from the interstitial fluid.

3. Additional review questions

<u>Filtration</u>:

1. Consider that not all substances are filtered by the glomerular capillaries. What factors determine what is or is not filtered?

Answer: Size is the main determinate. Relatively small substances (water and solutes) can pass through the filtration barrier which consists of the endothelial fenestrations, basement membrane, and the filtration slits of the podocytes. Blood cells and most plasma proteins are too large to cross the filtration barrier. Also, plasma proteins are negatively charged and cannot easily pass through the barrier because they are repelled by the negatively charged endothelium, basement membrane, and podocyte epithelium.

continued on next page

Reabsorption:

2. Glucose is 100% reabsorbed from the proximal tubule unless the glucose concentration exceeds the transport maximum. How is glucose transported here? How does the transport mechanism cause a limitation for how much glucose can be transported per unit time?

Basic answer: Glucose is transported by transmembrane transporters. The transporters become saturated when the tubular fluid load exceeds a threshold. Therefore, the number of available transporters is rate limiting.

More detailed answer: At the apical (luminal) membrane of the tubular cells, glucose is transported by secondary active transport (sodium-glucose transporters, SGLT). At the basolateral membrane, glucose is transported by passive, facilitated transport (GLUT transporters). At the peritubular capillaries, glucose is transported by bulk flow.

3. Beyond the proximal tubule, where else along the nephron is Na+ reabsorbed without the action of hormones?

Answer: At the ascending (thick) limb of the loop of Henle.

4. Beyond the proximal tubule, where else along the nephron is K+ reabsorbed?

Answer: At the ascending (thick) limb of the loop of Henle

Secretion:

5. Urea is filtered, reabsorbed, and secreted by the nephron. Explain the physiological basis of urea handling.

Answer: Urea's small size allows for pressure filtration. Then, it is passively reabsorbed from the tubular fluid as it moves down its concentration gradient. Urea is then secreted as a waste product to prevent its toxic accumulation in the blood.

6. In what context would you expect K+ to be secreted?

Answer: There are several answers. Examples include 1) increased K+ in the blood because of dietary intake; 2) low plasma volume which causes aldosterone release through the renin-angiotensin-aldosterone system (RAAS) and thus promotes K+ secretion; 3) hyperkalemia, high levels of excess K+ in the blood, associated with kidney disease.

continued on next page

7. Which nephron segments can secrete H+?

Answer: proximal tubule, loop of Henle (thick ascending limb), collecting duct

8. In what context would you expect H+ to be secreted?

Answer: H+ is secreted as part of bicarbonate handling by the nephron, which is important for maintaining acid-base balance.

Countercurrent mechanism:

9. In the renal medulla, what structure is the countercurrent multiplier?

Answer: The loop of Henle, specifically the hairpin arrangement of the two limbs.

10. Define countercurrent multiplication.

Answer: This is the mechanism that generates the vertical osmotic gradient in the renal medulla. It involves the active transport of sodium and cotransport of other ions into the renal medulla.

11. In the renal medulla, what structure is the countercurrent exchanger?

Answer: The vasa recta, specifically the parallel arrangement of the descending and ascending vessels.

12. What is the effect of countercurrent exchange for the renal medulla?

Answer: Countercurrent exchange preserves the vertical osmotic gradient.

■

Back to TOC

PERSPECTIVES ON TEACHING

EDUCATOR
Journal of the Human Anatomy and Physiology Society

Volume 29, Issue 2 | August, 2025
https://doi.org/10.21692/haps.2025.012

Effectiveness of Low-Stakes Online Quizzes on Engagement and Performance in Sophomore-Level Anatomy and Physiology

Nicholas Hobbs, PhD and Janet Steele, PhD

Department of Biology, University of Nebraska at Kearney, Kearney, NE, USA

Corresponding author: steelej@unk.edu

Abstract

Formative assessment is a type of low-stakes assessment that can help reduce student anxiety and encourage students to engage in the course regularly. Formative assessment can also help instructors identify knowledge gaps. Students who had completed the first semester of a two-semester sophomore-level anatomy and physiology sequence were surveyed to better understand how they perceived and prepared for these unsupervised, online lecture quizzes. There were 11 quizzes administered over the course of the semester, and the best 10 quiz scores accounted for 9.5% of the students' final grade in the course. Over 70% of the 79 participants indicated that the online quizzes helped them learn the material and prepare for summative assessment. Most (82%) of the respondents used course resources when taking the quizzes, and 43% indicated they worked with peers when taking the quizzes. Quantitative analysis showed a positive correlation between online quiz performance and final grade in the class. Our results suggest that the online quizzes met the goals of formative assessment by engaging students with the course, familiarizing them with course expectations, and providing feedback so that students could adjust their study habits prior to summative assessments. https://doi.org/10.21692/haps.2025.012

Key words: formative assessment, low-stakes assessment, summative assessment, online quizzes

Introduction

Formative assessment helps students learn under low stress conditions (Black & William, 2009). To be most effective, formative assessment must be available early and frequently and take on multiple forms (Evans et al., 2014). Formative assessment should have minimal impact on a student's grade as a form of low-stakes assessment. In addition, formative assessment can help students understand if they are meeting course objectives, giving them confidence when preparing for summative assessment, which can in turn impact future student success (Martin, 2023; Kulasegaram & Rangachari, 2018). Furthermore, formative assessment provides information on student knowledge gaps, enabling the instructor to modify future instruction (Evans et al., 2014). Incentives may be necessary in introductory courses, however, because a true interest in learning, as opposed to motivation to earn a particular grade, is significantly related to students' engagement in coursework (Gasiewski et al., 2012).

Introductory STEM courses are often viewed as "gatekeeper" courses, and students who find the courses challenging often do not have the motivation and study skills necessary to succeed (Tracy et al., 2022). In addition, students who struggle in introductory STEM courses often have less precollege preparation than those who do not (Tracy et al., 2022). Therefore, the use of formative assessment to correct misconceptions, identify knowledge gaps, and provide feedback to both the student and instructor is an essential component of the learning process.

Consistent use of formative assessment has been shown to improve students' engagement with the course (Martin, 2023; Ogochukwu, 2024). Frequent, unsupervised, low-stakes online quizzes are an ideal avenue for formative assessment due to their flexibility, ease of administration, and provision of immediate feedback (Kibble, 2007). In addition, instructors can allow students to choose whether they complete the quizzes on their own or in collaboration with peers. Group work has been shown to have a positive impact on student

continued on next page

engagement with courses (Gasiewski et al., 2012). This sense of community and engagement are particularly important for students whose previous schooling did not adequately prepare them for introductory STEM college courses (Brown et al., 2018; Tracy et al., 2022).

Context of Study

In the fall of 2022, we started incorporating small, unsupervised, online quizzes into a sophomore-level anatomy and physiology course (BIOL 225 Anatomy and Physiology I) with the goals of incentivizing students to keep current in the class and providing students with feedback concerning their understanding of class concepts. BIOL 225 is a 16-week course with three hours of lecture and two hours of laboratory weekly, and its only prerequisite is sophomore standing. Each quiz contributed less than 1% to the student's final grade in the class, and the quizzes collectively contributed 9.5% to the final grade. Final letter grades of students who earned 100% on the online quizzes ranged from A to F, which prompted examination of how the students approached the quizzes. The purpose of this study was to investigate students who had completed the class to better understand how they viewed the purpose of the quizzes and how they prepared for the quizzes. Results from this study can help other instructors understand how students engage with formative assessment in the form of unsupervised, online quizzes.

Methods

Online Quizzes

A total of 11 quizzes were given during the semester and the best 10 scores counted toward the students' final grades. Each quiz consisted of six matching questions (1 point each; Table 1) and two multiple-choice questions (2 points each; Table 2) for a total of 10 points. The matching questions had three distractors, and the multiple-choice questions had two distractors and were selected from a pool of three questions. Students had six minutes to complete the quiz. Quizzes emphasized core concepts of physiology (Michael & McFarland, 2020), followed class objectives (Table 3), and were administered during weeks that the class did not have a summative exam. Quizzes were available Friday afternoon through Monday prior to class.

Term	Description
Hypertonic	Solution with a greater solute concentration compared to another solution
Hydrostatic pressure	Driving force for filtration
Collagen	Protein that withstands stretching; binds structures together
Desmosome	Intercellular junction holding cells of epidermis together
Papillary layer	Relatively superficial region of dermis made of areolar connective tissue
Stratum spinosum	Active, thriving layer of epidermis; cells make keratin
Distractors	
Solution with a greater water concentration compared to another solution	
Fibrous protein that gives skin its waterproof property	
Relatively deep region of dermis made of irregular dense fibrous connective tissue	

Table 1. *Sample matching question from an online quiz for students enrolled in BIOL 225 Anatomy and Physiology I. Each question was worth 1 point.*

continued on next page

Peak depolarization of +30 mV ensures that: A. adequate charge will remain to open voltage-gated channels at the next node. B. enough sodium ions will enter the axon to reverse the gradient for sodium. C. the impact of potassium lead channels won't offset the entry of sodium ions.
Threshold is reached in the trigger zone by: A. summation of excitatory and inhibitory postsynaptic potential in cell body. B. excess sodium from previous action potential triggering adjacent sodium channels. C. the sodium-potassium pump restoring normal ion distribution following hyperpolarization.
When a neural impulse reaches the synaptic knob: A. voltage-gated calcium channels open, triggering neurotransmitter secretion. B. the impulse seems to jump across the cleft, which is called saltatory conduction. C. the combination of excitatory and inhibitory postsynaptic potentials determines the response.

Table 2. *Sample multiple choice questions from an online quiz for students enrolled in BIOL 225 Anatomy and Physiology I. Students were randomly presented with 2 of the 3 questions and each question was worth 2 points.*

1.	Define tissue.
2.	List and describe three different types of intercellular junctions.
3.	Describe general characteristics of epithelial tissue.
4.	Describe general characteristics of connective tissues.
5.	Describe general characteristics of muscle tissue.
6.	Describe the functions of the integumentary system.
7.	Describe the characteristics of the following tissues associated with the skin: simple cuboidal epithelium, stratified squamous epithelium, adipose tissue, areolar (loose) connective tissue, dense fibrous connective tissue, smooth muscle.
8.	List the five layers of the epidermis and describe the anatomy of each layer.
9.	Explain the importance of keratinization.
10.	Discuss the role of melanocytes in skin color.
11.	Describe the anatomy and physiology of the dermis.
12.	Describe the anatomy and physiology of the subcutaneous layer (hypodermis).
13.	Describe the following accessory organs of the skin: hair follicles, sebaceous glands, and sudoriferous (sweat) glands.
14.	Discuss the importance of regulating body temperature, including potential consequences of hyperthermia and hypothermia.
15.	Describe a homeostatic mechanism to regulate body temperature.

Table 3. *Sample course objectives (Tissues and Skin) from BIOL 225 Anatomy and Physiology I. Students are provided with clear objectives for each topic throughout the semester.*

continued on next page

Qualitative Survey

This study was approved by the University of Nebraska at Kearney Institutional Review Board (#030624-2). An anonymous Quatrics® survey was distributed to students who had been enrolled in BIOL 225 Anatomy and Physiology I during the fall 2022 or fall 2023 semesters. The survey consisted of six statements that participants rated on a five-point scale ranging from strongly agree to strongly disagree concerning their approach to and perception of the low-stakes online quizzes in BIOL 225. The survey was distributed to a total of 303 students during the spring 2024 semester. Students demonstrated their consent to participate in the study by completing the survey.

Quantitative data analysis

The relationship between total points earned on the online quizzes and final course grades was explored using Pearson's correlation analysis. Significant correlation was further investigated to explore how much variance it contributed to students' final grade using the coefficient of determination r^2 analysis. The correlation between performance on quizzes and final grade in the class was calculated with significance ascribed to $p < 0.05$. Statistical analysis was carried out using Excel® and bar graphs were created using Excel®.

Results

Qualitative Survey

A total of 79 students completed the survey (26% response rate). Most of the respondents (70-74%) agreed or strongly agreed that the quizzes had a valuable purpose (Figure 1, questions 1-3). In addition, most (82%) indicated that they used their notes or other materials during the quizzes, and over 40% indicated that they collaborated with others to complete the quizzes (Figure 1, questions 4 and 5, respectively). Less than half of the respondents (43%) indicated that they prepared for at least an hour prior to taking the quizzes (Figure 1, question 6).

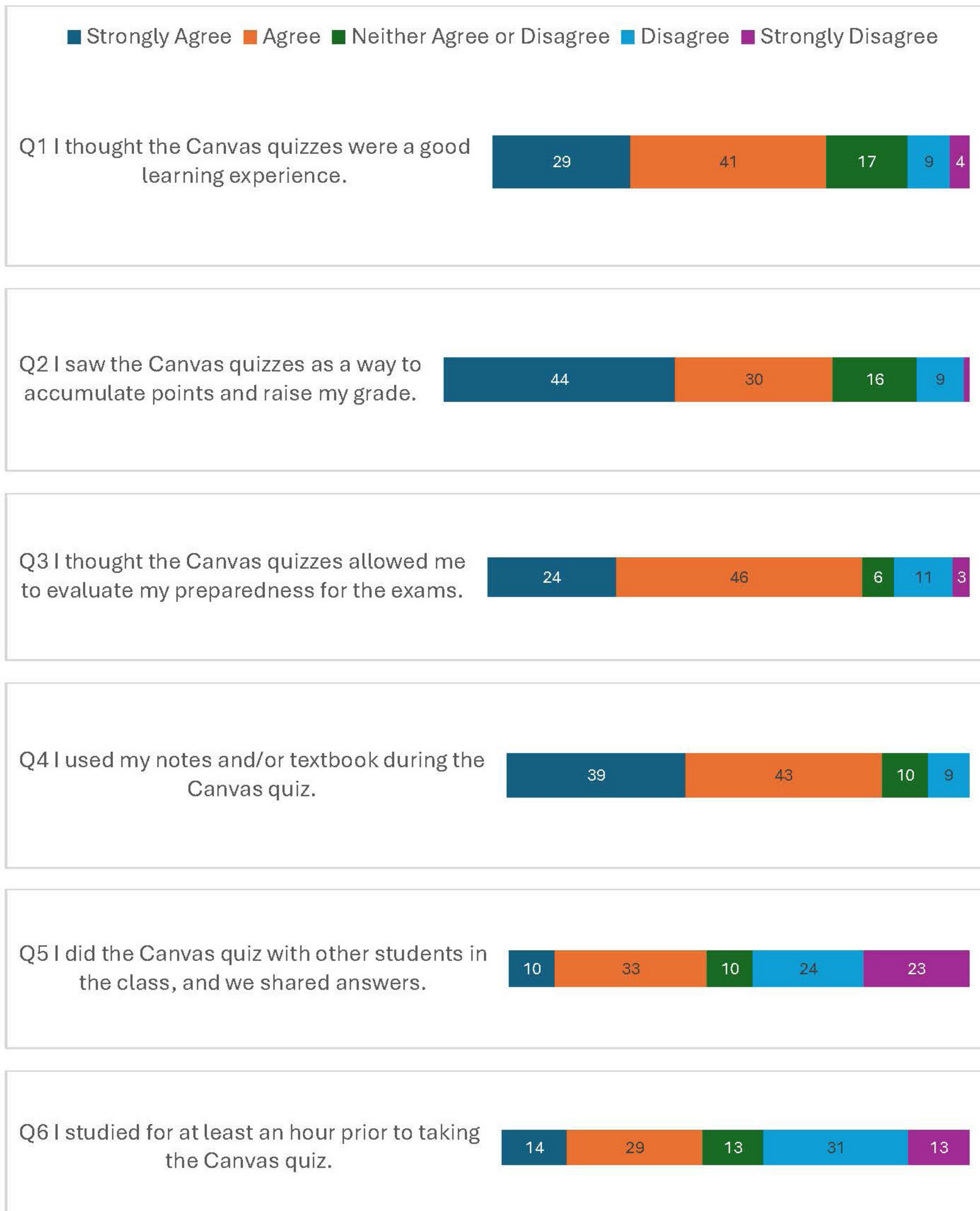


Figure 1. *Student ratings (n = 79) of statements related to how they perceived the purpose of and prepared for the online quizzes in BIOL 225 Anatomy and Physiology I. Embedded numbers represent percent of respondents.*

continued on next page

Correlation and Regression

There was a significant positive correlation between quiz total points and final overall percentage in the class (n=303, Pearson's correlation coefficient 0.76, $r^2 = 0.58$, $p < 0.0001$).

Discussion

Our results indicate that students found the online quizzes a valuable learning experience and an important self-evaluation of their preparedness for an upcoming exam. It has been shown that grades are very important to undergraduate students (Defeo et al., 2017), and 74% of the respondents viewed getting the questions correct and the accumulation of points a benefit of the quizzes. Using course resources during the quiz is consistent with Kibble (2007) and, coupled with the observation that 43% of the respondents took the quizzes with peers, indicates the quizzes promoted engagement with the course materials and classmates. Active learning with peers has been shown to enhance students' connection and engagement with the class, leading to more positive outcomes (Gasiewski et al., 2012).

Hora and Oleson (2017) found only 26% of undergraduate students in STEM courses studied throughout the semester, and this was primarily driven by specific factors such as weekly quizzes. Bickerdike et al. (2016) found just under 50% of medical students reported studying consistently throughout the year. One goal of instituting these quizzes was to encourage students to keep current on lecture topics in the class and study throughout the semester. Our results are consistent with previous studies, however, suggesting that less than half of the respondents thoroughly reviewed class material regularly on their own. The observation that larger percentages of respondents indicated they used course materials during the quizzes and took the quizzes with peers, however, suggest that the quizzes accomplished our goal of encouraging student engagement with the course materials on a regular basis.

Statistical analysis indicated that there was a high positive relationship between online quiz performance and final grade in the class. In addition, 58% of the variability in final grades can be explained by students' performance on the online quizzes. These results are consistent with the observation that some students who earned 100% on the online quizzes did not earn a high overall grade in the class. In fact, some students who earned 100% on the online quizzes did not pass the course. As stated by Evans et al. (2014), formative assessment will be of benefit only if students recognize their knowledge gaps and take steps to address those issues. In addition, students may struggle in introductory STEM courses due to both external factors (prior coursework, resources, interactions with instructor) and internal factors (motivation, time management, anxiety) that cannot easily be overcome (Tracy et al., 2022). Therefore, low-stakes formative assessment is a tool for but not a solution to improving student engagement and performance.

Not knowing what to study or how to prepare for summative assessment is a challenge for many students (Tracy et al., 2022). As reported by Hartwig and Dunlosky (2012), college students often employ ineffective study habits. For example, Dunlosky et al. (2013) found that most undergraduates prefer to study by re-reading texts and notes, both of which have been shown to be ineffective. Self-regulated learners have higher rates of academic achievement (Boekaerts & Corno, 2005), but the student must make the decision to be a self-regulated learner on their own (Cassidy, 2011). Students with lower incoming preparation and those from underrepresented groups are disproportionately likely to struggle in introductory STEM courses due to lack of good study habits and engagement (Tracy et al., 2022). Anatomy and physiology courses attract students with varied backgrounds and students interested in many different career paths. Due to the competition for admission to professional programs, students pursuing health care have been shown to avoid interactions with peers whom they view as competitors (Gasiewski et al., 2012). This competition further isolates those with less precollege preparation in "gateway" courses. Implementation of regular low-stakes formative assessment like unsupervised online lecture quizzes could be a means to encourage collaboration and engagement with both the course materials and with peers. In addition, the immediate feedback provided with low-stakes formative assessment could help both students and instructors identify knowledge gaps and help students improve their study habits.

Limitations

The response rate to the survey was low (26%) and it is unknown how the quiz scores for this subset of students correlated with their final grades in the course. The influence of formative assessment via online quizzes on summative assessment from this small sample therefore limits the generalizability of the results.

continued on next page

Conclusions

Formative assessment has been shown to have many benefits for both students and their instructors. The results of our study demonstrate that respondents found unsupervised, online, low-stakes quizzes to be a valuable learning tool in preparation for summative assessment. Respondents indicated they engaged with both the course material and peers while taking the quizzes. These outcomes are consistent with the first two assessments for learning as identified by Kulasegaram and Rangachari (2018) by helping learners identify where they are in the learning process and what they need to do to be successful. Moreover, engagement in "gatekeeping" introductory STEM courses is critical to retain students from underrepresented groups and students with lower precollege preparation (Gasiewski et al., 2012; Tracy et al., 2022). Our results suggest low-stakes online quizzes are a means of improving student engagement without adding a large administrative burden while still providing valuable feedback to both student and instructor.

Acknowledgment

The authors thank Dr. Melissa Wuellner, associate professor of Biology at the University of Nebraska at Kearney, for her assistance with statistical analysis.

About the Authors

Nicholas Hobbs earned his doctorate in biology from the University of Memphis and completed his postdoctoral fellowship in neuroscience at Michigan State University. He joined the faculty at the University of Nebraska at Kearney (UNK) in 2017 and is an associate professor of biology. Janet Steele earned her doctorate in zoology from Miami University and completed her postdoctoral fellowship in physiology and biophysics at the Medical College of Ohio. She joined the Biology faculty at UNK in 1993, holds the rank of full professor, and has been awarded the Leland Holt Distinguished Faculty Member award from UNK.

References

Bickerdike, A., O'Deasmhunaigh, C., O'Flynn, S., & O'Tuathaigh, C. (2016). Learning strategies, study habits and social networking activity of undergraduate medical students. *International Journal of Medical Education 7*, 230-236. https://doi.org/10.5116/ijme.576f.d074

Black, P., & William, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation, and Accountability 21*, 5-31. https://doi.org/10.1007/s11092-008-9068-5

Boekaerts, M., & Corno, L. (2005). Self-regulation in the classroom: A perspective on assessment and intervention. *Applied Psychology 54*(2), 199-231. https://doi.org/10.1111/j.1464-0597.2005.00205.x

Brown, S. J., Power, N., Bowmar, A., & Foster, S. (2018). Student engagement in a human anatomy and physiology course: A New Zealand perspective. *Advances in Physiology Education 42*(4), 636-643. https://doi.org/10.1152/advan.00035.2018

Cassidy, S. (2011). Self-regulated learning in higher education: Identifying key component processes. *Studies in Higher Education 36*, 989-1000. https://doi.org/10.1080/03075079.2010.503269

DeFeo, D. J., Tran, T. C., & Gerken, S. (2021). Mediating students' fixation with grades in an inquiry-based undergraduate biology course. *Science and Education 30*, 81-102. https://doi.org/10.1007/s11191-020-00161-3

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D.T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest 14*(1), 4-58. https://doi.org/10.1177/1529100612453266

Evans, D. J. R., Zeun, P., & Stanier, R. A. (2014). Motivating student learning using a formative assessment journey. *Journal of Anatomy 224*(3), 296-303. https://doi.org/10.1111/joa.12117

Gasiewski, J. A., Eagan, M. K., Garcia, G. A., Hurtado, S., and Chang, M. J. (2012). From gatekeeping to engagement: a multicontextual, mixed method study of student academic engagement in introductory STEM courses. *Research in Higher Education 53*, 229-261. https://doi.org/10.1007/s11162-011-9247-y

Hartwig, M. K., & Dunlosky, J. (2012). Study strategies of college students: Are self-testing and scheduling related to achievement? *Psychonomic Bulletin & Review 19*, 126-134. https://doi.org/10.3758/s13423-0181-y

Hora, M. T., & Oleson, A. K. (2017). Examining study habits in undergraduate STEM courses from a situative perspective. *International Journal of STEM Education 4*, Article e1. https://doi.org/10.1186/s40594-017-0055-6

continued on next page

Kibble, J. (2007). Use of unsupervised online quizzes as formative assessment in a medical physiology course: Effects of incentives on student participation and performance. *Advances in Physiology Education 31*(3), 253-260. https://doi.org/10.1152/advan.00027.2007

Kulasegaram, K., & Rangachari, P. K. (2018). Beyond "formative": assessments to enrich student learning. *Advances in Physiology Education 42*(1), 5-14. https://doi.org/10.1152/advan.00122.2017

Martin, D. J. (2023, September 20). Are your assessments fair and balanced? Retrieved from https://www.facultyfocus.com/articles/educational-assessment/are-your-assessments-fair-and-balanced/

Michael, J., & McFarland, J. (2020). Another look at the core concepts of physiology: Revisions and resources. *Advances in Physiology Education 44*(4), 752-762. https://doi.org/10.1152/advan.00114.2020

Ogochukwu, O. (2024). Use of consistent formative assessments to engage students in a second semester human anatomy and physiology course. *HAPS Educator 28*(1), 46-57. https://doi.org/10.21692/haps.2024.002

Tracy, C. B., Driessen, E. P., Beatty, A. E., Lamb, T., Pruett, J. E., Botello, J. D., et al. (2022). Why students struggle in undergraduate biology: Sources and solutions. *Life Sciences Education 21*(3), 1-14. https://doi.org/10.1187/cbe.21-09-0289

Back to TOC

PERSPECTIVES ON TEACHING

HAPS EDUCATOR
Journal of the Human Anatomy and Physiology Society

Volume 29, Issue 2 | August, 2025
https://doi.org/10.21692/haps.2025.013

Harnessing AI for Physiology Education: Teaching Acid-Base Balance Disorders Using AI

James T. Davis, PhD

Indiana University School of Medicine, Department of Anatomy, Cell Biology, and Physiology, Bloomington, IN, USA

Corresponding author: jidavi@iu.edu

Abstract

The use of large language models (LLMs) in education is often debated, but when used effectively, they can enhance learning. LLMs can be particularly useful for reinforcing physiology concepts, such as diagnostic reasoning in acid-base balance disorders. Traditional case-based learning is limited by the number of instructor-provided cases, whereas students can learn how to use LLMs to generate novel practice problems. By refining their prompts, students learned to create concise cases with essential clinical data (pH, $PaCO_2$, HCO_3^-) while omitting the diagnosis, encouraging independent problem-solving. Student feedback suggests this approach improved their understanding of acid-base disorder compensation mechanisms and was applicable to other physiology topics. This experience demonstrates how instructor-led engagement with LLMs can support active learning, enhance problem-solving skills, and serve as an ongoing academic resource. Future exploration should consider how AI can be further integrated into physiology education while maintaining critical thinking and academic integrity. https://doi.org/10.21692/haps.2025.013

Key words: AI, case-based learning, physiology education, critical thinking

Introduction

The use of large language models (LLMs) can be a polarizing topic in education (Bopp et al., 2024; Meyer et al., 2023; Van Wyk, 2024). As a result, students are often unsure whether they are allowed to use generative artificial intelligence (AI) tools like ChatGPT or Copilot for assignments or exam preparation. While many instructors discourage LLMs for writing-intensive tasks, these models have the potential to provide significant benefits in courses that emphasize conceptual understanding rather than writing skills (Kasneci et al., 2023; Wang et al., 2024).

In anatomy and physiology, particularly in large-enrollment lecture courses, LLMs can be valuable learning tools (Nazar et al., 2024). These courses emphasize memorization, pattern recognition, and conceptual reasoning skills that align well with case-based problem-solving. Dismissing LLMs outright can be a missed opportunity. Instead, we should teach students how to use LLMs effectively and responsibly, helping them leverage this technology as a powerful educational aid rather than a shortcut (Kasneci et al., 2023; Wang et al., 2024).

A well-established challenge in physiology instruction is providing students with sufficient opportunities to apply conceptual knowledge to clinical-style problems (Armbruster et al., 2009; Dong et al., 2024; Michael, 2007). While case-based learning promotes active engagement, instructors are often constrained by the time and effort required to design individualized cases (Armbruster et al., 2009; Michael, 2007). LLMs offer a scalable solution to this gap, allowing students to generate customized practice problems aligned with course objectives. When guided appropriately, this use of AI can support diagnostic reasoning, promote pattern recognition, and deepen conceptual mastery.

Using LLMs for Acid-Base Balance Problems

Building on the rationale for using LLMs in case-based learning, I focused their implementation on a particularly challenging topic in physiology: acid-base balance. Some instructors provide students with normal values for pH, arterial carbon dioxide ($PaCO_2$), and bicarbonate (HCO_3^-) and

continued on next page

then present a case study requiring students to diagnose an acid-base disorder. These problems demand critical thinking across multiple physiological variables; however, students are often provided with a limited number of case studies due to the considerable time required to create them *de novo*.

Several students visited my office seeking additional practice problems on acid-base disorders. This led me to explore the potential of LLMs as a scalable solution, that is a tool that can provide individualized learning support to multiple students simultaneously, without requiring a proportional increase in instructor time or resources. LLMs can generate an endless number of practice cases on demand, allowing each student to work at their own pace and receive instant feedback (Dong et al., 2024). Initially, when students asked an LLM for an acid-base case, the model provided a detailed patient history and extensive lab results—useful for a physician but far beyond what most introductory physiology students need. Additionally, the LLM included the diagnosis, which eliminated the opportunity for students to reason through the problem independently.

By iteratively refining the prompts, we were able to get the LLM to produce a concise 2-3 sentence patient history along with pH, $PaCO_2$, and HCO_3^- values—but without revealing the diagnosis. This allowed students to analyze the information independently, apply physiological reasoning to determine the acid-base disorder, and then discuss their reasoning and the feedback from the LLM with their instructor to confirm accuracy and reinforce understanding. To support this process, I developed a structured, instructor-guided approach that helped students refine prompts in real-time.

Teaching Students to Generate Effective Prompts

Refining LLM-generated case studies was an iterative process that occurred over the course of a few minutes during one-on-one office visits. During that time, students generated an initial prompt, reviewed the LLM's output, and–through guided discussion–revised the prompt. These sessions provided real-time support and created a structured opportunity for students to develop effective prompts tailored to learning acid-base balance. In cases where the LLM output was inaccurate or overly detailed, we worked together to revise the prompt so that the output better aligned with specific learning outcomes.

During these sessions, we discussed what made a case study useful. Effective cases include only the most clinically relevant information (e.g., brief patient history and arterial blood gas or ABG values), while omitting the diagnosis. This information encouraged students to reason through the case. Prompts that provided complex histories, irrelevant lab values, or the diagnosis were considered less effective. This helped students learn to craft prompts that would foster learning.

When first introducing LLM-generated cases to students, I started with a broad prompt:

Please generate a case study for a patient with an acid-base disorder.

The LLM produced a detailed clinical scenario, including patient history, vitals, lab results, diagnosis, and treatment plan. The output included more information than necessary for my introductory physiology class, where students are expected to diagnose acid-base disorders based solely on arterial blood gases (ABGs).

Through instructor-guided trial and error, students worked with me to refine their prompts. After generating a prompt and reviewing the LLM-generated output, we discussed what made the prompt effective or ineffective. The student would then revise their prompt based on this discussion. An example of a refined prompt was:

Please create a case study where I can diagnose a patient with an acid-base disorder but only provide the most relevant information needed to diagnose.

This prompt generated an output containing the most pertinent diagnostic information, including basic patient information, symptoms, and ABGs. However, the LLM still included the diagnosis, which prevented students from thinking through the case independently.

A final refinement –developed with instructor guidance– explicitly requested:

Generate a case study with pH, $PaCO_2$, and HCO_3^- values, but do not provide a diagnosis.

This instructor-guided, student-driven learning experience allowed students to develop their own diagnostic reasoning before consulting the LLM for feedback. The structure came from a clear sequence of steps: students generated a prompt, evaluated the LLM's output, discussed the strengths and weaknesses of the output with the instructor to determine its accuracy and usefulness, and, if necessary, revised their prompt. In just a few minutes, students learned how to create custom case studies, test their knowledge with novel problems, and refine their understanding of acid-base balance.

Encouraging Active Engagement with LLMs

While LLMs can quickly generate responses, one goal of education is to cultivate reasoning and understanding. It is not enough for students to arrive at the right answer. How they determined that answer matters. Being correct for the wrong reasons does not constitute success—students need to arrive at the correct answer through sound reasoning. When students articulate their thought process (e.g., This patient has respiratory acidosis with partial metabolic compensation because pH is low, CO_2 is high, and

continued on next page

bicarbonate is high), the LLM can provide feedback on the user's answer and reasoning, which can reinforce conceptual understanding rather than rote memorization.

A common perspective on the use of LLMs in education is to treat it like a smart friend. A smart friend can be an invaluable resource, but only if you interact with them productively. Simply asking for answers doesn't help—especially since students cannot consult their friend (or an LLM) during exams. However, engaging in a discussion about their reasoning allows students to strengthen their foundational knowledge and critical thinking skills. Student feedback provides insight into how they used LLMs in this way, and what they found most beneficial during the prompt refinement process.

Feedback from Students

Many students have expressed appreciation for the opportunity to incorporate LLMs into their study routine (Table 1). Several students reported that using LLMs to generate and analyze acid-base cases helped them recognize patterns in compensation mechanisms and solidify their diagnostic reasoning. Multiple students have shared specific insights on how AI helped make their studying more effective.

One student said she liked learning how "*to ask AI to give me exam-like practice problems to master concepts!*" Another student noted that using AI to create case studies provided a valuable resource for independent study, saying she learned how "*to use AI when studying at home and didn't have access to extra problems*."

Beyond acid-base balance, students have stated they have applied the same approach to other physiology topics and classes (Wang et al., 2024), which demonstrates the broader adaptability and usefulness of AI-assisted learning. These insights highlight the potential of LLMs not just as an occasional study tool, but as an ongoing resource for active learning and critical thinking development.

While no formal assessment data were collected during this initial implementation, qualitative feedback from students (Table 1) indicates how they used LLMs to practice diagnostic reasoning, identify knowledge gaps, and support independent study, specifically during the acid-base balance unit.

Theme	Student Comment
Practicing diagnostic reasoning	*"I would ask AI to give me a scenario where the pH, CO_2, and HCO_3 were altered… to decipher what condition they were experiencing and what is causing it."*
Learning to craft effective prompts	*At first, it was a learning experience working with AI, giving it the correct prompts and information it needed to give me appropriate questions."*
Identifying knowledge gaps	*"It helped me apply what I already knew, along with the ability to ask for clarification when I did not know the answer, giving me awareness of any knowledge gaps."*
Studying independently	*"AI was helpful when I was studying at home and/or I did not have access to extra problems."*

Table 1. *Representative Student Feedback on LLM Use During Acid-Base Balance Instruction*

continued on next page

Recognizing the Limitations of LLMs

Both students and instructors must recognize that LLMs, like human tutors, are not infallible. They can, and do, make mistakes, sometimes with unwarranted confidence, which can confuse students (Memarian & Doleck, 2023). Teaching students to critically evaluate LLM-generated content rather than accept it at face value is a valuable skill.

By fostering a mindset of active engagement rather than passive reliance, we can encourage students to use LLMs to enhance their learning rather than undermine it.

Conclusion and Future directions

With support from an internal grant from the Digital Gardeners program at Indiana University Bloomington, I am developing a Canvas module to teach students how to use LLMs effectively and ethically. This module will introduce responsible AI use, including prompt refinement strategies and critical evaluation of AI-generated content

While tailored to physiology, the module will introduce strategies that can be applied across disciplines, encouraging students to approach AI as an interactive learning partner rather than a passive answer generator.

Based on these classroom experiences, I have begun to consider broader questions about the future of AI in education:

Could structured LLM-assisted case studies enhance learning outcomes in medical and health science programs?

How might AI-generated case banks complement traditional instructor-designed materials?

What are the long-term effects of AI-assisted learning on retention and problem-solving skills?

To integrate AI into physiology education, it is important to explore these possibilities and assess how AI tools affect retention and problem-solving over time. Thoughtful integration of LLMs into the curriculum offers opportunities to deepen conceptual understanding and extend learning beyond the classroom.

About the Author

Jim Davis, PhD, is an assistant professor in the Department of Anatomy, Cell Biology, and Physiology in the School of Medicine at Indiana University, Bloomington (IUSMB). Prior to joining the faculty at IUSMB in July 2022, Dr. Davis worked for 6 years as an assistant professor in the Department of Kinesiology, Recreation, and Sport at Indiana State University. During these seven years, Dr. Davis has primarily taught basic human physiology. His education research interests include exam retakes, use of a flipped classroom model in a large setting, and student engagement.

Literature Cited

Armbruster, P., Patel, M., Johnson, E., & Weiss, M. (2009). Active learning and student-centered pedagogy improve student attitudes and performance in introductory biology. *CBE—Life Sciences Education, 8*(3), 203–213. https://doi.org/10.1187/cbe.09-03-0025

Bopp, C., Foerst, A., & Kellogg, B. (2024). The case for LLM workshops. In *Proceedings of the 55th ACM Technical Symposium on Computer Science Education V. 1*. https://doi.org/10.1145/3626252.3630802

Dong, B., Bai, J., Xu, T., & Zhou, Y. (2024, April). Large language models in education: A systematic review. In *2024 6th International Conference on Computer Science and Technologies in Education (CSTE)* **(pp. 131-134).** IEEE. https://doi.org/10.1109/CSTE62025.2024.00031

Kasneci, E., Sesler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., et al. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences, 103*, 102274. https://doi.org/10.1016/j.lindif.2023.102274

Memarian, B., & Doleck, T. (2023). ChatGPT in education: Methods, potentials, and limitations. *Computers in Human Behavior: Artificial Humans, 1*(2), 100022. https://doi.org/10.1016/j.chbah.2023.100022

Meyer, J. G., Urbanowicz, R. J., Martin, P. C., O'Connor, K., Li, R., Peng, P.-C., Bright, T. J., Tatonetti, N., Won, K. J., & Gonzalez-Hernandez, G. (2023). ChatGPT and large language models in academia: Opportunities and challenges. *BioData Mining, 16*(1), 20. https://doi.org/10.1186/s13040-023-00339-9

Michael, J. (2007). Faculty perceptions about barriers to active learning. *College Teaching, 55*(2), 42–47. https://doi.org/10.3200/CTCH.55.2.42-47

continued on next page

Nazar, A. M., Selim, M. Y., Gaffar, A., & Ahmed, S. (2024). Revolutionizing undergraduate learning: CourseGPT and its generative AI advancements. *arXiv*. 2407.18310. https://doi.org/10.48550/arXiv.2407.18310

Van Wyk, M. M. (2024). Is ChatGPT an opportunity or a threat? Preventive strategies employed by academics related to a GenAI-based LLM at a faculty of education. *Journal of Applied Learning and Teaching, 7*(1), 35–45. https://doi.org/10.37074/jalt.2024.7.1.15

Wang, S., Xu, T., Li, H., Zhang, C., Liang, J., Tang, J., et al. (2024). Large language models for education: A survey and outlook. *arXiv*, 2403.18105. https://doi.org/10.48550/arXiv.2403.18105

Back to TOC

HAPS EDUCATOR
HUMAN ANATOMY & PHYSIOLOGY SOCIETY
Journal of the Human Anatomy and Physiology Society
PERSPECTIVES ON TEACHING
Volume 29, Issue 2 | August, 2025
https://doi.org/10.21692/haps.2025.014

Bridging the Gap Between 3-D Digital and Traditional Physical A&P Lab Models Using Smart Tablet Technology

Luis D. Rosado, PhD

Department of Biology, Worcester State University, Worcester, MA, USA

Corresponding author: lrosado@worcester.edu

Abstract

A disadvantage that A&P classrooms using 3-D digital computer anatomical models experience is the disconnect between digital and physical models traditionally used. This project used 3-D digital models on smart tablets and traditional physical models for in-class A&P lab practical exam studying and preparation. We focused on the hands-on experiential learning component seen in both constructivist and ecological pedagogical approaches. Digital models labeled by students served as "answer keys" to the physical models and "mock practical" exam questions were also created by the students on the physical models as well. We found students largely agreed that the smart tablet approach was an effective supplement to traditional physical models. We also found that most students preferred the 3-D models over the traditional physical models. Students reported that digital models "helped them make connections," "were a good reference," and "they preferred using the tablets in combination with the physical models." https://doi.org/10.21692/haps.2025.014

Key words: course design, 3-D computer models, student-to-student interactions, anchor-student

Introduction

The lab component is an important element of the A&P learning experience and has been widely supported for many decades as vital in science classrooms overall (Hofstein et al., 1982, 2007). However, the increasing availability of new digital and online technologies has changed how we teach A&P labs (Hofstein et al, 2004; Brinson, 2015; Rosado, 2023; Stokes et al.,2021;). Many worry that moving away from hands-on physical learning experiences to virtual digitally based learning modalities could negatively impact our students' learning (Brinson, 2015; Massey et al., 2021). Interestingly, this gap between digital and physical structural models initially emerged in the field architectural design. Traditionally, building designs had been developed and visualized in scaled physical models but as virtual and computer-aided design (CAD) technology developed, architects moved toward digital three-dimensional (3-D) building designs. The differences between these approaches in architecture has not, surprisingly, been studied for their individual strengths and weaknesses (Sun et al., 2014; Seichter, 2007). Even a previous HAPS position statement regarding cadaver use emphasized the strengths of constructivist and kinesthetic learning, often referred to as hands-on learning, and warned that some technologies should not be considered "equivalent alternatives" (Human Anatomy and Physiology Society, 2024).

Human beings are not architectural buildings, but the study of human anatomical structure has traditionally been through still 2-D images and physical anatomical models. However, three-dimensional (3-D) digital technology to study human anatomical structures is now advancing rapidly allowing for CAD like 3-D models of human structures of ever improving detail and accuracy (Rosado, 2023).

The disconnect between these two learning modalities appears to be particularly evident in the time honored "lab practical exam," which traditionally utilizes physical models and specimens in a timed, station-by-station, assessment of student A&P knowledge. Many of our students communicate the "gap" between these technologies manifests for them when they use new digital or online learning modalities and then are evaluated on traditional physical lab models and anatomical specimens. Taken together, this led us to ask if a

continued on next page

perceived gap exists in A&P study methods that translates to a reliance on one method to support the other and do students have a preference between the two modalities? To address this question, this project utilized 3-D digital anatomical models combined with traditional physical models to prepare students for lab practical exams. We hypothesize that through the self-report of their preferences, students perceive greater effectiveness in a combined study modality (reducing the gap) and prefer the digital 3-D models to traditional physical anatomical models.

Methods

The Institutional Review Board at Worcester State University approved this project as exempt, and informed consent was obtained from all participants. The classes surveyed were the standard two-semester sequence of A&P. Of the 30 respondents to the overall questionnaire, 12 reported they had used the combined iPad-to-physical model learning approach to practical exam preparation. While other students may have used their own tablets or smart devices, our results for "Combined Effectiveness" are based only on the responses from those 12 students. Additionally, while not the focus of this project, our survey also asked about the effectiveness of the Visible Body® online digital platform (see questionnaire) and those results agree with previously reported outcomes (Rosado, 2023) and are not reported here.

The in-class instruction varied slightly by instructor style but followed the basic process described here.

In-class procedures:

1. Labeling and tagging on an iPad all exam relevant structures on a 3-D digital model of an anatomical structure. E.g., the human skull (Figure 1).
2. Using the labeled structures on the iPad (Figure 1) as a type of "legend" or "key" students then labeled portions of the traditional physical model (Figure 1).
3. Students were then instructed to work in groups to create "mock" lab practical-type exam questions referencing labeled structures on the traditional models (Figure 1).
4. Once each group in class had completed a set of 3-5 questions per "station", a mock lab practical exam was conducted.
5. Students experienced the timed station-by-station assessment based on questions they and their peers had developed utilizing both the 3-D digital and traditional anatomical models.

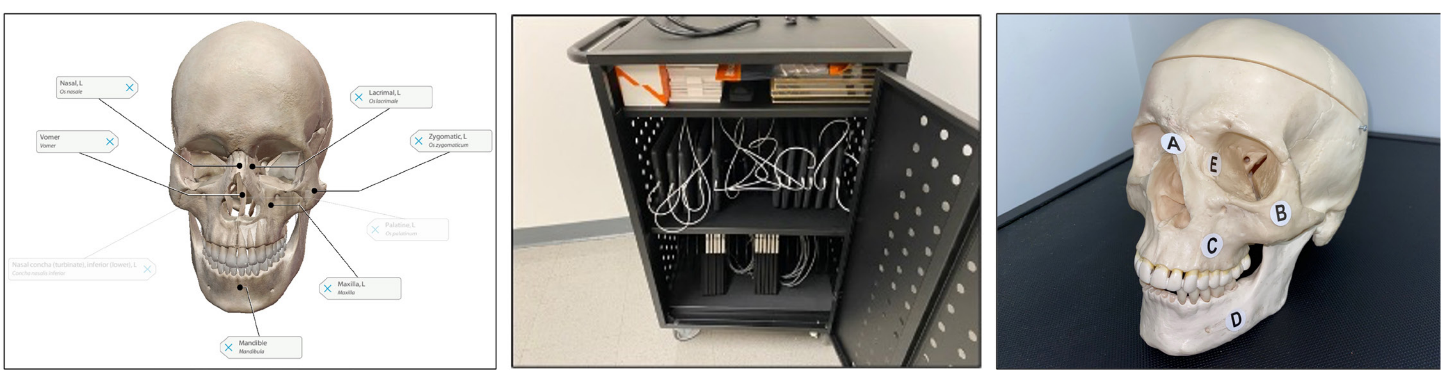


Figure 1. *A 3-D digital model of the skull tagged and labeled. Image courtesy of Visible Body® (left), twelve iPads used for tagging digital anatomical models (middle), and a traditional physical model of the skull with labeled structures (right).*

continued on next page

For self-reported modality effectiveness respondents replied with: Strongly Agree, Agree, Neutral, Disagree or Strongly Disagree to the following statements regarding:

Overall Modality Effectiveness:

1. The iPads helped me make connections with the physical models in lab.
2. The digital models were a good reference for the physical models in lab.
3. I preferred using the iPads in combination with the physical anatomical models in lab.
4. I preferred using the digital models only.
5. I preferred using the traditional physical anatomical models only.

A Chi-squared goodness of fit analysis was conducted to determine if all Likert responses were used evenly or if some responses were used at a greater frequency than a uniformed frequency distribution that assumed an equal response distribution cutoff of 12 (Figure 2). None of our students responded with "Strongly Disagree" responses therefore, the 5-point Likert scale responses were collapsed into three categories: Agree, which included "Strongly Agree" and "Agree" responses; Neutral for all the "Neutral" responses; and Disagree, which contained no "Strongly Disagree" responses (Figure 3). This ensured the interpretability and preserved ordinal directionality of the results. A Chi-squared test of independence was then conducted to determine the frequency of responses for each question using a 3X5 contingency table at $\alpha = 0.05$.

For preferences between the two modalities, respondents replied either: "3-D computer models", "traditional physical models", or the neutral response "no difference between the two learning modalities" to the following questions:

Modality Preference:

1. Which helped most in improving your knowledge of the human skeletal system?
2. Which helped most in improving your understanding of the human muscular system?
3. Which helped most in improving your understanding of macroanatomy?
4. Which helped most in improving your understanding of microanatomy?
5. Which did you use the most?

To statistically determine a significant two-way directional preference between either modality, all neutral responses were excluded and evaluated if the responses favored either 3-D models or the traditional physical models. This test was based on if the responses favoring 3-D digital models differed from chance established as a 50% likelihood. A two-tailed binomial test was therefore conducted, in which the number of "successes" represented the count of participants selecting 3-D models over the traditional physical models. The statistical significance was also evaluated at $\alpha = 0.05$ to determine whether responses for 3-D models were significantly different from a 50% cutoff (0.5). For each question with a p-value less than 0.05, a difference greater than 50% indicated a significant preference for 3-D models over the traditional physical models (Figure 4).

Results

For overall modality effectiveness, the Chi-squared goodness of fit analysis across all five responses (X^2 (4, N = 60) = 19.67, p<0.01) indicated that participants showed a clear preference for certain responses over others (Figure 2). To further clarify response preference, the Chi-squared test of independence of the three collapsed categories Agree, Neutral, and Disagree resulted in significant differences in response category (X^2 (8, N = 60) = 15.92, $p = 0.0436$) wherein question-1 through question-3 had higher agreement and question-4 indicated more disagreement (Figure 3).

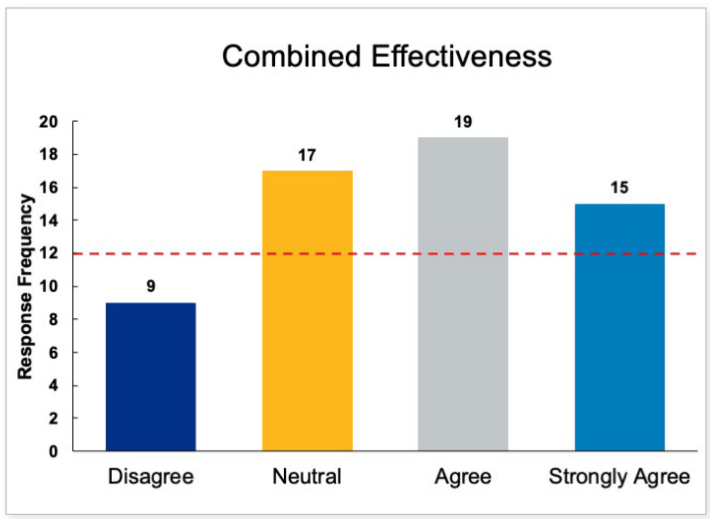
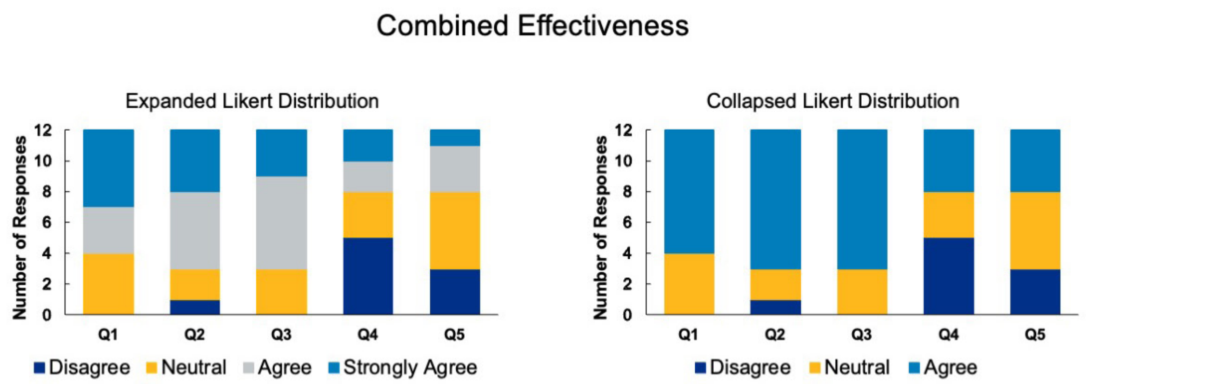


Figure 2. *Observed frequencies of participant responses across a 5-point Likert scale compared to an expected uniform distribution (cutoff set at 12). Responses of "Agree" and "Neutral" were more frequent than expected, while "Strongly Disagree" was never selected. A Chi-squared goodness-of-fit test indicated a statistically significant deviation from uniformity.*

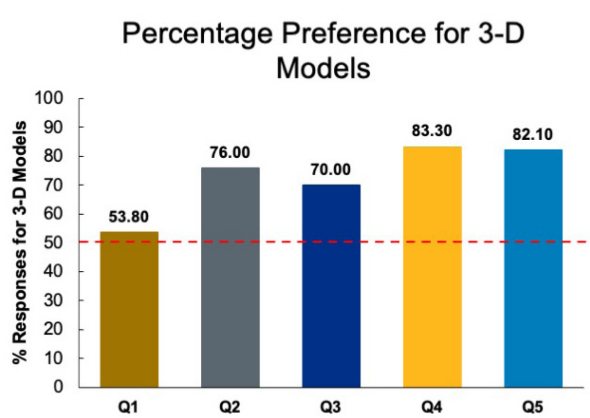
continued on next page



Question	Strongly Disagree	Disagree	Neutral	Agree	Strongly agree
Q1 = The iPads helped me make connections with the physical models in lab.	0	0	4	3	5
Q2 = The digital models were a good reference for the physical models in lab.	0	1	2	5	4
Q3 = I preferred using the iPads in combination with the physical anatomical models in lab.	0	0	3	6	3
Q4 = I preferred using the digital models only.	0	5	3	2	2
Q5 = I preferred using the traditional physical anatomical models only.	0	3	5	3	1

Figure 3. *The expanded 5-point Likert response distribution for each of the five survey questions (left). Likert response data collapsed into three categories: "Disagree," "Neutral," and "Agree" (right). A Chi-squared test of independence revealed that response patterns varied significantly across responses. The accompanying table presents raw Likert response frequencies for each question.*

The directional two-tailed binomial test between participant preference for 3-D models or traditional models showed that overall, they preferred the 3-D models over the traditional physical models (Figure 4). However, only 3 of the 5 questions showed a significant preference for the 3-D models. Questions 1 and 3 did not show a preference at p = 0.85 and $p = 0.12$, respectively. However, questions 2 showed 3-D models were preferred 76% of the time (n = 25, $p = 0.01$), question 4 preferred 3-D models 83.3% of the time (n = 24, $p < 0.01$), and question 5 preferred 3-D models 82.1% of the time (n=28, $p \leq 0.01$).



Question	N	3-D Models Selected	Percentage for 3-D Models	p-value
Q1 = Which helped most in improving your knowledge of the human skeletal system?	26	14	53.8	0.845
Q2 = Which helped most in improving your understanding of the human muscular system?	25	19	76	0.0146
Q3 = Which helped most in improving your understanding of macroanatomy?	20	14	70	0.1153
Q4 = Which helped most in improving your understanding of microanatomy?	24	20	83.3	0.0015
Q5 = Which did you use the most?	28	23	82.1	0.0009

Figure 4. *Proportion of participants preferring the 3-D models for each question, expressed as percentages. The dashed line at 50% indicates no preference. The accompanying table reports sample sizes (N) excluding neutral responses, counts for and percentage of responses for 3-D models, and corresponding p-values for each test.*

continued on next page

Discussion

Our study indicates our students perceive a gap between the 3-D anatomical models and the traditional physical models alone. While they prefer using the 3-D models, they also prefer using them combined with the traditional physical models. Intuitively, an exam that primarily uses physical models would likely skew the agreement toward only using the physical models to study and prepare for said exam. However, the effectiveness of the tablets in supplementing the physical models is shown in the student responses to the first three questions in that question set (Figure 3), indicating that students agreed the 3-D digital content helped make connections, acted as a good reference, and worked best in combination with the physical models. Likewise, the weaker response to questions 4 and 5 indicates that the combination of the two modalities was perceived as more desirable (Figure 3). This further indicates a desire to narrow the perceived gap between only using 3-D or traditional physical models without the other study modality. These are compelling results that indicate the nature of the gap and the strength of using both modalities to prepare for the lab practical exams. A potential limitation to these responses is that the in-class mock lab practical exam preparation forced students to use both modalities. While that is certainly the case, the 3-D digital models can be used at any time of day and in any location in which students have access to their own laptops or smart tablet technology to study. While we did not control for it in this study, presumably, students also studied for their exams while outside of class using the digital models they had created while in the lab setting as a reference. Future studies should explore to what extent students rely on the 3-D models outside of class time to prepare for lab practical assessments.

Our results also indicate that between the two modalities students preferred to use the 3-D models over the traditional models (Figure 4) albeit at varying amounts depending on anatomical system. It appears that the 3-D models were most preferred for the muscular system and microanatomy (Q2 & Q4, respectively) and less strongly preferred for the skeletal system and microanatomy (Q1 & Q3, respectively) while remaining the preferred modality overall (Q5). These data may shed some light on which systems have a stronger reliance on the traditional physical models to supplement studying and preparations. One possibility is that many of the skeletal system models such as fully articulated skeletons and disarticulated bone sets may already sufficient information to associate with the 3-D digital models. This may also be the case for any other macroanatomy models that can be easily supplemented with the 3-D digital content. However, it could also be that the age and the overuse of our current muscle models and lack of robust microanatomy physical models may lack sufficient context driving the increased preference for using the information rich 3-D digital models themselves.

Overall, respondents answered favorably regarding the 3-D digital model's effectiveness but it is unclear if the other portions of the in-class modality, such as the "mock exam" practice component, influenced the learning efficacy for our students. It is worth exploring to what extent the practice "mock exam" itself impacts student learning, especially for students that have never experienced a timed station-by-station practical lab exam. Ultimately, it appears that for students the combination of technology-based and physical learning modalities is more beneficial than relying too heavily on only one or the other.

This project was conducted on a small scale and while it provides interesting results it may not translate to an overall A&P program evaluation or a larger population of students. These reported results suggest there is something worth exploring further on a larger scale and has formed the basis for an ongoing A&P program evaluation at our institution. The ongoing investigation examines lab practical performance compared to overall class grade. This study aims to provide the added component of a performance-based assessment of this type of lab practical preparations using 3-D and traditional models during mock lab practical assessments. While we do think it is important to give attention to performance measures on assessments, we also believe it can be limiting to rely too heavily on these outcomes alone as measures of successful learning of A&P material. Therefore, our current ongoing project will include an assessment of fundamental A&P concept retention measuring baseline knowledge before and retained knowledge after the two A&P course sequence. In this ongoing research project, my colleagues and I hope to add important performance metrics to the outcomes described in this paper.

Acknowledgments

I would like to thank the financial support provided by the Worcester State University's Teaching and Learning Innovation Grant (TLIG) program. This teaching and technology focused award have made the implementation of the iPads for our A&P courses possible. I would also like to thank the WSU A&P teaching team of instructors who delivered the survey to their students and who utilized the iPads and this learning approach in their classrooms. Lastly, I would like to thank our Lab Technician II, Ashley Landry, M.S., for her tireless work providing technical support to the many lab sections at WSU.

About the Author

Luis Rosado, PhD, is an assistant professor and the Anatomy & Physiology course coordinator in the Department of Biology at Worcester State University. He teaches human biology, A&P, basic kinesiology, endocrinology, and human movement & perception. His research goals aim to improve the integration of technology in all human A&P-related courses. His broader research interests include human visual perception from Gibson's Ecological Approach. He is also very involved in the HAPS DEI committee and actively promotes diversity, equity, and access in the A&P two-course sequence at Worcester State University.

continued on next page

Literature Cited

Brinson, J. R. (2015). Learning outcome achievement in non-traditional (virtual and remote) versus traditional (hands-on) laboratories: A review of the empirical research. *Computers & Education, 87*, 218–237. https://doi.org/10.1016/j.compedu.2015.07.003

Hofstein, A., & Lunetta, V. N. (1982). The role of the laboratory in science teaching: Neglected aspects of research. *Review of Educational Research, 52*(2), 201–217. https://doi.org/10.3102/00346543052002201

Hofstein, A., & Lunetta, V. N. (2004). The laboratory in science education: Foundations for the twenty-first century. *Science Education, 88*(1), 28–54. https://doi.org/10.1002/sce.10106

Hofstein, A., & Mamlok-Naaman, R. (2007). The laboratory in science education: The state of the art. *The Royal Society of Chemistry, 8*(2), 105–107.

Human Anatomy and Physiology Society. (Adopted January 13, 2001, revised April 19, 2024). *Human Anatomical/ Body Donor Position Statement*. https://www.hapsweb.org/about-us/position-statements/

Massey, A., Zhang, W., & Amar, A. (2021). A comparison of non-traditional online and traditional wet-lab experiences in human anatomy and physiology: An innovative approach for pre-licensure nursing education. *Nurse Education Today*. https://doi.org/10.1016/j.nedt.2021.105149

Rosado, L. D. (2023). Utilizing 3-D digital models in synchronous blended anatomy & physiology courses during the COVID-19 pandemic. *HAPS Educator, 27*(3), 29–38. https://doi.org/10.21692/haps.2023.019

Seichter, H. (2007). Augmented reality and tangible interfaces in collaborative urban design. In *Computer-aided architectural design futures* (Vol. 7, pp. 3–16). Springer. https://doi.org/10.1007/978-1-4020-6528-6_1

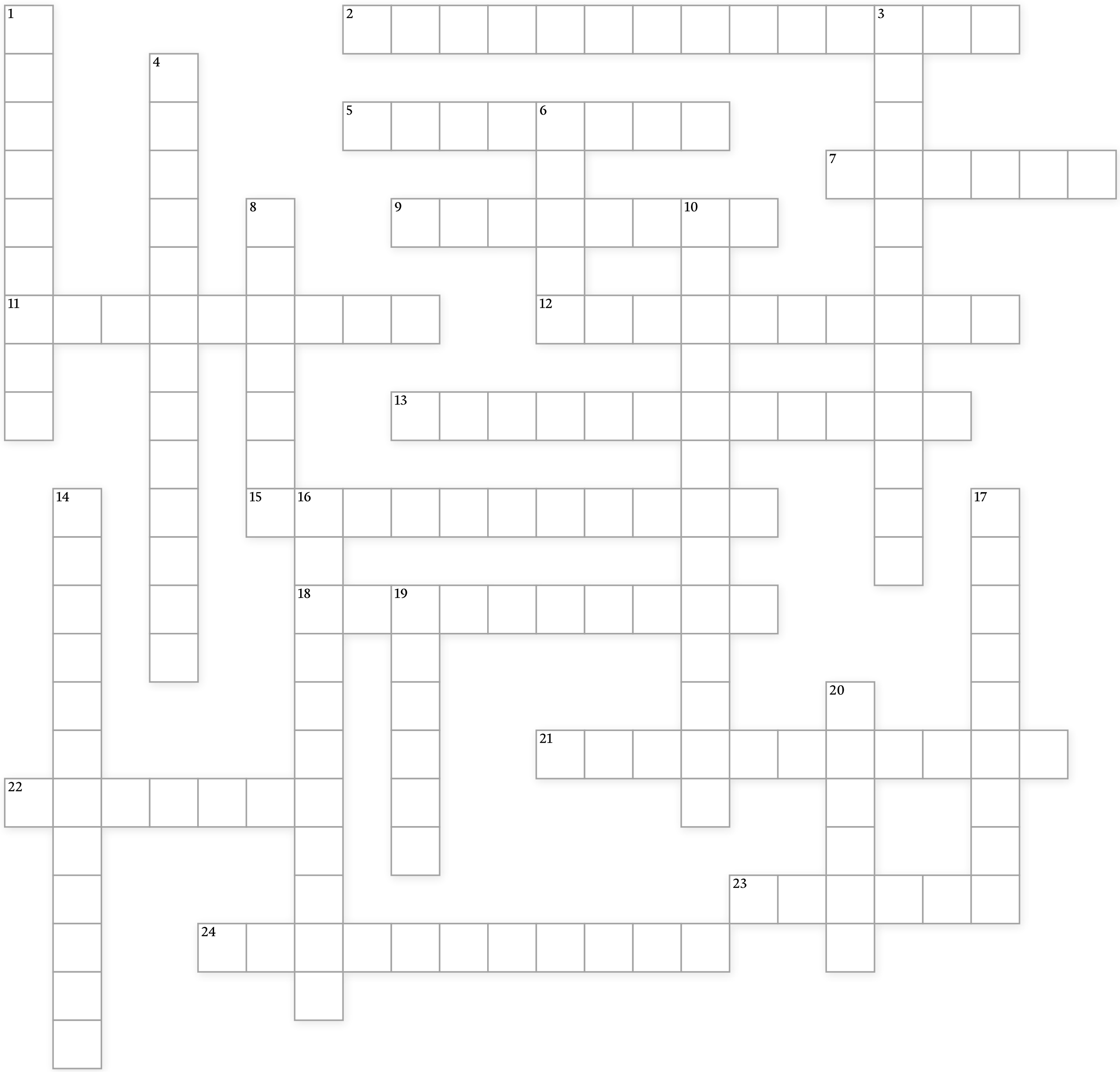
Stokes, J. A., & Silverthorn, D. U. (2021). Updating anatomy and physiology lab delivery: Shifting from a paper-based to an online lab instruction platform, just in time for a global pandemic. *Advances in Physiology Education, 45*, 290–298. https://doi.org/10.1152/advan.00190.2020

Sun, L., Yu, Z., Zhang, X., & Lee, H. (2014). Differences in spatial understanding between physical and virtual models. *Frontiers of Architectural Research, 3*, 28–35. https://doi.org/10.1016/j.foar.2013.11.005

Back to TOC

HAPS Educator Crossword 5: Digestive System

(For an online version CLICK HERE.)



ACROSS

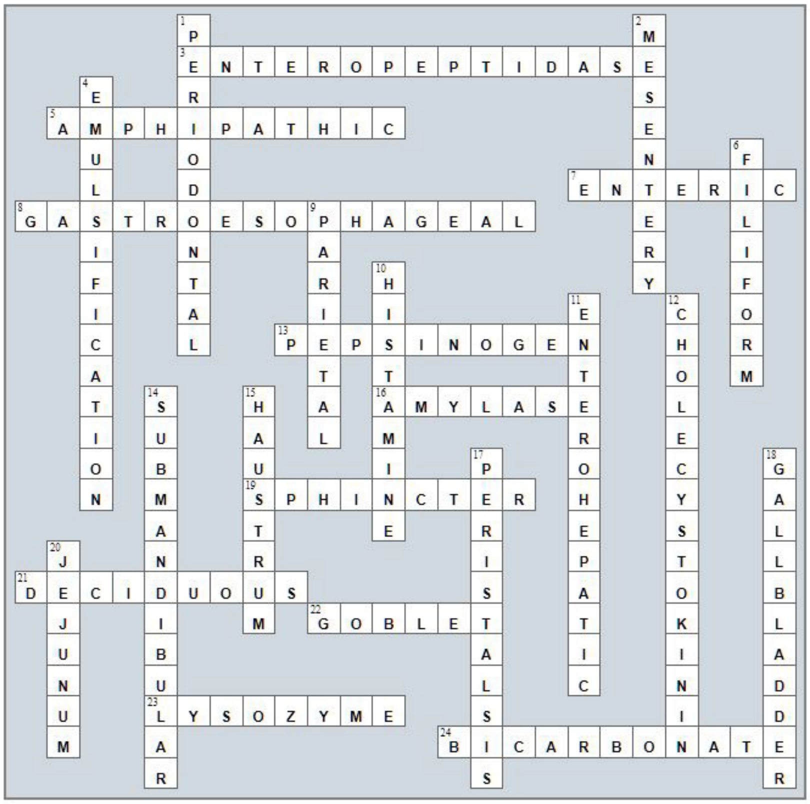
3. Brush border enzyme located in the duodenum that activates trypsin.
5. Bile salts are .... molecules, meaning that they have polar and nonpolar regions.
7. The .... nervous system consists of the submucosal and myenteric nerve plexuses.
8. Sphincter at the entrance to the stomach.
13. Enzyme precursor produced by gastric chief cells.
16. Starch digesting enzyme produced by the salivary glands and the pancreas.
19. A ring of smooth muscle at the entrances and exits of GI organs.
21. "Baby" teeth are more correctly referred to as .... teeth.
22. Cell type along the length of the GI tract that produces mucus.
23. Antibacterial enzyme found in saliva.
24. Pancreatic duct cells secrete this important component of pancreatic juice.

DOWN

1. The .... ligament attaches each tooth to the bony material of the jaw.
2. Double layer of peritoneum that allows blood vessels access to digestive organs.
4. Process of physically breaking up large fat droplets into smaller droplets.
6. The .... papillae provide friction on the surface of the tongue.
9. Gastric cell that produces both HCl and intrinsic factor.
10. Gastric acid production is stimulated by gastrin, acetylcholne, and ....
11. Bile salts are returned to the liver via the .... circulation.
12. Intestinal hormone that stimulates gallbladder contraction.
14. Parotid, sublingual and .... are the three pairs of salivary glands.
15. Pouch of the large intestine due to the teniae coli.
17. Directional motility pattern in GI tract.
18. Site of storage and concentration of bile.
20. Middle region of small intestine.

CLICK HERE for Answer Key

Answer key for: Crossword 5. Digestive System (from previous page)



GO BACK to the puzzle

Back to TOC

HAPS COMMITTEES AND BOARDS

Standing Committees:

2026 ANNUAL HOST COMMITTEE

Todd Gordon

This committee is in charge of coordinating the 2026 Annual Conference to take place in Kansas City, Kansas

ANATOMICAL DONOR STEWARDSHIP

Jeremy Grachan

This committee is charged with developing, reviewing, and recommending policies and position statements on the use of cadavers for human anatomy and physiology education in colleges, universities and related institutions.

AWARDS & SCHOLARSHIPS

Gilbert Pitts

This committee recognizes talented HAPS Members through achievement awards, conference travel awards, and scholarships for HAPS-I.

COMMUNICATIONS

Caitlin Burns

This committee is tasked with helping HAPS establish its voice in a technological landscape shaped by social media. Committee members work closely with the Marketing Committee to facilitate connections within HAPS as well as recruiting potential members via social media.

CONFERENCES

Beth Eischen

This committee actively encourages HAPS members to consider hosting an Annual Conference. We provide advice and assistance to members who are considering hosting an annual conference.

CURRICULUM & INSTRUCTION

Abbey Breckling

This committee develops and catalogs resources that aid in anatomy and physiology course development and instruction.

DIVERSITY, EQUITY, AND INCLUSION

Jennifer Stokes

This committee has the goal of creating spaces of belonging and accessibility for all members by embracing diversity and promoting equity and inclusion.

FUNDRAISING

Stacey Dunham

This committee supports HAPS and its members by seeking donations from those within the organization as well as external funding sources.

Click here to visit the HAPS committees webpage.

Special Committees and Programs:

WELCOMING AND BELONGING

Chasity O'Malley & Caitlin Hyatt

This committee identifies opportunities for member recruitment, retention, and engagement that foster an inclusive and welcoming environment promoting professional and personal growth.

STEERING

Chasity O'Malley

This committee consists of all committee chairs. It coordinates activities among committees and represents the collective committee activity to the HAPS BOD.

HAPS EDUCATOR

Jackie Carnegie, Editor-in-Chief
Brenda del Moral, Managing Editor

This committee is responsible for publishing spring, summer and winter editions of the HAPS Educator, the journal of the Human Anatomy and Physiology Society. The committee works closely with the Steering Committee and the President of HAPS.

EXAM PROGRAM LEADS

Valerie O'Loughlin
Dee Silverthorn
Janet Casagrand

This committee is a closed program managed by three Program Leads (Anatomy, Physiology, Anatomy & Physiology) that is charged with developing, maintaining, securing, and managing the HAPS standardized exams.

EXECUTIVE

Rachel Hopp

Composed of the HAPS President, President-Elect, Past President, Treasurer and Secretary

FINANCES

Ron Gerrits

NOMINATING

Larry Young

This committee recruits nominees for HAPS elected offices.

PRESIDENTS EMERITI ADVISORY COMMITTEE

Melissa Quinn

This committee consists of an experienced advisory group including all Past Presidents of HAPS. The committee advises and adds a sense of HAPS history to the deliberations of the BOD

Back to TOC