Established in 1989 by Human Anatomy & Physiology Teachers

HUMAN ANATOMY & PHYSIOLOGY SOCIETY

EDUCATOR

Journal of the Human Anatomy and Physiology Society

VOLUME 22, ISSUE 3 • DECEMBER 2018

Promoting Excellence in the Teaching of Human Anatomy & Physiology

HAPS Educator

WINTER 2018 TABLE OF CONTENTS

EDUCATIONAL RESEARCH

CURRENT TOPICS IN ANATOMY AND PHYSIOLOGY

PERSPECTIVES IN TEACHING

Cover art: Synapse

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 submissions are disseminated to non-HAPS members one year post-publication via the Life Sciences Teaching Resource Community database.

The HAPS Educator is published in April, August and December. The deadlines for submission are March 15, July 15 and November 15.

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 Links

Use the **Manuscript Submission form** for *HAPS Educator* submissions and the **Teaching Tips for A&P (Snippets)** form for shorter teaching tips.

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.

Medical Student Study Strategies in Relation to Class Size and Course Length

Polly R. Husmann, PhD
Indiana University, Jordan Hall 112, School of Medicine, 1001 E. 3rd St. Bloomington, IN 47405
phusmann@indiana.edu

Abstract
The influence of class size and course length on student academic achievement has been evaluated. However, most of the studies have been done at the primary education level and their influence on study strategies has not previously been assessed. The data for this study were collected via surveys at the beginning and end of a first year medical school anatomy course. Students in the study were divided into cohorts based on class size and course length and their study strategies were examined. The differences among cohorts were related to the discrepancy between the study strategies students initially thought they would use in the course and the study strategies students reported actually using. This article describes specific differences in study strategies that are related to class size and course length and examines how diverse study strategies may affect the long-term retention of knowledge. http://doi.org/10.21692/haps.2018.024

Key words: study strategies, course duration, large class, medical education

Introduction
A great deal of learning occurs outside of the classroom, yet aside from anecdotal evidence, not much is known about how students study on their own (Dunn-Lewis *et al.* 2016). If significant learning takes place outside of the classroom, determining the best way to do it is of paramount importance to students who want to master the course material and become successful learners. The first step in this process is determining how students are currently studying and what influences their studying decisions. Previous research has found that students who view the material as relevant to their future tend to do better in the course (Selvig *et al.* 2015) and that student confidence, motivation, and time on task correlate with higher grades (Pizzimenti and Axelson 2015, Dunn-Lewis *et al.* 2016, Husmann *et al.* 2016). Previous literature has also demonstrated the importance of an overarching approach (Tan and Thanaraj 1993, Papinczak *et al.* 2008, Ward 2011) as well as more specific study strategies (Prince 2004, Ward and Walker 2008, Selvig, *et al.* 2015).

Overarching approaches to studying may be divided into three categories. A surface or superficial approach to studying emphasizes recreating the content exactly as it was presented to the learner. A deep overarching approach attempts to integrate the new information with previous knowledge that the student brings to the classroom. A third overarching approach employs both superficial and deep learning as required by the course and its assessments (Papinczak, *et al.* 2008). Research has shown that first year medical students often take a surface approach to studying (Tan and Thanaraj 1993) and that use of this approach increases throughout the first year of medical school (Martenson 1986, Tooth *et al.* 1989, Papinczak *et al.* 2008, Ward 2011).

Individual study strategies can be divided into passive techniques, which emphasize the student receiving information from an official source (e.g. professor, notes, textbook), and active techniques, which emphasize the student personally engaging with the material (Prince 2004). Previous studies have shown that first year medical students tend to study anatomy more than any other subject (Malleson 1967) and that they generally utilize passive study techniques (e.g. viewing podcasts, re-reading or re-writing notes, cramming) (Entwistle 1960, Shatin 1967, Crombag *et al.* 1973, Ward and Walker 2008, Selvig *et al.* 2015). The use of passive study techniques during the first year of medical school may be correlated with lower grades (Selvig *et al.* 2015).

In a previous study examining the way that students report studying for anatomy and physiology during their first year of medical school, we found that there were very few significant relationships among specific study strategies (e.g. attending lecture or reading the text) and course outcomes (Husmann *et al.* 2016), which was consistent with earlier studies from the 60s and 70s (Shatin 1967, Malleson *et al.* 1968, Crombag *et al.* 1973). Yet students that reported using different study habits for anatomy and physiology tended to have lower final grades than those who used fewer consistent strategies. However, questions remained about how much change in study habits was actually occurring within each individual course (anatomy or physiology) versus between the two courses. Thus, one aim of this project was to evaluate how and how much students are changing their study strategies during a single course (Gross Human Anatomy) in the first year of medical school. The present research attempts to address this question by asking first year medical students to complete a pre-course survey on how they plan to study for the course, followed by

continued on next page

a post-course survey to assess how they actually studied. By assessing both pre- and post-course surveys, it is possible to gain insight into what knowledge or expectations of study strategies our medical students bring to the course and how that knowledge changes by the end of the course.

A second aim of this project was to evaluate the influence that course logistics have on study strategies and the changes in these strategies that may occur during the course. For example, with more and more medical schools converting away from the traditional undergraduate model of multiple subject-based classes (e.g. anatomy, physiology, biochemistry) towards an integrated model of a single condensed system or case-based blocks, how might student study strategies be different with a longer block versus a shorter block?

One additional factor that may affect these study strategies is the size of the class itself. For instance, in a smaller class, students may feel more connection with the instructor and thus, feel more motivation to perform well in the class. On the other hand, in a larger class, there may be less opportunity to interact with the instructor. This may force students to interact with the material more themselves (more active techniques) and thus improve student understanding of the material. Thus, this study will compare pre- and post-course study strategies among three cohorts with varying class sizes and course lengths to determine the commonalities and the differences in the study strategies employed for each context. This new information will then enable us to better advise our students on effective study strategies and to assess the impact of course logistics on these study choices, which in turn affect student success in our classrooms and beyond.

Class Size

Educational literature, largely with a primary school context, suggests that the overall relationship (across all ages, all courses, all students) between class size and academic achievement is unclear. Some studies have shown a positive relationship between class size and achievement, some studies have reported a negative relationship, and other studies have reported no significant relationship at all (Glass *et al.* 1982, Fleming *et al.* 2002). However, most research shows that smaller classes do seem to benefit students in the earliest grades and disadvantaged students (Fleming *et al.* 2002). Glass *et al.* (1982) reported that while the relationship between smaller classes and academic achievement is unclear, the relationship between smaller class sizes and affective variables is much more straightforward. For example, students and teachers generally prefer smaller class sizes as the smaller classes increase time on task (due to less time spent waiting for help, grade checks, etc.), decrease inattentiveness, improve faculty ability to adjust to student interest and learning speeds, boost morale, and increase non-content focused interactions between students and teachers (Glass *et al.* 1982). It is not much of a stretch to consider that noted differences in these affective components of a class may well translate into differences in study strategies as well. For example, if students have better rapport with a faculty member or interest in a class, this may influence their motivation for the class and thus their study strategies.

In 2005, the American Association of Medical Colleges (AAMC) called for a thirty percent increase in medical student enrollment to help prevent a future physician shortage (Erikson *et al.* 2014). Since that time, medical student enrollment has been increasing regularly (Schieffler *et al.* 2012, Association of American Medical Colleges 2015b, Association of American Medical Colleges 2015a) with two-thirds of that growth occurring at schools that were already accredited in 2002. The remaining third of the growth results from new medical schools (Erikson *et al.* 2014). In fact, Schieffler *et al.* (2012) reported that 83% of 125 medical schools had increased their enrollment as of 2009, yet few of these schools were concerned about additional financial resources required for this expansion, which suggests that more students were to be added with the extant personnel and resources.

Despite rising class size, research on the effects of class size in medical schools is incredibly scarce. Fifty years ago, Sanazaro (1966) found no evidence that larger class size decreased academic achievement. Yet a lot has changed in our medical schools (and our students) since that time and these changes not been evaluated. Brady and Eisler (1999) found that smaller class sizes were generally more interactive at the college level, which would suggest more active learning. Mahler and Neumann (1986) found more activity in smaller college classes, though even their "large" classes were only 17-50 students. Yet none of this previous work includes any discussion of outside study strategies. Thus, additional research is clearly needed on class size and its relationship to student study strategies, particularly in higher education.

Course Length

The basis for questioning the importance of course length for learning may be found in the literature on the "spacing effect" or "distributed practice". Dobson *et al.* (2017) explain that "distributed practice refers to spacing out one's practice or relearning material intermittently over time…" (p. 340) while Verkoeijen *et al.* (2004) define the "spacing effect" as "…the phenomenon that repeated items induce better recollection if both occurrences are separated by time or other targets (i.e. spaced presentation), compared with a situation in which repetitions occur in immediate succession (i.e. massed presentation)" (p. 796). The importance of spacing material over time for long-term retention has been established in the study literature (Greene 1989, Verkoeijen *et al.* 2004, Pashler *et al.* 2007, Rohrer and Pashler 2007, Rohrer and Pashler 2010). In the retrieval literature the importance of spacing study material was emphasized by Roediger and Karpicke 2006, Karpicke and Roediger 2007, Karpicke and Bauernschmidt

continued on next page

2011, Soderstrom *et al.* 2016, Dobson *et al.* 2017, and Whiffen and Karpicke 2017. in the feedback literature this concept was supported by Butler *et al.* 2007 and Pasher *et al.* 2007.

In their studies on optimal intersessional study intervals and retention intervals, Rohrer and Pashler (2007) have repeatedly found that long-term retention is more likely if studying is widely spaced. They have also noted that delaying feedback (as opposed to immediate feedback) as well as interleaving (as opposed to blocking) different types of skills or information encourage long-term retention (Pashler *et al.* 2007, Rohrer and Pashler 2007, Rohrer and Pashler 2010). Karpicke and Roediger (2007) have also repeatedly found the benefits of spaced or distributed practice in their studies on retrieval. In particular, they found the absolute spacing (multiple practice sessions spaced out over longer periods of time) to be beneficial for long-term retention of the material regardless of the relative spacing between the study sessions (Roediger III and Karpicke 2006, Karpicke and Bauernschmidt 2011). Dobson *et al.* (2017), in a study with sixty undergraduate students, further demonstrated the retention benefits of distributed practice, specifically for anatomy content. The practical application of these studies suggests that more condensed courses covering a large volume of information may not lead to the best long-term retention of the material (Pashler *et al.* 2007).

Unfortunately, previous studies into course logistics at the medical school level have largely focused on the number of course hours, hours spent in the classroom and/or laboratory during regularly scheduled class time (Drake *et al.* 2002, Drake *et al.* 2009, Cuddy *et al.* 2013). However, few researchers have evaluated the effects of the span of time in which those hours occurred, specifically the course length or course duration. Holla and colleagues (2009) did an evaluation of an 18-month curriculum versus a 12-month curriculum for gross anatomy. They found that the majority of students would prefer additional time to master gross anatomy, however they did not compare the academic achievement of the groups.

In business education, there is some evidence that students in short term, intensive courses perform better on course examinations than students in longer, more traditional courses (Van Scyoc and Gleason 1993, Austin and Gustafson 2006). However, the students' long-term retention of the material showed no significant differences based on the length of the course (Van Scyoc and Gleason 1993). Another study with business undergraduates found that students in a longer class did perform better overall, but that this effect was lost when only multiple choice questions were considered (Rayburn and Rayburn 1999). Thus, the type of questions, and possibly the level of critical thinking required for the questions, can also make a difference in what effects are seen.

Methodology

Class set-up

At Indiana University School of Medicine (IUSM), there are nine different campuses throughout the state of Indiana. Admissions procedures are identical for all IUSM students. After admission, students are assigned to one of the nine campuses based on their personal preferences. This process should minimize inter-campus variability in students. Faculty from all nine campuses are evaluated using the same criteria, though individual differences in background and experience could not be controlled for this study.

Prior to the 2016-2017 academic year, each campus of the Indiana University School of Medicine was run slightly differently. Table 1 gives an overview of the similarities and differences among the different campuses. Class sizes ranged according to campus from 23 to 158 students. All campuses had gross anatomy in the first year of medical school starting in the fall semester. However, the length of the course varied across the campuses between ten weeks and thirty weeks (two semesters or one academic year).

Gross anatomy on all campuses included both a classroom component and a laboratory component with cadaveric dissection. For the purposes of this project six of the campuses were chosen based on using a predominantly lecture format for the classroom component. Thus, the following campuses were excluded: one that used a predominantly team-based learning curriculum, one that used a predominantly problem-based learning curriculum, and one that had an instructor just visiting to present the content during the 2015-2016 academic year (and thus the instructor was not available on a regular basis as would be the case at the other campuses).

Table 1. *Similarities and differences in gross anatomy course across campuses of Indiana University School of Medicine prior to 2016-2017 academic year.*

Statewide Course Similarities across campuses	Statewide Course Differences among campuses
80% Core Content (as determined by session-level learning objectives) Start in Fall of first year Primarily lecture-based Cadaveric Dissection NBME shelf exam (at least 20% of final grade) 70% pass cut off	20% Discretionary Content Course length Class size Assessments for each content block

continued on next page

Each campus was required to cover all the same course-level learning objectives and 80% of the common session-level learning objectives, but was then provided some leeway in where they wanted to add the additional 20% of material. Each campus was also allowed some variation in the number and types of assessments, though both written (predominantly multiple-choice questions) and laboratory practical examinations (fill-in-the-blank) were required. In addition, all first year medical students at the Indiana University School of Medicine, regardless of campus, were required to take the National Board of Medical Examiners (NBME) Gross Anatomy "shelf" examination. This examination contains clinically-based multiple-choice questions and contributed a minimum of 20% to each student's final grade. All campuses agreed upon a pass cut-off of 70% for the course.

Study Strategies Survey

Study strategies surveys were administered to first year medical students at the Indiana University School of Medicine. The first page of the survey included a box for students to sign indicating their informed consent, and included a Family Educational Rights and Privacy Act (FERPA) release so that the author could obtain grades. The surveys included Likert scale questions (generally 5=always, 1=never), categorical response questions, and open-ended comment boxes (though these were sparsely used). The questions spanned three general topics:

1. Student study strategies (*e.g*. I plan to use the main course textbooks or textbook websites for studying by reviewing the figures.)
2. Class attendance and attitudes (*e.g*. I feel I have studied enough for the upcoming exam.)
3. Basic demographics.

The survey was designed using methods established by Fowler (1995). The survey was then discussed and evaluated with other anatomy educators, piloted with a small group, and then administered to a larger group of students for greater validity. Cronbach's alpha was previously calculated at .767, indicating a good reliability as well. For additional information on survey design and sample survey questions, see Husmann, *et al.* (2016).

Survey Administration

Participation in this study was completely voluntary and no incentives were offered. The survey was given to all first year medical students at eight of the nine Indiana University School of Medicine campuses (though only six were ultimately used for greater consistency in class format) in the first two weeks and again in the final two weeks of their first year gross anatomy course. Students who did not have the author as an instructor received a link to the survey via e-mail and three e-mail reminders to complete the survey. The survey was administered via Qualtrics (Qualtrics, Provo, UT) and accessed behind a Central Authentication Service (CAS) login to verify the identity of the person completing the survey. Students that had the author as an instructor received a paper copy of the survey from a staff member immediately following a class period and were asked to complete the survey and return it to the staff member who kept the surveys until the course grades were finalized. After all of the courses were completed, assessment scores were obtained from the instructors, including scores for the laboratory examinations, lecture examinations, NBME raw examination scores, and overall score (percentage) for the course. This was completed per Institutional Review Board protocol #1507250684A001.

The paper surveys were then returned to the author following submission of grades and the responses were manually added to the Excel file that was downloaded from the Qualtrics system, which included the responses from each of the other campuses. The only differences between the pre-course survey and the post-course survey were:

1. Future versus past tense language. (i.e. "I plan to look over the figures in the textbook" versus "I looked over the figures in the textbook")
2. One additional Likert scale item was added to the post-course survey that stated: "I feel that I have studied enough for the upcoming exam."

Following administration of the survey, study strategy questions were then condensed into seven categories to help minimize the number of statistical tests to be run on the data. These categories were:

1. Text-based Resources: These questions focused on the use of textbooks for the class, including diagrams, tables, or full-text use.
2. Lab-based Resources: These questions focused on the use of resources in or related to the laboratory component of the course, including dissectors and atlases.
3. Making Study Resources: These questions asked if students personally made any resources with which to study, such as tables, drawings, or flashcards.
4. Web-based Resources: These questions focused on how students used the internet to assist in studying, including both the website for the course and other sites that students found on their own or at instructor recommendation.
5. Studying with Others: These questions focused on how often students studied with one or more of their classmates.
6. Self-Quizzing: These questions focused on how often students participated in self-quizzing specific behaviors, such as using review questions from the text, old examinations, or flashcards.
7. Attendance: These questions focused on attendance for both lecture and laboratory components of the course.

continued on next page

Questions dealing with the number of hours of studying in the week preceding an examination and grade expected in the course were also included in the survey, but were kept separate from these larger categories. A copy of the survey is available on the HAPS website, Here.

Cohort set-up

Students were separated into three cohorts based on the size and length of their gross anatomy course.

Cohort #1 was compromised of students who all participated together in a one-semester (approximately fifteen weeks) gross anatomy course with 158 first year medical students per class.

Cohort #2 included students that all participated together in a two semester (a full academic year, approximately thirty weeks) gross anatomy course with approximately thirty-six first year medical students per class.

Cohort #3 included students that participated in a gross anatomy course lasting between ten and sixteen weeks (one semester) in a class with twenty-three to thirty-two first year medical students at one of four other campuses with similar primarily lecture curricula.

Academic standards for incoming students in all three cohorts should be roughly equivalent having all come through the same admissions procedures for the school. Additional demographic information will be presented for these cohorts below.

Analysis

Initial analyses (Brown-Forsythe 10.51, $p<0.000$) indicated that equal variance was not present in all of the variables being considered and that skewed distributions were present. Thus, non-parametric analyses were used. First, grade data were compared within Cohort #3 to see if grades were significantly different among the four campuses that were included in that cohort. NBME examination data was also compared among all three cohorts to test for significant differences in performance on this single common examination. For each of these analyses of variance tests, both Welch's F and Kruskal-Wallis tests were used as Welch's F has been shown to fair better when standard deviations are variable and the Kruskal-Wallis analysis has been shown to fair better when a non-normal distribution is present (McDonald 2014).

Welch's F and Kruskal-Wallis comparison of means assessments were also completed among all three cohorts for each of the study strategy categories in both pre-course surveys and post-course surveys. These tests were also completed to assess how much change had occurred within each study strategy category. Finally, Wilcoxon paired t-tests were run between pre- and post-course surveys within each cohort to assess which study habits changed throughout the course. All statistics were run using SPSS 24.0 (IBM, Armonk, NY).

Results

One hundred and four students (36.2%) completed the pre-course survey while eighty-six students (30.0%) completed the post-course survey. Within Cohort #3 (n=41), no significant difference in overall grades (Pre: $p = 0.611$, Post: $p = 0.854$) or in final NBME scores were found among the four campuses that were included in this cohort. Demographic comparisons among the three cohorts (Table 2) show no significant differences among the cohorts. However, there were significant differences in final grades ($p = 0.001$) among the three cohorts for those individuals that completed the post-course survey.

Cohort #1 had the highest final grades with a mean of 86.7% (n = 51), followed by Cohort #2 with a mean of 83.8% (n=34), and finally Cohort #3 with a mean of 80.8% (n=41). However, when comparing the NBME examination, the only consistent examination across all three cohorts, there were no significant differences among the three cohorts ($p=0.084$). When comparing each Cohort sample represented in this study with the larger class at their specific campus, there were no statistically significant differences for Cohorts #2 ($p=0.859$) or #3 (p values for each campus range from 0.474 to 0.997). The sample from Cohort #1 did show higher NBME scores than their larger class ($p=0.007$), indicating that our sample of Cohort #1 did better on this particular examination than was normal for their larger class.

Results for each analysis are shown in the tables, while results below are discussed by cohort. Table 3 shows the differences among cohorts on the pre-course survey while Table 4 shows differences among cohorts on the post-course survey. Table 5 illustrates statistically significant differences between the pre-course and post-course surveys within each cohort. It is interesting to note that, with no exceptions, all of the significant differences represent a decrease between the study techniques students initially thought they would use (pre-course survey) and the study techniques students reported they actually used (post-course survey). Finally, Table 6 shows differences among cohorts in the average amount of change that occurred for each category between the pre-course survey and the post-course survey.

Cohort #1 (large class, one semester)

On the pre-course survey, Cohort #1 planned to attend class the least, though they still planned to attend most classes, and also had the lowest grade expectations. This lower attendance pattern was also seen on the post-course survey for this cohort. Significant differences that occurred between the pre-course and post-course surveys for this cohort include reporting lower incidences of studying with others and fewer self-quizzing behaviors. These students also had the least amount of change in the use of text-based resources. The combination of less self-quizzing and studying with others, but more text use, may be an indication of more passive studying techniques.

continued on next page

Table 2. *Demographic comparisons among the three cohorts*

	Cohort #1 (large class, 1 semester)	Cohort #2 (small class, 2 semester)	Cohort #3 (small class, 1 semester or less)	Chi-square	Statistical significance (p value)
Gender				.609	.737
Male	37	21	27		
Female	16	13	14		
Race				6.578	.583
American Indian	0	0	0		
Asian-American	10	5	3		
African-American	3	1	2		
Hispanic	2	1	1		
White	32	22	18		
Age				6.487	.593
Under 22	2	3	0		
22-23	25	20	19		
24-25	14	7	5		
26-27	4	1	1		
28 or older	2	1	2		
Previous Anatomy or A&P course				1.508	.470
Yes	21	18	19		
No	32	16	22		
Parent Education				2.878	.237
Advanced degree	37	23	22		
No advanced degree	16	11	19		
English first language				1.418	.492
Yes	40	30	24		
No	7	2	3		

continued on next page

Table 3. *Significant differences in pre-course surveys between cohorts*

Question	Cohort 1 mean (n=46)	Cohort 2 mean (n=31)	Cohort 3 mean (n=27)	ANOVA (P value)	Kruskal-Wallis P value	Welch's F (P value)
Text-based Resources	3.18	3.45	3.58	2.375 (.098)	.136	2.547 (.086)
Lab-based Resources	3.79	4.59	3.89	**10.826 (.000)**	**.000**	**10.892 (.000)**
Making Study Resources	3.82	3.99	3.94	.396 (.674)	.676	.390 (.679)
Studying With Others	3.66	4.00	3.28	**4.305 (.016)**	**.015**	**4.251 (.019)**
Web-based Resources	3.40	3.39	3.48	.155 (.857)	.939	.160 (.852)
Self-Quizzing	4.15	4.30	4.02	1.228 (.298)	.264	1.304 (.279)
Attendance	4.84	4.97	4.91	2.825 (.064)	**.008**	**4.541 (.015)**
How many hours…	4.77	4.34	4.63	**4.525 (.013)**	**.006**	**4.347 (.018)**
Grade Expected	1.87	2.28	2.30	**5.549 (.005)**	**.016**	**5.619 (.006)**

(Green highlighting = significantly more than average, red highlighting = significantly less than average)
(All means are based on a five-point Likert scale in which generally 5 = always or almost always and 1 = never or rarely.)

Table 4*. Significant differences in post-course surveys between cohorts*

Question	Cohort 1 mean (n=27)	Cohort 2 mean (n=32)	Cohort 3 mean (n=27)	ANOVA (P value)	Kruskal-Wallis P value	Welch's F (P value)
Text-based Resources	2.97	1.80	3.08	**14.691 (.000)**	**.000**	**18.369 (.000)**
Lab-based Resources	3.15	2.94	3.40	1.647 (.199)	.249	1.738 (.185)
Making Resources	3.25	3.38	2.95	1.474 (.235	.152	1.496 (.233)
Working with Others	2.90	3.12	2.74	.902 (.410)	.516	.850 (.433)
Web-based Resources	2.80	2.98	2.65	1.444 (.242)	.185	1.510 (.230)
Self-Quizzing	3.31	3.38	2.30	**20.444 (.000)**	**.000**	**25.764 (.000)**
Attendance	4.55	4.97	4.96	**18.214 (.000)**	**.000**	**9.416 (.000)**
How many hours…	4.59	4.25	4.77	**4.316 (.016)**	**.010**	**4.703 (.013)**
Studied enough	2.10	3.12	2.48	**7.623 (.001)**	**.002**	**7.601 (.001)**

(Green highlighting = significantly more than average, red highlighting = significantly less than average)
(All mean values are based on a five-point Likert scale in which generally 5 = always or almost always and 1 = never or rarely.)

continued on next page

Table 5. *Questions with significant differences between pre-course and post-course surveys for each cohort (Wilcoxon tests)*

	Cohort 1 (N = 23)	Cohort 2 (N=32)	Cohort 3 (N=13)
Text-based Resources	.219	**.000**	.010
Lab-based Resources	.004	**.000**	.371
Making Resources	.009	.026	.004
Studying with Others	**.001**	**.000**	.088
Web-based Resources	.025	.018	.052
Self-Quizzing	**.001**	**.000**	**.001**
Attendance	.032	.564	.157

(Correction for multiple tests results in a required p value of .0028 or less to achieve statistical significance.)

Table 6. *Significant differences in the amount of change between pre- and post-surveys found in each cohort*

Question	Cohort 1 mean (n=23)	Cohort 2 mean (n=32)	Cohort 3 mean (n=13)	ANOVA (P value)	Kruskal-Wallis P value	Welch's F (P value)
Text-based Resources	.39	1.69	.79	**9.373 (.000)**	**.001**	**8.901 (.001)**
Lab-based Resources	.63	1.75	.27	**9.959 (.000)**	**.000**	**9.233 (.001)**
Making Resources	.52	.54	1.02	1.102 (.338)	.464	1.134 (.334)
Studying with Others	.85	.93	.61	.399 (.672)	.507	.295 (.747)
Web-based Resources	.35	.44	.49	.135 (.874)	.749	.165 (.848)
Self-Quizzing	.83	.92	1.87	**7.763 (.001)**	**.002**	**8.801 (.001)**
How many hours…	.22	.14	.08	.121 (.886)	.804	.151 (.861)
Attendance	.26	.02	.08	**5.782 (.005)**	**.016**	**3.612 (.040)**

(Green highlighting = significantly more than average change, red highlighting = significantly less than average change)
(All mean values are reported as the absolute value for the amount of change that occurred.)

continued on next page

Cohort #2 (small class, two semester)
On the pre-course survey, this cohort planned to use lab-based resources and studying with others the most. They also responded with the lowest numbers for the number of hours they were planning to spend studying in the week leading up to an exam. In the post-course analysis, these students had the lowest numbers for the use of text-based resources and number of study hours leading up to the exam, but the highest numbers for confidence going in to the exam. In the pre-course and post-course survey comparison, this group saw decreasing numbers for text-based resources, lab-based resources, studying with others, and self-quizzing. These drops were significantly larger than the other cohorts for text-based resources and lab-based resources. This cohort also saw the least change in attendance, though this result was not statistically significant from Cohort #3.

Cohort #3 (small class, one semester or less)
On the pre-course survey, Cohort #3 had the lowest number of plans for studying with others while the post-course survey showed the lowest numbers for self-quizzing and the highest numbers for use of text-based resources. The pre-course and post-course survey comparisons showed a statistically significant decrease in self-quizzing practices with this cohort, demonstrating the greatest drop in these numbers among all three cohorts.

Discussion

In general, these results show that medical students are initially very ambitious with their study plans. Unfortunately, they are unable to complete all of the plans they initially make and are forced to decrease their use of some techniques. This may well be due to lack of time combined with an overabundance of information from the proverbial "fire hose" and is supported by previous literature showing an increase in the surface approach to learning across the first year of medical school (Martenson 1986, Tooth *et al.* 1989, Papinczak *et al.* 2008, Ward 2011). However, exactly which resources are abandoned varies within the three contexts evaluated here.

Class Size Comparison
Students in the large class (Cohort #1) show different expectations and attitudes towards the class from the beginning as has been previously documented in the primary school literature (Glass *et al.* 1982). The students in the large class do not plan to attend class as regularly as students in the other groups (Cohorts #2 and #3), though attendance was not recorded or explicitly required at any of the campuses. The lower attendance reported by the larger class size may be related to the anonymity that is granted with larger class sizes and the effects of larger class size on student interest and motivation that have previously been noted at primary education level (Glass *et al.* 1982).

Students in the larger classes do not have grade expectations as high as students in the smaller classes. While there was ultimately no difference in the NBME scores and the difference in final grades was in favor of these students, the difference in grade expectations may be related to the affective differences that have previously been reported for larger class sizes (Glass *et al.* 1982). Glass *et al*. (1982) documented that there is a relationship between class size and affective components of the classroom such as morale, student satisfaction, and teacher satisfaction. Thus, while lower attendance rates and initial grade expectations may not be of immediate concern, they may have repercussions in future inter-professional relationships and/or in the parts of the curriculum that are not assessed by examinations, such as the implicit or hidden curriculum, which consists of the social, and cultural messages that are often communicated in schools. Furthermore, these affective components may lower student expectations and/or motivation for the course. Lower motivation may then decrease the effectiveness of student studying and thus have an indirect effect on the students' future outcomes.

Course Length Comparison
There were a number of interesting trends in the comparison of the longer (approximately thirty weeks) course and the shorter (10-16 weeks) courses. Firstly, the students in the longer, two semester course (Cohort #2) do not use the text-based resources as much as other students and experienced the largest drop in use of lab-based resources. One reason for this may be accessibility of the faculty and time in which to develop of a rapport with them. In the longer course and the smaller class size, students may have developed a closer relationship with their instructor and were more comfortable using resources provided by the instructor or using the instructor as a resource rather than looking up information in a textbook or atlas. This idea is supported by anecdotal evidence, though this is merely one hypothesis.

Another interesting trend with the students in the longer course concerns the lower scores for hours spent studying in the week preceding the examination. These lower scores were seen both in the pre-course survey planning and in the post-course survey. There are a few potential explanations for this. One explanation might be that there was a greater amount of time between examinations because the course was longer and thus more time to spread the studying out. If this explanation is true, this would further suggest that long-term (i.e., multi-year) retention for this material may be better with this cohort due to the use of spaced retrieval. Karpicke and his colleagues have completed multiple studies that show the importance of repeated retrieval of information. In particular, they have documented that spaced retrieval is more beneficial for long-term retention than repeatedly testing on material in short succession (Karpicke and Roediger 2007, Karpicke and Bauernschmidt 2011).

continued on next page

In the longer course, it is also likely that there was less material on each examination since there were more unit examinations in this group than in the other courses (six unit examinations across thirty weeks versus three or four unit examinations across ten to sixteen weeks). This trend also likely relates to the fact that students in Cohort #2 had more courses going on at the same time, which means there were more courses other than anatomy that required tending to.

Yet, despite reports of studying less leading up to the examination, the students in the longer course also reported the highest scores for confidence going into the examination. These trends may also relate back to having more time between examinations and/or less material per examination. However, the second highest scores for confidence going into examinations are seen in Cohort #3. The students in Cohort #3 would have had the same amount of time with the materials as Cohort #1, or potentially even less since some of their courses were a bit shorter. Thus, confidence going into the examination may also relate to the rapport between faculty and students, which may be more common in smaller classes (Glass *et al.* 1982). This theory would explain why the highest examination confidence scores were seen in Cohort #2 (smaller class size, more time), with the second highest scores seen in Cohort #3 (smaller class size, but less time), and the lowest examination confidence scores in Cohort #1 (large class size, at least as much time as Cohort #3).

Yet another possible factor in the lower confidence scores for Cohort #1 may relate to their greater use of passive study techniques, which are less likely to help them gauge their progress than studying with others or self-quizzing. As mentioned in the previous section, there were ultimately no significant differences between cohorts with the NBME standardized examinations. Yet, it should be noted that previous studies have found that differences in scores between shorter and longer courses may be lost with multiple choice tests such as the NBME (Rayburn and Rayburn 1999). Thus, while it may be difficult to demonstrate retention differences with these examinations, differences in long-term (particularly multi-semester or multi-year) retention and understanding may still exist.

Limitations

The first limitation that must be acknowledged is that both pre- and post-course survey data are all student reported. The accuracy of the data may be questionable. This may be especially true for Cohort #2 students who were taking courses with the faculty member conducting the study. However, every effort was made to assure students that the faculty member would have no knowledge of participation rates or responses until after the course was completed. The faculty member was not in the room when the surveys were administered.

It must also be acknowledged that the sample sizes for each cohort are not ideal. The study could be improved with another ten to twenty students (at least) in each cohort. Unfortunately, this was not possible since it was a voluntary study. The curriculum has since changed to an integrated form, which does not allow for future cohorts of medical students to be added.

Finally, the largest limitation is the number of additional confounding variables that must be acknowledged when looking across six different campuses of a medical school. These include, but are not limited to, availability and quality of some academic resources (*e.g*. old exams, power points), and the variability in instructors who may differ in approachability, pedagogical beliefs, organization and materials, among other factors. That said, these confounding variables are somewhat more controlled by looking at multiple campuses across a single medical school than they would be evaluating multiple medical schools that may have competing policies, different course objectives, and/or different admissions criteria and policies.

Conclusions

In general, medical students plan to study all of the resources that are available to them, but time constraints generally do not allow this. Something has to give. In larger classes, students are less likely to attend class and generally have lower grade expectations from the beginning (though not lower grades at the end of the course). Longer courses may provide opportunities for increased rapport with faculty and more spaced practice of the material (less "cramming") since there is more time and potentially less material for each examination. So these students may give up textbooks and atlases in favor of the resources provided directly by their instructors. In this study, these benefits were seen in the form of higher confidence going into the examination, despite reporting less time spent studying in the week preceding the examination. The higher confidence scores and more spaced studying may lead to better long-term retention of the material (Roediger and Karpicke 2006, Karpicke and Bauernschmidt 2011), though further research is necessary to confirm this theory. If true, this could have substantial implications for medical education, particularly at schools that have blocked curricula, which may only allow a week or two between examinations and sometimes involve entire courses that are only four to six weeks in length.

continued on next page

Acknowledgements

I would like to thank all of the students who participated in this study. I would also like to thank Valerie O'Loughlin and Jim Brokaw for allowing me to bounce around some ideas for this research and for their guidance in tracking down some of the data. Finally, I would like to thank Jackie Cullison for administering and holding the surveys for the Bloomington campus.

About the Author

POLLY R. HUSMANN, Ph.D. is an assistant professor of anatomy and cell biology at the Bloomington campus of the Indiana University School of Medicine. She teaches anatomy to medical, graduate, and undergraduate students and conducts educational research on factors outside of the classroom that affect student learning.

Literature Cited

Association of American Medical Colleges (2015a) Medical School Applicants, Enrollees Reach New High. In: Bergen B (Editor). https://www.aamc.org/newsroom/newsreleases/446400/applicant-and-enrollment-data.html#.Wig-SmX16Yc.

Association of American Medical Colleges (2015b) Medical School Enrollment to Approach 30 Percent Increase by 2019. In: Bergen B (Editor). https://www.aamc.org/newsroom/newsreleases/431036/20150430.html#.Wig-NSPLjRl.

Austin AM, Gustafson L (2006) Impact of Course Length on Student Learning. *Journal of Economics and Finance Education* 5: 26-37.

Brady KL, Eisler RM (1999) Sex and Gender in the College Classroom: A Quantitative Analysis of Faculty-Student Interactions and Perceptions. *J Educ Psychol* 91: 27-145.

Butler AC, Karpicke JD, Roediger III HL (2007) The Effect of Type and Timing of Feedback on Learning From Multiple-Choice Tests. *J Exp Psychol* Appl 13: 273-281.

Crombag HF, Gaff JG, Chang TM (1973) Student Characteristics and Academic Performance in a Medical School: Differences that do not make a Difference. *Br J Med Educ* 7: 146-151.

Cuddy MM, Swanson DB, Drake RL, Pawlina W (2013) Changes in Anatomy Instruction and USMLE Performance: Empirical Evidence on the Absence of a Relationship. *Anatomical Sciences Education* 6: 3-10.

Dobson JL, Perez J, Linderholm T (2017) Distributed Retrieval Practice Promotes Superior Recall of Anatomy Information. *Anatomical Sciences Education* 10: 339-347.

Drake RL, Lowrie JDJ, Prewitt CM (2002) Survey of Gross Anatomy , Microscopic Anatomy, Neuroscience, and Embryology Courses in Medical School Curricula in the United States. *The Anatomical Record (New Anat)* 269: 118-122.

Drake RL, McBride JM, Lachman N, Pawlina W (2009) Medical Education in the Anatomical Sciences: The Winds of Change Continue to Blow. *Anatomical Sciences Education* 2: 253-259.

Dunn-Lewis C, Finn K, FitzPatrick K (2016) Student Expected Achievement in Anatomy and Physiology Associated with Use and Reported Helpfulness of Learning and Studying Strategies. *HAPS-Educator* 20: 27-37.

Entwistle GR, ED (1960) Study-skills Courses in Medical Schools? J Med Educ 35: 843 - 848.

Erikson C, Whatley M, Tilton C (2014) Results of the 2013 Medical School Enrollment Survey. Association of American Medical Colleges.

Fleming T, Toutant T, Raptis H (2002) Class Size and Effects: A Review. Bloomington, Indiana: *Phi Delta Kappa Educational Foundation*. 34 p.

Fowler FJJ (1995) Improving Survey Questions: Design and Evaluation. 1st Ed. Edition. Thousand Oaks, CA: Sage Publications. 193 p.

Glass GV, Cahen LS, Smith ML, Filby NN (1982) School Class Size: Research and Policy. Beverly Hills: Sage Publications. 160 p.

Greene RL (1989) Spacing Effects in Memory: Evidence for a Two-Process Account. J *Exp Psychol Learn Mem Cogn* 15: 371-377.

Holla SJ, Ramachandran K, Isaac B, Koshy S (2009) Anatomy Education in a Changing Medical Curriculum in India: Medical Student Feedback on Duration and Emphasis of Gross Anatomy Teaching. *Anatomical Sciences Education* 2:179-183.

Husmann PR, Barger JB, Schutte AF (2016) Study Skills in Anatomy and Physiology: Is There a Difference. *Anatomical Sciences Education* 9:18-27.

Karpicke JD, Bauernschmidt A (2011) Spaced Retrieval: Absolute Spacing Enhances Learning regardless of Relative Spacing. *J Exp Psychol Learn Mem Cogn* 37: 1250 - 1257.

Karpicke JD, Roediger III HL (2007) Expanding Retrieval Practice Promotes Short-Term Retention, but Equally Spaced Retrieval Enhances Long-Term Retention. *J Exp Psychol* 33: 704-719.

Mahler S, Neumann L (1986) The Class-size Effect upon Activity and Cognitive Dimensions of Lesson in Higher Education. *Assessment and Evaluation in Higher Education* 11: 43-59.

Malleson N (1967) Medical Students' Study: Time and Place. *Br J Med Educ* 1: 169-177.

Malleson N, Penfold DM, Sawiris MY (1968) Medical Students' Study: The Way They Work. *Br J Med Educ* 2:11-19.

Martenson DF (1986) Students' Approaches to Studying in Four Medical Schools. *Med Educ* 20: 532-534.

continued on next page

McDonald JH (2014) Handbook of Biological Statistics. 3rd Edition. Baltimore, Maryland: Sparky House Publishing.

Papinczak T, Young I, Groves M, Haynes M (2008) Efffects of a Metacognitive Intervention on Student's Approaches to Learning and Self-efficacy in a First Year Medical Course. *Advances in Health Science Education* 13: 213-232.

Pashler H, Rohrer D, Cepeda NJ, Carpenter SK (2007) Enhancing learning and retarding forgetting: Choices and consequences. *Psychonomic Bulletin & Review* 14: 187-193.

Pizzimenti MA, Axelson RD (2015) Assessing Student Engagement and Self-Regulated Learning in a Medical Gross Anatomy Course. *Anatomical Sciences Education* 8:104 - 110.

Prince M (2004) Does Active Learning Work? A Review of the Research. *Journal of Engineering Education* 93: 223-231.

Rayburn LG, Rayburn JM (1999) Impact of Course Length and Homework Assignments on Student Performance. *Journal of Education for Business* 74: 325-331.

Roediger III HL, Karpicke JD (2006) Test-Enhanced Learning: Taking Memory Tests Improves Long-Term Retention. *Psychol Sci* 17: 249-255.

Rohrer D, Pashler H (2007) Increasing Retention Without Increasing Study Time. *Curr Dir Psychol Sci* 16: 183-186.

Rohrer D, Pashler H (2010) Recent Research on Human Learning Challenges Conventional Instructional Strategies. *Educational Researcher* 39: 406-412.

Sanazaro PJ (1966) Class Size in Medical School. *The Journal of Medical Education* 41: 1017-1029.

Schieffler Jr. DA, Azevedo BM, Culbertson RA, Kahn MJ (2012) Financial Implications of Increasing Medical School Class Size: Does Tuition Cover Cost? *The Permanente Journal* 16:10-14.

Selvig D, Holaday LW, Purkiss J, Hortsch M (2015) Correlating Students' Educational Background, Study Habits, and Resource Usage with Learning Success in Medical Histology. *Anatomical Sciences Education* 8: 1-11.

Shatin L (1967) Study Skills in Medical Education: A Report and Analysis. *J Med Educ* 42: 833-840.

Soderstrom NC, Kerr TK, Bjork RA (2016) The Critical Importance of Retrieval - and Spacing - for Learning. *Psychol Sci* 27: 223-230.

Tan CM, Thanaraj K (1993) Influence of Context and Preferred Learning Environments: Approaches to Studying Physiology. *Med Educ* 27: 143-159.

Tooth D, Tonge K, McManus IC (1989) Anxiety and Study methods in Preclinical Students: Causal Relation to Examination Performance. *Med Educ* 23: 416-421.

Van Scyoc LJ, Gleason J (1993) Traditional or Intensive Course Lengths? A Comparison of Outcomes in Economics Learning. *Journal of Economic Education* 24: 15-22.

Verkoeijen PPJL, Rikers RMJP, GSH (2004) Detrimental Influence of Contextual Change on Spacing Effects in Free Recall. *J Exp Psychol Learn Mem Cogn* 30: 796-800.

Ward PJ (2011) First Year Medical Students' Approaches to Study and their Outcomes in a Gross Anatomy Course. *Clin Anat* 24: 120-127.

Ward PJ, Walker JJ (2008) The Influences of Study methods and Knowledge Processing on Academic Success and Long-term Recall of Anatomy Learning by First-year Veterinary Students. *Anatomical Sciences Education* 1: 68-74.

Whiffen JW, Karpicke JD (2017) The Role of Episodic Context in Retrieval Practice Effects. *J Exp Psychol Learn Mem Cogn* 43: 1036-1046.

■

Back to TOC

The Impact of Physiology Prerequisites on Future Anatomy and Physiology Courses

Justin F. Shaffer[1]*, Samuel E. Schriner[2], Catherine Loudon[3], Samantha J. Dacanay[4], Usman Alam[4], Jennifer V. Dang[4], Nancy Aguilar-Roca[3], Pavan Kadandale[4], Brian K. Sato[4]*
[1]Department of Chemical and Biological Engineering, Colorado School of Mines, Golden, CO 80401
[2]Department of Pharmaceutical Sciences, University of California Irvine
[3]Department of Ecology and Evolution, University of California Irvine
[4]Department of Molecular Biology and Biochemistry, 2238 McGaugh Hall MC3900, University of California Irvine, CA 92697
* Corresponding authors: jshaffer@mines.edu, bsato@uci.edu

Abstract
Universities use prerequisites to regulate the path of students through a program or major. However, the impact of prerequisites on follow-up courses in anatomy and physiology is not well understood. If success in follow-up courses depends on completion of prerequisite courses, then students should earn higher grades on exam questions that assess prerequisite knowledge. To test this hypothesis, we investigated the potential impact of a required prerequisite human physiology lecture course on a follow-up human physiology laboratory course and a follow-up molecular pharmacology course. We also investigated the potential impact of a recommended prerequisite human physiology laboratory course on a follow-up human anatomy course. We assessed student exam performance in the follow-up courses based on their familiarity with the material from the required prerequisite course or the recommended prerequisite course. Our results were mixed and demonstrated limited performance gains in the follow-up courses despite overlap of material among the courses. These results suggest that prerequisite courses may not have a significant impact on the outcome of future related courses and that individual academic programs should evaluate the effectiveness of prerequisite courses in a local context. http://doi.org/10.21692/haps.2018.025

Key words: physiology, undergraduate, curriculum, familiarity, prerequisites

Introduction
Prerequisite courses are widespread in Science, Engineering, Technology, and Math (STEM) majors at colleges and universities and students must often successfully complete one or more prerequisite courses in order to advance to the next course in a series. Prerequisite courses may exist for a variety of reasons, including content-specific reasons (e.g. the next course in a series directly builds on the material or skills students learn in the prerequisite course), other academic reasons (e.g. students should not take a certain course until they have reached upper-division standing even if the course content does not directly build on the material from the prerequisite course) or logistical reasons (e.g. future course enrollment can be predicted based on enrollment in the prerequisite). Despite the specific reason for a particular prerequisite course, for thorough programmatic assessment it is important to evaluate whether prerequisite courses directly impact student success in follow-up courses.

In the biological sciences, human physiology courses are often paired with human anatomy courses either by requiring that one be taken before the other, or by combining human anatomy and physiology in an Anatomy and Physiology I and Anatomy and Physiology II series. Additionally, human physiology lecture courses may serve as prerequisites for more advanced physiology lecture courses or physiology laboratory courses. In theory, the linked prerequisite and follow-up courses should be designed in a manner to allow students to scaffold their knowledge, guided by the core set of physiology concepts established by Michael and McFarland (2011).

To a minimal degree, the impact of prerequisite courses and prior student knowledge on later success has been examined in the context of physiology education. While the number of past science courses taken and university grade point average (GPA) have been found to positively correlate with passing grades in physiology courses (McCleary *et al.* 1999), success is also impacted by faulty mental models or misconceptions, derived from personal experience with the subject matter, or related to prior classroom instruction (Michael 1998). For example, student misconceptions in cardiovascular physiology (Michael 1998) and respiratory physiology (Michael 1998, Michael *et al.* 1999) have been described, which may impact student success in subsequent physiology courses. In an attempt to address this issue, Modell *et al.* (2000) used various laboratory protocols in an attempt to "repair" faulty mental models that students hold about how tidal volume responds to changes in minute ventilation (Michael 1998, Michael *et al.* 1999). They found that students "corrected" their misconceptions to the greatest extent when following laboratory protocols that required them to not only predict their experimental results but to verbally explain them to their instructor before attempting the experiment.

continued on next page

Regardless of where misconceptions came from, we assume that students apply the knowledge gained from prior physiology courses to subsequent ones, building on their existing knowledge. However, in a specific instance assessing student understanding of the cardiovascular system, Richardson (2000) found no impact of a recommended basic-level physiology course on performance on cardiovascular items in a subsequent upper-division physiology course. Additionally, Rovick *et al*. (1999) found that physiology instructors were poor judges of students' incoming knowledge (whether they had taken a physiology prerequisite course or not) and specifically overestimated students' ability to apply whatever prior knowledge they had. Taken together, these studies suggest that prerequisites course and misconceptions may have negative effects on future performance in physiology courses and that more work is needed to understand how these factors may influence student success.

Prerequisites can be either required or recommended for student enrollment in a subsequence course. If the prerequisite course is suggested but not required, comparing the performance of students who have or have not completed the prerequisite may assess its impact on the follow-up course. Indeed, several studies have explored this type of analysis. In a study of first-year medical students, Forester *et al.* (2002) found that students who had completed a gross anatomy course during their undergraduate program earned significantly higher grades in a gross anatomy course at the medical school level. McRae *et al.* (2010) found a similar result in that student grades in an undergraduate organic chemistry course were significantly correlated with grades earned in a chiropractic biochemistry course. While these studies showed positive connections between prerequisites and follow-up courses this may be due to the fact that these studies examined graduate students, a highly motivated population.

Other studies have reported a lack of correlation. Wright *et al.* (2009) examined performance in a biochemistry course and found that students performed equally well whether they had completed an organic chemistry prerequisite or not. Canaday and Lancaster (1985) found that there was no impact of undergraduate courses in biochemistry, vertebrate anatomy, histology, or embryology on the GPA of first-year medical students, courses that would presumably provide students with an advantage in medical school. Lastly, Steele and Barnhill (1982) found that an undergraduate genetics course had no impact on student performance in a medical school genetics course. Given the variability in these results, it is possible that prerequisites may (or may not) be important for discipline-specific reasons or for specific student populations.

On the other hand, if the prerequisite course is required before enrolling in the follow-up course, an alternative assessment approach must be taken due to the lack of a control group that has not completed the prerequisite. One method to examine the impact of prerequisites in this scenario is by assessing student performance on exam questions in the follow-up course that are based on the level of familiarity one has with the tested content from the prerequisite course (Shaffer *et al.* 2016, Sato *et al.* 2017). In this analysis, exam questions are coded (by instructors, students, or lecture slide analysis) as "very familiar" if the topic was previously covered extensively in the prerequisite, "familiar" if the topic was briefly introduced in the prerequisite course, and "not familiar" if the topic was not taught at all in the prerequisite course. Shaffer *et al.* (2016) examined the impact of a genetics prerequisite on performance in a molecular biology course and a human physiology prerequisite on performance in a human anatomy course. This study (Shaffer *et al.* 2016) demonstrated limited differences on performance among questions of varying familiarity, suggesting that students did not necessarily utilize or need their prerequisite knowledge in follow-up courses. Sato *et al.* (2017) found similar results using the familiarity analysis for a microbiology lecture course that served as a recommended prerequisite for a microbiology laboratory course.

As human physiology courses are often tightly coupled with advanced physiology lecture courses, physiology laboratory courses, or human anatomy courses, we sought to determine the impact of physiology courses on related subsequent courses. Since prior studies have shown little or no impact of prerequisites on student success on specific topics taught in physiology courses (Richardson 2000) it is important to know if this is also the case on a course wide basis. Specifically, this study addresses the following research questions in the context of a large research-intensive university:

1. Does content familiarity from a required undergraduate prerequisite human physiology course impact exam performance in a follow-up physiology laboratory course or a molecular pharmacology course?
2. Does enrollment in a recommended undergraduate physiology laboratory course predict higher exam scores in a follow-up human anatomy lecture course?

Materials and Methods

Study Context

This study was conducted at a large, PhD-granting research university in the western United States. Data were collected from two sections (Winter 2016 (n=218 students) and Spring 2016 (n=150)) of an upper division Biological Sciences physiology laboratory course (Bio PL, taught by author NAR), one section (Winter 2016, n = 139) of an upper division molecular pharmacology (MP, taught by author SES) course, and three sections (Spring 2015, Winter 2016, and Spring 2016, total n = 377 students) of an upper division human anatomy course (HA, taught by author JFS). These courses were associated with prerequisites that were examined in this study, specifically with the particular pathways illustrated in Figure 1 and course descriptive information in Table 1.

continued on next page

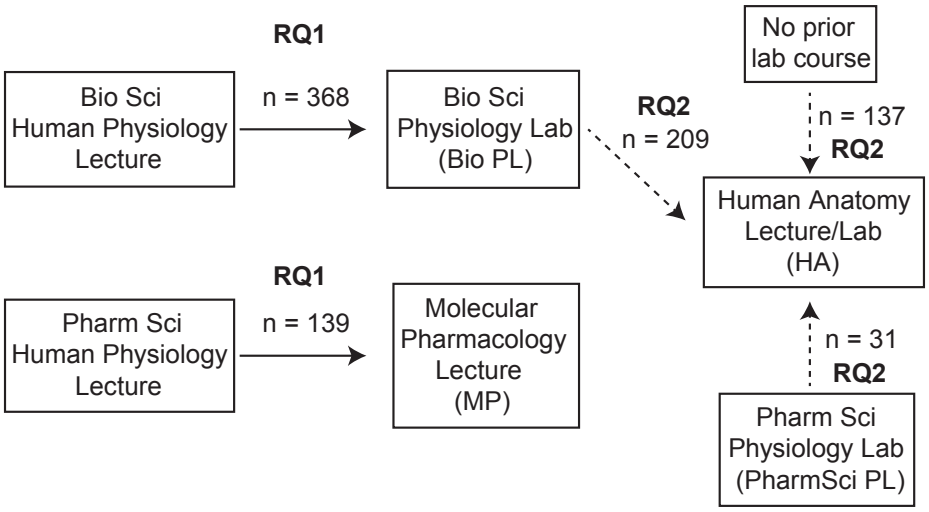

Figure 1. *Sequence of study courses. Performance in Biological Sciences Physiology Lab (Bio PL), Molecular Pharmacology (MP), and Human Anatomy (HA) courses were examined in the context of their respective prerequisites. An additional course, a pharmaceutical sciences physiology lab (PharmSci PL) served as a prerequisite for HA. Solid arrows indicate that the course to the left is a required prerequisite while a dotted arrow indicates that the course to the left is a recommended prerequisite.*

Table 1. *Study Courses*

Course	Course Description	Course Assessments
Bio Sci Human Physiology Lecture	How a human body works, focusing interactions between the digestive, muscular, endocrine, respiratory, cardiovascular, efferent nervous, and reproductive systems.	Weekly online quizzes, weekly homework sets, in-class "clicker" activities, 2 midterms, final exam
Pharm Sci Human Physiology Lecture	Function of the human body covering all major organ systems	3 midterms, final exam
Bio Sci Physiology Lab	Laboratory with a focus on the whole organism and its organ systems. Examples of structure-function relationships are drawn from both animal and human physiology. This courses also fulfills an upper-division writing requirement.	Weekly short writing assignments, 1 major lab report, 3 midterms, final exam
Molecular Pharmacology Lecture	Basic pharmacodynamics and pharmacokinetics, and drug mechanisms for cardiovascular disease and diabetes	3 midterms, final exam
Human Anatomy Lecture & Lab	A broad overview of the anatomy of the organ systems of the human body focusing on structure-function relationships	Online pre-class assignments, weekly review quizzes, in-class "clicker" activities, 3 lecture midterms, lecture final exam, 2 laboratory practical exams

All of these courses are taken by third through fifth year students (primarily fourth year students), except for the Spring sections of HA in which approximately one-third of the enrollment is made up of second-year nursing science students. Fifth year students include transfer students or students taking extra time to graduate due to making up courses, double majoring, etc. The majority of students enrolled in the Bio PL and HA sections were Biological Sciences majors, while the majority of students in MP were Pharmaceutical Sciences majors. Enrollment in Bio PL and HA required successful completion (a grade of C or better) of an upper division physiology lecture course taught within the School of Biological Sciences. Additionally, students enrolled in HA may or may not have successfully completed Bio PL prior to enrolling in HA. Enrollment in MP required successful completion (a grade of C or better) of a distinct upper division physiology lecture course taught within the Department of Pharmaceutical Sciences.

Data Collected

Data from Bio PL and MP, in the form of performance on individual exam questions, were collected. This included 51 questions from the Winter Bio PL exams, 57 questions from the Spring Bio PL exams, and 60 questions from the MP exams. The authors evaluated each question in terms of the Core Concepts in Physiology (12) in order to determine what physiology concepts the exams assessed. The majority of questions in the Bio PL course addressed Core Concepts

continued on next page

7 (Structure/Function) and 8 (Scientific Reasoning), whereas the majority of questions in the MP course addressed Core Concepts 1 (Homeostasis) and 2 (Cell Membrane). The questions in the Bio PL course covered cardiovascular physiology, respiratory physiology, muscle contraction, exercise physiology, and metabolism, while the questions in the MP course covered cell signaling, metabolism, and cardiovascular physiology. Final course grades in HA were evaluated which included performance on four lecture exams and two laboratory practical exams.

The Institutional Review Board of the University of California, Irvine approved this study. (HS# 2015-1455 and HS# 2013-9959).

Familiarity Designation

Exam question familiarity based on the material presented in the prerequisite human physiology courses for Bio PL and MP was assigned as previously described (Shaffer *et al.* 2016, Sato *et al.* 2017). Briefly, the instructors for the Biological Sciences and Pharmaceutical Sciences human physiology courses examined the Bio PL and MP exam questions and rated each as very familiar (VF), familiar (F), or not familiar (NF) based on the material taught in the physiology courses. A "VF" question is one that could have been answered by a student who had enrolled in the human physiology prerequisite but had yet to be exposed to the Bio PL or MP course. An "F" question covered material taught in the prerequisite course, but not in enough detail to answer the question, and a "NF" question covered a topic not discussed in the prerequisite. In addition to instructor-designated familiarity, members of the research team (SJD, JVD, UA) characterized familiarity utilizing lecture slides from the prerequisite courses. Consensus was reached according to previously described methods (Welsh 2012, Heiner *et al*. 2014). Each team member characterized exam familiarity independently. This was followed by a group discussion to identify consensus. For areas where there was significant disagreement among team members, discussion continued until consensus was reached. A high level of agreement among the study team was reached, with two of the three authors agreeing 88% of the time for the winter Bio PL exam questions, 93% of the time for the spring Bio PL exam questions, 93% of the time for the spring Bio PL exam questions, and 98% of the time for MP exam question designation.

In addition, the Bloom's level of each question (Bloom *et al.* 1956) was determined by members of the research team (JS, NA, PK). Each individual rated the questions independently. Two of the three study authors agreed 81% of the time for the winter Bio PL exam questions, 83% of the time for the spring Bio PL exam questions, and 90% of the time for MP exam question designation. Question familiarity and Bloom's level designations can be seen in Table 2.

Table 2. *Exam question descriptive statistics*

Course	Familiarity	# of Questions based on Instructor	Average Bloom's level	# of Questions based on Lecture Slides	Average Bloom's level
Winter PL	VF	10	2.9	6	3.0
	F	13	2.5	17	2.8
	NF	28	2.6	28	2.5
Spring PL	VF	17	2.8	9	2.8
	F	17	2.9	15	3.0
	NF	23	2.7	33	2.7
MP	VF	11	1.2	6	3.0
	F	12	1.4	11	1.8
	NF	37	2.0	43	1.5

Exam questions from each course were categorized based on familiarity, according to either instructor of the prerequisite course or by an independent team who viewed the course lecture slides, as described in the methods. The number of questions in each familiarity bin are shown above, as are the average Bloom's level of the questions in each bin.

continued on next page

Data Analysis
The impact of the human physiology prerequisite courses on Bio PL and MP performance was determined as previously described (Shaffer *et al.* 2016, Sato *et al.* 2017). Briefly, Bio PL and MP exam questions were segregated in two different ways, one based on instructor-designated familiarity and one based on the lecture slide designated familiarity. A multiple regression analysis was then run in which performance on an exam question was examined using both familiarity and Bloom's level as variables. After running these models, it was determined that Bloom's level did not significantly impact performance. Thus, rather than utilize multiple regression analysis, we ran an ANOVA for each course to determine whether question familiarity impacted performance as well as individual t-tests with Tukey contrasts to determine whether performance differed in pairwise comparisons between familiarity groups (VF versus F, F versus NF, VF versus NF).

To highlight the similarity, or lack thereof, between exam question familiarity designations by instructor and lecture slide, we compared designations for each question. If the question designation was identical between the two methods, the level of agreement was characterized as "agree". It was labeled as "slightly agree" if the question designation varied by one level (i.e. very familiar by lecture slide but familiar by instructor). And it was labeled as "disagree" if the designation varied by two levels (i.e. very familiar by lecture slide but not familiar by instructor).

To determine the possible impact of a recommended human physiology laboratory course on student performance in HA, final grades in HA were compared for students who had completed a physiology laboratory course offered by the School of Biological Sciences (Bio PL; n = 209), a physiology laboratory course offered by the Department of Pharmaceutical Sciences (PharmSci PL; n = 31) or lack of either course (n = 137). Multiple linear regression models were developed using final grades (as a percentage out of 100) as the response variable and prior laboratory course completion (a three-level factor variable) and overall GPA as explanatory variables.

Results

Impact of content familiarity on PL and MP exam performance. We examined exam question performance in the Bio PL and MP courses in the context of how familiar the tested concepts were from their respective prerequisite physiology courses. If the material taught in the prerequisite course was important for future success, we predicted that student performance would be highest on Bio PL and MP VF questions, then F questions, and finally NF questions. When analyzing exam performance in the winter and spring Bio PL courses, we found variable results in regard to the impact of familiarity on performance. For the winter quarter exam data, instructor-designated familiarity had no impact on exam performance, while students actually performed slightly better on NF exam questions when familiarity was designated by lecture slides (Figure 2A, 2B). For the spring exam data, students earned significantly higher exam scores on VF questions designated by instructor, but there was no impact on familiarity as designated by lecture slides. Data from the MP course exams similarly demonstrated that concept familiarity does not have an obvious impact on performance, as there was no significant difference in exam score using either instructor or lecture slide to designate familiarity (Figure 3A, 3B).

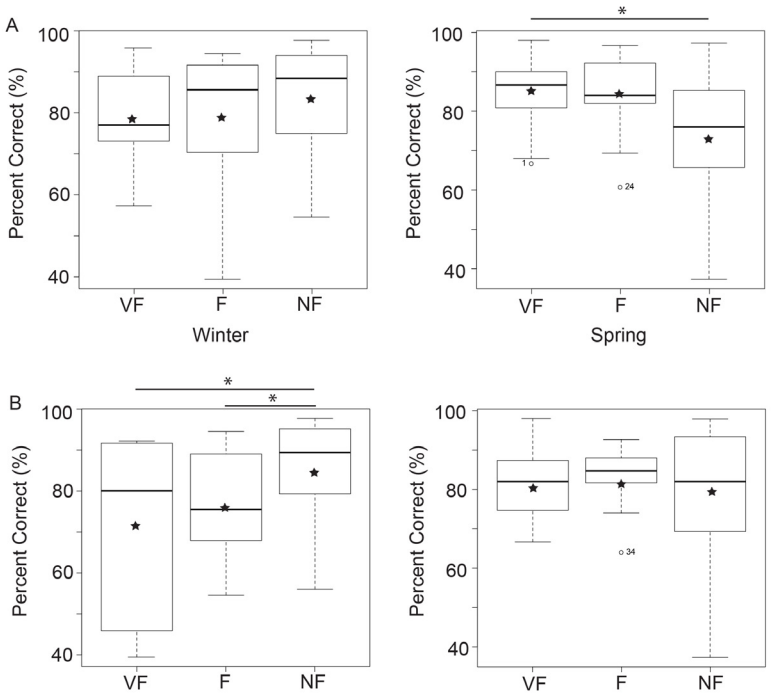

Figure 2. *Impact of prerequisite familiarity on Bio PL course exam performance. Exam questions from the winter and spring Bio PL courses were characterized based on familiarity as designated by either (A) the prerequisite course instructor or (B) the prerequisite course lecture slides. Box plots present the distribution of scores for questions of the indicated familiarity level. The stars indicate the mean score for that particular group of questions. VF = very familiar, F = familiar, NF = not familiar, * p = 0.04*

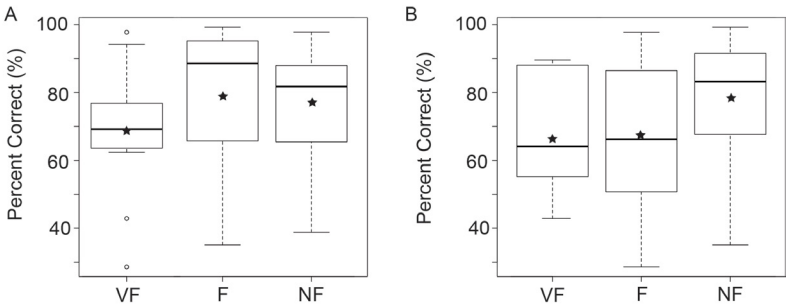

Figure 3. *Impact of prerequisite familiarity on MP course exam performance. Exam questions from the MP course were characterized based on familiarity according to either (A) the prerequisite course instructor or (B) the prerequisite course lecture slides. Box plots present the distribution of scores for questions of the indicated familiarity level. The stars indicate the mean score for that particular group of questions.*

continued on next page

We were curious as to whether the differences in prerequisite impact on the winter and spring quarters of Bio PL were due to different student demographics for those enrolled in the two quarters. We compared a variety of student characteristics by Chi-squared test and found that students did not differ by gender (p = 0.33), percentage of underrepresented minorities status (p = 0.22), low-income status (p = 0.48), first generation status (p = 0.87), and Math SAT score (p = 0.28 by ANOVA).

As the impact of concept familiarity on performance was variable depending on the method utilized to designate familiarity (lecture slides versus instructor), we wanted to examine how similar the familiarity designations were depending on which of these two methods was used. To accomplish this, we compared the familiarity characterization for each exam question, and found that there were identical designations for at least fifty percent of the questions from each course examined (Figure 4). Of the remaining questions, the vast majority differed only by one level of familiarity, highlighting that while each designation method produced a unique perspective in terms of what was taught in a prerequisite course, there was considerable overlap between these perspectives. This is a similar conclusion as was seen in a previously published study (Shaffer et al. 2016).

Impact of prior physiology laboratory on HA performance.

Another context in which we wanted to examine the impact of prerequisites was a human anatomy course with a recommended human physiology laboratory course prerequisite. Many students who enrolled in the HA course successfully completed a human physiology laboratory course, one offered by either the Biological Sciences or Pharmaceutical Sciences departments. Since the anatomy covered in the human physiology laboratory courses overlap with material covered in the HA course, we hypothesized that students who successfully completed a human physiology laboratory course would earn significantly higher grades in HA than students who did not take the prerequisite course. When controlling for student GPA, students who successfully completed the Biological Sciences human physiology course (Bio PL) earned modest, but significant, higher final grades in the anatomy course (model estimate of 1.58 +/- 2.02 out of 100 percent) than students who did not complete the prerequisite course (Table 3). This positive trend also existed for students enrolled in the other human physiology laboratory course (PharmSci PL) although the result was not statistically significant.

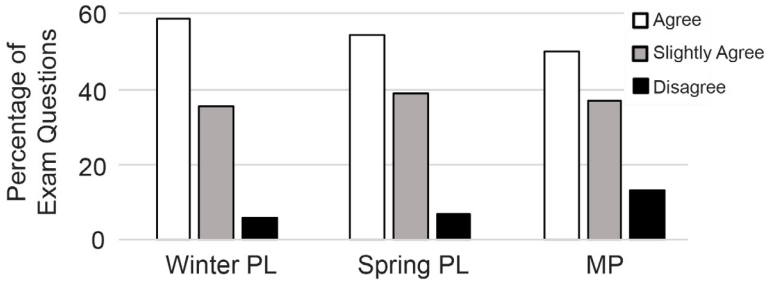

Figure 4. *Comparison of instructor and slide familiarity characterization. Familiarity characterization of exam questions from the indicated course were compared according to whether they were designated by the instructor or lecture slides. Agree refers to ratings of a particular question that were identical between the two designation methods. Slightly agree refers to ratings that were off by one familiarity level. Disagree refers to ratings that were off by two familiarity levels.*

Table 3. *Student performance in human anatomy (HA) based on prior physiology laboratory course completion*

	Final course grade	
Regression coefficient	Estimate ± SEM	*p* value
Model intercept	29.09 ± 3.56	5.0e-15
Lab (Bio PL)	1.58 ± 2.02	0.04
Lab (PharmSci PL)	0.78 ± 1.42	0.58
GPA	15.62 ± 1.05	< 2e-16
Adjusted R^2	0.3726	

The baseline for this model is students who did not successfully complete a human physiology laboratory course prior to enrolling in HA. Values for the estimates are provided as the mean +/- the standard error.

continued on next page

Discussion

In this study, we investigated the potential impacts of a required prerequisite human physiology lecture course on a follow-up human physiology laboratory course and a follow-up molecular pharmacology course using the previously established familiarity categorization method. We also investigated the potential impact of a recommended prerequisite human physiology laboratory course on a follow-up human anatomy course by comparing exam performance between students who did and did not successfully complete the recommended prerequisite. Overall, we found limited evidence that the prerequisite courses played an important role in student success in the follow-up courses, which is contrary to our hypothesis (Figures 2 and 3). These results are similar to those previously reported for a human anatomy course and a human physiology prerequisite (Shaffer et al. 2016), a molecular biology course and a genetics prerequisite (Shaffer et al. 2016), and a microbiology laboratory course and a microbiology lecture prerequisite (Sato et al. 2017)

A possible reason that we did not find significantly higher achievement on very familiar questions could be that students have difficulty applying knowledge in a different context. Even though the courses in this study were closely linked in subject matter, students often have difficulties applying knowledge acquired in one situation to another, even when similar concepts are involved (Bransford *et al.* 2000). Another possibility is that students simply are not retaining what they learned in the prerequisite course. As knowledge may decay over time or due to interference from learning new information (Tomlinson *et al.* 2009, Bunce *et al.* 2011), it is possible that students initially learned the material in the prerequisite when the material was needed, but then lost that information over the time between the courses.

Conversely, the instructors of the subsequent courses may anticipate this, providing a sufficient review of previously covered material, making the prerequisite unnecessary. It is also possible that students may focus more on the new, unfamiliar, material and less on the prerequisite, familiar material, which could lead to a leveling of performance results. Finally, it may be possible that the prerequisite course and the follow-up courses were not well aligned with one another. Even though very familiar material was taught in the prerequisite course, the follow-up course may not have explicitly built on this material and thus inadvertently did not show the students how this material was related. As repetition and retrieval practice has been shown to improve learning (Roediger and Karpicke 2006a, Roediger and Karpicke 2006b, Karpicke and Blunt 2011, Dobson and Linderholm 2015), prerequisite courses could point out material that students will encounter in the follow-up course. Likewise, follow-up courses could explicitly show how the course builds on previous material.

In the case of recommended prerequisites, if the follow-up course builds on the prerequisite course, then it could be expected that students who have successfully completed the prerequisite course would earn higher grades in the follow-up course compared to students who did not complete the prerequisite course. In this study, we found limited support of this hypothesis. While students who completed a recommended prerequisite human physiology laboratory course earned significantly higher grades in a follow-up human anatomy course compared to students who did not enroll in the prerequisite, the effect was small (Table 3). This result is similar to those previously reported (Steele and Barnhill 1982, Canaday and Lancaster 1985, Wright *et al.* 2009, Sato *et al.* 2017).

While there is considerable overlap between a physiology laboratory course and a human anatomy course, the prerequisite physiology laboratory may not be important for success in the human anatomy course if the physiology laboratory course emphasizes physiology over anatomy. For example, a physiology laboratory course may emphasize electromyography, length-tension relationships, and the cross-bridge cycle in a unit on the muscular system, whereas a human anatomy course may emphasize muscle structure, location, identification, and nerve innervations. So, while the same organ systems were taught, they were done so with different goals, and thus the concepts covered in the physiology laboratory course may not directly impact student performance in the follow-up human anatomy courses.

Limitations

While our results largely show that the impact of prerequisite courses on student success in subsequent physiology courses is minimal, there are several limitations to this study. First, we assessed a limited number of course sequences at a single large research-intensive institution. The courses that were studied were designed relatively independently of one another, and as such there is room for improvement on the deliberate linkage between the courses. To this end, bigger impacts of prerequisites may be seen in physiology courses that are more tightly aligned with their prerequisite courses.

Additionally, the courses in this study assessed a single student population and results may differ if different populations are studied based on type of institution, geographic location, and ethnic/gender composition. The results of this study could also have been confounded by the fact that we did not conduct detailed written surveys or interviews with students, which could have revealed whether students had acquired knowledge of physiology in previous settings such as in other institutions or in high school. Finally, we only assessed the impact of prerequisite courses on a limited number of Core Concepts and physiology topics (see Methods). While it is possible that our results extend to other Core Concepts and topics, it is also possible that the impact of the prerequisite

continued on next page

courses would have been larger for other areas. Future studies assessing a more thorough set of Core Concepts or physiology topics would be warranted to determine a broader impact of prerequisite courses on physiology education.

Conclusion

The results from this study highlight the importance of assessing course sequence and linkage during curricular design and reform since the results suggest that prerequisites may not be as critical for success as we thought (at least for the courses and students in this study).

We have two suggestions for increasing the alignment of courses for more positive outcomes. First, curriculum mapping could be used to identify common learning outcomes among prerequisite and follow-up courses (Uchiyama and Radin 2009), which could lead to improved linkage of theses courses and thus improved outcomes. Second, instructors of prerequisite and follow-up courses could meet regularly to discuss the content of both courses and how best to design the courses so that the follow-up courses build on the subject matter presented in the prerequisite courses.

As a central component of many health-related undergraduate and graduate programs, wide-scale exploration of the impact of physiology and anatomy and physiology courses would go a long way towards ensuring that we are providing a high-quality educational experience for the next generation of health-care practitioners.

Acknowledgements

We would like to acknowledge and thank the students in our courses that took part in this study.

About the Authors

Justin Shaffer is a Teaching Associate Professor at the Colorado School of Mines. He teaches introductory biology and chemical engineering courses.

Sam Schriner is an Assistant Teaching Professor at UC Irvine where he teaches physiology.

Catherine Loudon is a Teaching Professor at UC Irvine where she teaches physiology.

Samantha Dacanay, Usman Alam, and Jennifer Dang earned their bachelor's degrees from UC Irvine.

Nancy Aguilar-Roca is an Associate Teaching Professor at UC Irvine where she teaches physiology.

Pavan Kadandale is an Associate Teaching Professor at UC Irvine where he teaches biochemistry.

Brian Sato is an Associate Teaching Professor at UC Irvine where he teaches molecular biology.

Literature cited

Bloom BS, Krathwohl DR, and Masia BB (1956) Taxonomy of educational objectives : the classification of educational goals. New York: D. McKay, p 2 volumes (207, xiv, 196 pages).

Bransford J, National Research Council (U.S.). Committee on Developments in the Science of Learning. and National Research Council (U.S.). Committee on Learning Research and Educational Practice (2000) How people learn : brain, mind, experience, and school. Washington, D.C.: National Academy Press, p x, 374 p.

Bunce DM, VandenPlas JR, Cameron S (2011) Decay of Student Knowledge in Chemistry. *J Chem.l Educ.* 88(9): 1231-1237.

Canaday SD and Lancaster CJ (1985) Impact of undergraduate courses on medical student performance in basic sciences. *J Med Educ* 60: 757-763.

Dobson JL and Linderholm T (2015) Self-testing promotes superior retention of anatomy and physiology information. *Adv Health Sci Educ* 20: 149-161. doi: 10.1007/s10459-014-9514-8.

Forester JP, McWhorter DL, and Cole MS (2002) The relationship between premedical coursework in gross anatomy and histology and medical school performance in gross anatomy and histology. *Clinical anatomy* 15: 160-164. doi: 10.1002/ca.1114.

Heiner CE, Banet AI and Wieman C (2014) Preparing students for class: How to get 80% of students reading the textbook before class. *Am J Phys* 82: 989-996.
doi: 10.1119/1.4895008.

Karpicke JD and Blunt JR (2011) Retrieval Practice Produces More Learning than Elaborative Studying with Concept Mapping. *Science* 331: 772-775. doi: 10.1126/science.1199327.

McCleary VL, Aasen G, and Slotnick HB (1999) Predictors of success in undergraduate human physiology. *Am J Physiol* 277: S119-126. doi: 10.1152/advances.1999.277.6.S119.

McRae MP (2010) Correlation of preadmission organic chemistry courses and academic performance in biochemistry at a midwest chiropractic doctoral program. *J Chiropr Educ* 24:30-34.

Michael J and McFarland J (2011) The core principles ("big ideas") of physiology: results of faculty surveys. *Adv Physiol Educ* 35: 336-341. doi: 10.1152/advan.00004.2011.

continued on next page

Michael JA (1998) Students' misconceptions about perceived physiological responses. *Am J Physiol* 274: S90-98. doi: 10.1152/advances.1998.274.6.S90.

Michael JA, Richardson D, Rovick A, Modell H, Bruce D, Horwitz B, Hudson M, Silverthorn D, Whitescarver S, and Williams S (1999) Undergraduate students' misconceptions about respiratory physiology. *Am J Physiol* 277: S127-135. doi: 10.1152/advances.1999.277.6.S127.

Modell HI, Michael JA, Adamson T, Goldberg J, Horwitz BA, Bruce DS, Hudson ML, Whitescarver SA and Williams S (2000) Helping undergraduates repair faulty mental models in the student laboratory. *Adv Physiol Educ* 23: 82-90. doi: 10.1152/advances.2000.23.1.S82.

Richardson DR (2000) Comparison of naive and experienced students of elementary physiology on performance in an advanced course. *Adv Physiol Educ* 23: 91-95. doi: 10.1152/advances.2000.23.1.S91.

Roediger HL and Karpicke JD (2006a) The Power of Testing Memory Basic Research and Implications for Educational Practice. *Perspect Psychol Sci* 1: 181-210. doi: 10.1111/j.1745-6916.2006.00012.x.

Roediger HL and Karpicke JD (2006b) Test-enhanced learning - Taking memory tests improves long-term retention. *Psychol Sci* 17: 249-255. doi: DOI 10.1111/j.1467-9280.2006.01693.x.

Rovick AA, Michael JA, Modell HI, Bruce DS, Horwitz B, Adamson T, Richardson DR, Silverthorn DU, and Whitescarver SA (1999) How accurate are our assumptions about our students' background knowledge? *Am J Physiol* 276: S93-101. doi: 10.1152/advances.1999.276.6.S93.

Sato BK, Lee AK, Alam U, Dang JV, Dacanay SJ, Morgado P, Pirino G, Brunner JE, Castillo LA, Chan VW, and Sandholtz JH (2017) What's in a Prerequisite? A Mixed-Methods Approach to Identifying the Impact of a Prerequisite Course. *CBE Life Sci Educ* 16. doi: 10.1187/cbe.16-08-0260.

Shaffer JF, Dang JV, Lee AK, Dacanay SJ, Alam U, Wong HY, Richards GJ, Kadandale P and Sato BK (2016) A Familiar(ity) Problem: Assessing the Impact of Prerequisites and Content Familiarity on Student Learning. *Plos One* 11: e0148051. doi: 10.1371/journal.pone.0148051.

Steele MW and Barnhill, BM (1982) Lack of impact of undergraduate genetic courses on the teaching of medical genetics. *Am J Hum Genet* 34: 501-506.

Tomlinson TD, Huber DE, Rieth CA, and Davelaar EJ (2009) An interference account of cue-independent forgetting in the no-think paradigm. *Proceedings of the National Academy of Sciences of the United States of America* 106: 15588-15593. doi: 10.1073/pnas.0813370106.

Uchiyama KP and Radin JL (2009) Curriculum Mapping in Higher Education: A Vehicle for Collaboration. *Innovative Higher Education* 33:271-280.

Wright R, Cotner S, and Winkel A (2009) Minimal impact of organic chemistry prerequisite on student performance in introductory biochemistry. *CBE Life Sci Educ* 8: 44-54. doi: 10.1187/cbe.07-10-0093.

Welsh AJ (2012) Exploring undergraduates' perceptions of the use of active learning techniques in science lectures. *J Coll Sci Teach* 42: 80-87.

Back to TOC

Academic Anxiety in Higher Education: Causes, Implications, and Potential Solutions

July-Ann El-Baze[1], Skye Stowe[1], Suzanne Hood[2], Heather Lawford[2], Murray Jensen[3], Kerry Hull[*,2]
[1]Undergraduate students, Bishop's University, Sherbrooke, QC, Canada who participated equally in this work.
[2]Department of Biology, Bishop's University, Sherbrooke, QC, Canada
[3]University of Minnesota, Minneapolis, MN
*To whom correspondence should be addressed.
khull@bishops.ca

Abstract
Academic performance is not determined by cognitive abilities alone; student anxiety influences how students learn and how accurately evaluations reflect their learning. It is thus important to consider the emotional, as well as the cognitive, impact of changing teaching practices. This article reviews the contributions of individual differences and our classroom practices to student anxiety, and provides recommendations on how instructors can help students avoid the negative spiral of anxiety and poor performance. https://doi.org/10.21692/haps.2018.033

Key words: academic anxiety; personality measures; active learning; evidence-based teaching practices

Introduction
As instructors, we are increasingly called upon to consider *how* we are teaching our students as well as *what* we are teaching them. By adopting evidence-based classroom and evaluative practices, we anticipate that our students will perform better and that we will more accurately measure their acquired knowledge and skills. What may not be given adequate consideration in this process, however, is how these classroom practices affect the non-cognitive aspects of learning such as a student's emotional state. Classroom and evaluative practices impact student anxiety, and anxiety can interfere with learning and reduce the fidelity of evaluations. This article reviews the contributions of individual differences and our classroom practices to student anxiety, and provides recommendations on how instructors can help students avoid the negative spiral of anxiety and poor performance.

Academic Anxiety
Classroom practices can have emotional as well as cognitive impacts that carry implications for student mental health. In fact, students cite academic factors as significant stressors/causes of concern more frequently than non-academic factors such as finances and personal relationships (Kumaraswamy 2013, Deasy *et al.* 2014, Beiter *et al.* 2015). While a certain amount of stress is unavoidable and even potentially beneficial, a common response to stress is anxiety (Johnson 2009). Anxiety can be a temporary feeling of insecurity and physiological arousal triggered by a stimulus. This common and transient response to a stressful situation can be described as *state anxiety* (Spielberger 1966). In contrast, *trait anxiety* refers to a general tendency to perceive stimuli as threatening and to display disproportionate reactions. Thus, an individual with higher *trait* anxiety will experience *state* anxiety more frequently and/or intensely. Within trait anxiety, test anxiety is the form most relevant to this discussion.

Test anxiety is a form of trait anxiety triggered by evaluative situations (Zeidner 1998). While most commonly studied in the context of traditional exams, test anxiety can extend to any form of academic evaluation, including essays and oral presentations, and is thus often described as evaluative anxiety. Students with higher test anxiety focus on avoiding failure and not appearing incompetent (Sarason 1984). As with other forms of anxiety, test anxious individuals can experience cognitive symptoms, such as worry and fear; physiologic (bodily-affective) responses, such as elevated heartbeat; and behavioural responses, such as nervousness and uneasiness (Sarason 1984, Zeidner and Mendel 2007). While estimates vary, studies report that approximately 25-40% of the population suffers from test anxiety (Cassady 2010), with a higher prevalence in women and minority groups (Hembree 1988, Wong *et al.* 2006, Bayram and Bilgel 2008).

Of particular relevance to anatomy and physiology instructors, especially those in the allied health fields, is the phenomenon of *science anxiety*. This term describes students who experience heightened anxiety during science exams but are able to remain calm and productive in non-science exams (Mallow 2006). A major contributing factor to science anxiety is the difficulty and rigor of post-secondary science classes, for which students feel underprepared by their high

continued on next page

school science classes (Udo *et al.* 2004). For instance, students with high science anxiety often note negative experiences in past science classes and unhelpful instructors (Brownlow *et al.* 2000). The emphasis on "making science fun" and memorization, at the cost of teaching important analytical and critical skills, may have contributed to students' feelings of unpreparedness (Mallow 2006). The lack of appropriate science role models may also be a factor, particularly in women and minorities (Hembree 1988, Mallow 1994). Bryant *et al.* (2013) also note that instructor bias against women as a contributing cause for increased science anxiety in women.

Anxiety and Academic Performance

Apart from its influence on student well-being, anxiety also impacts academic performance. As reviewed by Hembree (1988) and others (Baumeister 1984, Cassady and Johnson 2002, Ramirez and Beilock 2011) test anxiety correlates negatively with academic achievement, SAT scores, IQ test scores, problem solving ability, memory, and grades. High anxiety students experience self-deprecatory preoccupations and intrusive thoughts that reduce goal-focussed attention and impair task-relevant cognitive activity (Mowbray 2012). Tasks requiring more cognitive resources are particularly susceptible to the negative impacts of anxiety (Beilock *et al.* 2004); for example, anxiety impairs deep-level processing tasks (e.g., critical thinking and integrating previous knowledge) and is positively correlated with surface-level processing (e.g., memorizing and repetition) (Rozendaal *et al.* 2001). Test anxiety is thus particularly detrimental for evaluations that emphasize the higher levels of Bloom's taxonomy, which prioritize analysis and synthesis over rote learning (Crowe *et al.* 2008).

Academic anxiety impacts all stages of the learning-testing cycle: planning and executing learning tasks, completing evaluations, and reflecting on performance (Cassady 2004, Bryant *et al.* 2013). In terms of exam preparation, evaluation anxiety is particularly problematic for students with poor study skills, since it interferes with new memory formation (Naveh-Benjamin *et al.* 1981). Highly test-anxious students have trouble identifying the main points of reading assignments and integrating the information (Naveh-Benjamin *et al.* 1987). For instance, the cheat sheet of highly anxious students is more likely to consist of verbatim reports of lectures (Cassady 2004). Even test-anxious students with good study skills can still perform badly on exams, since anxiety impedes information retrieval (Naveh-Benjamin *et al.* 1987). Indeed, student grades may reflect their ability to cope with evaluation anxiety as much as their skills and knowledge (Zeidner and Mendel 2007). In accordance with the negative impact of anxiety on higher-level processing skills, highly anxious students do worse on evaluations requiring high cognitive involvement, such as essay questions and take home examinations (Benjamin *et al.* 1981). Students with higher test anxiety report increased feelings of helplessness post-test (Cassady 2004). They are more likely to attribute failure to factors outside of their control, either external (e.g., unfair test) or internal but fixed (e.g., low intelligence). Students holding these views may be less motivated to improve their study practices, reducing the chance of success in future evaluations (Covington 1985). As discussed later in this review, the personality measures associated with these feelings of helplessness (such as locus of control and academic self-efficacy) are potential targets for instructional interventions.

It should be noted that academic anxiety is not always detrimental (Cooper *et al.* 2018a). As suggested by recent adaptations of the Yerkes-Dodson law (Teigen 1994) increased arousal increases the learning rate for easy tasks. Hence, anxiety may be a useful motivator for student completing easy activities. An intermediate level of arousal is optimal for learning difficult tasks; however, there is a "tipping point" in arousal level beyond which performance is impaired (Teigen 1994). Thus, providing motivation by purposely increasing student anxiety during the completion of cognitively demanding tasks would be beneficial for low-anxiety students but detrimental for higher anxiety students. Since anxiety is often higher in women and minorities and in science classes generally (Hembree 1988, Wong *et al.* 2006, Bayram and Bilgel 2008, Cooper *et al.* 2018a), employing tactics to boost arousal level with the goal of increasing student engagement could instead result in an overall decrease in performance, particularly in groups underrepresented in STEM fields.

Anxiety and Stable Personality Measures

The impact of a student's anxiety on academic performance reflects the interaction between the student's personality and the characteristics of the learning situation. *Personality traits* are relatively stable individual differences in behavioural patterns (Costa and McCrae 1992, Hogan *et al.* 1996). The measurable personality supertraits of *conscientiousness*, *neuroticism*, and *openness to experience* are particularly relevant to learning (Lowe *et al.* 2008, De Feyter *et al.* 2012). Conscientious individuals are characterized by their self-discipline, persistence, order, dutifulness, compliance, and imperturbability (Furnham *et al.* 2003). Unsurprisingly, conscientiousness is a strong predictor of good academic performance (Diseth 2003, Noftle and Robins 2007, Poropat 2009, De Feyter *et al.* 2012). In contrast, neuroticism, characterized by emotional instability and anxiety, can be negatively or positively correlated with academic performance (Chamorro-Premuzic and Furnham 2004). The nature of this relationship with academic performance reflects the complexity of this trait. For instance, *self-handicapping* students use their anxiety as an excuse for poor performance; they may avoid studying all together so that failure cannot be attributed to lack of ability (De Feyter *et al.* 2012). In contrast, *defensive pessimists* protect their self-esteem by setting unrealistically low expectations, but also intensify their efforts to prevent failure and thereby improve performance

continued on next page

(Norem and Cantor 1986). As discussed below, behaviors and attitudes such as academic self-efficacy, mindset, and self-regulation appear to influence the impact of neuroticism on performance.

Finally, the impact of the *openness to experience* trait is task-specific. In comparison with their lower-scoring counterparts, students reporting higher levels of openness perform better in creative problem-solving tasks but worse in memorization tasks (Furnham *et al.* 2003). As such, increased test anxiety in these individuals may impair performance on more rote, memorization-based evaluations.

It is important to note that conscientiousness, neuroticism, and openness to experience are very broad supertraits. Academic performance may only relate to certain aspects of each, such as dutifulness and achievement striving (both facets of conscientiousness) and anxiety, which can differ in individuals with identical supertrait scores (Furnham *et al.* 2003). While instructors should be aware of the impact of these traits on student anxiety and performance, they are relatively stable and have a strong biological basis, and are thus resistant to change.

Anxiety and Academic Attitudes and Behaviors

In contrast to the broad supertraits, other measurable indices of student personality such as self-regulation, intelligence mindset, academic self-efficacy, and academic self-concept are relatively malleable and can provide targets for potential interventions (Bidjerano and Yun Dai 2007). We begin by briefly explaining each of these terms before describing the impact of various classroom practices.

Self-Regulation

Academic self-regulation describes the student's ability to control their learning environment without external intervention. Learners with high self-regulation skills have the ability to constantly adapt their goals and strategies to the demands of each learning situation (Bidjerano and Yun Dai 2007). This construct may mediate the relationship between the personality supertrait of conscientiousness and academic performance: that is, highly conscientious individuals may excel academically because they have strong self-regulatory skills (Snow *et al.* 1996, Bidjerano and Yun Dai 2007).

The components and sub-components of self-regulation are illustrated in Figure 1. In order to be effective learners, students need to possess the necessary study skills (cognition), be aware of how and when to use particular skills (metacognition), and have the will to do so (motivation) (Schraw *et al.* 2006). As reviewed by Schraw *et al.* (2006), Cognitive skills can be further divided into learning techniques (e.g., mnemonics, concept maps), problem-solving skills (e.g., predict-observe-explain), and critical thinking strategies (e.g., evaluating the reliability of an information source). Metacognition, in turn, can be divided into three stages: planning, monitoring, and evaluation. In the planning stage, students set goals and select appropriate strategies. Next, they monitor and adjust their strategies as they are implemented, and, finally, they evaluate the effectiveness of the strategies in respect to goals. The final component, motivation, refers to beliefs and attitudes that determine how the student will acquire and use their cognitive and metacognitive skills. The two components of motivation - mindset and academic self-efficacy - are discussed further below.

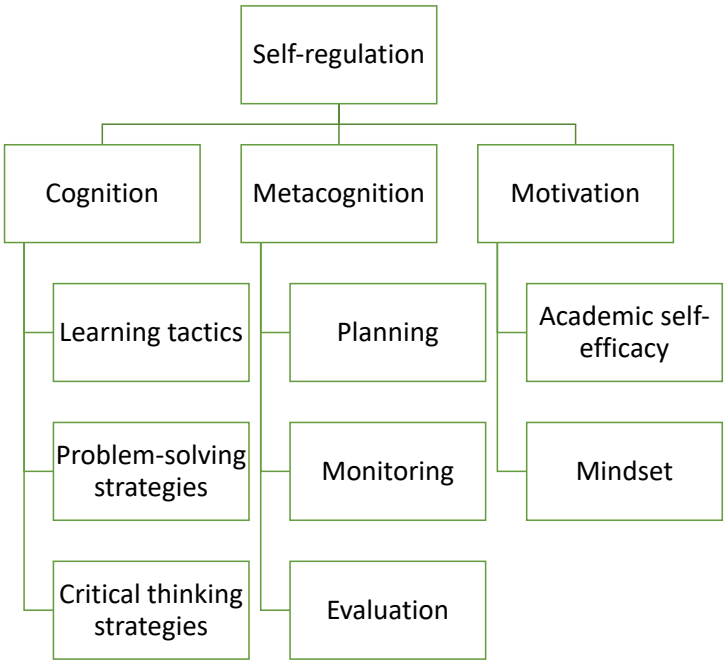

Figure 1. *Components of Self-regulation. Based on information from Shraw (2006).*

Intelligence Mindset

Students' beliefs about the malleability of their own intelligence can significantly impact their anxiety and performance. A possible area of intervention is thus to encourage students to shift from an entity theory of intelligence (fixed mindset) in which intelligence is seen as fixed and independent of effort, to an incremental theory of intelligence (growth mindset) in which intelligence is seen as malleable and directly related to effort and persistence (Dweck and Leggett 1988, Komarraju and Nadler 2013). The fixed mindset is correlated with high anxiety: students reporting higher worry and emotionality are concerned about not appearing incompetent and avoiding failure. In other words, they focus on performance, a characteristic of the fixed mindset. In contrast, students reporting lower anxiety levels are more likely to focus on mastering goals, which is associated with having a growth mindset: specific behaviors exhibited by these students include developing skills and mastering the

continued on next page

content (Stan and Oprea 2015). While the causal mechanisms underlying this relationship are unclear, it is well-accepted that student mindset is a malleable construct (Dweck and Leggett 1988) and thus could serve as a potential target for both increasing performance and reducing anxiety.

Academic Self-Efficacy

Academic self-efficacy refers to students' subjective beliefs about their ability to cope with academic challenges (Bandura 1997, McIlroy *et al.* 2000). Possessing the necessary skills and knowledge to complete the task is not sufficient for academic success; students must also believe that they can be successful under the challenging circumstances associated with evaluation (Bandura 1997, Artino 2012). As such, academic self-efficacy is both situational and task-specific (Artino 2012). Improving academic self-efficacy may also directly impact academic performance, since many studies reveal strong correlations between these two parameters (see, for instance, (Richardson *et al.* 2012, Honicke and Broadbent 2016). Indeed, Lent *et al.* (1987) found that self-efficacy is the most useful predictor of both grades and persistence in STEM majors. The meta analysis of Talsma *et al.* 2018 (2018) took the additional step of examining causal relationships between performance and self-efficacy, and observed a reciprocal relationship: their data supported the validity of the statements "I believe therefore I achieve" and "I achieve therefore I believe", at least for adult learners.

The task-specific nature of academic self-efficacy has implications for how we measure it (Bandura 2012). For instance, biology self-efficacy is a measure of self-efficacy that focuses on skills, concepts, and knowledge specific to biology (Ainscough *et al.* 2016). Moreover, this specificity suggests that self-efficacy should be addressed consistently throughout a student's career; gains in self-efficacy in one course may not necessarily transfer to the next course. Aiscough *et al.* (2016) observed lower science efficacy in females than in males. While this gender difference was not associated with poorer performance in their study, it did appear to influence their intention to remain in science or even their career choice (Ainscough *et al.* 2016); this is especially pertinent as women are still a minority in STEM (Hango 2013).

Fostering academic self-efficacy can help alleviate the deleterious impact of anxiety on performance. High-anxiety, low self-efficacy students demonstrate poorer performance compared to students with similar anxiety levels but higher self-efficacy (Raufelder and Ringeisen 2016). Strong academic self-efficacy can improve performance by protecting against cognitive and bodily aspects of anxiety resulting from a lack of confidence (Raufelder and Ringeisen 2016), while weak self-efficacy can decrease performance and as a result cause students to withdraw from completing STEM majors (Dweck and Leggett 1988, Zimmerman *et al.* 1992).

Academic calibration refers to the relative accuracy of a student's belief in their ability to achieve and their actual achievement (Glenberg *et al.* 1987). Self-regulation plays a major role in calibration (Stone 2000). Poor academic calibration occurs in two forms: overconfidence (when one's self-assessed knowledge is greater than their actual knowledge), and underconfidence (when one knows more than one believes) (Stone 2000). Importantly, each of these forms of miscalibration influences students' studying behavior. Underconfidence, which can result from high levels of neuroticism or the combination of low self-efficacy with emotional stability, can provide motivation for increased studying to avoid failure and thus be beneficial (De Feyter *et al.* 2012). Poor academic calibration in the form of moderate overconfidence has the potential to improve student persistence and motivation in challenging academic tasks (Stone 2000). This finding suggests that poor academic calibration is not always detrimental to student learning. In contrast, overconfidence resulting from the combination of high academic self-efficacy and high emotional stability can result in overconfidence and low academic motivation (Furnham *et al.* 2003). Training students in metacognitive study strategies such as self-testing and self-monitoring facilitates better self-efficacy calibration (Stone 2000, Hattie and Timperley 2007).

Interestingly, college self-efficacy scales taken at the beginning of the first semester may not predict academic performance accurately. Moreover, self-calibration is generally lower in first year university students because they are not yet familiar with the academic expectations at the university level (Gore 2006) and their self-efficacy beliefs primarily depend on prior high school experiences (Ainscough *et al.* 2016). Students need feedback on their performance in order to accurately calibrate their self-efficacy beliefs (Gore 2006).

A related construct known as *locus of control* indicates the degree to which one attributes the outcome to external forces, such as the instructor, or to internal forces, such as one's own actions (Hrbáčková *et al.* 2012). An internal locus of control has been linked to greater academic achievement, particularly in males (Findley and Cooper 1983). However, the association between locus of control and academic performance is weaker than that of academic self-efficacy.

Academic self-concept

If academic self-efficacy indicates self-perceived ability to achieve a particular goal, *academic self-concept* is a more general measure of self-perceived ability in a particular domain (Jansen *et al.* 2015, Raufelder and Ringeisen 2016, Cooper *et al.* 2018b). Importantly, academic self-concept is influenced by the perception of other student's academic abilities (Cooper *et al.* 2018b). Self-concept is very pertinent in the context of active learning classrooms because students interact more with one another and hence have more

continued on next page

opportunity to compare their perceived abilities with others (Cooper *et al.* 2018b).

The process by which students build self-concept is influenced by mindset (Dweck and Leggett 1988). For example, a student categorized as a fixed mindset learner would "feel smart" in comparison to peers through attaining *performance* goals, such as finishing an assignment first, regardless of the process or the difficulty of the assignment. In contrast, a student categorized as a growth mindset learner would "feel smart" by achieving *learning* goals, such as answering a difficult question independently by using complex or work-intensive strategies (Dweck and Leggett 1988).

Compared with academic self-efficacy, academic self-concept is less effective for predicting a student's current ability; however, it is more useful as a measure of student's motivation to pursue a career in the chosen domain (Jansen *et al.* 2015). Moreover, academic self-concept is not as strong as self-efficacy as a predictor of anxiety (Bong and Skaalvik 2003, Jansen *et al.* 2015).

Table 1. *Academic attitudes relevant to anxiety and performance Please see the narrative for references.*

	What is it?	What does it predict or influence?	Examples	Can it be improved? How?
Academic self-regulation	The student's ability to control a learning environment without external intervention; includes cognition, metacognition, and motivation.	A student's ability to adapt their goals and strategies according to the demands of specific learning situations.	See Figure 1 for more information.	Instructors can teach cognitive and metacognitive strategies. See "intelligence mindset" and "academic self-efficacy" below for strategies to address motivation.
Intelligence mindset	The student's belief that intelligence is a malleable trait that can be improved (growth mindset) or is predetermined and cannot be changed (fixed mindset).	A growth mindset is associated with greater effort and persistence.	Fixed: "I am not smart, this is not going to change, why should I invest time in something which cannot be changed?" Growth: "I can't do it yet, but there is hope…."	Instructors can encourage students to focus on goal mastery rather than performance.
Academic Self-efficacy	The student's belief that they possess the skills and knowledge necessary to meet a specific performance goal.	Higher academic self-efficacy is associated with lower anxiety and higher performance.	High: "I am confident that I can achieve good results if I work hard." Low: "No matter how hard I work, I can't master the course topics."	Instructors can provide opportunities for students to master challenging and meaningful tasks, and provide detailed, individual feedback focussing on strategies for the future. Group work is very helpful.
Academic Self-concept	The student's perception of their general abilities in a particular domain, often in comparison with others.	Students with a higher academic self-concept are more likely to pursue a career in the relevant domain.	High: "I am smart so I will finish my degree" Low: "I am not as smart as my peers so I will never finish my degree"	Instructors can foster students' academic self-concept by fostering a supportive and less competitive classroom environment, and by providing career advice.
Locus of control	The degree to which the student attributes the outcome of a situation to internal or external factors.	Students with an internal locus of control show greater persistence and motivation.	External: "The teacher/test was unfair" Internal: "I didn't study enough, my notes weren't good enough"	Teaching students metacognitive strategies can foster the development of an internal locus of control.

continued on next page

Personality and Classroom Interventions

The components of self-regulation are potential targets for classroom interventions to reduce anxiety and improve performance (Mallow 2006). These measures are not independent entities, so targeting one can have positive or negative impacts on others.

Tanner (2012) explains how metacognitive strategies can be adapted to various learning situations: entire courses, individual lectures, and specific exams. For instance, the monitoring stage of metacognition can be addressed by the "Muddiest Point" technique, in which students take a few minutes to write about a confusing aspect of the class material. The planning and evaluation stages can also be fostered by asking students to complete pre-assessments (What do I already know about this topic?) and retrospective post-assessments (How has my understanding of this topic changed as a result of the learning activity?) (Tanner 2012).

Academic self-efficacy can also be improved (or worsened) by instructor practices. Instructors can foster academic self-efficacy by creating opportunities for students to master challenging and meaningful tasks (Pajares 1996). As discussed below, active learning strategies such as question and answer, collaborative and inquiry exercises, student response systems, and conceptual problem assignments provide students with these potential mastery experiences (Bong and Skaalvik 2003, Usher and Pajares 2009). Regardless of the activity type, clear instructions are of primordial importance in determining student success (Bong and Skaalvik 2003, Usher and Pajares 2009).

Appropriate student feedback is also key to promoting academic self-efficacy (Hattie 1999, Hattie and Timperley 2007). Providing detailed, individual feedback and recommending strategies for future iterations can help students estimate their future success (Margolis and Mccabe 2006) and thus also promote self-efficacy calibration (Stone 2000). However, feedback can harm academic self-efficacy if students perceive a high level of threat to their self-esteem or if the feedback does not help students understand the cause of poor performance (Hattie and Timperley 2007). The positive impact of feedback can be maximized by acting on students' mindset and self-regulation. Focussing on how the student approached a learning goal and their learning strategies, rather than the outcome, promotes a growth mindset and fosters self-regulation. By understanding the causes of both their successes and their failures, students are more apt to set goals that are challenging yet achievable, and to invest more effort and commitment in their attempts to meet the goals (Locke and Latham 2002, Hattie and Timperley 2007).

During new experiences, for which students do not have prior experience, watching peers succeed can be very helpful (Usher and Pajares 2009). While not as effective as performance accomplishments, vicarious learning can also promote the development of academic self-efficacy and appears to augment the impact of the performance accomplishments (Hackett *et al.* 1992, Bandura 1997). To be effective, students must perceive the peer model as similar to themselves, but also credible, competent, and enthusiastic. Instructors can play a role in ensuring that student models are enthusiastic at or slightly above the skill level of the others, training them to effectively model the experience as credible and competent (Artino 2012).

Anxiety and Active Learning Practices

Active learning (constructivist) practices have been shown to increase examination performance and positively impact skill development compare with lecture-based (expositional) practices (Freeman *et al.* 2014). This shift in classroom practices substantially alters how students interact with their peers, with the instructor, and with the material, and thus can impact anxiety. With respect to academic anxiety, Harper and Daane (1998) observed a positive impact of active learning practices, concluding that mathematics anxiety largely "stemmed from rigid and structured classroom instructional methods" (p. 34) and often decreased with the implementation of active learning methods. The interactive nature of active learning pedagogies tends to alleviate the unfriendly atmosphere of large lecture classrooms by increasing students' sense of belonging. This change has been credited with improving science self-efficacy and performance (Ballen *et al.* 2017). Since underrepresented minority students are most affected by traditional lecture-based classrooms, active learning pedagogies may act to decrease academic performance gaps between non-minority and minority students and improve the retention of under-represented groups in STEM majors (Lent *et al.* 1987, Ballen *et al.* 2017). Active learning pedagogies may similarly reduce the achievement gap between males (the majority gender) and females (the underrepresented gender) in STEM classrooms (Lorenzo *et al.* 2006).

It should be noted that the establishment of a safe and equitable classroom environment is of primary importance in reducing student anxiety (Mallow 2006, Rocca 2010, Tanner 2013). Tanner (2013) outlines 21 practices that promote the engagement of all students, regardless of gender, ethnicity, personality, background knowledge, but does not directly consider the potential of each practice to foster or reduce anxiety.

Indeed, some students identify active learning as a source of anxiety, particularly in relation to answering verbal questions (England *et al.* 2017). More recent work has started to identify characteristics of different active learning classroom practices that can foster or reduce anxiety (Cooper *et al.* 2018a). Cold calling (asking a student a question who has not volunteered) is often rated by students as the most anxiety inducing

continued on next page

practice, followed by volunteering to answer a posed question, completing worksheets, working in groups, and using student response clickers (England *et al.* 2017). We address cold calling, group work, and student response systems in greater detail below.

Cold Calling

Cold calling has been described as a highly effective method of formative assessment (Broeckelman-Post *et al.* 2016). If all students are called upon equally, this technique can ensure that all student voices are heard and create a more equitable classroom environment (Eddy *et al.* 2014). Random calling, in which the instructor chooses a student to share by drawing a popsicle stick or index card, can help guard against selective cold calling (Tanner 2013). However, cold calling is also notorious for provoking student anxiety (Broeckelman-Post *et al.* 2016, England *et al.* 2017, Cooper *et al.* 2018a). Students commonly feel anxiety about knowing the answer, fear negative evaluation by peers and teachers, and are apprehensive about speaking in front of others (Broeckelman-Post *et al.* 2016, Cooper *et al.* 2018a). Mallow (2006) suggests that reducing the wait time for a response from a student before asking another student can help reduce anxiety, but others note that increasing the wait time to 3-5 seconds is more equitable for introverted students (Tanner 2013). The amount of anxiety associated with cold calling is directly correlated with the amount of motivation it provides; students reporting that the practice motivated them to pay attention and attend class were the same students who reported increased anxiety (Broeckelman-Post *et al.* 2016). This correlation suggests that it could be the anxiety, rather than the technique itself, that is creating the motivation, in accordance with the Yerkes-Dodson law. Considering that repeated use of cold calling can reduce attendance of students with anxiety disorders (Broeckelman-Post *et al.* 2016), one could conclude that cold calling will only be beneficial for low-anxiety students.

It has been suggested that desensitizing students by regularly implementing cold calling practices can reduce this anxiety (Mccroskey 2009; Rocca 2010). In support of this practice, (Dallimore *et al.* 2013) observed that consistent use of cold calling increased voluntary participation over time. However, this finding has been challenged by (Cooper *et al.* 2018a), who postulated that the consistently negative experiences of cold calling reported in their study would exacerbate the anxiety invoked by the technique in the future. They support an alternate method, in which instructors gather group answers by walking around the classroom and sharing the ideas with the entire class.

Student Response Systems

An increasingly popular classroom technique involves the use of student response systems, such as "clickers" or online quizzing engines. Clickers are often used in conjunction with Peer Instruction, a method in which students answer questions individually, discuss their answer with a peer, and then answer the question a second time (Crouch and Mazur 2001). The instructor subsequently provides feedback. Peer instruction has well-documented positive effects on student performance (Crouch and Mazur 2001).

However, the implementation of peer instruction and student response systems in general has been shown to vary considerably between instructors. Schell and Butler (2018) and England *et al.* (2017) noted that the impact of using student response systems on student anxiety was quite sensitive to differences in implementation. Clicker use tended to increase anxiety if the accuracy of student responses had an impact on final course grades (England *et al.* 2017) and/or if the instructor imposed significant time constraints on student answers (Cooper *et al.* 2018a). Conversely, student response systems may decrease anxiety and increase academic self-concept if students are provided with adequate time to respond to questions and if the accuracy of their responses does not impact (or minimally impacts) their final grade (Cooper *et al.* 2018a).

Group Work

Working in small groups is a common feature of many active learning pedagogies, including process-oriented guided inquiry learning (POGIL), problem-based learning (PBL), and case-based learning (CBL) (Eberlein *et al.* 2008, Jensen 2016, Hopper 2018). By facilitating student understanding of class material, students report that group work decreases their overall anxiety about their abilities to succeed in sciences (Cooper *et al.* 2018a). Small-group discussions provoke equally low levels of anxiety in both males and females, whereas large-group discussions are considered more anxiety provoking by females than males (Eddy 2015).

However, group work can also induce anxiety if students fear negative evaluation from their peers; encouraging students to get know each other and allowing them to choose their own groups can mitigate this response (Cooper *et al.* 2018a). It is also important to note that background noise, such as discussions from nearby groups, may impair the ability of some students to perform tasks involving working memory (Beaman 2005). Moreover, students may direct their attention away from the task at hand because of the cognitive demands of ignoring the sound burden (Klatte *et al.* 2013). Since students differ in their ability to ignore extraneous stimulation, instructors may want to consider the placement of noisy groups vs. noise-sensitive groups to mitigate the impact of noise on particularly noise-sensitive students.

Anxiety and Evaluation Practices

Since academic anxiety is often focussed around assessments, evaluation practices can have significant emotional impacts that, in turn, influence performance. Even students with

continued on next page

adequate or exemplary study skills can experience retrieval failure as a result of debilitating evaluation anxiety, thereby reducing the assessment's fidelity. Since students often judge the difficulty of a test based on the first few questions, high-anxiety students may form a premature judgment which then fosters the self-deprecating thoughts and reactions that interfere with retrieval (Sarason *et al.* 1986, Zeidner 1995, Deasy *et al.* 2014). Including straight-forward questions at the beginning of an exam is a simple but effective way to reduce student anxiety. Instructors using multiple-choice questions with reasonable distractors should also be aware that anxiety can foster undecisiveness (Bar-Tal *et al.* 2013). Encouraging students to stick with their first answer in cases of indecision can be helpful.

While exam situations are often believed to be the most anxiety inducing for students, open-ended assessments such as papers and oral presentations are often perceived as equally or more stressful (Deasy *et al.* 2014, Pitt *et al.* 2017). Anxiety is influenced by the students' perception of the assessment's accuracy and fairness (Pitt *et al.* 2017). The use of rubrics has been shown to help students better understand teacher expectations and to reduce anxiety (Andrade and Du 2005, Panadero and Jonsson 2013). If accompanied by appropriate feedback, rubric use can improve academic self-efficacy (Panadero and Jonsson 2013) and facilitate the planning, monitoring, and evaluation phases of metacognition (Andrade and Du 2005).

Instructors can also address the physiological and cognitive symptoms of test anxiety by using relaxation techniques (Margolis and Mccabe 2006, Embse *et al.* 2013). Doherty and Wenderoth (2017) refined a novel approach pioneered by Ramirez and Beilock (2011), in which students spent five minutes writing about their testing worries at the beginning of an exam. Students then crumpled up the paper and threw it into the classroom corridor. This approach was shown to increase student performance and decrease student anxiety. However, it should be noted that, at moderate levels, the increased cognitive arousal associated with test anxiety appears to enhance performance (Cassady and Johnson 2002) by triggering adaptive self-regulation strategies such as greater use of metacognitive skills (Schutz and Davis 2000).

Conclusion

Academic performance is not determined by cognitive abilities alone; student anxiety influences how students learn and how accurately evaluations reflect their learning. While the inclusion of active learning practices generally appears to reduce anxiety, the impact on student emotions and attitudes may depend on how the practice is implemented as well as student personality characteristics. While more research remains to be done, instructors can have a positive impact on their students' learning, evaluation, and mental health by implementing classroom measures to improve student academic self-efficacy and the classroom environment.

About the Authors

July-Ann Baze and Skye Stowe are undergraduate students at Bishop's University in Sherbrooke, Quebec, who participated equally in this study. Suzanne Hood, Heather Lawford, and Kerry Hull teach at Bishop's University. Murray Jensen teaches at the University of Minnesota.

Literature cited

Ainscough L, Foulis E, Colthorpe K, Zimbardi K, Robertson-Dean M, Chunduri P and Lluka L (2016) Changes in Biology self-efficacy during a first-year university course. CBE—Life Sci Educ 15:ar19. doi: 10.1187/cbe.15-04-0092

Andrade HL and Du Y (2005) Student perspectives on rubric-referenced assessment. Pract Assess Res Eval 10:12.

Angelo T and Cross K Classroom assessment techniques: A handbook for college teachers, 2nd edition. Jossey-Bass, San Francisco, CA.

Artino AR (2012) Academic self-efficacy: From educational theory to instructional practice. Perspect Med Educ 1:76–85. doi: 10.1007/s40037-012-0012-5

Ballen CJ, Wieman C, Salehi S, Searle JB and Zamudio KR (2017) Enhancing diversity in undergraduate science: Self-efficacy drives performance gains with active learning. CBE—Life Sci Educ 16:ar56. doi: 10.1187/cbe.16-12-0344

Bandura A (1997) Self-efficacy: The exercise of control. W H Freeman/Times Books/ Henry Holt & Co, New York, NY, US.

Bandura A (2012) On the functional properties of perceived self-efficacy revisited. J Manag 38:9–44. doi: 10.1177/0149206311410606

Bandura A (1986) Social foundations of thought and action: A social cognitive theory.

Bar-Tal Y, Shrira A and Keinan G (2013) The effect of stress on cognitive structuring: A cognitive motivational model. Personal Soc Psychol Rev 17:87–99. doi: 10.1177/1088868312461309

Baumeister RF (1984) Choking under pressure: Self-consciousness and paradoxical effects of incentives on skillful performance. J Pers Soc Psychol 46:610–620. doi: 10.1037/0022-3514.46.3.610

Bayram N and Bilgel N (2008) The prevalence and socio-demographic correlations of depression, anxiety and stress among a group of university students. Soc Psychiatry Psychiatr Epidemiol 43:667–672. doi: 10.1007/s00127-008-0345-x

continued on next page

Beaman CP (2005) Auditory distraction from low-intensity noise: a review of the consequences for learning and workplace environments. Appl Cogn Psychol 19:1041–1064. doi: 10.1002/acp.1134

Beilock SL, Kulp CA, Holt LE and Carr TH (2004) More on the fragility of performance: Choking under pressure in mathematical problem solving. J Exp Psychol Gen 133:584–600. doi: 10.1037/0096-3445.133.4.584

Beiter R, Nash R, McCrady M, Rhoades D, Linscomb M, Clarahan M and Sammut S (2015) The prevalence and correlates of depression, anxiety, and stress in a sample of college students. J Affect Disord 173:90–96. doi: 10.1016/j.jad.2014.10.054

Benjamin M, McKeachie WJ, Lin Y-G and Holinger DP (1981) Test anxiety: Deficits in information processing. J Educ Psychol 73:816–824. doi: 10.1037/0022-0663.73.6.816

Bidjerano T and Yun Dai D (2007) The relationship between the Big-Five model of personality and self-regulated learning strategies. Learn Individ Differ 17:69–81. doi: 10.1016/j.lindif.2007.02.001

Bong M and Skaalvik EM (2003) Academic self-concept and self-efficacy: How different are they really? Educ Psychol Rev 15:40. doi: 10.1023/A:1021302408382

Broeckelman-Post M, Johnson A and Reid Schwebach J (2016) Calling on students using notecards: Engagement and countering communication anxiety in large lecture. J Coll Sci Teach 45:27–33.

Brownlow S, Jacobi T and Rogers M (2000) Science anxiety as a function of gender and experience. Sex Roles 42:119–131. doi: 10.1023/A:1007040529319

Bryant FB, Kastrup H, Udo M, Hislop N, Shefner R and Mallow J (2013) Science anxiety, science attitudes, and constructivism: A binational study. J Sci Educ Technol 22:432–448. doi: 10.1007/s10956-012-9404-x

Cassady JC (2010) Anxiety in schools: The causes, consequences, and solutions for academic anxieties. Peter Lang, New York, NY

Cassady JC (2004) The influence of cognitive test anxiety across the learning–testing cycle. Learn Instr 14:569–592. doi: 10.1016/j.learninstruc.2004.09.002

Cassady JC and Johnson RE (2002) Cognitive test anxiety and academic performance. Contemp Educ Psychol 27:270–295. doi: 10.1006/ceps.2001.1094

Chamorro-Premuzic T and Furnham A (2004) A possible model for understanding the personality-intelligence interface. Br J Psychol 95:249–264. doi: 10.1348/000712604773952458

Chen PP (2002) Exploring the accuracy and predictability of the self-efficacy beliefs of seventh-grade mathematics students. Learn Individ Differ 14:77–90. doi: 10.1016/j.lindif.2003.08.003

Cooper KM, Downing VR and Brownell SE (2018a) The influence of active learning practices on student anxiety in large-enrollment college science classrooms. Int J STEM Educ. doi: 10.1186/s40594-018-0123-6

Cooper KM, Krieg A and Brownell SE (2018b) Who perceives they are smarter? Exploring the influence of student characteristics on student academic self-concept in physiology. Adv Physiol Educ 42:200–208. doi: 10.1152/advan.00085.2017

Costa PT and McCrae RR (1992) Normal personality assessment in clinical practice: The NEO Personality Inventory. Psychol Assess 4:5–13. doi: 10.1037/1040-3590.4.1.5

Covington MV (1985) Test anxiety: Causes and effects over time. Adv Test Anxiety Res 4:55–68.

Crouch CH and Mazur E (2001) Peer instruction: Ten years of experience and results. Am J Phys 69:970–977. doi: 10.1119/1.1374249

Dallimore EJ, Hertenstein JH and Platt MB (2013) Impact of cold-calling on student voluntary participation. J Manag Educ 37:305–341. doi: 10.1177/1052562912446067

De Feyter T, Caers R, Vigna C and Berings D (2012) Unraveling the impact of the Big Five personality traits on academic performance: The moderating and mediating effects of self-efficacy and academic motivation. Learn Individ Differ 22:439–448. doi: 10.1016/j.lindif.2012.03.013

Deasy C, Coughlan B, Pironom J, Jourdan D and Mannix-McNamara P (2014) Psychological distress and coping amongst higher education students: A mixed method enquiry. PLOS ONE 9:e115193-23. doi: 10.1371/journal.pone.0115193

Diseth A (2003) Personality and approaches to learning as predictors of academic achievement. Eur J Personal 17:143–155. doi: 10.1002/per.469

continued on next page

Doherty JH and Wenderoth MP (2017) Implementing an expressive writing intervention for test anxiety in a large college course †. J Microbiol Biol Educ 18:3. doi: 10.1128/jmbe.v18i2.1307

Dweck CS and Leggett EL (1988) A social-cognitive approach to motivation and personality. Psychol Rev 95:256. doi: 10.1037/0033-295X.95.2.256

Eberlein T, Kampmeier J, Minderhout V, Moog RS, Platt T, Varma-Nelson P and White HB (2008) Pedagogies of engagement in science: A comparison of PBL, POGIL, and PLTL. Biochem Mol Biol Educ 36:262–273. doi: 10.1002/bmb.20204

Eddy SL, Brownell SE, Thummaphan P, Lan M-C and Wenderoth MP (2015) Caution, Student Experience May Vary: Social Identities Impact a Student's Experience in Peer Discussions. CBE Life Sci Educ 14:ar45. doi: 10.1187/cbe.15-05-0108

Eddy SL, Brownell SE and Wenderoth MP (2014) Gender gaps in achievement and participation in multiple introductory biology classrooms. CBE Life Sci Educ 13:478–92. doi: 10.1187/cbe.13-10-0204

Embse N von der, Barterian J and Segool N (2013) Test anxiety interventions for children and adolescents: A systematic review of treatment studies from 2000–2010. Psychol Sch 50:57–71. doi: 10.1002/pits.21660

England BJ, Brigati JR and Schussler EE (2017) Student anxiety in introductory biology classrooms: Perceptions about active learning and persistence in the major. PLOS ONE 12:e0182506. doi: 10.1371/journal.pone.0182506

Findley MJ and Cooper HM (1983) Locus of control and academic achievement: A literature review. J Pers Soc Psychol 44:419–427. doi: 10.1037/0022-3514.44.2.419

Freeman S, Eddy SL, McDonough M, Smith MK, Okoroafor N, Jordt H and Wenderoth MP (2014) Active learning increases student performance in science, engineering, and mathematics. Proc Natl Acad Sci 111:8410. doi: 10.1073/pnas.1319030111

Furnham A, Chamorro-Premuzic T and McDougall F (2003) Personality, cognitive ability, and beliefs about intelligence as predictors of academic performance. Learn Individ Differ 14:47–64. doi: 10.1016/j.lindif.2003.08.002

Glenberg AM, Sanocki T, Epstein W and Morris C (1987) Enhancing calibration of comprehension. J Exp Psychol Gen 116:119–136. doi: 10.1037/0096-3445.116.2.119

Gore PA (2006) Academic self-efficacy as a predictor of college outcomes: Two incremental validity studies. J Career Assess 14:92–115. doi: 10.1177/1069072705281367

Hackett G, Betz NE, Casas JM and Rocha-Singh IA (1992) Gender, ethnicity, and social cognitive factors predicting the academic achievement of students in engineering. J Couns Psychol 39:527–538. doi: 10.1037/0022-0167.39.4.527

Hango D (2013) Gender differences in science, technology, engineering, mathematics and computer science (STEM) programs at university. Statistics Canada.

Harper NW and Daane CJ (1998) Causes and Reduction of Math Anxiety in Preservice Elementary Teachers. Action Teach Educ 19:29–38. doi: 10.1080/01626620.1998.10462889

Hattie J and Timperley H (2007) The power of feedback. Rev Educ Res 77:81–112. doi: 10.3102/003465430298487

Hembree R (1988) Correlates, causes, effects, and treatment of test anxiety. Rev Educ Res 58:47–77. doi: 10.2307/1170348

Hogan R, Hogan J and Roberts BW (1996) Personality measurement and employment decisions: Questions and answers. Am Psychol 51:469–477. doi: 10.1037/0003-066X.51.5.469

Honicke T and Broadbent J (2016) The influence of academic self-efficacy on academic performance: A systematic review. Educ Res Rev 17:63–84. doi: 10.1016/j.edurev.2015.11.002

Hopper M (2018) Alphabet Soup of Active Learning: Comparison of PBL, CBL, and TBL. HAPS Educ 22:144–149. doi: 10.21692/haps.2018.019

Hrbáčková K, Hladík J and Vávrová S (2012) The Relationship Between Locus of Control, Metacognition, and Academic Success. Procedia - Soc Behav Sci 69:1805–1811. doi: /10.1016/j.sbspro.2012.12.130

Jansen M, Scherer R and Schroeders U (2015) Students' self-concept and self-efficacy in the sciences: Differential relations to antecedents and educational outcomes. Contemp Educ Psychol 41:13–24. doi: 10.1016/j.cedpsych.2014.11.002

Jensen M (2016) Curriculum for the active learning classroom - the POGIL method. HAPS Educ 20:93–97. doi: doi: 10.21692/haps.2016.020

continued on next page

Johnson M (2009) Community college students' perceptions of stress. J Biol Exerc 5:1–14. doi: 10.4127/jbe.2009.0022

Klatte M, Bergström K and Lachmann T (2013) Does noise affect learning? A short review on noise effects on cognitive performance in children. Front Psychol. doi: 10.3389/fpsyg.2013.00578

Komarraju M and Nadler D (2013) Self-efficacy and academic achievement: Why do implicit beliefs, goals, and effort regulation matter? Learn Individ Differ 25:67–72. doi: 10.1016/j.lindif.2013.01.005

Kumaraswamy N (2013) Academic stress, anxiety and depression among college students- A brief review. Int Rev SocSciHum 5:1–9.

Lent RW, Brown SD and Larkin KC (19871101) Comparison of three theoretically derived variables in predicting career and academic behavior: Self-efficacy, interest congruence, and consequence thinking. J Couns Psychol 34:293. doi: 10.1037/0022-0167.34.3.293

Locke EA and Latham GP (2002) Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. Am Psychol 57:705–717. doi: 10.1037//0003-066X.57.9.705

Lorenzo M, Crouch CH and Mazur E (2006) Reducing the gender gap in the physics classroom. Am J Phys 74:118–122. doi: 10.1119/1.2162549

Lowe PA, Lee SW, Witteborg KM, Prichard KW, Luhr ME, Cullinan CM, Mildren BA, Raad JM, Cornelius RA and Janik M (2008) The test anxiety inventory for children and adolescents (TAICA): Examination of the psychometric properties of a new multidimensional measure of test anxiety among elementary and secondary school students. J Psychoeduc Assess 26:215–230. doi: 10.1177/0734282907303760

Mallow JV (2006) Science Anxiety: Research and Action. Handbook of College Science Teaching: 3-14.

Mallow JV (1994) Gender-related science anxiety: a first binational study. J Sci Educ Technol 3:227–238. doi: 10.1007/BF01575898

Margolis H and Mccabe PP (2006) Improving self-efficacy and motivation: What to Do, what to Say. Interv Sch Clin 41:218–227. doi: 10.1177/10534512060410040401

Mccroskey JC (2009) Communication apprehension: What have we learned in the last four decades. Hum Commun 12:157–171.

McIlroy D, Bunting B and Adamson G (2000) An evaluation of the factor structure and predictive utility of a test anxiety scale with reference to students' past performance and personality indices. Br J Educ Psychol 70:17–32. doi: 10.1348/000709900157949

Mowbray T (2012) Working memory, test anxiety and effective interventions: A review. Aust Educ Dev Psychol 29:141–156. doi: 10.1017/edp.2012.16

Naveh-Benjamin M, McKeachie W and Lin Y (1987) Two types of test-anxious students: Support for an information processing model. 79:131–136.

Naveh-Benjamin M, McKeachie WJ and Lin Y (1981) Test anxiety: Deficits in information processing. 73:816–824.

Noftle EE and Robins RW (2007) Personality predictors of academic outcomes: big five correlates of GPA and SAT scores. J Pers Soc Psychol 93:116–130. doi: 10.1037/0022-3514.93.1.116

Norem JK and Cantor N (1986) Defensive pessimism: harnessing anxiety as motivation. J Pers Soc Psychol 51:1208–1217.

Pajares F (1996) Self-efficacy beliefs in academic settings. Rev Educ Res 66:543–578. doi: 10.3102/00346543066004543

Panadero E and Jonsson A (2013) The use of scoring rubrics for formative assessment purposes revisited: A review. Educ Res Rev 9:129–144. doi: 10.1016/j.edurev.2013.01.002

Pitt A, Oprescu F, Tapia G and Gray M (2017) An exploratory study of students' weekly stress levels and sources of stress during the semester. Act Learn High Educ 36:146978741773119–15. doi: 10.1177/1469787417731194

Poropat AE (2009) A meta-analysis of the five-factor model of personality and academic performance. Psychol Bull 135:322–338. doi: 10.1037/a0014996

Ramirez G and Beilock SL (2011) Writing about testing worries boosts exam performance in the classroom. Science 331:211–213. doi: 10.1126/science.1199427

Raufelder D and Ringeisen T (2016) Self-perceived competence and test Anxiety. J Individ Differ 37:159–167. doi: 10.1027/1614-0001/a000202

continued on next page

Richardson M, Abraham C and Bond R (2012) Psychological correlates of university students' academic performance: A systematic review and meta-analysis. Psychol Bull 138:353–387. doi: 10.1037/a0026838

Rocca KA (2010) Student Participation in the College Classroom: An Extended Multidisciplinary Literature Review. Commun Educ 59:185–213. doi: 10.1080/03634520903505936

Rozendaal J, Minnaert A and Boekaerts M (2001) Motivation and self-regulated learning in secondary vocational education: information-processing type and gender differences. Learn Individ Differ 13:273–289. doi: 10.1016/S1041-6080(03)00016-5

Sarason IG (1984) Stress, anxiety, and cognitive interference: reactions to tests. J Pers Soc Psychol 46:929–938.

Sarason IG, Sarason BR, Keefe DE, Hayes BE and Shearin EN (1986) Cognitive interference: Situational determinants and traitlike characteristics. J Pers Soc Psychol 51:215–226. doi: 10.1037/0022-3514.51.1.215

Schell JA and Butler AC (2018) Insights From the science of learning can inform evidence-based implementation of peer instruction. Front Educ. doi: 10.3389/feduc.2018.00033

Schraw G, Crippen KJ and Hartley K (2006) Promoting self-regulation in science education: Metacognition as part of a broader perspective on learning. Res Sci Educ 36:111–139. doi: 10.1007/s11165-005-3917-8

Schutz PA and Davis HA (2000) Emotions and self-regulation during test taking. Educ Psychol 35:243–256. doi: 10.1207/S15326985EP3504_03

Snow RE, Corno L and Jackson III D (1996) Individual differences in affective and conative functions. Handb Educ Psychol 243–310.

Spielberger C (1966) Theory and research on anxiety. Anxiety and Behavior. Academic Press, San Diego, CA, pp 3–20.

Stan A and Oprea C (2015) Test Anxiety and Achievement Goal Orientations of Students at a Romanian University. Procedia - Soc Behav Sci 180:1673–1679. doi: 10.1016/j.sbspro.2015.05.066

Stone NJ (2000) Exploring the relationship between calibration and self-regulated learning. Educ Psychol Rev 12:39. doi: 10.1023/A:1009084430926

Talsma K, Schüz B, Schwarzer R and Norris K (2018) I believe, therefore I achieve (and vice versa): A meta-analytic cross-lagged panel analysis of self-efficacy and academic performance. Learn Individ Differ 61:136–150. doi: 10.1016/j.lindif.2017.11.015

Tanner KD (2012) Promoting student metacognition. CBE—Life Sci Educ 11:113–120. doi: 10.1187/cbe.12-03-0033

Tanner KD (2013) Structure matters: Twenty-one yeaching dtrategies to promote student engagement and cultivate classroom equity. CBE—Life Sci Educ 12:322–331. doi: 10.1187/cbe.13-06-0115

Teigen KH (1994) Yerkes-Dodson: A law for all seasons. Theory Psychol 4:525–547. doi: 10.1177/0959354394044004

Udo MK, Ramsey GP and Mallow JV (2004) Science anxiety and gender in students taking general education science courses. J Sci Educ Technol 13:435–446. doi: 10.1007/s10956-004-1465-z

Usher EL and Pajares F (2009) Sources of self-efficacy in mathematics: A validation study. Contemp Educ Psychol 34:89–101. doi: 10.1016/j.cedpsych.2008.09.002

Wong JGWS, Cheung EPT, Chan KKC, Ma KKM and Wa Tang S (2006) Web-based survey of depression, anxiety and stress in first-year tertiary education students in Hong Kong. Aust N Z J Psychiatry 40:777–782. doi: 10.1080/j.1440-1614.2006.01883.x

Zeidner M (1998) Test anxiety: The state of the art. Plenum Press, New York, NY, US

Zeidner M (1995) Coping with examination stress: Resources, strategies, outcomes. Anxiety Stress Coping 8:279–298. doi: 10.1080/10615809508249379

Zeidner M and Mendel G (2007) Test anxiety in educational contexts: Concepts, findings, and future directions. Emot Educ 165–184. doi: 10.1016/B978-012372545-5/50011-3

Zimmerman BJ, Bandura A and Martinez-Pons M (1992) Self-motivation for academic attainment: The role of self-efficacy beliefs and personal goal setting. Am Educ Res J 29:663–676. doi: 10.2307/1163261

Back to TOC

Anorexia Nervosa: Pathophysiology, Treatment, and Genetic Considerations

Sarah Cooper, MEd[1], Alicia Marrone[2], Randy Tammara, PharmD[3], Jennifer Wood, PhD[4]
[1]Arcadia University, Biology Department, 450 S. Easton Road, Glenside, PA 19038
[2]Arcadia University, BA, Biology '18
[3]Arcadia University, Masters in Public Health Program, 450 S Easton Road, Glenside, PA 19038
[4]University of Maryland University College and Certified Integrative Health Coach, Life Space Health, LLC.
coopers@arcadia.edu, amarrone@arcadia.edu, rantam111@gmail.com, jen.j.wood@gmail.com

Abstract
Anorexia is an eating disorder that is associated with self-imposed starvation, negative body image, stress, anxiety, and depression. When body weight drops below 60% of an individual's recommended normal weight, all body systems are affected. A growing body of research suggests that individuals with anorexia may have an underlying genetic predisposition for developing the disease. Genome-Wide Association Studies (GWAS) have identified three specific genes have been correlated with the development of anorexia: a variant of the *Brain Derived Neurological Factor* (*BDNF*), *Transcription factor AP-2ß* (*TFAP2B*), and *Potassium Channel Tetramerization Domain Containing 15 gene* (*KCTD15*). This article examines the pathophysiology and treatment of anorexia and the possibility that a genetic predisposition for developing anorexia may underlie the psychological symptoms that typically characterize the disease.
https://doi.org/10.21692/haps.2018.028

Key words: anorexia nervosa**,** body mass index, *BDNF, TFAP2B*, *KCTD15*

Introduction
Anorexia Nervosa (AN) is a disease that is characterized by an extremely low body mass index, which is self-induced by caloric restriction and high intensity exercise (Gorwood *et al.* 2016). Although anyone can develop AN, it is more common in females and among teenagers and rarely diagnosed in people over the age of 40 (Madra and Zeltser 2016, Mayo Clinic (a) 2018). Anorexia affects 0.3 to 3.0% of women worldwide and in Western societies it is the third most common disease affecting adolescent girls. The disease is associated with a suicide rate that is 56 times greater than that which would be expected in a healthy population of comparable age and sex (Miller 2011).

There are two distinct subtypes of anorexia; one is characterized as a restrictive disorder and the other is a binge/purge disorder (Miller 2011). Individuals who have restrictive anorexia are typically described as being highly self-disciplined, which gives them the determination and self-control to carry out a type of self-starvation. Those who are classified with binge/purge anorexia typically purge after eating to alleviate the fear of gaining weight (Miller 2011). Both types of anorexia have similar symptoms, including the irrational fear of weight gain, distorted body image, failure to maintain body weight greater than 85% of their ideal weight, and abnormal eating habits (Miller 2011, Wang *et al.* 2011).

The binge/purge subtype of anorexia may be distinguished from bulimia in that bulimics are not typically underweight and are not troubled with adverse body image thoughts, the urge to exercise excessively, or the compulsion to greatly reduce calorie intake. Bulimics eat large quantities of food and purge, typically with vomiting and laxatives, in order to maintain normal, or nearly normal weight (Bech-Duus 2016). Fifty percent of people with anorexia become bulimics over time (Miller 2011).

Anorexia has historically been characterized as a psychological disorder since factors including negative body image, the influence of the thinness culture in the media, and peer pressure, can contribute to the onset of the condition (Miller 2011, Wang *et al.* 2011). However, in a position paper published in 2008, the Academy for Eating Disorders made the case for acceptance of a biological basis for anorexia and bulimia and the classification of these disorders and their variants as serious mental illnesses (Klump *et al.* 2008). The position paper advocated for this classification because of medical and scientific evidence that suggests that eating disorders have a heritable basis, possess the ability to influence brain function, are known to impair cognitive function and emotional stability, and can negatively impact the quality of life of those afflicted with these disorders (Klump *et al.* 2008).

continued on next page

Anorexia is currently regarded as a psychiatric disorder along with Alzheimer's disease, attention-deficit hyperactivity disorder, alcohol dependence, autism spectrum disorder, bipolar disorder, major depressive disorder, nicotine dependence and schizophrenia (Sullivan *et al.* 2012). Psychiatric disorders are among the most difficult conditions that modern medicine attempts to treat. Collectively, they are characterized by high morbidity, high mortality rates, and heavy costs to society and the affected individuals and their families (Miller 2011, Sullivan *et al.* 2012).

Little is known about the etiology of anorexia and in the last decade genetics has become a major area of investigation. Interest in a genetic basis for anorexia is supported by the observation that people who have a genetic tendency towards perseverance, perfectionism, and sensitivity are at greater risk for developing AN (Mayo Clinic (a) 2018, Sullivan *et al.* 2012). This article describes the pathophysiology and treatment for anorexia and the emerging body of research into the genetic basis for this disease.

Pathophysiology of Anorexia

In patients with anorexia, the body is deprived of adequate levels of macromolecules and essential nutrients and enters a state of starvation. When body weight drops below 60% of an individual's recommended normal weight, all body systems are affected (Casper 1986).

Gastrointestinal System

In the gastrointestinal system, there is a marked decrease in stomach mobility, delayed stomach emptying, and atrophy of the stomach due to the persistent reduced calorie load of a starvation diet, which may consist of as few as 200 calories per day (Rudnick 2014). Constipation, bloating and nausea are common upper digestive complaints that correlate with delayed gastric emptying (Mayo Clinic (a) 2018, Rigaud *et al.* 1988). Riguad (1988) found that delayed gastric emptying can sometimes return to normal without pharmacological intervention if psychological assistance is accepted by the patient and malnutrition is reversed.

Cardiovascular System

The cardiovascular system may be especially hard hit by anorexia and cardiovascular complications are very common. Complications often present as chest pain, which can be secondary to gastroesophageal reflux, muscle strain from vomiting, or anxiety (Rudnick 2014). In cases where chest pain is related to a cardiac event, the problem is most often associated with the loss of cardiac muscle mass that accompanies severe weight loss. Mitral valve prolapse, accompanied by sharp pain under the sternum, is an example of a benign complication of anorexia that is related to a loss of cardiac muscle mass; it usually improves on its own when weight gain is resumed (Casper 1986, Rudnick 2014, Schocken *et al. 1989*).

Heart failure, a very serious complication of anorexia, presents with chest pain and a weakened, failing heart may result in congestive heart failure. Heart failure may also occur as a result of low levels of serum phosphorus in the refeeding syndrome. (The refeeding syndrome, a risk factor for recovering anorexics, is described in this article under Complications.) Other symptoms of heart failure may include fatigue, shortness of breath, edema, and tachycardia (Rudnick 2014).

Fifty percent of deaths from anorexia are sudden cardiac deaths that are brought about as a result of cardiac arrhythmias (Rudnick 2014). The electrocardiogram of patients with anorexia typically shows a prolonged QT interval, which predisposes them to a type of arrhythmia known as *torsades de pointes,* a form of polymorphic ventricular tachycardia characterized by rapid QRS complexes (Mitchell BL 2018, Rudnick 2014). *Torsades de pointes* may generate ventricular fibrillation, which is a deadly condition. A prolonged QT interval can be the result of severe weight loss or very fast weight loss accompanied by electrolyte imbalances such as low potassium and low magnesium. The variations in QT intervals in anorexic patients are believed to be associated with myofibrillar degeneration, which is a common finding in anorexia patients (Rudnick 2014).

Cardiomyopathy, which literally means "sick heart", is another cardiovascular complication of anorexia. It is caused by the effects of starvation or the long-term use of emetics to induce vomiting (Rudnick 2014). Chronic alcohol use, a comorbidity with anorexia, can also cause cardiomyopathy. Cardiomyopathy leads to heart failure and life threatening arrhythmias.

Other abnormalities in the electrocardiograms of patients with anorexia may include low voltage, sinus bradycardia, T-wave flattening, and S-T segment elevation. All may be attributed to the effects of severe weight loss and myofibrillar degeneration (Casper 1986, Schocken *et al.* 1989). Bradycardia and blood pressure less that 90/50 are common in anorexia when starvation has decreased the body mass to less than 80% of the ideal body weight. Bradycardia is caused by an overactive parasympathetic nervous system as the body attempts to conserve energy. Low blood pressure (hypotension) is most often due to a combination of weakened cardiac muscle and dehydration resulting from frequent vomiting (Casper 1986, Rudnick 2014, Schocken *et al.* 1989).

Blood pressure in patients with anorexia is typically very low, below 100/50 mmHg. The heart rate is slow, often less than 55 beats per minute (Casper 1986, Schocken *et al.*

continued on next page

1989). In advanced cases of anorexia, there may be evidence of decreased left ventricular mass, which is associated with abnormal systolic functioning (Casper 1986, Schocken *et al. 1989*). Cardiac events, including life-threatening arrhythmias, are most common during the first two weeks of refeeding in the treatment stage of anorexia (see below) (Schocken *et al.* 1989).

Reproductive System

In the reproductive system, abnormalities in menstruation are associated with disturbed secretion of gonadotropin-releasing hormone (GnRH) (Usdan 2008). Extreme calorie reduction decreases the release of GnRH, which causes abnormalities in the release of lutenizing hormone (LH) and follicle stimulating hormone (FSH). LH and FSH revert to an immature secretion pattern that causes an increased FSH to LH ratio and decreased frequency of LH release. Consequently, the amount of gonadotropin stimulation on the ovary is inadequate to cause ovulation (Miller 2011, Usdan 2008). As a result of these hormonal imbalances, levels of estradiol are typically low in anorexia, and progesterone may fall to such low levels that it cannot be measured (Miller 2011). Amenorrhea in anorexia is a protective adaptation to prevent pregnancy during a time of starvation. Following weight gain that reaches 90% of the ideal weight gain for a person's height, menses usually resume within one year (Usdan 2008).

Thyroid Metabolism

Patients with anorexia often show signs of hypothyroidism, which may include hypothermia, bradycardia, dry skin, low blood pressure, a decrease in overall metabolism and low cholesterol levels (Usdan 208). There are many factors that can influence the altered thyroid hormone levels that are seen in anorexia. Anorexia patients commonly have decreased peripheral deiodination of T_4 (thyroxin) to T_3 (Triiodothyronine) and increased conversion to reverse T_3 (Casper 1986, Miller 2011, Usdan 2008). Carbohydrates, which are often at low levels in anorexia, are important in stimulating the peripheral conversion of T_4 to active T_3. The release of thyrotropin-releasing hormone from the hypothalamus may be decreased in anorexia, preventing a thyroid-stimulating hormone response to low peripheral thyroid hormone levels (Usdan 2008). Low insulin-like growth factor levels secondary to malnutrition may cause atrophy of the thyroid gland, which exacerbates on-going starvation (Usdan 2008). Fortunately, the biochemical problems associated with anorexia often correct themselves without treatment when weight gain is accomplished (Casper 1986, Miller 2011, Usban 2008).

Skeletal System

The health of bone is compromised early in the progression of anorexia. Peak bone mass is normally accumulated during adolescence. Over 50% of anorexic adolescents have osteopenia, while 25% have osteoporosis. These conditions prevent the normal accumulation of bone mass in adolescence (Usdan 2008). In adult females with anorexia, 90% suffer from reduced bone density and 38% have osterporosis (Usdan 2008). The damage to bone is greater in those who have the binge/purge type of anorexia compared to restrictive anorexia. Even with improved eating, osteopenia may persist for up to ten years after the initial diagnosis, along with an increased risk of bone fracture (Usdan 2008).

For reasons that are not yet understood, trabecular bone appears to be more affected by anorexia than compact bone. It also seems to recover more rapidly from anorexia-induced damage. Decreased bone density is attributed to estrogen deficiency and reduced bone size is attributed to the cumulative effects of malnutrition (Usdan 2008). Both of these conditions are hallmarks of anorexia.

Abnormalities in osteoblast activity and severe nutritional deficiency account for most of the abnormal bone metabolism seen in anorexia and the generalized decrease in bone remodeling processes that characterizes the disease (Usdan 2008). Low levels of progesterone, which is needed for bone remodeling, lead to a decrease in the production of insulin-like growth factor-1 (IGF-1) and the biosynthesis of type 1-collagen declines (Casper 1986, Miller 2011). Excessive exercise perpetuates hyopthalmic amenorrhea, which is also associated with bone loss. Once body weight improves to normal levels, bone density may increase in as little as three months (Casper 1986, Miller 2011, Usdan 2008).

Other conditions associated with anorexia

Other conditions associated with anorexia include generalized muscle weakness, which is the most common neurological symptom. The hair may be brittle and as the disease progresses, fine body hair known as lanugo may cover the body (Casper 1986, Miller 2011).

Stress and anxiety are also thought to play a major role in anorexia and it has been proposed that restrictive eating is a reward that helps alleviate anxiety in anorexia patients. Anxiety is often sustained in anorexia patients because of a fear of weight gain and negative social pressure (Lloyd *et al.* 2018).

Potassium is usually the most affected electrolyte in anorexia but disturbances in calcium, magnesium and phosphorus are also common. Electrolyte disturbances are attributed to frequent vomiting, dehydration, and loss of muscle mass (Casper 1986, Miller 2011).

Anorexia is a life-threatening disease

Severe anorexia is a life-threatening illness that may result in death due to prolonged calorie restriction. Anxiety, stress, and depression may contribute to fatalities and depression may lead to suicidal thoughts and actions. Thornton *et al.* (2016) reported that individuals with anorexia have a higher incidence of Major Depression Disorder that correlates with a higher risk of suicide when compared to females in a similar

continued on next page

age group of 15 to 34 years old who did not have anorexia. Fatalities associated with anorexia are most frequently linked to cardiac arrhythmias, electrolyte imbalances, and kidney damage (Mayo Clinic (a) 2018).

Complications

Refeeding syndrome

Restoration of normal body weight is the most important treatment goal for anorexia but it carries with it the risk of developing refeeding syndrome. Slow, careful reintroduction of nutrients is called for during recovery from anorexia in order to maintain normal electrolyte levels and avoid drastic shifts in body fluids. Overall, the reintroduction of proteins and fats is less problematic than the reintroduction of carbohydrates. Anorexia patients have low glucose and insulin levels and high glucagon levels throughout the disease and abnormalities in glucose metabolism persist into recovery (Usdan 2008). Typically there is a reduced insulin response to the intake of calories and a metabolic preference to metabolize glucose over fats. Cholesterol levels are often elevated and should be monitored (Usdan 2008).

During recovery, carbohydrate ingestion stimulates the release of insulin and the reintroduction of insulin into a system that has been accustomed to starvation can stimulate a decrease in the excretion of sodium in the urine (antinatriuresis) secondary to greater reabsorption of sodium from the distal convoluted tubules (DeFronzo 1981, Usdan 2008). The reintroduction of insulin can also cause an influx of electrolytes into the intracellular fluid, particularly potassium, magnesium, and phosphorus. Sodium retention caused by an increase in insulin levels can lead to hypervolemia, which is followed by cardiac arrest and respiratory failure (Usdan 2008). Increased amounts of potassium, magnesium, and phosphorus can lead to cardiac arrhythmias, muscle weakness, and respiratory failure (Mayr *et al.* 2015, Rohrer and Dietrick 2014).

Superior Mesenteric Artery Syndrome

Excessive weight loss may put people at risk for Superior Mesenteric Artery Syndrome (SMAS), a rare gastrointestinal obstruction in which the distal portion of the duodenum is subjected to vascular pressure as it is compressed between the aorta posteriorly and the superior mesenteric artery anteriorly (Casper 1986, Ugras *et al.* 2017). SMAS typically leads to vomiting and abdominal pain secondary to a partial or complete blockage of the duodenum. Most cases can be treated medically but approximately 30% of cases require surgery to relieve the pressure (Casper 1986, Ugras *et al.* 2017).

Pharmacological Treatments

Ideally, pharmacological options for the treatment of anorexia should bring about a remission of the acute symptoms of anorexia, prevent the recurrence of symptoms over time, and be able to treat the most common psychiatric manifestations of the disease. Currently, there are no drugs that are capable of acting on all of these conditions. Fluoxetine (Prozac), a selective serotonin reuptake inhibitor (SSRI), used to treat depression, panic attacks, and obsessive-compulsive disorder, is the only drug that has been approved by the FDA for treating eating disorders (Milano *et al.* 2013, medlineplus 2018). Fluoxetine is approved specifically for the treatment of bulimia nervosa (BN).

Antipsychotic (AP) drugs, major tranquilizers such as Risperdal and Clozaril, which focus on the interaction between dopamine and serotonin, are used cautiously in the treatment of anorexia because they have been associated with weight gain in patients suffering from schizophrenia and bipolar disorder (Milano *et al.* 2013). AP drugs are not associated with weight gain in patients who have anorexia but, for many patients, they seem to reduce the degree of some of the primary symptoms of the disease such as poor body image, fear of weight gain, and obsessive-compulsive tendencies (Milano *et al.* 2013, Aigner *et al.* 2011).

Antidepressant drugs (AD), such as Zoloft, Paxil and Prosac, may be helpful in managing depression as a comorbidity of anorexia (Milano *et al.* 2013, Aigner *et al.* 2011). AD drugs may also be effective in some patients in balancing levels of norepinephrine and serotonin in the early stages of the disease (Milano *et al.* 2013, Aigner *et al.* 2011>

The lack of definitive dose standardization for eating disorders, the presence of frequent psychiatric comorbidities (such as major depression, obsessive-compulsive disorders, and alcohol dependence) and medical comorbidities (such as osteoporosis and infertility) in anorexia patients, and the possible side effects associated with AP drugs, have prevented pharmacotherapy from being the treatment of choice for eating disorders (Milano *et al.* 2013, Aigner *et al. 2011*). Patients with anorexia are considered to be fragile candidates for drug treatment since ongoing malnutrition can influence the anticipated pharmacokinetics and pharmacodynamics of the treatment molecules, which may expose patients to unanticipated, individualized side effects. Consequently, psychotherapies, such as cognitive-behavioral therapy and family therapy remain the mainstay of treatment for eating disorders (Milano *et al.* 2013, Aigner *et al.* 2011).

Mindfulness as a Supplemental Treatment

Mindfulness, "the awareness that arises from paying attention, on purpose, to the present moment, without judgment" (Kabat-Zinn 2017), may be considered a supplemental treatment to support more traditional protocols for treating anorexia (Beccia *et al.* 2018, Butryn *et al.* 2013, Lattimore *et al.* 2017). Along with psychotherapy, nutritional counseling, and medication, on-going support is often needed for people with anorexia, especially during periods of high stress or when an individual's unique triggers are present, since the likelihood for relapse is considered to be high (Mayo Clinic (b) 2018). Individuals who suffer from anorexia may be particularly vulnerable to daily stressors and may have trouble

continued on next page

short-circuiting their typical coping mechanism, which is to deny their body much needed food and nourishment (Gleissner 2017). Providing patients with a protocol for managing stressful thoughts and impulses, which is the role of mindfulness, can help bridge the time between therapy sessions and help the patient find ways to cope with their condition and build resilience to setbacks.

Mindfulness can help to ground an individual in the present and provide the opportunity to become more familiar with the thoughts and feelings that give rise to the behaviors that lead to anorexia (Kabat-Zinn 2017). Once an individual arrives at a place of awareness of these thoughts and feelings, a pause for reflection may be possible (Kabat-Zinn 2017). It is this pause that can create an opportunity for a patient suffering from anorexia to make a change. The pause may reveal, for example, an intensely self-critical inner voice or the recognition of unrealized pain or emotional needs. Once a patient is aware of what is triggering the need to deprive the body of food, there is an opportunity to decide between reacting to this internal condition in a way that perpetuates the current status (anorexia) or responding in a way that is more healthy and nurturing (Kabat-Zinn 2017).

Research suggests that patients with an eating disorder who are trained in mindfulness experience greater post-intervention self-esteem and improved treatment outcomes (Beccia *et al.* 2018, Butryn *et al.* 2013, Lattimore *et al.* 2017). This might result in a greater awareness of the protective factors that are associated with eating disorders. For example, having greater self-esteem might help patients pursue a more positive body image. It might also help reduce some of the known risk factors for eating disorders, such as feelings of inadequacy, pessimism, self-doubt and cynicism (Beccia *et al.* 2018, Butryn *et al.* 2013).

The Genetic Considerations Associated with AN

Rask-Andersen *et al.* (2010) have proposed that controlled Genome-Wide Association Studies (GWAS) can be used to determine if there is a genetic predisposition to developing anorexia. For example, genes that control the way one interprets different tastes and flavors may play a role in anorexia; influencing the foods that anorexia patients eat as well as the diet they restrict themselves to. It is possible that mothers who have anorexia while they are pregnant may expose their children to abnormal feeding behavior *in utero*. As a consequence, these babies may be born with a low body mass index, possibly making them susceptible to getting anorexia in later life (Gorwood *et al.* 2016).

There is growing evidence that suggests that there is a genetic component to the development of anorexia. Wang *et al.* (2011) conducted GWAS to identify specific genes, single-nucleotide polymorphisms (SNPs), and genotypes associated with anorexia. Wang *et al.* (2011) found that there is significant association between genes such as HTR1D and the subtype known as restrictive anorexia (2011). However, this study did not find associations with these genes and binge/purge anorexia. SNPs within the HTR1D gene were also found to be associated with restrictive anorexia. A second gene, and its associated SNPs within the OPRD1 locus, was found to be associated with a general condition of anorexia (Wang *et al.* 2011).

Boraska *et al*. (2014) conducted GWAS that identified SNPs associated with anorexia and other diseases including Alzheimer's disease and Schizophrenia. Several anorexia associated SNPs were identified through these studies but they were not useful in distinguishing between restrictive anorexia and the binge/purge subtype. Further analysis of anorexia-associated SNPs is needed (Boraska *et al.* 2014).

Candidate genes

GWAS have identified specific candidate genes that are correlated with anorexia and are believed to support the biological systems that maintain the anorexia phenotype. Candidate genes include specific genes in the reward system and genes that govern mood and appetite (Rask-Andersen *et al.* 2010).

Five candidate genes that have a strong association with anorexia regulate theprocesses that govern mood, appetite, and reward. These genes are BDNF, SK3, COMT, AGRP, and OPRD1 (Rask-Andersen *et al.* 2010), which have the following associations:

> BDNF encodes for brain-derived neurotropic factor that promotes the survival of neurons by promoting their growth and differentiation.
>
> The SK3 gene is associated with the relaxation of smooth muscle tissue. Mutations of SK3 may play a role in neurological disorders such as anorexia, Alzheimer's and schizophrenia.
>
> COMT endcodes for a gene that produces an enzyme that helps control the levels of certain hormones and parts of the prefrontal cortex of the brain linked to inhibition of behaviors and emotion.
>
> AGRP encodes for a receptor antagonist that regulates hypothalamic control of feeding behavior and thus plays a role in weight homeostasis. Mutations of this gene may be associated with obesity.
>
> OPRD1 plays a role in the perception of pain and developing tolerance to morphine. (Genomics Home Reference 2018)

continued on next page

The current research examined in this article centers on a gene that is strongly associated with aneroxia (BDNF), for its effect in the CNS, a weakly associated gene (KCTD15), for its association with obesity, and a third gene known as (TFAP2B), for its effect on endocrine glands and the neural crest (Genomics Home Reference 2018, Rask-Andersen *et al.* 2010).

Variants of brain-derived neurotrophic factor (BDNF) are known to induce stress and anxiety in humans (Madra and Zeltser 2016). In anorexia patients, this factor increases the predisposition to maintain anorexia and increases the likelihood of relapse after undergoing rehabilitation. A specific variant BDNF, Val66Met, increases the likelihood of anxiety in young people, which may increase the prevalence of anorexia in susceptible individuals (Madra and Zeltser 2016). Studies of animal models have been conducted with genotypes of this variant to determine how severe the genetic predisposition is to AN with this gene (Madra and Zeltser 2016).

Until recently, genes KCTD15 and TFAP2B have been investigated in obese individuals, but not in those with eating disorders. Recent GWAS have confirmed that KCTD15 has SNPs that are responsible for regulating BMI (Albuquerque *et al.* 2017). Similar studies have found that TFAP2B is associated on its own with binge-eating disorder. TFAP2B and KCTD15 together have homologues that control food regulation and reward behaviors (Albuquerque *et al.* 2017). These genes have all been identified by GWAS as possibly being important in the etiology of anorexia.

Recent Research Focused on Genes Associated With AN

Studies of *Transcription factor AP-2ß* (*TFAP2B*), *Potassium Channel Tetramerization Domain Containing 15 gene* (*KCTD15*) and *Brain Derived Neurological Factor* (*BDNF*) have added to the body of knowledge that supports a genetic predisposition for developing anorexia.

Study of TFAP2B and KCTD15

Alburquerque *et al*. (2017) conducted a study to verify the link between Transcription Factor AP-2 Beta (TFAP2B), Potassium Channel Tetramerization Domain Containing (KCTD15), and eating disorders. The study included genotyping human blood samples from 425 participants and determining SNP selections of specific genes. Psychiatrists and psychologists adhering to strict guidelines had previously diagnosed all of the participants. One hundred and six participants had been diagnosed with anorexia, 63 had been diagnosed with bulimia nervosa (BN), 181 participants were classified as healthy, and 75 participants were classified as being obese. All of the study participants were white Spanish women (Alburquerque *et a*l. 2017).

In this study, the TFAP2B and KCTD15 genes were examined to determine their role in the development of anorexia, bulimia, and obesity. All patients were tested for body mass index (BMI). The patients with anorexia had a significantly lower BMI than healthy patients, while bulimic patients had a lower BMI than those struggling with obesity (Alburquerque *et al.* 2017). The highest BMI to the lowest BMI in terms of patient-condition were as follows: obese, BN, healthy, and anorexia respectively. Along with BMI, the study identified specific SNPs within the genes of interest that corresponded to an increased risk of developing an eating disorder. A specific T variant allele in KCTD15 (rs287103) corresponded to BN risk. In this study, a significant 63.5% of the participants were carriers for the allele. The T variant allele in TFAP2B (rs760900) was found to predispose individuals to anorexia, and a low BMI (Alburquerque *et al.* 2017).

Alleles as well as haplotypes of these genes were analyzed to establish a potential genetic risk factor for anorexia. KCTD15 haplotype *6 was demonstrated to be associated with healthy patients (Alburquerque *et al.* 2017). It was also found to be more prevalent in anorexia patients than in obese patients. Haplotype *2 of the TFAP2B gene was found to be strongly correlated with anorexia patients and their need and motivation to be thin. Haplotypes *3 and *4 were strongly associated with BN (Alburquerque *et al.* 2017). The epistasis of TFAP2B and KCTD15 carried a risk for developing anorexia that was dependent on the age of the patient. SNPs and gene-gene interaction analyses demonstrated genetic components that corresponded with personality traits of those with AN. For example, anorexia patients experience moods of perfectionism, maturity fears, or social insecurity that are affected by the gene-gene interactions (Alburquerque *et al.* 2017).

Study of BDNF gene variant, Val66Met

Madra and Zeltser (2016) studied the BDNF gene variant, Val66Met, in a mouse model. Mice resulting from intercrosses of *h*BDNF$^{Val/Met}$ were kept in a controlled environment with access to a diet of chow and water. The mice were then randomly selected for various housing arrangements to test for specific environmental interactions. Young mice were separated into groups of three to five mice or housed separately, one mouse per cage (Madra and Zetlser 2016). Animals were given either a) standard feeding of diet and water or b) a 20-30% calorie restrictive diet, presented twice daily. For the restricted group, calorie restriction occurred for 10-11 days beginning at seven weeks of age. Animals were monitored a few times per week for weight gain, weight loss, and anxious behavior (Madra and Zetlser 2016).

PCR was performed on mouse DNA from hypothalamic and pituitary samples to genotype the BDNF locus and other gene expressions. Mice were placed in a room with a photobeam-activity monitoring system, a Feeding Monitor, and TSE ActiMotsystem. These systems measured and assessed central and peripheral locomotor activity and anxiety behavior after a 24-hour period of adjustment (Madra and Zetlser 2016).

continued on next page

Corticosterone levels were measured to analyze stress. Mice that were singly housed, whether calorie restricted or standard diet, were given attention in the early stages of puberty for a short period of time each day during weeks six and seven, while the other animals did not receive any attention.

All groups of mice experienced aphagia episodes (AEs), periods of time when they did not eat. Mice with genetic susceptibility (Val66Met carriers), an environmental stressor (social isolation), and a long duration of caloric restriction experienced the greatest number of AEs (Madra and Zeltser 2016). Some mice continued to experience AEs even after caloric restriction was terminated. Female mice had more AEs than males. The results of this study suggest that genetics and environmental influences such as social isolation, social interaction, group housing or individual housing may interact to affect the body and how it responds to disease such as AN.

Conclusion

Anorexia is a psychiatric disorder that leads to malnutrition. Extreme malnutrition and attempted refeeding can have severe consequences in anorexia. The disease ultimately affects all organs and it can lead to electrolyte imbalances, multi-organ failure, and death. Many of the endocrine abnormalities associated with anorexia constitute protective, energy conserving measures that the body undertakes to endure periods of starvation. Anorexia has far-reaching effects on bone growth, endocrine function, and the metabolism of nutrients. Some abnormalities, such as those that affect the thyroid gland and the reproductive system, typically reverse in the absence of malnutrition; long-lasting consequences of anorexia may include infertility, osteoporosis, short stature, and lingering psychological distress.

Current investigations in genetics are highly suggestive of a correlation between specific haplotypes and the development of anorexia. These studies also demonstrate the relationships between the disease and environmental stressors that can cause the patient's condition to persist.

To date, anorexia is analyzed primarily in the field of psychiatry and recognized in biology primarily in terms of pathophysiology. Therefore, primary biological research is still scarce. Identifying the genes that contribute to anorexia may lead to novel treatment options for those who suffer with this devastating and potentially life-threatening condition.

About the Authors

Sarah Cooper is an Associate Professor of Biology at Arcadia University and Managing Editor of the *HAPS Educator*. She has taught human anatomy and general biology at Arcadia University since 1981. She is the pre-nursing adviser and the coordinator of the interdisciplinary science courses at Arcadia and she has served as a member of the Arcadia University Judicial Board since 1984.

Alicia Marrone graduated from Arcadia University with a BA in Biology in the spring of 2018. She will continue her education and is enrolled in an accelerated nursing program at Gwynedd Mercy University for the spring of 2019. Upon completion of the BSN program, she plans on becoming a Certified Registered Nurse Practitioner (CRNP) specializing in Hematology/Oncology.

Randy Tammara is a pharmacist who has been working in the Philadelphia area for over twenty-five years. He is also a certified diabetes educator who gives presentations to caregivers through Northhampton Community College in Bethlehem, PA, where he is an adjunct faculty member in the Community Education Division. He is currently enrolled in the Masters in Public Health program at Arcadia University.

Jennifer Wood, PhD, teaches non-profit management at University of Maryland University College where she is an Associate Professor. She is a certified integrative health coach with a family practice, Life Space Health, LLC, specializing in stress management and anxiety reduction.

Literature cited

Aigner M, Treasure J, Kaye W, Kasper S, and The WFSBP task force on eating disorders (2011) World Federation of Societies of Biological Psychiatry (WFSBP) Guidelines for the Pharmacological Treatment of Eating Disorders. *The World Journal of Biological Psychiatry* 12:400-443.

Albuquerque D, Carillo JA, Flores I, Gamero-Villarroel C, Garcia-Herraiz A, Gervasini G, Gonzalez LM, Rodriguez-Lopez R (2017) Influence of TFAP2B and KCTD15 genetic variability on personality dimensions in anorexia and bulimia nervosa. *Brain and Behavior.* 7(9): e00784.

Beccia AL, Dunlap C, Hanes D, Courmeene B, Zwickey H (2018) Mindfulness-based eating disorder prevention programs: A systematic review and meta-analysis. *Mental Health & Prevention* 9:1-12. https://doi.org/10.1016/j.mhp.2017.11.001

Bech-Duus, MJI (2016) https://www.quora.com/What-is-the-difference-between-binge-purge-anorexia-and-bulimia. Retrieved November 3, 2018.

continued on next page

Boraska V, Franklin CS, Thorntor LM, Huskins LM, Southam L … Bulik CM (2014) A genome-wide association study of anorexia nervosa. *Molecular Psychiatry* 19(10): 1085-1094.

Butryn ML, Juarascio A, Shaw J, Kerrigan S, Clark V, O'Planick A, Forman E (2013) Mindfulness and its relationship with eating disorders symptomatology in women receiving residential treatment. *Eating Disorders.* 14(1):13016. https://doi.org/10.1016/j.eatbeh.2012.10.005

Casper, RC (1986) The Pathophysiology of Anorexia Nervosa and Bulimia Nervosa. *The Annual Review of Nutrition*. (6)299-316
doi: org/10.1146/annurev.nu.06.070186.001503

DeFronzo RA (1981) The effect of insulin on renal sodium metabolism. A review with clinical implications. *Diabetologia*. 21(3):165-71.

Genomics Home Reference (2018) https://ghr.nlm.nih.gov/gene/
Retrieved November 3, 2018.

Gorwood P, Blanchet-Collet C, Chartrel N, Duclos J, Dechelotte P, Hanachi M, … Epelbaum J (2016) New Insights in Anorexia Nervosa. *Frontiers in Neuroscience*. *10*:256.

Gleissner, Greta (2017) Mindful Techniques in Eating Disorder Recovery. *Eating Disorder Specialists*. http://eatingdisorderspecialists.com/mindfulness-techniques-eating-disorder-recovery/ Accessed September 10, 2018.

Kabat-Zinn, Jon (2017) Defining Mindfulness. *Mindful*: *healthy mind, healthy life.* (https://www.mindful.org/jon-kabat-zinn-defining-mindfulness/). Accessed Sept. 10, 2018

Klump, Kelly L, Bulik CM, Kaye HW, Treasure J, Tyson E (2008) Academy for eating disorders position paper: Eating disorders are serious mental illnesses. *Academy for Eating Disorders Position Paper*. https://doi.org/10.1002/eat.20589

Lattimore P, Mead B, Irwin L, Grice L, Carson R, Malinowski P (2017) 'I can't accept that feeling': Relationships between interoceptive awareness, mindfulness and eating disorder symptoms in females with, and at-risk of an eating disorder. *Psychiatry Research*. 247:163-171. https://doi.org/10.1016/j.psychres.2016.11.022

Lloyd CE, Haase AM, Verplanken B (2018) Anxiety and the development and maintenance of anorexia nervosa: protocol for a systemic review. *Systemic Reviews.* 7(14):1-5.

Madra M and Zeltser L (2016) BDNF-Val66Met variant and adolescent stress interact to promote susceptibility to anorexic behavior in mice. *Translational Psychiatry.* 6(4):e776-e776.

Mayo Clinic (a)
https://www.mayoclinic.org/diseases-conditions/anorexia/symptoms-causes/syc-20353591?p=1 Accessed July 4, 2018.

Mayo Clinic (b)
https://www.mayoclinic.org/diseases-conditions/anorexia-nervosa/diagnosis-treatment/drc-20353597
Accessed September 20, 2018

Mayr M, Imgart H, Skala K, Karwautz A (2015) Quality management in weight restitution in Anorexia nervosa – pathophysiology, evidence-based practice and prevention of the refeeding syndrome. *Neuropsychiatr*. 29(4):200-11. doi: 10.1007/s40211-015-0165-7

Medlineplus (2018) https://medlineplus.gov/druginfo/meds/a689006.html
Accessed September 14, 2018.

Miller, KK (2011) Endocrine dysregulation in anorexia nervosa update. *The Journal of Clinical Endocrinology & Metabolism*. 96(10):2939-2949.
https://doi.org/10.1210/jc.2011-1222

Milano W, DeRosa M, Milano L, Saneverino B, Capasso A (2013) The Pharmacological Options in the Treatment of Eating Disorders. *ISRN Pharmacol.*
2013:352865. doi: 10.1155/2013/352865

Rask-Andersen M, Olszewski PK, Levine AS, Schioth HB (2010) Molecular mechanisms underlying anorexia nervosa: Focus on human gene association studies and systems controlling food intake. *Elsevier.* 62:147-164.

Rigaud D, Bedig G, Merrouche M, Vulpillat M, Bonfils S, Apfelbaum M (1988)
Delayed gastric emptying in anorexia nervosa is improved by completion of a renutrition program. *Digestive Diseases and Sciences*. 33(8):919-925.

Rohrer S and Dietrick JW (2014) The Refeeding Syndrome – A Literature review.*Z Gastroenterol.* 50(6):593-600. doi: 10.1055 / s-0034-1366430
(Translated from German)

Thornton LM, Welch E, Munn-Chernoff MA, Lichtenstein P, Bulik CM (2016) Anorexia Nervosa, Major Depression, and Suicide Attempts: Shared Genetic Factors. *Suicide & Life-Threatening Behavior*. 46(5):525–534.

continued on next page

Rudnick C (2014) Cardiovascular complications of eating disorders. https://www.edcatalogue.com/cardiovascular-complications-eating-disorders/ Retrieved November 3, 2018.

Schocken D, Holloway LD, Powers PS (1989) Weight loss and the heart Effects of anorexia nervosa. *Arch Intern Med,* 149(4):877-881. doi:10.1001/archinte.1989.00390040085017

Sullivan, Partick F, Daly MJ, O'Donovan M (2012) Genetic architectires of psychiatric disorders: the emerging picture and its implications. *Nature Reviews Genetics.* 13: 537-551.

Ugras M, Suat B, Fatma TC, Endi R, Baki E (2017) Superior mesenteric artery syndrome: A rare but life threatening disease. *Turk J Emerg Med*. 17(2):70-72 doi: 10.1016/j.tjem.2017.01.001

Usdan LS, Khaodhiar L, Apovian CM (2008) The endocrinopathies of anorexia nervosa. *Endocr Pract*. 14(8): 1055-1063. doi: [10.4158/EP.14.8.1055]

Wang K, Zhang H, Bloss CS, Duvvuri V, Kaye W, Schork NJ, Berrettini W, Hakonarson H (2011) A genome-wide association study on common SNPs and rare CNVs in anorexia nervosa. *Molecular Psychiatry,* 16:949-959.

Back to TOC

Rethinking How We Talk About and Teach Muscle Fatigue

James E Clark[1,2], Krista Rompolski, PhD[3], Brett A Comstock, PhD[4]
[1]Manchester Community College, STEM Division, Department of Biology, Manchester, CT
[2]Scientific Health: Education and Human Performance, PO Box 1331, Vernon-Rockville, CT 06066
[3]Drexel University, Philadelphia, PA
[4]Bloomsburg University of Pennsylvania, Bloomsburg, PA
JClark@manchestercc.edu, klr94@drexel.edu, bcomstock@bloomu.edu

Abstract
Significant developments in the understanding of muscle physiology have occurred in the last century. However, misconceptions are still finding their way into undergraduate and graduate instruction. This may leave students with a limited comprehension of muscle recruitment, strength, and why skeletal muscles fatigue during sustained (or repeated) contraction. The purpose of this article is to present instructors of human biology, physiology, or anatomy and physiology with an up-to-date understanding of muscle fatigue, and to dispel misconceptions that remain in the available educational resources. This article examines the role of the size principle in allowing muscles to become excited, the role of muscle metabolism in the ability of skeletal muscle to perform sustained activity, and the role of central drive in initiating contraction and sustaining muscle tension.
https://doi.org/10.21692/haps.2018.030

Key words: muscle physiology, muscle fatigue, central drive, muscle metabolism

Introduction
A common challenge in the teaching of human physiology is helping students to understand and appreciate the complexity of how body tissues function (Feder 2005, Hasan and Sequeira 2012, McFarland and Pape-Lindstrom 2016, Michael 2007). Moreover, there is a need to appreciate that the regulation of most physiologic processes typically do not follow the linear cause and effect mechanism laid out in many textbooks and accompanying instructional materials. Often, a host of factors that interact to regulate a physiologic process are not mentioned in the texts or other educational resources. This leaves instructors (or resource authors) who are not experts in the given field no choice but to rely on older interpretations or misconceptions. Unfortunately, misconceptions are then perpetuated through rote regurgitation of information and provide an impression to the student that the simplified description provided by the instructor (or resource material) is correct. One such phenomenon that follows this pattern of instruction is muscle fatigue.

Muscle fatigue is defined as the loss of contraction strength from the point of maximal strength achieved within a contraction cycle. Despite the simple definition, there are a variety of factors that contribute to the onset of fatigue. The factors that determine the onset of muscle fatigue fall into two categories: modifiable (trainable) or non-modifiable (inherent or innate) that interact with each other to not only impact the length of time that a contraction can be maintained (i.e. time under tension), but also the level of strength that is attainable within that contraction (Aagaard 2004, Flück 2006, Keener and Sneyd 2009, Payne 2006, Pette 2006, Sandri 2008, Schiaffino *et al.* 2006). The total time under tension a muscle can achieve will be dependent upon several interacting factors: the ability of the muscle fiber to undergo depolarization (excitation-contraction coupling), metabolite availability of the materials for ATP regeneration, availability of ions necessary for depolarization and muscle contraction (muscle kinematics), the ability of the muscle to dissipate heat, and the amount of pain as it relates to the overall psychological drive to continue the activity (Arendt-Nielsen and Graven-Nielsen 2008, Graven-Nielsen *et al.* 2003, St Clair Gibson *et al.* 2013).

Despite advances in the understanding of the complex interaction of mechanisms driving muscle recruitment and fatigue, many misconceptions have survived to the present day. Therefore, the primary purpose of this review is to update instructors on the anatomical and physiological mechanisms underlying muscle fatigue. Secondly, this review aims to encourage instructors (and resource authors) to rethink how they discuss muscle fatigue with their students. This update is important for several reasons. Without accurate information on muscle fatigue, students studying skeletal muscle physiology will have a limited ability to discuss how muscles function with colleagues, or with those in their future charge *e.g.* clients, patients or students. In addition, uninformed students may further perpetuate the misconceptions about fatigue. The better-informed instructors are, the more students can be informed about the mechanisms by which muscle strength is established and maintained and how fatigue is avoided.

continued on next page

Terminology Related To Muscle Contraction And Fatigue

To understand how muscles function and ultimately reach fatigue, there are several key concepts in muscle kinematics that must first be clarified since they are often mistakenly interchanged with each other (Table 1) (Knuttgen and Kraemer 1987, Winter *et al.* 2015). These terms must be addressed individually in the context of the development of muscle fatigue and changes in the level of muscle contraction.

Table 1. *Definitions, explanations, and mathematical representations of terms associated with muscle function.*

Kinematic Term (unit representation)	Explanation and Definition	Mathematical Representation
Force (Newton, N)	Total load that has been applied to the muscle	Force in: F, Newton (N) =mass (kg)*acceleration (m/s^2) F=m*a
(Inertia, I)	The muscle force that must be generated to cause movement at the boney articulation	Inertia (I) =mass (kg)*angular acceleration (Θ/s^2) I=m*α
Work (Joule, J)	Distance that the muscle moves the load	Work in: Joule (J) = Force (N)*change in distance (m) J=F*d
(Torque, τ)	The perpendicular distance from the pivot of the articulation that the force was applied	Torque (τ) = Inertia (I)*distance of lever (m) τ=I*d
Power (Watts, W)	Rate at which work is completed by the muscle moving the load	Power in: Work (Watt, W)= Joule (J)/time (second) W=J/second Watt (W)= Torque (τ)/angular speed(Θ/s) W=τ/ω
Energy (Joule, J, kilowatt - hr, kWH)	Capacity for completion of work	Energy in: Joule (J) = Force (N)*change in distance (m) J=F*d
(Calorie, Cal)	Heat Produced during execution of work equivalent to a rise of 1 ^oC in 1 L H_2O	Cal= amount of energy required to change 1L H2O 1^oC (Δ Temperature (^oC) * L H_2O in body) Conversion between Joules 4.184 kJ/ 1 Cal
Strength (Pascal, Pa Tensile torque, τ)	Tensile load that can be withstood per unit of cross-sectional area	Strength in: Tension (Pa)= Force (N)/cross-sectional area (m^2) Pa=N/m^2 Tensile Torque (τ)=Inertia (I)/cross-sectional area (m^2) τ=I/m^2
Intensity (Candela, I)	Power, or Energy of work, performed per square unit of area	Intensity in: Candela (I) = Watt (W)/ area (m^2) I= W/m^2
Fatigue	Relative change in strength that occurs over the time of a muscle contraction	Fatigue as: Fatigue=change in Stress (Pa)/change in time (second) Fatigue=ΔPa/second Fatigue=change in Stress (τ)/change in time (second) Fatigue=Δτ/second
Muscle Kinematic Term	**Explanation and Definition**	
Twitch	A single contraction of a single muscle fiber. The end result of a single excitation contraction couple. Typically contraction contained to the fiber itself and has no visible morphological changes in the muscle.	
Treppe/ Treppe Effect	Response of secondary excitations that occur during repolarization but prior to returning to a resting potential. Results in a step-wise increased strength of contraction in each subsequent excitation contraction couple.	
Tetany	Commonly seen as the "full muscle" contraction. Results from the summation of recruitment across the entirety of the muscle due to the rapidity of the excitation contraction couple that does not allow for repolarization to occur prior to subsequent excitation signals. Always associated with morphological changes to the muscle as a whole during contraction.	
Tonic Contraction	Muscle contraction where the level of strength and recruitment (i.e. tone) remains stable throughout the contraction cycle. Contraction that is normally seen over long periods of time that gets associated with smaller muscle fibers.	
Phasic Contraction	Muscle contraction where the level of strength and recruitment (i.e. tone) rapidly changes to meet a new demand. Highly variable in the level of strength and duration of time of the contraction. Associated with larger muscle fibers and more apt to produce fatigue within the muscle.	

continued on next page

Force is the load (in Newtons, N) that is being applied to the muscle. **Strength** is the amount of tension that can be generated within the muscle per cross-sectional area of the fiber to resist the load's force. Muscle strength is directly related to the rate at which muscles fatigue. **Work** is the distance over which a force is being applied. **Power** describes the speed by which the work is completed. **Energy** is the capacity to do work and **intensity** examines the area of the muscle where this energy is expended. Each of these is used to determine how long tension (strength) can be maintained at any given level of load resistance. **Fatigue** is described as the loss of strength from the desired (or optimal) point of contraction over a period of time. It is key to remember that all muscles will undergo fatigue. The onset of fatigue is based on the relative load, the strength achieved within the contraction, and the type of muscle fiber involved with developing the strength to resist the load (Allen *et al*. 2008, Enoka and Stuart 1992, Russ *et al*. 2002).

With respect to muscle fiber types, it is important to remember that a faster onset of fatigue occurs within fibers that are able to reach a greater maximum strength (*e.g*. Type IIX) than those that reach a lesser maximum (*e.g*. Type I) (Carpinelli 2008, Flück 2006, Harrison *et al.* 2011, Narici and Maganaris 2006, Schiaffino *et al.* 2006, Schiaffino and Reggiani 2011). The influence of fiber types on muscle fatigue will be examined in detail in the section "The Size Principle: Muscle As Excitable Tissue And Its Role In The Onset Of Fatigue". As noted in Figure 1, a continuum exists in which maximum absolute (or relative) strength of muscle contraction decreases as fiber size decreases. Additionally, because of the fiber size and type, metabolic pathways involved with adenosine triphosphate (ATP) regeneration will also vary (Enoka and Duchateau 2008, Russ *et al.* 2002, Schiaffino and Reggiani 2011). Metabolic pathways will be discussed further in the section "Metabolism and the Energetics of Fatigue".

The Size Principle: Muscle As Excitable Tissue And Its Role In The Onset Of Fatigue

There are two key concepts that explain the development of strength during the muscle contraction, both of which are related to the cross-sectional area of the fiber (Hamilton and Booth 2000, Henneman and Olson 1965, Jungblut 2009, Russell *et al.* 2000). First is the ability to depolarize to a threshold potential and second is the active contractile strength of the fiber once depolarized (Figure 1). In order for skeletal muscle contraction to occur, muscle fibers must be recruited via stimulation from a somatic motor neuron, whether a person is consciously thinking about this stimulation or not (Figure 2). A volitional contraction occurs because skeletal muscle is an excitable tissue where a membrane potential exists and an external stimulus leads to a depolarization of the tissue. Depolarization eventually reaches a threshold that will induce a physiological change in the tissue leading to a cytoskeletal change in response to the depolarization. In the case of skeletal muscle, this is the shortening of sarcomere. For muscle, this external stimulus comes from the efferent (lower motor) neuron. Therefore, skeletal muscle contraction cycles only occur when the muscle is stimulated to threshold level at the motor-end plate (neuromuscular junction).

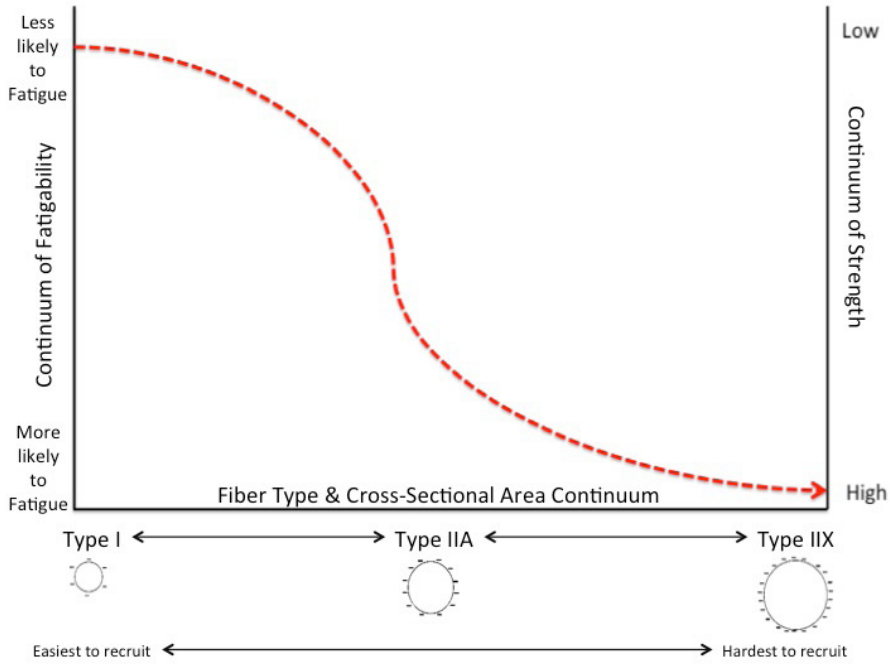

Figure 1. *Size Principle as it relates to fatigue and strength of a muscle contraction. Note the direct relationship between size with ease of recruitment (i.e. Henneman's Size Principle) and the inverse relationship between fatigability and strength.*

continued on next page

Once threshold has been reached, excitation-contraction coupling (ECC) occurs and tension is produced within the muscle (Carpinelli 2008, Henneman and Olson 1965, Henneman *et al.* 1965, Jungblut 2009). The overall strength of contraction is determined by a combination of active (actin and myosin interaction) and passive (connective proteins) tension within the individual fibers of the muscle fascicles. Furthermore, the more proteins involved in the contraction, the greater the amount of strength within the contraction. Simply put, larger fibers have more proteins and thus generate greater strength. This action occurring within the fibers ultimately provides for the visible morphological and kinematic changes of a muscle, allowing for movement of the limb or stabilization of the body (Harris *et al.* 2007, Karandikar and Vargas 2011, Russell *et al.* 2000). The sequence of developing strength via excitation and contraction coupling can also lead to an ever greater response from the muscle, depending on the rapidity of stimulus and amount of fascicles within the muscle undergoing recruitment. Repeated stimulation will cause the summation of muscle fiber twitches, progressing through treppe and eventual tetany of the muscle (Table 1) (Keener 2009, Sandercock 2005, Schiaffino 2006).

While muscle tissue is excitable, the excitability of the individual fibers and fascicles within the muscle will vary (Carpinelli 2008, Clamann 1993, Henneman and Olson 1965, Henneman, Somjen and Carpenter 1965, Jungblut 2009). The variability is not due to the absolute threshold of membrane potential, but instead the ability of the muscle fiber to reach this threshold potential. This ability is directly related to the diameter of the fiber (the amount of protein available for contractile strength) and thus the amount of membrane that must undergo an alteration in electrical potential to allow for functional changes to take place. This process is known as Henneman's Size Principle (Carpinelli 2008, Clamann, 1993, Henneman and Olson 1965, Henneman, Somjen and Carpenter 1965). According to this principle, there is a pattern that proceeds from small fibers, (those that are easily recruited due to the smaller amount of membrane) to larger diameter fibers (those that are harder to recruit due to the amount membrane requiring change) (Figure 1), which produces a large enough strength of contraction to meet the demand (Carpinelli 2008, Clamann 1993, Henneman and Olson 1965, Henneman, Somjen and Carpenter 1965, Jungblut 2009). Therefore, the relative strength of a muscle is determined by the ability of the muscle fibers to reach threshold. The ability of muscle fibers to reach threshold is based on the size of the fiber being depolarized (Carpinelli 2008, Farina *et al.* 2002, Jungblut 2009, Schiaffino *et al.* 2006). Simply stated, larger diameter muscle fibers produce greater levels of contraction

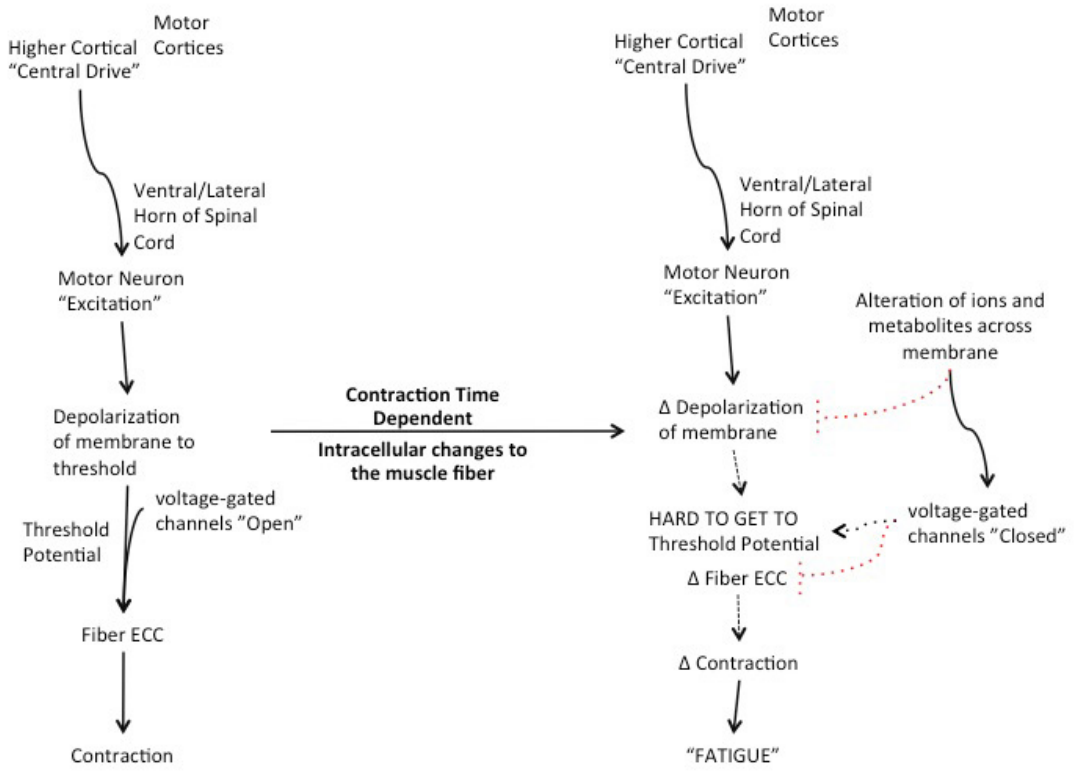

***Figure 2**. The excitation contraction coupling (ECC) is based on the ability for the muscle fiber to reach threshold.*

continued on next page

strength than their smaller fiber counterparts (Carpinelli 2008, Jungblut 2009).

By describing muscle contraction (total time under tension) using the ECC and Henneman's Size Principle, we can see that strength and fatigue of the skeletal muscle is based on the ability of the muscle to reach and maintain threshold potential. However, skeletal muscle and its fascicles are a mixture of fiber types and diameter, regardless of a history of exposure to tensile stresses or a genetic predisposition for fiber type and hypertrophication (Aagaard 2004, Flück 2006, Sale 1988, Sandri 2008, Schiaffino and Reggiani 2011, Schiaffino *et al.* 2007). Henneman's Size Principle explains that since there is less total membrane surface area, smaller fibers will more easily reach threshold potential and produce less tension compared to their larger counterparts in a given skeletal muscle (Henneman and Olson 1965, Henneman, Somjen and Carpenter 1965, Jungblut 2009). Additionally, because of the kinematics of the myosin power-stroke, the strength of the smaller fibers will be much slower to develop than the strength of the larger muscle fibers. This means that as recruitment progresses through fiber size, the strength and speed by which that strength is developed both increase (Carpinelli 2008, Clamann 1993, Harris *et al.* 2007, Jungblut 2009, Pette 2006, Sale 1988).

When slow tonic contractions of low strength are required, smaller (Type I) fibers will be recruited in higher proportion than larger (Type II) fibers (Clamann 1993, Harrison *et al.* 2011, Jungblut 2009, Payne 2006, Pette 2006, Schiaffino 2007). Smaller fibers are not only easier to recruit, but the duration of time necessary to return to a resting state is also shorter. Shorter resting time means smaller fibers will be able to quickly depolarize again and thus be able to be recruited again for continuous contraction. However, when contractions are rapid or phasic, they require more strength and a greater proportion of Type II fibers will be recruited.

Unfortunately, this change in recruitment pattern comes at a cost to the muscle (Allen *et al.* 2008, Carpinelli 2009, Enoka 2008, Harrison *et al.* 2011, Russell *et al.* 2000, Schiaffino and Reggiani 2011, Schiaffino 2007). The rapid contraction cycle of larger fibers leads to large-scale changes in metabolites in and around muscle tissue membranes that impacts depolarization (Figure 2). Larger fibers not only need greater stimulus to reach threshold, but differences in the anatomy of the tissue, primarily related to lower rates of perfusion and reliance on anaerobic metabolic pathways, mean that larger fibers experience greater changes in metabolites that can alter membrane potentials relative to the smaller fibers. (The metabolites of muscle fatigue are described in detail in the next section.) As a result, larger fibers may require more time to re-establish the concentration of metabolites and ions across the membrane, which are necessary for the ECC from subsequent (or continuous) stimulus by the motor neuron (Figure 2). Thus a higher rate of fatigue is established in larger fibers. Differences in recruitment time and strength will ultimately lead to variability in fatigue rate, the amount of strength developed, and the total time under tension that is produced by the entirety of the skeletal muscle during a sustained contraction.

Therefore, muscle fatigue can be induced via an inability to polarize the membrane of the muscle fiber so as to pair subsequent excitation with depolarization to threshold potential. As summarized in Figure 2, when muscle fibers undergo contractions intracellular actions lead to alterations of ions and metabolites across the membrane that alters the electrochemical potential and increases the amount of excitation necessary to open voltage gated channels and achieve threshold. The inability to get to ECC in the fiber leads to a change in the contraction capacity of the muscle as a whole and leads to the onset of fatigue within the muscle. Considering the variability of fiber size and resulting kinematics, Henneman's Size Principle helps to explain part of the mechanisms for first gaining and then losing strength throughout a contraction cycle (Carpinelli 2008, Clamann 1993, Harris *et al.* 2007, Jungblut 2009, Pette 2006, Sale 1988). This principle may underlie the immediate reduction in muscle strength due to the inability of larger fibers to undergo the ECC necessary for maximal contraction to be continued. The reduction in strength is more pronounced in larger muscle fibers than smaller ones, leaving only the smaller fibers and motor units available to recruit for longer duration muscular activities. The reduction will be seen in both strength produced and the electric activity of the muscle registered via electromyography (EMG) readouts (Farina *et al.* 2002, González-Izal *et al.* 2012, Halperin *et al.* 2014).

Metabolism and Energetics of Fatigue

Muscle metabolism frequently dominates the discussion of fatigue. This discussion suggests that fatigue is due to the accumulation of the glycolytic byproduct lactate (mistakenly referred to as lactic acid) that accumulates after a prolonged state of anaerobic activity within the muscle fiber (Enoka and Duchateau 2008, Gladden 2004, Jorfeldt *et al.* 1978, Sahlin *et al.* 1987). Hull (2016) offers more detail on the role of lactate in muscle fatigue and is a highly recommended resource on this topic.

Skeletal muscle is comprised of distinct fiber types (*e.g.* Type I, IIA, IIX) that utilize varying energetic pathways for regenerating the required ATP for a sustained muscle contraction. Type IIX, and to a lesser extent type IIA fibers, are primarily reliant upon the phosphocreatine and glycolytic pathways for the reformation of ATP. Type I and Type IIA fibers however, rely on either glycolytic or lipolytic pathways (Harrison *et al.* 2011, Hawley *et al.* 2006,Russell *et al.* 2000, Schiaffino and Reggiani

continued on next page

2011). Differences in the reliance on distinct pathways for ATP regeneration directly contribute to the difference in strength capacity, and thus, fatigability of these fibers.

Due to the speed by which the fibers utilize ATP, Type IIX fibers are more reliant upon the phosphocreatine pathway for regeneration of ATP than other metabolic pathways. While the phosphocreatine pathway is able to rapidly regenerate ATP for use by myosin, the capacity to do so is limited by the short duration that the limiting enzyme creatine kinase (CK). The exhaustion of the CK within six seconds reduces the total regeneration of ATP necessary for continuation of the contraction for the largest muscle fiber (Figure 3). Given the limited time available for the phosphocreatine pathway to replenish the required ATP for Type IIX fibers, these fibers are the first to undergo fatigue. Moreover, since these fibers have the greatest strength capacity of the various fibers in the muscle, the loss of the ability to have Type IIX fibers involved in the contraction will result in the early loss of strength in a sustained contraction. This decrease in strength is particularly pronounced if contraction is at a sustained maximal volitional contraction, as typically examined in learning laboratory experiments.

The next fibers to fatigue are the medium and remaining large fibers (Type IIA), which are more reliant on a combination of phosphocreatine and glycolytic pathways. The smaller fibers, which are more reliant on oxidative pathways in the regeneration of ATP, are the last to fatigue (Boonyarom and Inui 2006, Harrison *et al.* 2011, Russell *et al.* 2000, Schiaffino *et al.* 2007). When we look at the series of energetic pathways used to reform ATP (the phosphocreatine system, glycolysis, Krebs cycle (TCA), and oxidative phosphorylation) each generates several metabolites that can ultimately interfere with the reformation of ATP, alter membrane potential, or competitively inhibit myosin-actin interactions. All of these metabolites ultimately interfere with muscle kinematics. Additionally, as muscles undergo continual contraction, stresses placed on the cell leads to intracellular damage. Tissue damage from contraction leads to the accumulation and release of metabolites, ions, and metabolic wastes. These metabolites (Table 2) include lipopolysaccharides (LPS), Na^+, creatine phosphate (cP), ATP and ATP metabolites, and reactive oxidative species (ROS). The various metabolites interfere with membrane dynamics and the ability for ECC to occur in muscle fibers. Thus, the combination of reduced ATP and the alteration of membrane dynamics (accumulation of ions and metabolites alter the electrochemical charge of the membrane) induce the onset of fatigue (Figure 4). Therefore, we see several factors interacting on the cellular level that interfere with the ability of myosin filaments to perform a power stroke due to reduced ATP. Additionally, accumulating metabolites around the membrane change the relative electrical potential and also contribute to muscle fatigue. If several metabolites interfere with muscle function and possibly induce fatigue, why is it that lactate gets all of the blame?

To address the misconception of lactate accumulation as the primary cause of muscle fatigue, one must start at the origin of the concept, which has been discussed at length in reviews by Allen and Enoka (Allen *et al.* 2008, Enoka and Duchateau 2008) and Hull and Marx (2016). This misconception is based on *in situ* muscle studies dating back to the turn of the last century (Allen *et al.* 2008, Di Giulio *et al.* 2006, Enoka and Duchateau 2008, Lombard 1892), and subsequent studies of the 1920's, that determined that *in situ* skeletal muscle function began to degrade as lactic acid was infused into the isolated muscle and diffused into the sarcomere. This led to the speculation that lactic acid inhibits several steps in the sliding filament theory and thus is the cause for fatigue (Hogan *et al.* 1998, Pette 2006, Schiaffino *et al.* 2007). There were two key ways suggested that lactic acid can inhibit normal skeletal muscle contraction. First, there is inhibition of the energetic reactions leading to the reformation of ATP within the final stages of pyruvate formation in the glycolysis pathway. Second, there may be competitive inhibition at the myosin ATPase binding sites, limiting the ability for the ATPase to hydrolyze ATP to ADP. This would minimize the ability of the myosin globular head to perform its ratcheting power-stroke-release-catch that is necessary to move actin across myosin allowing for contraction.

In muscle tissue *in vivo*, intracellular mechanisms tend to slow the accumulation of lactate [lactic acid], and thus minimize the impact that accumulation can have in muscle tissue (Gladden 2004, Jorfeldt *et al.* 1978 Hull and Marx 2016). Several studies have indicated that the nervous system has the ability to excite muscle tissue following the induction of metabolic fatigue (Allen *et al.* 2008, Enoka and Duchateau 2008, Enoka and Stuart 1992, Hogan *et al.* 1998, Russ *et al.* 2002). The classic studies cited did note that as lactic acid was washed out of the muscle with Ringer's solution, the degradation of muscle contraction was reduced and muscle strength capacity returned to the pre-infusion level (Allen *et al.* 2008, Enoka and Duchateau 2008). This process is similar to what muscle tissue actually does by shuttling lactate out of the muscle during activity. This suggests that the accumulation of lactate in muscle fibers may not be the underlying factor in muscle fatigue.

Central Drive And Muscle Fatigue

Another incorrect observation commonly attributed to lactic acid accumulation is the reported sensation of "heat and pain in the muscle" during and following prolonged, or repeated, muscle contractions. While it might be simple to conclude that acid accumulating within tissue might lead to heat and pain in the tissue, we must remember that lactate does not actually accumulate (Allen *et al.* 2008). What then creates the heat and pain sensations associated with the personal experience of musculoskeletal fatigue during activity? Before we move further, it is important to remember that heat and pain are perceived sensations and are linked to higher cortical functions. Central drive is a highly complex phenomenon that

continued on next page

Table 2. *The metabolites, electrolytes, and chemical signals known to impact muscle function and induce fatigue with prolonged contraction time.*

Metabolite Electrolyte Enzyme	Impact on Muscle Function	Mechanism Contributing to Muscle Fatigue
Lactate (Lactic Acid, LA)	Impact on enzyme (PDH, LDH) activity Competitive inhibition at actin-myosin binding site	Energetics Myofibril interference
Reactive Oxidative Species (ROS)	Impact membrane dynamics & osmolarity conditions around intramuscular nociceptors Impact enzyme function within energetic pathways	Alter membrane potentials Interference with transmembrane transporters Energetics
ATP metabolites (AMP/ADP & Pi/ATP, AMPk)	Impact membrane dynamics Impact phosphocreatine cycling and all energetic pathways Competitive inhibition at actin-myosin binding site Impact osmolarity conditions around intramuscular nociceptors (via metabolite and acid specific ion channels)	Alter membrane potentials Competitive interference on regulatory enzymes Competitive interference at myosin ATPase Activation of spinothalamic afferents inducing sensation of "Pain"
Creatine Kinase (CK)	Impact of ATP-cP cycling	Energetics Myofibril interference (IIX>IIA>I)
Mg^{++} ions	Competitive inhibit Ca^{++} at troponin C and CaM/CaMKII pathway Alter binding interaction between myosin head and actin chain Impact membrane dynamics & depolarization	Alter troponin/tropomyosin kinetic Myofilament interference Alter excitation-contraction coupling
Ca^{++} ions	Impact regulatory proteins Impact neuromuscular junction dynamics & depolarization CaMk energetic pathways	Alter troponin/tropomyosin kinetic Alter excitation-contraction coupling Altered ATP production rate & oxidation of metabolites
K^+ ions	Impact membrane dynamics & depolarization Impact osmolarity conditions around intramuscular nociceptors	Alter excitation-contraction coupling Activation of spinothalamic afferents inducing sensation of "Pain"
Na^+ ions	Impact membrane dynamics Impact osmolarity conditions around intramuscular nociceptors	Alter excitation-contraction coupling Activation of spinothalamic afferents inducing sensation of "Pain"
Cl^- ions	Impact membrane dynamics Impact osmolarity conditions around intramuscular nociceptors	Alter Excitation-Contraction Coupling Activation of spinothalamic afferents inducing sensation of "Pain"
H^+ ions	Impact glycolysis activity by impact on the formation of pyruvate pH alteration around intramuscular nociceptors (via metabolite and acid specific ion channels)	Energetics Activation of spinothalamic afferents inducing sensation of "Pain"
Bradykinin	Alter vascular dynamics (increase inflammation) Impact osmolarity conditions around intramuscular nociceptors	Intertissue inflammation leading to secondary activation of pressure receptors & spinothalamic afferents via alteration of osmotic conditions
Serotonin (5-HT)	Alter vascular dynamics (increase inflammation) Induce pain sensation	Intertissue inflammation leading to activation of pressure receptors cFOS production and activation of spinothalamic afferents
Prostaglandin E_2 (PGE_2)	Induce pain sensation	cFOS production and activation of spinothalamic afferents
NF-κB	Intracellular inflammation	Intracellular inflammation Activation of secondary inflammatory and pain markers (COX2, Il-1)

continued on next page

suggests an integrated relationship between changes within the peripheral tissues and alterations in central drive (higher cortical) or spinal cord motor output (Arendt-Nielsen and Graven-Nielsen 2008, Graven-Nielsen and Mense 2001, Lambert 2005, Light *et al.* 2008, St Clair *et al.* 2004; Figure 5). While the concept of central drive is traditionally applied to pacing strategies and fatigue onset during long-duration endurance exercise, the hallmark principles of central drive can be integral in explaining the fatigue from single contraction efforts.

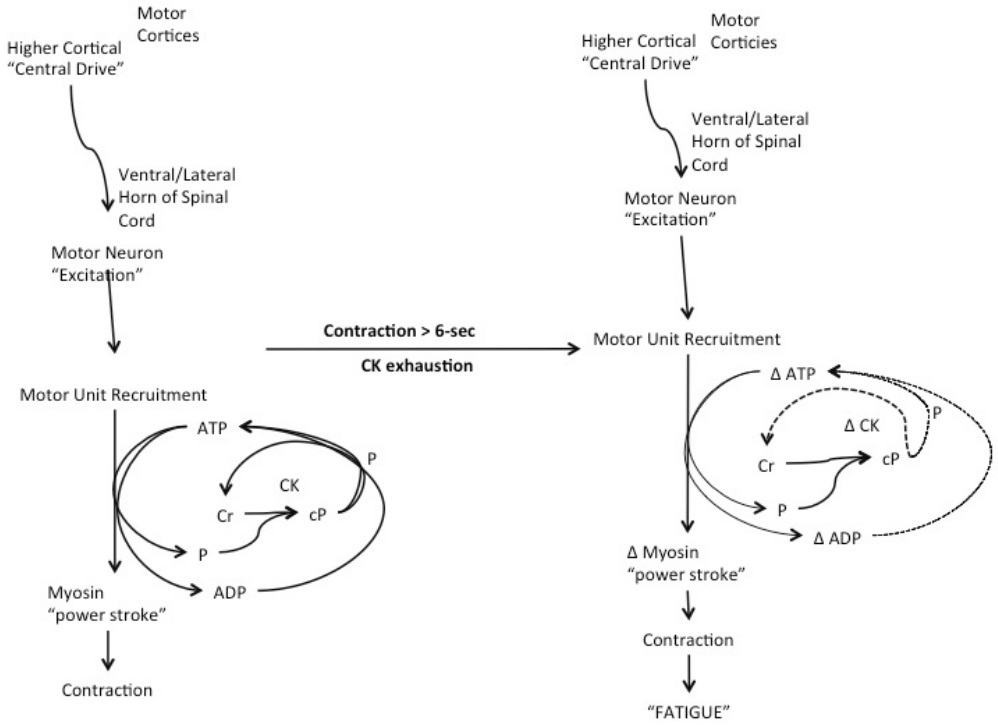

Figure 3. *The poor efficiency of the phosphocreatine system reduces ATP levels and the onset of fatigue starts within seconds of contraction.*

Central drive can be seen as a cognitive regulation of strength development that interplays with the recruitment pattern, via Henneman's Size Principle, in such a way as to maximize recruitment of the muscle for the demand being placed on it. While at the same time central drive minimizes any excessive damage to tissues that can result from a combination of increased cellular damage or accumulation of metabolites interfering with the cellular membrane functions. Thus, we can see the interplay between the higher cortical regions and the activation of motor output that influences the ability to a) reach threshold and b) produce appropriate strength for the demand (Carpinelli 2009, Clamann 1993, Henneman and Olson 1965, Henneman, Somjen and Carpenter 1965, Jungblut 2009, St Clair *et al.* 2004).

Outside of kinematics of the muscle fiber, central drive ensures that recruitment to the demand can be countered by the ability to dissipate excessive free-energy and heat from the contraction to the surrounding tissues or external environment. At the same time, an immune-mediated inflammatory response occurs to adapt the stresses being placed on the tissue during the sustained contraction (Graven-Nielsen and Mense 2001, Lambert 2005, St Clair *et al.* 2004). The combination of peripheral changes, along with alteration in cortical function and attention to muscle activation, alters regulatory signals and induces a sense of fatigue, known as central fatigue (Halperin *et al.* 2014, Halperin *et al.* 2015,

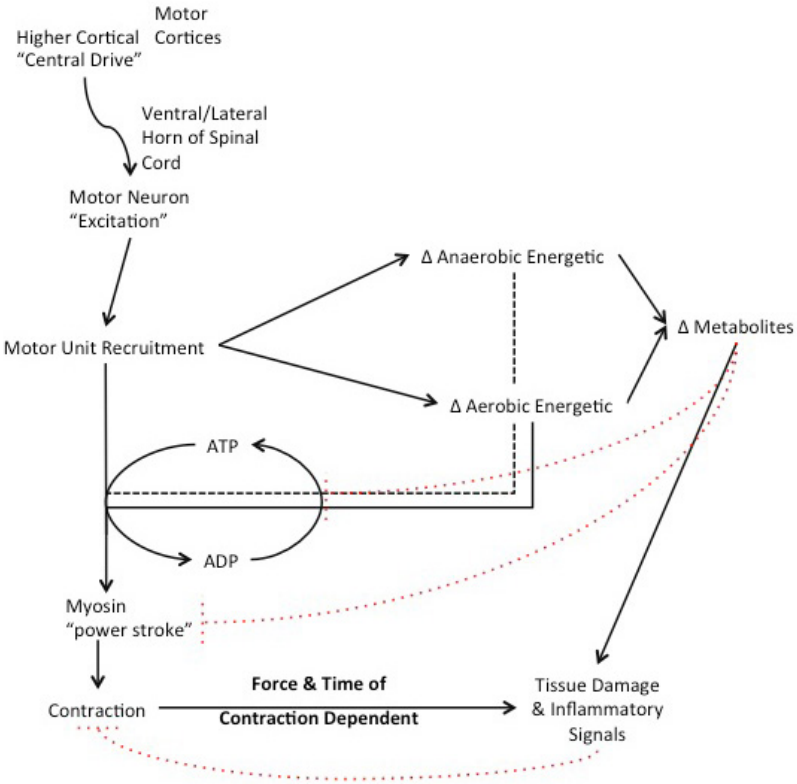

Figure 4. *Changes in metabolic pathways and intracellular damage induce fatigue within a minute of contraction.*

continued on next page

Nybo 2008) or a "loss" of central drive. Therefore, we must look at what is triggering the afferent pathways that carry sensations of heat and pain to the central nervous system, which leads to reduction in the motor output, loss of the larger fibers involved in the contraction, and the sensations of fatigue (Figure 5).

The combination of heat and pain that is felt during prolonged muscle activity is most likely tied to the release of metabolites (Table 2) into the surrounding interstitial tissue as damage to muscle tissue occurs. The degree of damage and overall response of pain is related to, and dependent on, the level of strength produced (*i.e.* amount of tissue being recruited), the fiber type being recruited, and the length of contraction necessary for the strength produced during activity. The accumulation of metabolites surrounding the muscle tissue not only impacts the membrane potential of the muscle cell, impacting the cells ability to reach threshold and thus contract, but also excites the nociceptors within the skeletal muscle (Graven-Nielsen *et al.* 2003, Graven-Nielsen and Mense 2001; Figure 5). Additionally, immune cells that respond to tissue injury release cytokines such as serotonin and histamine, which initiate the swelling and inflammatory responses necessary for repair following the stress of the contraction. The inflammatory response is meant to instigate a self-splinting of the tissue, so as to minimize any further damage to the traumatized muscle.

Finally, the alteration in local heat accumulation changes both local tissue and whole-body temperatures, which may alter the regulatory signals and induce a sense of fatigue (Nybo 2008). Nociceptive and thermoreceptive input, whether driven by tissue metabolites or immune responses, are carried via sensory neurons supplying the working muscle to the central nervous system, ultimately resulting in the perception of heat and pain in the cerebral cortex (Figure 5). In turn, areas of higher cortical function reduce motor output signals that lead to lessened stimuli on the muscle for contraction. Less muscle activation and reduced muscle strength are indications of fatigue (Arendt-Nielsen and Graven-Nielsen 2008, Graven-Nielsen and Mense 2001, Lambert 2005, Light *et al.* 2008). Clearly, a simple explanation of muscle fatigue as a "buildup of lactic acid" fails to acknowledge the interplay between the afferent and efferent pathways that connect working muscles with the central nervous system (Enoka and Duchateau 2008, Halperin *et al.* 2015, Hamilton and Booth 2000, Lambert 2005, Nybo 2008, St Clair *et al.* 2004). Thus, central fatigue, or the loss of central drive, creates the psychological sense of fatigue and promotes cognitive thoughts of "I don't want to do this anymore," or "I can't do this anymore," that are frequently reported during sustained and prolonged muscle activity.

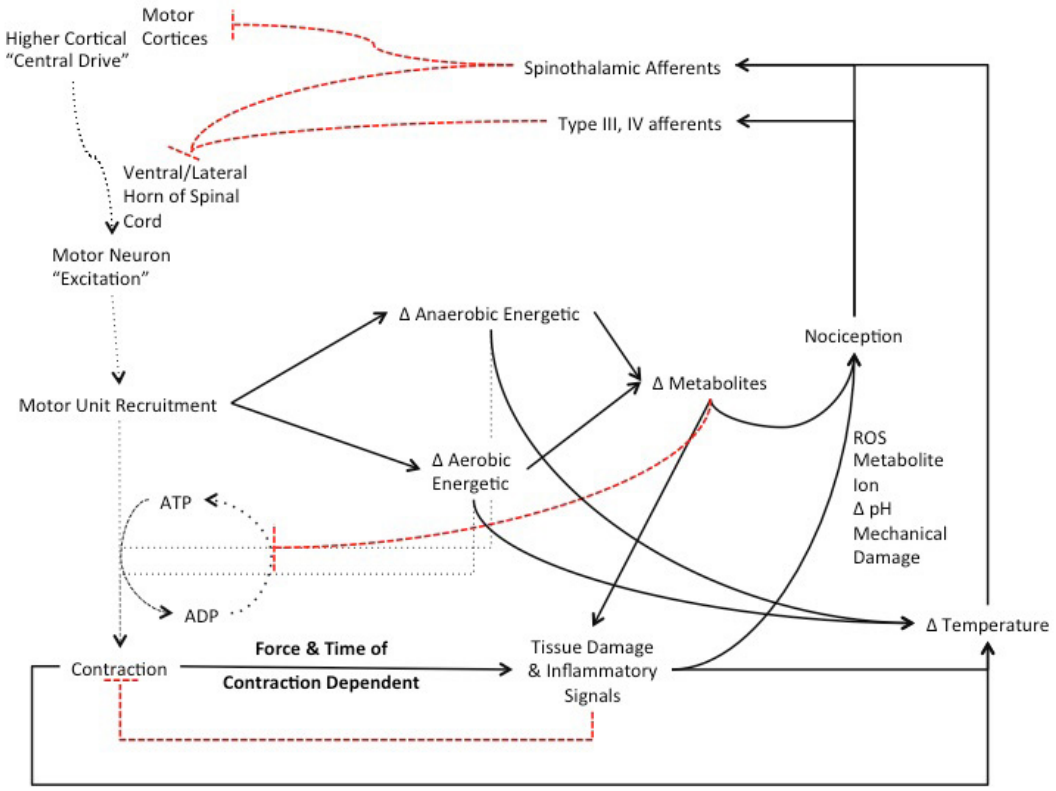

Figure 5. *Central drive and central fatigue are muscle fatigue feedback loops.*

continued on next page

Adaptations of Muscle Recruitment and Impact to Fatigue

We tend to think of both contraction and fatigue as instantaneous phenomena wherein the entire muscle is being recruited or the entire muscle is lost. However, it is the motor units within a muscle that are being recruited, and recruitment is progressive in onset during the volitional muscle contraction (Carpinelli 2008, Clamann 1993, Harris *et al.* 2007). This progression is related to type of muscle tissue being recruited (Type I, IIA, or IIX) and the total number of muscle fibers necessary to meet the demand for tension within the muscle, which is attributable to the interplay between central drive with Henneman's Size Principle. This interplay means that adaptations in how motor recruitment occurs are an important aspect of muscle physiology that can be altered by training (Aagaard 2004, Flück 2006, Halperin *et al.* 2014, Hood *et al.* 2006, Payne and Delbono 2006, Sale 1988, Schiaffino *et al.* 2007, Schiaffino *et al.* 2006).

When untrained muscle is stimulated, fibers are sequentially recruited as the level of stimulus increases until there are enough muscle fibers to meet (or overcome) the tension of the resistance. As the muscle is recruited more regularly, one is able to learn the most efficient pattern of recruitment required for a sustained level of activity (Aagaard 2004, Flück 2006, Halperin *et al.* 2014, Hood *et al.* 2006, Payne and Delbono 2006, Sale 1988, Schiaffino *et al.* 2007, Schiaffino *et al.* 2006). This adaptation ensures that the appropriate number of large, moderate, and small muscle fibers are recruited to sustain the level of contraction necessary to meet tension being placed on the muscle. Moreover, one can learn how to oscillate the total amount of muscle being recruited at any instant so as to be able to maintain the necessary level of tension for ever-increasing durations (Pette 2006;Sale 1988, Schiaffino *et al.* 2007, Schiaffino *et al.* 2006). The resultant change leads to the ability to produce the same level of tension with less overall recruitment coming from the medium and smaller fiber types (Hood *et al.* 2006, Narici and Maganaris 2006, Schiaffino *et al.* 2007). This allows for a longer sustained contraction at the optimal level of tension while at the same time minimizing the amount of damage to the tissue or optimally dissipating energy and heat from the activated muscle tissue.

As the ability to recruit the muscle changes, there are changes in the metabolic pathways utilized by the muscle fibers to regenerate ATP. Specifically, the plasticity of Type II fibers leads to a modification where Type IIX fibers shift towards Type IIA metabolic characteristics (Aagaard 2004, Flück 2006, Hood *et al.* 2006, Narici and Maganaris 2006, Schiaffino *et al.* 2007, Schiaffino *et al.* 2011). This change in metabolism that leads to increased reliance on glycolytic and oxidative pathways within the trained fiber allowing for longer time under tension than was allowed for in the untrained muscle. At the same time, the stress of producing tension progresses beyond biochemical changes and induces genetic regulation, causing greater contractile and non-contractile protein synthesis (Boonyarom and Inui 2006, Flück 2006, Sandri 2008, Schiaffino *et al.* 2007, Schiaffino *et al.* 2011). Training adaptation results in greater fiber diameters (more pronounced in the Type II fibers than Type I fibers) and thus more strength, as measured by a greater capacity to have prolonged time under tension (Halperin *et al.* 2014, Hood *et al.* 2006, Payne and Delbono 2006, Sale 1988, Schiaffino *et al.* 2007). These alterations are the key aspects of muscle adaptation that occur with training, and illustrate the concept of the SAID (Specific Adaptations to Imposed Demands) principle (Sale 1988). These adaptations allow for variance in intramuscular responses to effort based on psychological desire to instigate, and then continue, the effort being undertaken beyond what would have been possible by reducing possible inhibition signals that once caused fatigue.

A final, but potentially highly influential adaptation that helps training prevent fatigue is the cognitive desire for continuation of contraction that is externally derived *e.g*. environmental conditions, social cueing, and peer pressure (Andreacci 2002, Bigliassi 2013, Hutchinson 2014, McNair 1996). External cues may allow one to override inhibitory signals and continue to exercise well past the initial indications of fatigue. This allows for an increased level of strength to be generated within the muscle even when fatigue has been indicated either metabolically, kinetically, or by the subject indicating a reduced central drive to continue. Environmental cues are the reason for the selection of music played in weight-rooms and the verbal encouragement that is heard from observers and spotters during maximal effort activities, or the verbal encouragement in the delivery room during childbirth (Andreacci 2002, Bigliassi 2013, Hutchinson 2014, McNair 1996). These situations are easily replicated during physiology laboratory exercises of EMG recordings of handgrip dynamometer or arm-wrestling. The instructor can run replications of time to fatigue in environments with no external stimulus versus one with the loud music and verbal encouragement for the student performing the test.

Conclusions and Reflections on Teaching Muscle Fatigue

Our understanding of the physiological regulation of skeletal muscle contraction and fatigue continues to develop. However, previously held theories, such as lactate accumulation causing muscle fatigue, are still finding their way into the discussion of fatigue. Therefore, it is important for anatomy and physiology instructors to revisit these concepts regularly. Muscle fatigue is a complex experience involving multiple factors that include the impact of Henneman's Size Principle, metabolic changes, immune responses, and a loss of central drive. We must remember that fatigue is the reduction of the strength from the optimal (or maximal) level within a sustained contraction. Fatigue is due to one or combinations of the various factors discussed in this article and the traditional teaching of a single cause and effect relationship for muscle fatigue is an oversimplification. As instructors of muscle physiology, we

continued on next page

must abandon the *in situ* cellular studies of the late 1800's and early 1900's and expand our view of what fatigue is and how fatigue develops. Furthermore, teaching of fatigue should address the complex network of interacting factors that influence a muscle's ability to first develop tensile strength and then regulate the development of fatigue to maintain that level of strength.

To help students better understand muscle contraction and the onset of fatigue, several key factors must be addressed. First, given that muscle physiology goes hand-in-hand with the physical principles of biomechanics and kinematics, it would be prudent to be consistent in the terminology in lectures and associated laboratory exercises. Secondly, while many educational resources discuss the general relationship between cross-sectional area and strength, very few (if any) address the concept as it relates to recruitment as detailed by Henneman's Size Principle. In this, it is important to detail the interaction of recruitment, energetics, and inflammatory responses that vary by stress placed on the individual fibers that inhibit the ability to a) recruit the muscle fiber and b) have kinematic changes within the muscle fiber following recruitment. Therefore, it would be prudent to include such information in human biology, physiology, or human anatomy and physiology texts for undergraduate students as well as the ancillary materials (*e.g.* computer simulations and review texts) and laboratory exercises used to assist with their learning of the topics.

About the Authors

James E Clark is a professor and the course coordinator of Human Anatomy and Physiology at Manchester Community College. He has two decades of experience in education and sports medicine as a strength and conditioning coach and athletic trainer. His undergraduate education and degrees are in Biology (Physiology and Pre-Medical emphasis) and Psychology and he has graduate education and advanced degrees in Physical Therapy, Education, Integrative and Evolutionary Biology, and Exercise Physiology. James is involved with research related to the means by which exercise and diet modification is beneficial for changing health status and resolving issues of overfatness along with improving educational performance for students.

Krista L Rompolski, PhD received a BS in Exercise Science and an MS in Exercise Physiology from Bloomsburg University, and a PhD in Exercise Physiology from the University of Pittsburgh. She is a Certified Exercise Physiologist through the American College of Sports Medicine. Dr. Rompolski is an active member of the Human Anatomy and Physiology Society and the American Academy of Anatomists. Her teaching interests include pathophysiology, gross anatomy, and anatomy and physiology.

Brett A Comstock, PhD is a professor of exercise science at Bloomsburg University in Pennsylvania and an active member of the National Strength and Conditioning Association.

Literature cited:

Aagaard P(2004) Making muscles "stronger": exercise, nutrition, drugs. J *Musculoskelet Neuronal Interact* 4: 165-174.

Allen DG, Lamb GD, Westerblad H (2008) Skeletal muscle fatigue: cellular mechanisms. *Physiol Rev* 88: 287-332.

Andreacci JL, LeMura LM, Cohen SL, Urbansky EA, Chelland SA, Von Duvillard SP (2002) The effects of frequency of encouragement on performance during maximal exercise testing. *J Sports Sci*. 20: 345-352.

Arendt-Nielsen L, Graven-Nielsen T (2008) Muscle pain: sensory implications and interaction with motor control. *Clin J Pain* 24: 291-298.

Bigliassi M, Estanislau C, Carneiro JG, Kanthack TFD, Altimari LR (2013) Music: a psychophysiological aid to physical exercise and sport, *Arch Med Deporte* 30: 311-320

Boonyarom O, Inui K (2006) Atrophy and hypertrophy of skeletal muscles: structural and functional aspects. *Acta Physiol (Oxf)* 188: 77-89.

Carpinelli RN (2008) The size principle and a critical analysis of the unsubstantiated heavier-is-better recommendation for resistance training. *J Exerc Sci Fit* 6: 67-86.

Clamann HP (1993) Motor unit recruitment and the gradation of muscle force. *Phys Ther* 73: 830-843.

Di Giulio C, Daniele F, Tipton CM (2006) Angelo Mosso and muscular fatigue: 116 years after the first Congress of Physiologists: IUPS commemoration. *Adv Physiol Educ* 30: 51-57.

Enoka RM, Duchateau J (2008) Muscle fatigue: what, why and how it influences muscle function. *J Physiol* 586: 11-23.

Enoka RM, Stuart DG (1992) Neurobiology of muscle fatigue. *J Appl Physiol* 72: 1631-1648.

Farina D, Fosci M, Merletti R (2002) Motor unit recruitment strategies investigated by surface EMG variables. *J Appl Physiol* 92: 235-247.

Feder ME (2005) Aims of undergraduate physiology education: a view from the University of Chicago. *Adv Physiol Educ* 29: 3-10.

Flück M (2006) Functional, structural and molecular plasticity of mammalian skeletal muscle in response to exercise stimuli. *J Exp Biol* 209: 2239-2248.

Gladden LB (2004) Lactate metabolism: a new paradigm for the third millennium. *J Physiol* 558: 5-30.

González-Izal M, Malanda A, Gorostiaga E, Izquierdo M (2012) Electromyographic models to assess muscle fatigue. *J Electromyogr Kinesiol* 22: 501-512.

continued on next page

Graven-Nielsen T, Jansson Y, Segerdahl M, Kristensen JD, Mense S, Arendt-Nielsen L, Sollevi A (2003) Experimental pain by ischaemic contractions compared with pain by intramuscular infusions of adenosine and hypertonic saline. *Eur J Pain* 7: 93-102.

Graven-Nielsen T, Mense S (2001) The peripheral apparatus of muscle pain: evidence from animal and human studies. *Clin J Pain* 17: 2-10.

Halperin I, Aboodarda SJ, Basset FA, Byrne JM, Behm DG (2014) Pacing strategies during repeated maximal voluntary contractions. *Eur J Appl Physiol* 114: 1413-1420.

Halperin I, Chapman DW, Behm DG (2015) Non-local muscle fatigue: effects and possible mechanisms. *Eur J Appl Physiol* 115: 2031-2048.

Hamilton MT, Booth FW (2000) Skeletal muscle adaptation to exercise: a century of progress. *J Appl Physiol* 88: 327-331.

Harris N, Cronin J, Keogh J (2007) Contraction force specificity and its relationship to functional performance. *J Sports Sci* 25: 201-212.

Harrison BC, Allen DL, Leinwand LA (2011) IIb or not IIb? Regulation of myosin heavy chain gene expression in mice and men. *Skeletal Muscle* 1: 5.

Hasan Z, Sequeira R (2012) Challenges of teaching physiology in an integrated system-based curriculum. *Can Med Educ J* 3: e73-e76.

Hawley JA, Hargreaves M, Zierath JR (2006) Signaling mechanisms in skeletal muscle: role in substrate selection and muscle adaptation. *Essays Biochem* 42: 1-12.

Henneman E, Olson CB (1965) Relations between structure and function in the design of skeletal muscles. *Journal of Neurophysiology* 28: 581-598.

Henneman E, Somjen G, Carpenter DO (1965) Excitability and inhibitability of motoneurons of different sizes. *Journal of Neurophysiology* 28: 599-620.

Hogan MC, Ingham E, Kurdak SS (1998) Contraction duration affects metabolic energy cost and fatigue in skeletal muscle. *Am J Physiol* 274: E397-402.

Hood DA, Irrcher I, Ljubicic V, Joseph A-M (2006) Coordination of metabolic plasticity in skeletal muscle. *J Exp Biol* 209: 2265-2275.

Hutchinson JC, Sherman T (2014) The relationship between exercise intensity and preferred music intensity. *Sport, Exercise, and Performance Psychology* 3: 191-202.

Hull K, Marx P(2016) Lactate: Toxic villain or metabolic superhero? *HAPS Educator* 20: 38-47.

Jorfeldt L, Juhlin-Dannfelt A, Karlsson J (1978) Lactate release in relation to tissue lactate in human skeletal muscle during exercise. *J Appl Physiol* 44: 350-352.

Jungblut S (2009) The correct interpretation of the size principle and its practical application to resistance training. *Medicina Sportiva* 13: 203-209.

Karandikar N, Vargas OOO (2011) Kinetic chains: a review of the concept and its clinical applications. *PM&R* 3: 739-745.

Keener J, Sneyd J (2009) Muscle. In: Keener J, Sneyd J eds. *Mathematical Physiology II: Systems Physiology*: Springer, 717-772.

Knuttgen HG, Kraemer WJ (1987) Terminology and measurement in exercise performance. *Journal of Applied Sports Science Research* 1: 1-10.

Lambert EV (2005) Complex systems model of fatigue: integrative homoeostatic control of peripheral physiological systems during exercise in humans. *Br J Sports Med* 39: 52-62.

Light AR, Hughen RW, Zhang J, Rainier J, Liu Z, Lee J (2008) Dorsal Root Ganglion Neurons Innervating Skeletal Muscle Respond to Physiological Combinations of Protons, ATP, and Lactate Mediated by ASIC, P2X, and TRPV1. *Journal of Neurophysiology* 100: 1184-1201.

Lombard WP (1892) Some of the influences which affect the power of voluntary muscular contractions. *J Physiol* 13: 1-58.

McFarland J, Pape-Lindstrom P (2016) The pipeline of physiology courses in community colleges: to university, medical school, and beyond. *Adv Physiol Educ* 40: 473-476.

McNair PJ, Depledge J, Brettkelly M, Stanley SN (1996) Verbal encouragement: effects on maximal effort voluntary muscle action. *Br J Sports Med* 30: 243-245

Michael J (2007) What makes physiology hard for students to learn? Results of a faculty survey. *Adv Physiol Educ* 31: 34-40.

Narici M, Maganaris C (2006) Muscle Architecture and Adaptations to Functional Requirements. In: Bottinelli R, Reggiani C eds. *Skeletal Muscle Plasticity in Health and Disease, From Genes to Whole Muscle*: Springer, 265-288.

Nybo L (2008) Hyperthermia and fatigue. *J Appl Physiol* 104: 871-878.

Payne AM, Delbono O (2006) Plasticity of Excitation-Contraction Coupling in Skeletal Muscle. In: Bottinelli R, Reggiani C eds. *Skeletal Muscle Plasticity in Health and Disease, From Genes to Whole Muscle*: Springer, 173-211.

Pette D (2006) Skeletal Muscle Plasticity-History, Facts and Concepts. In: Bottinelli R, Reggiani C eds. *Skeletal Muscle Plasticity in Health and Disease, From Genes to Whole Muscle*: Springer, 1-27.

continued on next page

Russ DW, Vandenborne K, Walter GA, Elliott M, Binder-Macleod SA (2002) Effects of muscle activation on fatigue and metabolism in human skeletal muscle. *J Appl Physio* 92: 1978-1986.

Russell B, Motlagh D, Ashley WW (2000) Form follows function: how muscle shape is regulated by work. *J Appl Physiol* 88: 1127-1132.

Sahlin K, Katz A, Henriksson J (1987) Redox state and lactate accumulation in human skeletal muscle during dynamic exercise. *Biochem J* 245: 551-556.

Sale DG (1988) Neural adaptation to resistance training. *Med Sci Sports Exerc* 20: S135-145.

Sandercock TG (2005) Summation of motor unit force in passive and active muscle. *Exerc Sport Sci Rev* 33: 76-83

Sandri M (2008) Signaling in muscle atrophy and hypertrophy. *Physiology* 23: 160-170.

Schiaffino S, Reggiani C (2011) Fiber types in mammalian skeletal muscles. *Physiol Rev* 91: 1447-1531.

Schiaffino S, Sandri M, Murgia M (2007) Activity-dependent signaling pathways controlling muscle diversity and plasticity. *Physiology* 22: 269-278.

Schiaffino S, Sandri M, Murgia M (2006) Signaling Pathways Controlling Muscle Fiber Size and Type in Response to Nerve Activity. In: Bottinelli R, Reggiani C eds. *Skeletal Muscle Plasticity in Health and Disease, From Genes to Whole Muscle*: Springer, 91-119.

St Clair Gibson A, De Koning JJ, Thompson KG, Roberts WO, Mickelwright D, Raglin J, Foster C (2013) Crawling to the finish line: Why do endurance runners collapse? Implications for understanding mechanisms underlying pacing and fatigue. *Sports Med* 43: 413-424.

St Clair Gibson A, Noakes TD (2004) Evidence for complex system integration and dynamic neural regulation of skeletal muscle recruitment during exercise in humans. *Br J Sports Med* 38: 797-806.

Winter EM, Abt G, Brookes FBC, Challis JH, Fowler NE, Knudson DV, Knuttgen HG, Kraemer WJ, Lane AM, Mechelen Wv, Morton RH, Newton RU, Williams C, Yeadon MR (2016) Misuse of "Power" and other mechanical terms in Sport and Exercise Science Research. *J Strength Cond Res* 1: 292-300 DOI: 10.1519/JSC.0000000000001101.

■

Back to TOC

Academic Performance and Time Allocation of Athletes at a NCAA Division III Women's University

John Pellegrini, PhD[1] and Rosina Hesla[2]
[1]St. Catherine University, 2004 Randolph Ave, St Paul, MN 55105
[2]Richfield Middle School, 7461 Oliver Ave S, Richfield, MN 55423

Abstract
Prior investigations of the academic performance of student athletes have yielded mixed results: while the NCAA's large scale surveys point to classroom success, other researchers have documented academic underperformance by student athletes. The purpose of this study was to examine the grades and the time budgets of student athletes at a NCAA Division III university for women. Results indicated that athletes earned higher grades in anatomy and physiology and had higher cumulative grade point averages than their non-athlete peers. Surveys suggested that, during the season, these student athletes typically spent about 20 hours per week on athletics and a similar amount of time on academics. While they viewed this time allocation as nearly ideal, faculty members thought students should be spending twice as much time on academics as they do on sport. Implications for student advising and faculty attitudes toward athletes are discussed. https://doi.org/10.21692/haps.2018.026

Key words: athletics, athletes, grades, NCAA

Introduction
Many college educators are concerned about resource allocation to athletic programs and the ability of student athletes to balance their efforts between sports and academics. Much of the media attention and research on these topics have focused on the athletically elite, the National Collegiate Athletic Association (NCAA) Division I. However, there are more athletes in NCAA Division III, and athletes make up a much higher percentage of enrolled students at Division III schools (26%) than at Division I institutions (4%) (NCAA 2018a). Division III schools are not allowed to offer financial scholarships for athletics, but some educators worry that the emphasis on sports at many smaller schools is influencing admissions unfairly and compromising the academic enterprise (Strauss 2017, Beaver 2014, Bowen and Levin 2003). Others argue that athletics is helping a variety of colleges draw and retain disciplined, capable, diverse students (Miles 2015, Horton 2009, Melendez 2006, Mendoza *et al.* 2012).

The effects of intercollegiate sports participation on a student's academic work are debatable. Involvement in athletics might enhance academic achievement by fostering disciplined study habits as well as providing the cognitive and emotional benefits that have been frequently demonstrated for exercise (see Buckworth *et al.* 2013). Indeed, some researchers report higher grade point averages (GPAs) or higher retention rates among athletes compared to their non-athlete peers (Horton 2009, Mendoza *et al.* 2012, Baucom and Lantz 2001). Also, the NCAA regularly publishes reports pointing to success among student athletes, claiming "…on average NCAA student-athletes graduate at a higher rate than the general student body" (NCAA 2018a). However, the NCAAs methods have been criticized since their sample of the general student body includes part-time students while their sample of athletes does not (Eckard 2010). Furthermore, some studies have found evidence for negative effects of athletic participation on academics.

The College Sports Project, a large scale, multiyear study, found significant academic underperformance by Division III athletes, even when factoring in disparities such as incoming test scores (Emerson 2012). Similarly, other researchers have found lower GPAs among athletes (Bowen and Levin 2003, Maloney and McCormick 1993) or mixed results (Robst and Keil 2000). Some of the variation in results stems from the observation that athletes in higher profile sports (*e.g.* football and basketball) show more substantial academic underachievement (Maloney and McCormick 1993, Emerson 2012). Another source of variation in athlete to non-athlete comparisons is that, unlike male athletes, female athletes have been found to perform academically as well as non-athlete students (Johnson *et al.* 2010, Kane *et al.* 2008). In addition, examining cumulative GPAs can be problematic, as there are also reports that student athletes cluster within certain majors (Malekoff 2005), and it is possible that the coursework taken by athletes varies in difficulty from that taken by non-athletes, hence the pejorative description of introductory geology as "rocks for jocks".

To understand how athletic participation might influence academic achievement, investigators examined student athletes' schedules for school and sports. Athletes often report that they are better students during their highly structured sports season than they are out of season, but there is little evidence to support this (see Scott *et al.* 2008). Instead, it might be argued that their large time expenditure on athletics during their sport's season could negatively impact their grades. These effects can be difficult to study, as student-athletes might register for fewer classes during semesters when their sports are in season.

continued on next page

There have been several surveys examining student-athletes' balancing of academics and sport. In reply to the NCAA "GOALS" surveys, many student-athletes report spending upwards of 30 hours per week on athletic commitments; they also report missing slightly more than one class per week on average (NCAA 2016). A high percentage of athletes (a majority in some surveys) indicate that they do not spend as much time on academic work as they would like, and that they at least "somewhat agree" that they see themselves more as athletes than as students (Jolly 2008). However, the NCAA's survey data indicate that student athletes spend more time on academics than athletics (NCAA 2016), and some researchers find that athletes spend more time on coursework than non-athlete peers (Miles 2015). In addition, Umbach *et al.* (2016) found that student athletes were equally, if not more, academically engaged than non-athlete students.

How should we view our institutions' and our students' investments in intercollegiate sports? How should we advise students pursuing challenging academic programs while participating in athletics? To help inform these discussions, this article examines the grades and time budgets of student athletes at a women's university that competes in NCAA Division III. By comparing female student athletes to their all-female peers, our sample (admittedly one of availability) eliminates the confounding effect of gender on academic achievement, which is that on average females academically outperform males (Johnson *et al.* 2010). We made cumulative GPA comparisons of all university athletes and non-athletes at our university to enhance the ability to generalize from our study by providing a large sample size. In addition, we make detailed analyses of a large set of anatomy and physiology grades. Most of the student athletes at St. Catherine University list the allied health professions or exercise science as their major on their university webpage biography. We chose to examine the grades in anatomy and physiology because our two-semester sequence of integrated anatomy and physiology is a foundation course for the health care professions and as such, is taken by the majority of student athletes. This provides us with a healthy sample size while controlling for course difficulty and makes it possible for grades to be compared for athletes versus non-athletes within the same classes. The frequent quizzes and tests in our anatomy and physiology classes also allow us to analyze athletes' performance during, and outside of, their sport's season.

To determine if recruitment of athletes was diluting the academic strength of our university's student body we examined athletes' American College Test (ACT) scores. To help determine if participation in college sports influenced academic performance, we compared college grade point average (GPA) to ACT scores for athletes and non-athletes. Finally, to examine time budgeting, we surveyed athletes, coaches, and faculty members with respect to student athletes' time spent on sports and on academic study.

Methods

Academic Performance Study Methods

After obtaining approval of the study from the St. Catherine University Institutional Review Board (IRB Protocol #488), one of the author's (JP) anatomy and physiology class rosters from 2009-2015 (data available from ten semesters) were compared with the published rosters of all the university's athletic teams, which were archived on teams' websites. To identify students who participated in athletics for those semesters, anyone listed on the roster of a team that was active during a given semester was categorized as an athlete. Generally students take these classes during the two semesters of their sophomore year. Athletes' final course grades (on a 100% scale) were then compared to those of students who were not active in athletics that semester. Ninety-four course grades for athletes were compared to 761 grades for students who were not active athletes during those terms.

A broader, university-wide comparison of athletes' grades to non-athletes grades in all coursework was also conducted. Using a large data set provided by the University's Institutional Research Office (which regularly tracks athletic status and grades), we examined the cumulative GPAs (on a 4.0 scale) at the end of the last semester available for all bachelor's degree seeking students who began their first year of college at our university. GPAs for student athletes who transferred to our university or who discontinued athletic participation were not considered. GPAs from fall 2009 through spring 2017 were analyzed (n = 2770 non-athletes, 432 athletes) and, when available, their American College Test (ACT) composite scores were also compared (n = 2584 non-athletes, 415 athletes).

More temporally detailed analysis examined if student athletes' academic performance changed when they were in season, versus out of season, for their sport. Since our anatomy and physiology class involves one or more assessments each week, the grades of 35 student athletes for whom we had two or more test or quiz scores in both the in-season and out-of-season conditions were examined in a within-subjects manner. Averages of all of the students' scores in-season were compared to averages of all their scores out of season, with each quiz and test score normalized to the mean score for the entire class on that particular assessment (in other words, if there score was exactly equal to the class average for a given test, it was scored as 1.0 in our analysis). Seasons were defined by the detailed archived athletic schedules on each team's website. These data were from 2013-2017 (4 academic years).

continued on next page

Time Allocation Study Methods
During the 2015-16 academic year, we surveyed community members (26 athletes, 6 coaches, 10 faculty members) about their perceptions of student athletes' time commitments to athletics and academics. Consent was obtained verbally and paper surveys were distributed on campus, as per the IRB-approved protocol. Using a fill-in-the-blank form, we asked respondents to estimate the number of hours a student athlete devoted to athletics and academics when that athlete's sport was in season and when the sport was out of season. We also asked what they thought the ideal number of hours spent on sport and academics would be.

Results

Grades
Student athletes' course grades in anatomy and physiology were slightly but significantly higher than those of non-athletes as shown in Figure 1A (two-tailed t-test $p = 0.038$; Cohen's effect size, $d = 0.24$). The mean course average for athletes was 82.0% (s.d. = 9.81) while that of non-athletes was 79.5% (s.d. = 10.87). This finding was corroborated and extended by looking at the cumulative GPAs (all coursework; see Figure 1B) of 3,202 students, where the mean athlete GPA of 3.35 (s.d. = 0.53) was significantly higher than that for non-athletes, which was 3.07 (s.d. = 0.73). This result was highly significant (t-test $p < 0.001$) and had a moderate effect size (Cohen's $d = 0.43$).

Student athletes' incoming ACT scores ($n = 416$) were slightly but significantly higher than non-athletes' ($n = 2{,}632$), with a mean composite score of 23.3 (s.d. = 3.58) for athletes versus 22.3 (s.d. = 3.82) for non-athletes (two-tailed t-test $p < 0.001$; Cohen's $d = 0.26$). Thus it did not appear that recruitment of students with athletic prowess weakened the academic preparedness of classes entering our university. When athletes' collegiate academic performance was expressed as cumulative GPA divided by ACT, their performance was still slightly better than non-athletes (Figure 2), indicating that athletic participation did not lead to lower than expected grades, and instead may have enhanced academic success; athlete mean = 0.147 (s.d. = 0.026), non-athlete mean = .140 (s.d. = 0.035; t-test $p < 0.001$; Cohen's $d = 0.20$).

Regarding academic performance in anatomy and physiology within, versus outside of, sports seasons, there appeared to be a slight trend toward higher grades when students were not in season. Twenty-one out of 35 student athletes performed better (relative to the class average) when they were outside of their season compared to when they were in season, but these data did not reach significance in a two-tailed, repeated measures t-test (Figure 3; $p = 0.099$). It should be noted that even when they were in-season, mean athletes' grades were higher than the overall class averages.

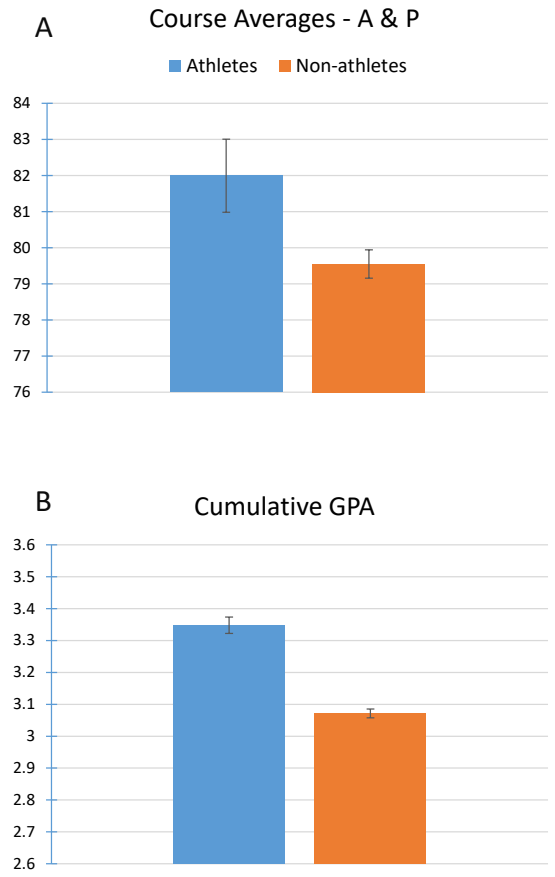

Figure 1. *A. Athletes' mean final course average for anatomy and physiology was 2.4 percentage points higher than that of non-athletes (81.9% vs. 79.4% with SE bars plotted; $p = 0.038$). B. Mean cumulative GPA in all coursework was also higher for athletes (3.35 vs 3.07 with SE bars plotted; $p < .001$).*

continued on next page

Time allocation

Perceptions of the way student athletes budget time are summarized in Figure 4. Student athletes estimated that they spent an average of 19.8 hours on athletics when they were in season and about one hour less than that per week on academics. When not in season, the mean number of hours athletes estimated they spent on sports was 9.0 and the mean number of hours estimated that they spent on academics was 22.2 (data are not graphed). Faculty members and coaches were fairly accurate in their estimates of student time dedication to sports. There were no significant differences between student, faculty, and coach estimates of time allocation for athletics in or out of season. However, coaches thought their student-athletes were spending significantly more time on academics than other respondent groups, as they estimated an average student allocation of 33 hours of schoolwork each week (ANOVA $F = 5.29$; $p = 0.009$; Tukey pairwise comparisons $p < .05$).

When responses to the questions about ideal time allocation were expressed as quotients (the number of hours spent on academics divided by the number of hours spent on athletics) significant differences emerged between groups (one-way ANOVA $F = 7.84$; $p = 0.001$). Post-hoc Tukey tests revealed that faculty's judgment that student athletes should ideally spend more than twice as much time on academics as they do on athletics was significantly ($p < .01$) larger than students' ideal ratio which was roughly 1:1. (Cohen's effect size $d = 1.09$).

Discussion

Within the anatomy and physiology class and across all coursework at our university, the academic performance of athletes in this study exceeded that of non-athletes. Past studies of this type have yielded varied results and have indicated that the academic success of athletes differed when competition level, sport, and gender were analyzed. Specifically, male, Division I athletes participating in high-profile sports (football and basketball) tended to have lower grades and graduation rates than athletes in other demographic groups (Bowen and Levin 2003, Johnson *et al.* 2010, Emerson 2012, Robst and Keil 2000). So it was perhaps not surprising that athletes at our institution, a women's university that competes in Division III, were academically successful. Another sampling issue that we considered was whether struggling students had been eliminated from our athlete group because they became ineligible, thereby skewing our data toward academically successful athletes. An interview with our university's athletic director suggested that this was unlikely, as very few of our athletes (under 1%) had been declared academically ineligible during the study's time period. Still, it was possible that some academically struggling athletes voluntarily left their sport in order to focus on schoolwork.

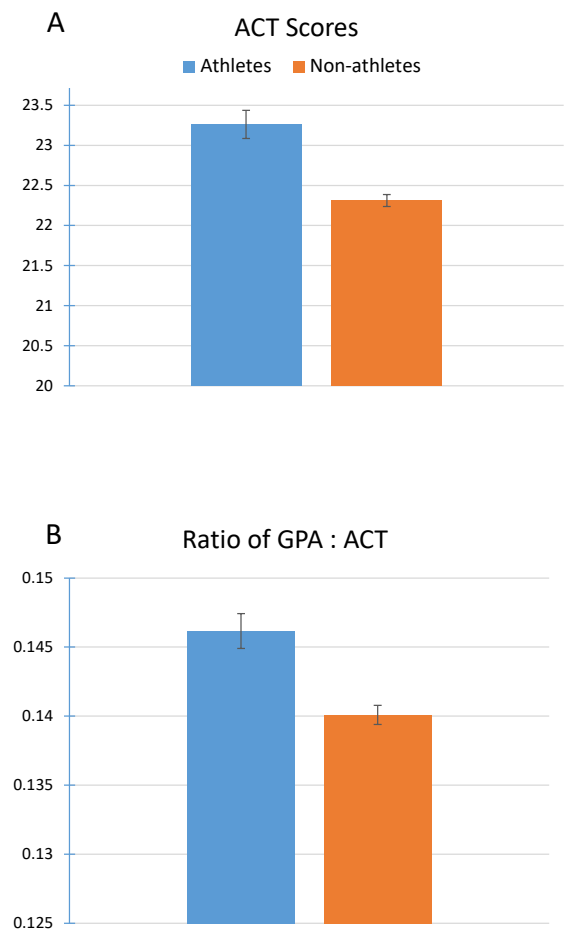

Figure 2. *A. The mean ACT score of athletes was higher than that of non-athletes (23.3 vs. 22.3 with SE bars plotted; $p < .001$). B. Mean cumulative GPA divided by ACT was also higher for athletes (0.146 vs. 0.140) with SE bars plotted; $p < .001$).*

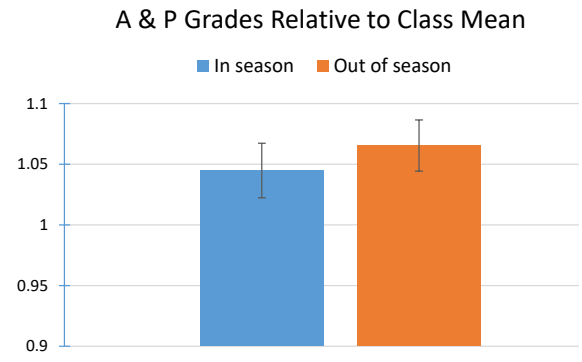

Figure 3. *A. Athletes' grades tended to be slightly higher when their sport was not in season (means and SE bars; $p = 0.99$).*

continued on next page

This study specifically sought to examine whether the decision to participate in athletics might influence a student's academic success, not just whether athletes were stronger students when they arrived at university. While athletes' mean ACT scores were higher than those of non-athletes, the ratio of their GPA-to-ACT was also significantly higher than that of non-athletes, suggesting that participation in athletics may have facilitated their scholastic achievement. Still, athletes should be advised to plan their course schedules carefully, as our findings with anatomy and physiology grades suggested there might be a slight decline in academic performance when their sport is in season. If we had chosen to run a one-tailed t-test, the different anatomy and physiology grades during athletes' competition season would have met the $p < .05$ criterion for statistical significance. Our findings of modestly lower grades in-season are consistent with those of Scott *et al.* (2008) for collegiate athletes and with those of Shultz (2017) for varsity high school athletes.

The high level of academic success of athletes seen in this study may be related to many factors, including the cognitive benefits of exercise, enhanced engagement in college life with team membership, or perhaps a correlation with higher economic status *i.e.* it is possible that throughout their development, athletes' families were better able to afford investments in athletics and education than families of non-athletes (Horton 2015). Future studies might examine these issues by asking athletes and non-athlete students about family income and the number of hours per week they spend at jobs to earn wages.

Ideally, all stakeholders would endorse the NCAA's motto that college athletes are "students first, athletes second" (NCAA 2018b, Vanover and DeBowes 2013). To put this philosophy to the test we should continually examine the time investments of student athletes. Although the NCAA has long had a 20-hour-per-week rule to limit the time spent on athletics, many activities (*e.g.* individual workouts) do not count toward that limit. In our results, student athletes reported spending about twenty hours per week on athletics when they were in-season, and this was one hour more per week than they reported spending on academics. While athletes appeared to be succeeding academically relative to peers, questions regarding how well they fulfilled their academic potential remained largely unaddressed, although the ACT data suggest they are not underachieving relative to peers. As indicated above, our student athletes appeared to score a little higher in anatomy and physiology when they were out of season. This might be related to the increased time they report spending on academics (22.2 hours/week out of season versus 18.8 hours/week in season). However, our survey was given at one point in time, and the respondents' recall of actual hours spent might not have been as accurate as we would hope. Regarding estimates of

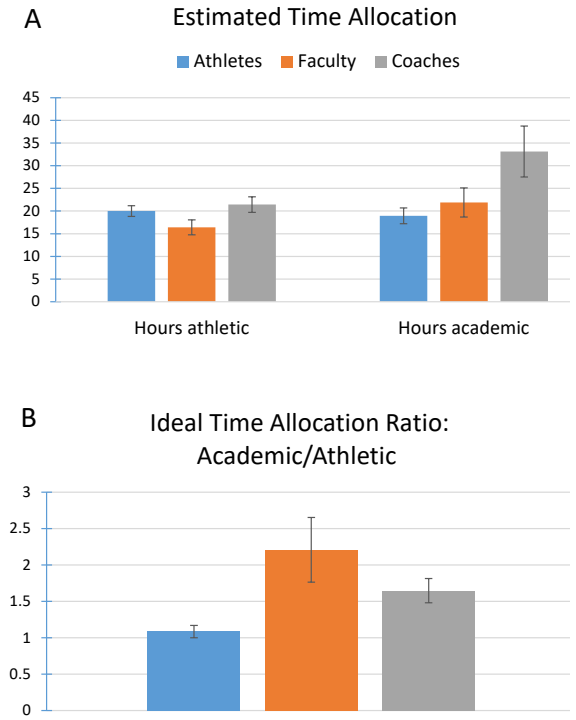

Figure 4. *A. Means of responses (with SE bars) to surveys regarding the number of hours per week student athletes spend on athletic and academic pursuits during their sports' seasons. B. Means of responses (with SE bars) to surveys regarding the ideal number of hours student athletes should spend on athletic and academic pursuits (expressed as ideal number of hours spent on academics divided by ideal number of hours spent on athletics). Faculty's ideal ratio was significantly larger than that of student athletes (p < 001).*

continued on next page

ideal time budgets, it was not surprising that faculty members indicated that it would be better if athletes spent more time on academics than they currently do.

In sum, educators would do well to reject stereotypes about the academic weakness of athletes (see Baucom and Lantz 2001). Many athletes succeed in challenging courses such as the anatomy and physiology courses examined in this study. Our analysis of a fairly large data set finds that the GPAs of Division III athletes across our university exceed those of non-athlete students. To facilitate continued success, instructors, advisors, administrators, coaches, and students should engage in data-rich discussions about time management.

Acknowledgement
The authors wish to thank Scott Pakudaitis for providing sorted institutional data on student grades and ACT scores.

About the Authors
John Pellegrini, PhD, is a Professor of Biology at St. Catherine University where he has taught Human Anatomy and Physiology for twenty-four years. His academic interests include pedagogy, the history of anatomy and physiology, and the biology of emotion.

Rosina Hesla teaches ecology and properties of matter at Richfield Middle School. She is an alumna of St. Catherine University where she competed in intercollegiate swimming.

Literature Cited
Baucom C and Lantz CD (2001) Faculty attitudes toward male Division II student-athletes. *J. of Sport Behav.* 24: 265-276.

Beaver W (2014) The changing nature of Division III athletics. *College & University* 89: 34-40.

Bowen WG and Levin SA (2003) Reclaiming the Game: College Sports and Educational Values. Princeton: Princeton University Press.

Buckworth, J, Dishman, R, O'Connor, P, Tomporowski, P (2013) *Exercise Psychology* (Second Edition) Champaign, IL: Human Kinetics.

Eckard EW (2010) NCAA Athlete graduation rates: less than meets the eye. *J. of Sport Mgmt.* 24: 45-58. doi: 10.1123/jsm.24.1.45

Emerson JD. (2012) The College Sports Project: What have we learned? In *JSM Proceedings*, Statistics in Sports Section Alexandria VA: American Statistical Association: 1121-1131.

Horton, D Jr. (2009) Comparative study of the persistence and academic success of Florida community college student-athletes and non-athlete students: 2004-2007. University of Florida dissertation retrieved from http://etd.fcla.edu/UF/UFE0024321/horton_d.pdf 8/1/18.

Horton, D Jr. (2015) Between a ball and a harsh place: a study of black male community college student-athletes and academic progress. *Comm. Coll. Rev.* 43: 287-305.

Johnson JE, Wessel RD, Pierce D (2010) Reexamining student-athlete GPA: traditional vs. athletic variables. *J. of Intercoll. Sport* 3: 234-255.

Jolly JC (2008) Raising the Question #9 Is the Student-Athlete Population Unique? And Why Should We Care? *Communication Educ.* 57: 145-151. doi: 10.1080/03634520701613676

Kane MJ, Leo P, Holleran LK (2008) Issues related to academic support and performance of Division I student-athletes: A case study at the University of Minnesota. *J of Inter Coll Sports.* 1: 98-130.

Malekoff R (2005) The College Sports Project and the Reform of Division III Athletics. In *Forum Futures 2005* Retrieved from http://forum.mit.edu/articles/the-college-sports-project-and-the-reform-of-division-iii-athletics/ 7/5/18

Maloney MT, McCormick RE (1993) An examination of the role that intercollegiate athletic participation plays in academic achievement: athletes' feats in the classroom. *J. of Hum. Resources.* 28: 555-570.

continued on next page

McCormick RA, McCormick AC (2006) The myth of the student-athlete: the college athlete as employee. *Wash. L. Rev.* 81: 71-157.

Melendez M (2006) The influence of athletic participation on the college adjustment of freshmen and sophomore student athletes. *Coll. Stu. Retention* 8: 39-55.

Mendoza P, Horton D, Mendez JP (2012) Retention among community college student-athletes. *Comm. Coll. J of Res. and Prac.* 36: 201-219. doi: 10.1080/10668921003677183

Miles, L. (2015) Academic achievement in student athletes versus non-student athletes. Rowan University Theses and Dissertations. https://rdw.rowan.edu/etd/475. Retrieved 7/15/18.

NCAA (National Collegiate Athletic Association) (2016) NCAA GOALS study of the student-athlete experience. Retrieved from http://www.ncaa.org/sites/default/files/GOALS_2015_summary_jan2016_final_20160627.pdf 7/9/18.

NCAA (2018a) NCAA Recruiting Fact Sheet. Retrieved from https://www.ncaa.org/sites/default/files/Recruiting%20Fact%20Sheet%20WEB.pdf 8/1/18.

NCAA (2018b) NCAA Amateurism Retrieved from http://www.ncaa.org/amateurism 8/2/18.

Robst J and Keil J (2000) The relationship between athletic participation and academic performance: evidence from NCAA Division III. *App. Econ.* 32: 547-558.

Scott BM, Paskus TS, Miranda M, Petr TA, McArdle JJ (2008) In-season vs. out-of-season academic performance of college student-athletes. *J. of Intercoll. Sport* 1: 202-226.

Shultz K (2017) Do High School Athletes Get Better Grades During the Off-Season? *J. Sports Econ.* 18: 182-208. doi: 10.1177/1527002514566279

Strauss B (2017) The liberal arts football factory. *Slate*. Retrieved from https://slate.com/culture/2017/12/wesleyan-university-football-is-good-business.html 6/21/2018

Vanover ET and DeBowes MM (2013) The impact of intercollegiate athletics in higher education. *Higher Educ. Pol & Econ* 1:1-24.

■

Back to TOC

Can Modern AI replace teachers? Not so fast! Artificial Intelligence and Adaptive Learning: Personalized Education in the AI age

Vasiliy Kolchenko, MD/PhD
Department of Biological Sciences, New York City College of Technology, The City University of New York, 300 Jay St., Brooklyn NY 11201
vkolchenko@gmail.com

Abstract
Artificial intelligence (AI) is transforming many fields, including education. Can modern AI replace teachers? We discuss some popular AI applications for adaptive learning, how they are related to personalized education, what makes these applications "intelligent," and how the quality and quantity of student-generated data changes the behavior of adaptive learning systems and the learning experiences of students. If the data are scarce and corrupt, the benefits of adaptive learning are minimal. Motivated and persistent students generate greater amounts of high quality data and, as a result, tend to have better learning experiences. This can increase the gap between high achieving and low achieving students. More realistic student models and better understanding of the pedagogical context are needed to improve the performance of the AI educational systems. These are some of the reasons why experienced teachers cannot be replaced by the current AI applications.
https://doi.org/10.21692/haps.2018.032

Key words: artificial, Intelligence, adaptive, learning, AI

Is AI everywhere?
We live in the age of artificial intelligence (AI). Pervasive, but mostly hidden, it is taking over many jobs, devices, technologies and media, promising to transform all fields, from industry and commerce to medicine and education (Hawkins and Blakeslee 2007, Nilsson 2009).

Increasingly, our actions and decisions are guided by AI (Bostrom 2014). It tells us how to drive to our destination (GPS), what to buy (online shopping and ads), what to read (online search) and whom to date (dating sites).

Can it also tell educators how to do their jobs better, or even replace teachers completely? After all, AI is about to start replacing many professionals, from truck drivers to radiologists.

Reliance on AI is inevitable because in many situations modern AI is more effective and much more efficient than people. Not only can it win in chess and Jeopardy, it can also evaluate radiological images better than radiologists and avoid traffic accidents better than truck drivers, saving lives in both cases.

In education, AI promises to deliver what has always been the highest goal of pedagogy, wise and caring guidance for each student, adapted to the individual's needs. AI would potentially provide customized learning resources and activities, combined with the pace and style of instruction that best suits each individual student (Luckin *et al*. 2016).

As a result, all learners would have their own highly intelligent digital tutor (Andriessen and Sandberg 1999, Beck *et al*. 1996) and the problem of personalized education could be solved once and for all. The remaining teachers would supervise and adjust this computational *adaptive learning*.

But can modern AI actually deliver this? What can AI do at the moment and what are the problems that slow down its progress?

What is AI?
AI is the computing that aspires to do what people do, only better. It encompasses perception (computer vision and hearing), reasoning (understanding and generating language in forms such as Siri and Alexa), learning, prediction, making decisions, and performing actions (robots and self-driving cars).

To do all of this intelligently, modern AI needs effective learning algorithms and a lot of high quality data for algorithm training (Russell and Norvig 2010, Domingos 2015). The most popular and successful algorithms today are *artificial neural networks*, biologically inspired programs based on the neural networks of the brain.

The data are specific to the applications. For example, all our choices online – what we click, read, watch, and buy – are valuable data for the AI algorithms that are trying to figure out how to serve us better (and also how to sell us more goods and services).

continued on next page

The data feed learning and learning improves the program's intelligence, which means that the AI makes better predictions and decisions. The more high-quality data are available for training, the better the outcome. This continuous cycle of self-improvement reinforces the program's abilities. It is similar to some positive feedback mechanisms in physiology.

But if the data are scarce or confusing, the AI program may perform progressively worse, becoming more clueless instead of smarter and smarter. For example, if you are clicking on the online ads at random, you are providing AI with meaningless, confusing data. In this case, AI needs to recognize this immediately. Otherwise it will be learning from corrupt data, which may lead to wrong predictions and bad decisions. Artificial intelligence specialists informally call this "garbage in, garbage out."

Something similar can happen in adaptive learning. In fact, it is quite likely to happen, considering what students are supposed to do while using many popular programs.

What is adaptive learning?

Adaptive learning (AL) is an educational approach that uses computing algorithms in order to organize the interactive learning experience and adapt to the individual needs of each student. Therefore, it is part of personalized education. Another type of personalized education is, for example, one-on-one tutoring, with no computers involved. This type of personalized education has been around since the times of Socrates, and probably even predates Socrates.

Modern AL is mostly intelligent. This means that it uses AI. There are some straightforward computer algorithms in educational programs that do not use AI, but they have become rare. Not all adaptive learning systems are intelligent, and not all intelligent learning systems are adaptive, but for the most part they intersect (Brusilovsky and Peylo 2003).

A simple example would be a computer program that provides various learning resources and engages students in the interactive learning (Sleeman and Brown 1982). When a student makes a mistake answering a question, the program does not simply give him another similar question. Instead, it searches for the patterns in all of the student's prior interactions with the program and analyzes them. Based on this, it creates a model of the student and uses it for the optimal response (Brusilovsky and Millan 2007).

This response, which might be a suggestion to review a relevant paragraph, to watch a short video, or to answer a series of increasingly difficult questions, depends not on just one wrong answer but on the complete knowledge about the student that the program has accumulated so far. It is different for each student because each student has interacted with the program in a different way. The program remembers all of these interactions and makes its own conclusions about the appropriate resources, pace, and style of learning for each student.

The objective is to stay in the zone of proximal development (Murray and Arroyo 2002), making the learning suggestions not too easy but also not too hard. If the learning suggestions are too easy, the student's experience gradually becomes boring. If they are too challenging, the student may become increasingly frustrated. In a way, these programs are similar to computer games, which are designed to keep the players interested, occupied, and even addicted to the game, all of this being critical for the game's success.

Many big publishing companies invest millions of dollars in adaptive learning programs that complement popular textbooks. These applications are available either online or on desktop computers (Daniel *et al*. 2015). The companies are compelled to invest heavily in adaptive learning programs, fearing competition and hoping for a high dividend. This is a new business model for publishing companies because highly profitable textbooks are not selling as much as they used to and AL access is the new source of profits. Access to adaptive learning programs cannot be stolen as easily as free PDFs of the textbook. More importantly, AL applications can eventually replace textbooks as the main educational resource of the future.

How do the students react? Two divergent scenarios.

For a few semesters, all of my anatomy and physiology students systematically used online AL resources provided by the publisher. Every semester my classroom observations and student comments were essentially the same. The feedback seemed consistently inconsistent and ambiguous. Then patterns started to appear.

As the class begins using AL resources, students usually respond in surprisingly different ways. Some love the experience and others hate it. The two common scenarios that suggest possible reasons for this wide range of reactions are described below.

Motivated, persistent students use the AL programs a lot, reading and answering many questions, reviewing all the suggested materials and thinking hard about the topic. In other words, they do what the program expects them to do. In the process, they supply the program with abundant data of good quality. The program uses the data to learn about the student, adapting to the student's individual learning needs. This makes the learning experience even more exciting and rewarding for the student, further motivating the learner to work hard.

continued on next page

Now let us consider another group of students who are easily distracted and bored. They are not very persistent and motivated in learning. These students tend to skip practicing altogether. Even if they start answering the questions, they may stop early. Overall, they generate much less data than motivated, persistent students. With scarce data, the AI program cannot help these students very much, which makes it even more difficult for students to complete the assignment.

Another problem with easily distracted students is that the quality of the data is usually not very high. If these students are forced to reluctantly answer all the questions, they may start clicking on the answers at random, without reading or thinking about them at all. This behavior is easy to anticipate since the students are trying to complete the assignment as fast as possible instead of trying to learn as much as possible. These students may quickly become bored and frustrated.

However, randomly clicking on answers is not what the program expects students to do. The program becomes confused. It gives the student new questions that make as little sense to the learner as the old ones did. The frustration grows. What was supposed to be an easy way out – clicking through the questions, as useless as that is – becomes an increasingly difficult path. This may account for some persistent and very personal student responses I had in my anatomy and physiology classes, for example, "I hate this program so much!"

In this case, the low quantity and poor quality of student data makes AI "blind." It becomes as dumb and ineffective as it can be ("garbage in, garbage out"). This also explains, to some extent, why there are always a few students who benefit from AI adaptive learning much more than others.

The paradox is clear: adaptive learning may help the most the ones who need this help least. It also may help the least the ones who need this help most. As a result, instead of decreasing the gap between the high achieving and low achieving students, these AI tools can actually increase the gap.

Interestingly, it is the same AL game for all, but for some it is exciting and even addictive and for others it is boring and tedious. It very much depends on how you play it from the start. The harder you try, the better it becomes. The more you skip and cheat, the harder it is to play it. The program can adapt to the learner, but the learner needs to adapt as well, in order to get the greatest benefits.

The AL experiences of my students were program-specific, but the AL achievement gap paradox is not. This hypothesis is rooted in the basic understanding of the modern data-driven AI, machine learning. The user provides the data. Good data means effective AI. Poor data means ineffective AI. It cannot be otherwise.

Unfortunately, there are no rigorous studies of the AL and achievement gap in the current literature. Hopefully, close attention to this problem will help improve performance of the AI educational systems. Students who need more assistance should get more, not less, from AL programs.

Can modern AI replace teachers?

The AL programs discussed in this article cannot replace an experienced educator. These programs can be useful learning tools, especially for some students. They can help teachers by providing additional student training that frees classroom time for higher level learning activities. However, adaptive learning programs can also be easily confused by corrupt or scarce student data. Adaptive learning programs do not understand the wide range of the all-important pedagogical contexts. They do not know how to deal with students who habitually avoid learning or are afraid of learning and have never had a positive learning experience. The student models created by adaptive learning programs may be too simplistic and they are often unrealistic.

Teachers, on the other hand, have always been engaged in their own version of adaptive learning. We adapt our teaching to the particular class and to individual students, as much as we can, based on our experience, intuition, testing results and the available "bag of tricks," which has the limited selection of educational resources and techniques we have at our disposal.

The feedback we use for creating our own "student models" includes very important non-verbal clues in the classroom, such as facial expression, posture, movement, immediate reactions, and interactions in groups. We build personal working relationships with many students and respond in class and online with good humor and grace.

Can AI emulate all of this in the future? The next generation adaptive learning programs are on the way (Lane *et al.* 2016, Luckin *et al.* 2016). They will be much more sophisticated in their understanding of pedagogical context and in the range of potential adaptations to the learning needs of the student. It is possible that future adaptive learning pedagogical programs will change the way we learn.

About the author

Vasiliy Kolchenko is a professor of biology at New York City College of Technology, The City University of New York. Vasiliy teaches anatomy and physiology, pathophysiology, and bioinformatics. He is developing a new introductory course in artificial intelligence and neuroscience. His research includes biosensor development and graphic representation in science education. He also writes and performs music. This is his Teaching Science song: https://www.youtube.com/watch?v=CpeI5wHvKE4

continued on next page

Literature cited

Andriessen J and Sandberg J (1999) Where is education heading and how about AI. *International Journal of Artificial Intelligence in Education* 10.2: 130-150.

Beck J, Mia S, Haugsjaa E (1996) Applications of AI in education. *Crossroads* 3.1: 11-15.

Bostrom N (2014) Superintelligence: Paths, Dangers, Strategies. Oxford University Press.

Brusilovsky P, Peylo C (2003) Adaptive and intelligent web-based educational systems. *International Journal of Artificial Intelligence in Education (IJAIED),* 13: 159-172.

Brusilovsky P, Millán E (2007) User models for adaptive hypermedia and adaptive educational systems. In: The Adaptive Web (pp. 3-35). Berlin, Heidelberg: Springer.

Daniel J, Esteban VC, Mercè GC (2015) The future of MOOCs: Adaptive learning or business model? *International Journal of Educational Technology in Higher Education* 12.1: 64-73.

Domingos P (2015) The Master Algorithm: How the Quest for the Ultimate Learning Machine will Remake Our World. Basic Books.

Hawkins J, Blakeslee S (2007) On Intelligence: How a New Understanding of the Brain will Lead to the Creation of Truly Intelligent Machines. Macmillan.

Lane C, Gover S, Roschelle J (2016) AI Applications in Education. *Menlo Park, CA, USA: The Center for Innovative Research in CyberLearning*.

Luckin R, Holmes W, Griffiths M, Forcier L (2016) Intelligence unleashed: An argument for AI in education. London: Pearson Education.

Murray T, Arroyo I (2002) Toward measuring and maintaining the zone of proximal development in adaptive instructional systems. *International Conference on Intelligent Tutoring Systems* (pp. 749-458). Berlin, Heidelberg: Springer.

Nilsson NJ (2009) The Quest for Artificial Intelligence. Cambridge University Press.

Russell SJ, Norvig P (2010) Artificial Intelligence: A Modern Approach (Third Edition). Pearson Education Limited.

Sleeman D, Brown JS (1982) Intelligent Tutoring Systems. London: Academic Press.

Back to TOC

"How Do You Know If They Help?" Implementing Multiple Student-Centered Learning Opportunities in Human Anatomy and Physiology Undergraduate Labs

Heather A. Rudolph, PhD[1], Anna Schwabe[2], Nastaran Soleimanibarzi[2]
[1]Gillette College, 300 West Sinclair, Gillette, WY 82718
[2]School of Biological Sciences, University of Northern Colorado, Greeley, CO 80639
hrudolph@sheridan.edu, schw0701@bears.unco.edu, sole9426@bears.unco.edu

Abstract
To improve the student learning outcomes of an introductory Human Anatomy and Physiology course we have implemented many changes over the last two years. In this study, the third of a three-part series, we present data from surveys collected and analyzed during the spring semester of 2018 at the University of Northern Colorado. The results reveal that the most appreciated change for most students, regardless of the final grade earned, was the affectionately named "Vanessa videos", which are described in parts one and two of this series. According to the students surveyed, the Exit Quizzes were the most emotionally distressing learning opportunity. Grade improvements during the semester ranged from 0.07-13.4%. Having access to a variety of learning opportunities, such as customized online videos, student-made study guides, and frequent low-stakes assessments appears to provide the greatest benefit to the widest range of students. https://doi.org/10.21692/haps.2018.027

Key words: student-centered learning, survey study, customized teaching, flipped learning, learner-centered teaching

Introduction
In 1925, Dewey described learning as:
"emerging through experience, that is, action within the world in which we live, action related to solving problems and meeting ends, and through that action, building new structures of knowledge and understanding" (as cited in Bresler 2004 p. 198).

Eighty years later, neuroscientists have provided tangible evidence that physical actions stimulate learning areas of the brain related to language and that language about physical actions stimulates areas of the brain that prepare the body to perform those actions (Pulvermüller 2005). However, Avraamidou and Osborne (2009) suggest that while scientific concepts are grounded in scientific inquiry, they are often presented in scientific language, creating comprehension challenges for non-scientists. Anatomy and physiology terminology could certainly be considered a language of its own. Despite almost 100 years of research supporting active hands-on experiences as effective ways of learning, many traditional human anatomy labs still rely on memorization, use of mass-marketed lab manuals, and few high stakes assessments to evaluate learning objectives (Gopalan 2016, McDaniel and Daday 2017, Rawson and Dunlosky 2012).

As a postdoctoral fellow hired for a 20/80 teaching/research position in 2015, I was enthusiastic to start my own research projects and spread my fascination with qualitative methods to others in my new department while improving my limited quantitative research skills. After a year of teaching introductory biology to incoming students majoring in the field, I transitioned to teaching my favorite subject, Introduction to Human Anatomy and Physiology. Aside from wanting to become a more adept researcher, I wanted to apply the research theory learned while earning my PhD through the inclusion of active learning techniques. Ultimately, I sought to make the labs and lectures a positive and effective learning experience for my students. I met with my graduate teaching assistants (GTAs) before the semester began and, together we sought opportunities to improve the structure of the labs and lectures. This study looks specifically at lab improvements. Fortunately, the three GTAs assigned to teach the labs that semester had extensive learning and teaching experiences specific to human Anatomy and Physiology courses. Through collaboration and peer learning, both of which are recognized in the literature as effective forms of active learning (Freeman *et al*. 2014, Hughes 2011, Moyer 2016), we made several changes to the structure of the lab that we hoped would also promote positive learning outcomes for our students (Rudolph *et al.* 2018, Rudolph and Schwabe 2017).

These changes have evolved over the last two years and have been applied to labs encompassing nearly 1,000 students. The purpose of this study was to examine how undergraduate students utilize these learning opportunities that are now embedded in the structure of Anatomy and Physiology labs at the University of Northern Colorado (UNC). We were particularly interested in (1) which learning opportunities do students report using most (2), if the use of learning opportunities is correlated with grades (3) if learning opportunity use changes over the semester, and if so, why and (4) if the number and/or types of learning opportunities relate to grade improvements in lab.

continued on next page

Materials and Methods

All students enrolled in introductory Anatomy and Physiology labs at UNC (N=327) were invited to participate in this study during the spring semester of 2018. There are no prerequisite courses or minimum GPA required for enrollment in the course. The enrolled students were provided a consent form during the first week of the semester to sign if they were willing to participate in the study. The researchers were in contact with the students throughout the semester, but completed surveys were stored and not analyzed until completion of the semester and final grades had been disseminated. One hundred and seventy-eight students gave their consent and participated in the first survey (Appendix A), while ninety-six students responded to the second survey (Appendix A). All participants were 18 years of age or older and did not represent any known, vulnerable populations. Random numbers were assigned to all participants and names were removed to protect identity before data analysis began.

The Internal Review Board of the University of Northern Colorado approved this project, IRB# 1195188-1, and informed consent was obtained from all participants. Informed consent allowed us to distribute the surveys, access lab grades for data analysis, and take pictures of students doing activities during labs. Pictures will be used in conference presentations and potential future articles.

Survey 1 was conducted after grades were recorded following the 1st practical exam. Surveys were provided to all students enrolled in the course, but only those surveys for consenting students were included in the analysis. Survey 1 included five questions. Question 1 provided a list of learning opportunities students may have participated in and asked them to check the boxes of the ones they used to prepare for their practical exam. Responses were recorded as presence-absence data to indicate which opportunities students utilized and which were used most often. The remaining survey questions were open ended.

The second survey, Survey #2, was administered after the completion of the final practical. Survey #2 included the original five questions plus one additional question that allowed us to assess changes the students made in study behavior over the course of the semester. Both surveys were conducted during scheduled laboratory periods and required less than ten minutes to complete (Appendix A).

Data for the first survey question were entered into Microsoft Excel spreadsheets for analysis using "1" to indicate learning opportunity used and "0" to indicate learning opportunities that were not used (*i.e.* binary format). Learning opportunity categories for the first survey question included the following:

Term sheets: This is the list of terms students are tested on. We included this as a learning opportunity option on the survey and nearly all students marked that choice. However, we did not include results of term sheet use in this paper because we felt that knowing the terms on the list is a requirement for the course and it is not truly a "learning opportunity."

Vanessa videos: These are online instructional videos of one of our GTAs, Vanessa Johnson, clearly walking the viewer through all of the terms on the list for each practical using the models available in our lab. Vanessa also provides tips and tricks to help students remember the information. These are referred to as the "Vanessa Videos."

Homework PowerPoints: These are student-made electronic photographic study guides in a PowerPoint format synthesized during lab times under the supervision of GTAs and/or undergraduate teaching assistants (UTAs).

Pre-Labs: These homework exercises are completed by the students prior to lab as a first learning opportunity to help students become familiar with terms.

Entrance Quizzes: Cumulative entrance quizzes are conducted at the beginning of each lab to reinforce and check retention of material from the previous weeks' labs in preparation for the next lab practical. The entrance quizzes consist of 10 one-point questions, and are presented in a PowerPoint slide show with photographs of the models used in lab. Quizzes are cumulative and include material from previous labs, but not from material that has not yet been covered.

Exit quizzes: These are five-question quizzes are given at the end of each lab to check for learning progress during the lab period that day. These quizzes are in the same format as the practical exams in that there is a station with a model indicating a structure that the student is asked to identify.

Practice Practical: This is a low-stakes mini practical exam the students take one week prior to the first practical exam. Many students have never taken a practical exam and this experience is designed to assess the effectiveness of their study techniques. The practice exam is peer-graded in class to provide immediate feedback, so students have time to adjust studying techniques prior to the high-stakes practical exam. Since a practice practical was only administered before practical #1, we do not include it in the survey #2 analysis.

Open Labs: GTAs hold office hours as "open labs" when labs are not scheduled, increasing the time available for students to study under the guidance of a knowledgeable supervisor. We divided this into two categories to explore whether students were making a point of going to open lab with their respective GTA, or any open lab, categorized as "Open Lab (any)" and "Open Lab (with your GTA)."

Mini Lecture PowerPoint Presentations: After entrance quizzes are conducted, GTAs give a short (~10 minutes) PowerPoint lecture. These provide labeled figures designed to help students identify and understand topics and structures for the current week's lab. The PowerPoints can be helpful to students

continued on next page

when they are looking for the structures to label while doing their Homework PowerPoints. They are available for students on the Learning Management System (LMS) for further study outside of lab times.

Meeting with GTA: This category refers to an appointment a student makes for additional one-on-one time. This was provided as a category as some students' schedules conflicted with official open-lab office hours. The mini-lecture is simply an overview of the current week's material. For this reason, we did not consider instructional differences among GTA's.

Class Lecture Notes: Since we have worked to align lecture and lab topics, we included this as a category as a learning opportunity.

Other: This category was added to learn what students were using to study that we might include at a later date. It included a prompt to explain what was used if the box was checked.

Quizlet: On Survey #1, we noticed several students writing in "Quizlet," which is a free online, flashcard program where students can make their own quizzes and can access quizzes other students have made. Given the high occurrence of student reference to this resource in the "Other" category, we created a category in the analysis specifically for Quizlet and collected data for Survey #1 and Survey #2.

Learning categories with positive values were summed to determine the frequency of usage for each learning opportunity. Scatterplots and R^2-values were generated in Microsoft Excel and used to address research questions 2 and 4. Excel generated the scatterplots for each learning opportunity and 12 graphs for the relationship between learning opportunities and final grade received in the lab were produced. The trend lines for the graphs were calculated and R^2 values are reported in Table 1. R^2 is a statistical measure of how close the data are to the fitted regression line and it is also known as the co-efficient of the regression. We chose not to use chi-squared test because our regression analysis was as informative as chi-squared analysis for our research. We also generated 11 scatterplots to determine the relationship between the learning opportunities and overall grade improvement in the class and inserted the trend line. R^2 was calculated and is shown in Table 2 for these scatterplots. To analyze question 4, "How often did you use the learning opportunity(ies) that you found most helpful?", we grouped students' responses to the question into six categories and created codes for each. The categories were: never, 1-2 times per week, 3-4 times per week, 5-6 times per week, every day, and vague (for example, "as often as I could). Question 5 on Survey 1 and question 6 on Survey 2 asked for suggestions to improve the labs. Question 5 on Survey 2 asked if students changed how they studied during the semester and if so, why?

We were interested in variables that are predictive of grade improvement. Therefore, individuals that did not represent our question of interest (grade improvement) were intentionally removed from the dataset. We calculated grade changes and removed those participants who specifically did not improve over time. We felt that including students whose grades did not improve would confound the results of interest (grade improvement) and the possible relationship with the various ancillary materials.

Results

The following reported results only include students who provided consent to participate in this study. By filling out the surveys we determined students were inherently reporting their preference of their study methods. The correlation, signified by R^2, measures the amount of linear association between each learning opportunity and their grade improvement. The higher the R^2 the more significant the relationship between the student's usage of learning opportunities and their grade improvement. Based on our sample size and the grade improvement, we considered an R^2 higher than 0.35 as showing a stronger relationship.

The learning opportunities students reported using to prepare for the 1st practical were the highest for Homework PowerPoints (76.40%), followed by the Practice Practical (47.75%), Pre-Labs (44.38%), and the Vanessa Videos (43.25%). To prepare for the 2nd practical, the highest reported usage by students was, again, Homework PowerPoints (73.96%). However, their second highest choice for studying was Vanessa Videos (56.25%), followed by Pre-Labs (38.54%), and Class Lecture Notes (32.29%) (Figure 1).

Table 1. *Correlation of learning opportunity with final grade received in the lab section of the course. Students who reported using the Vanessa Videos had the greatest positive correlation on the final grade received in lab, while Entrance Quizzes had the greatest negative correlation with the final grade received in lab. The Homework PowerPoints as a study tool did not seem to correlate with the final grade received in the lab.*

Opportunity Used	R^2
Vanessa Videos	0.563
GTA Specific Open Lab	0.432
Lab Mini-lecture PowerPoints	0.361
Open Lab	0.175
Other	0.150
Class Lecture Notes	0.080
Quizlet	0.072
Homework	0.000
Meeting with TA	-0.006
Exit Quizzes	-0.039
Pre-lab Exercises	-0.247
Entrance Quizzes	-0.671

continued on next page

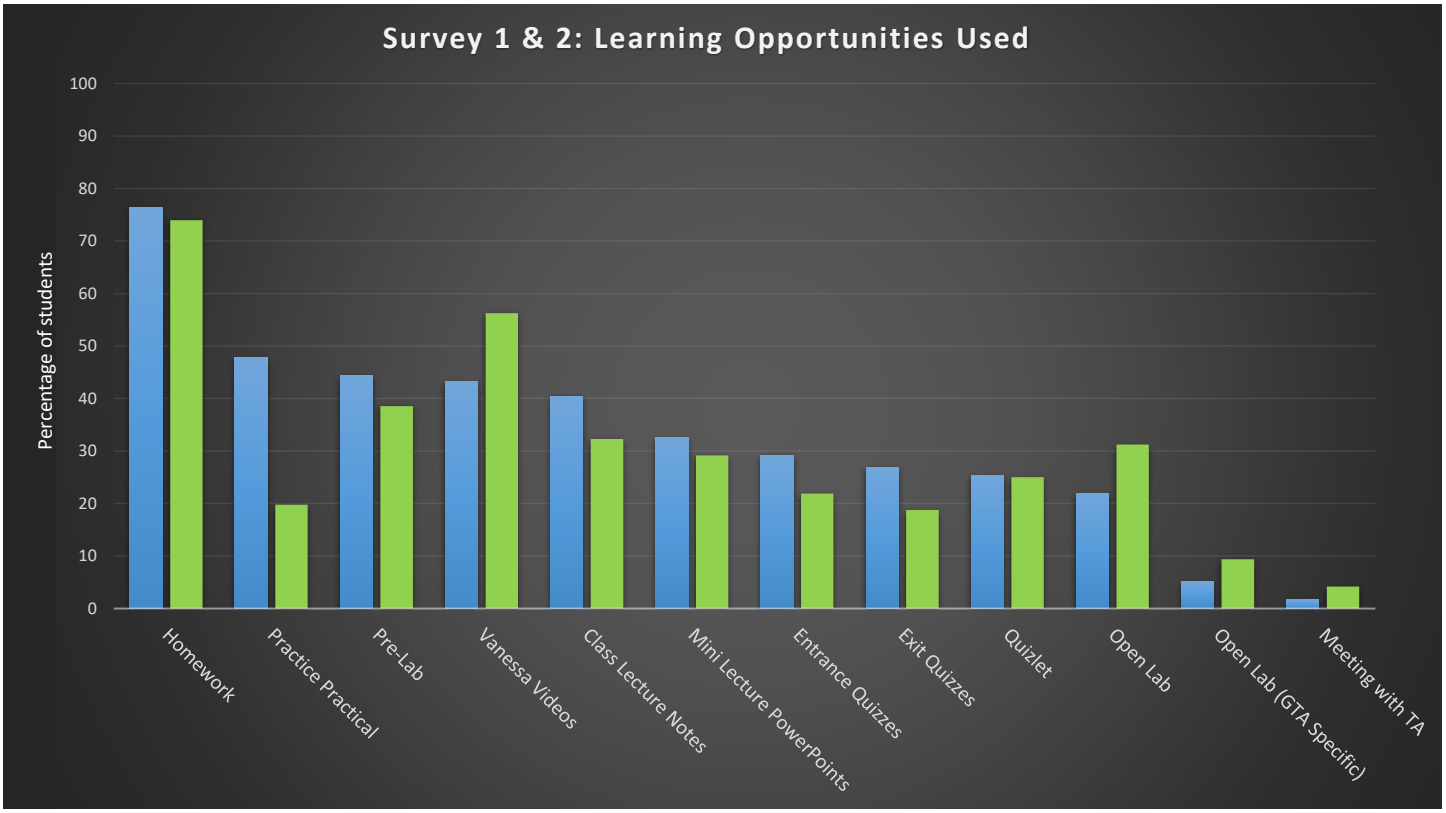

Figure 1. *Comparisons of Survey 1 (blue) and Survey 2 (green) with the learning opportunities students reported using as a study tool.*

In order to determine if the use of our learning opportunities was related to grade improvement, we implemented a dual strategy. First, the data from Survey 2 was used to compare the average number of opportunities used by students in each final grade category (A, B, C, D, F). Grades were categorized on the following scale: 90% (A), 80% (B), 70% (C), 60% (D), < 60 % (F). We found no relationship in average number of learning opportunities used and final grade earned. The second approach compared learning opportunities reported in Survey 2 with final grade, for which a strong relationship was observed (Table 2). In short, the Vanessa Videos, GTA-Specific Open Labs, and Lab Mini-lecture PowerPoints had the highest positive relationship with grades, with R^2 values of 0.563, 0.432, and 0.362 respectively. There was a slight positive relationship with grade improvement for students who sought out other learning opportunities ($R^2 = 0.150$) and/or those who went to non-GTA-specific Open Labs ($R^2 = 0.175$). Utilizing Homework PowerPoints as a learning tool had no measurable effect on the grade received ($R^2 = 0.000$). Several learning opportunities showed a negative relationship with the final grade ($R^2 = -0.671$ to -0.006), of which the Entrance Quizzes had the highest negative impact (-0.671).

Table 2. *Correlation between learning opportunity utilized with overall improvement (%) in the lab. Utilizing the Homework as a study tool was highly correlated with grade improvements between the first practical and the final grade received in the course. Students who reported watching the Vanessa Videos and visiting their GTA during open labs did not have overall improvements in their grades (discussed in text).*

Opportunity Used	R^2
Homework	0.929
Lab Mini-lecture PowerPoints	0.854
Other	0.789
Class Lecture Notes	0.288
Pre-lab Exercises	0.239
Entrance Quizzes	0.113
Quizlet	0.041
Exit Quizzes	0.009
Open Lab	0.004
GTA Specific Open Lab	-0.32
Vanessa Videos	-0.65

continued on next page

In addition to investigating whether there was a relationship between the final grade earned and learning opportunities used, we also calculated whether grade improvements correlated with the number and/or type of learning opportunities used in lab. We used total lab grades at two separate checkpoints: checkpoint 1 (following practical 1) and checkpoint 2 (end of semester). The difference in total grades between the two was calculated and used to determine grade improvement for each student who completed surveys 1 and 2. Grade improvements ranged from 0.07% - 13.4%. Approximately 70% of students who completed both surveys (n = 96) improved their grades during the semester. Students who did not complete both surveys were not included in this portion of our analysis because we would have no way to determine if their behavior changed during the semester. We divided grade improvements into two broad categories: major (5% and above) and minor (below 5%) (Figure 2). Since this question asks specifically about grade improvement, we did not include data from students whose grades decreased between the first practical and the end of the semester. When analyzed by number of learning opportunities used with grade improvement category, there was no significant difference between groups

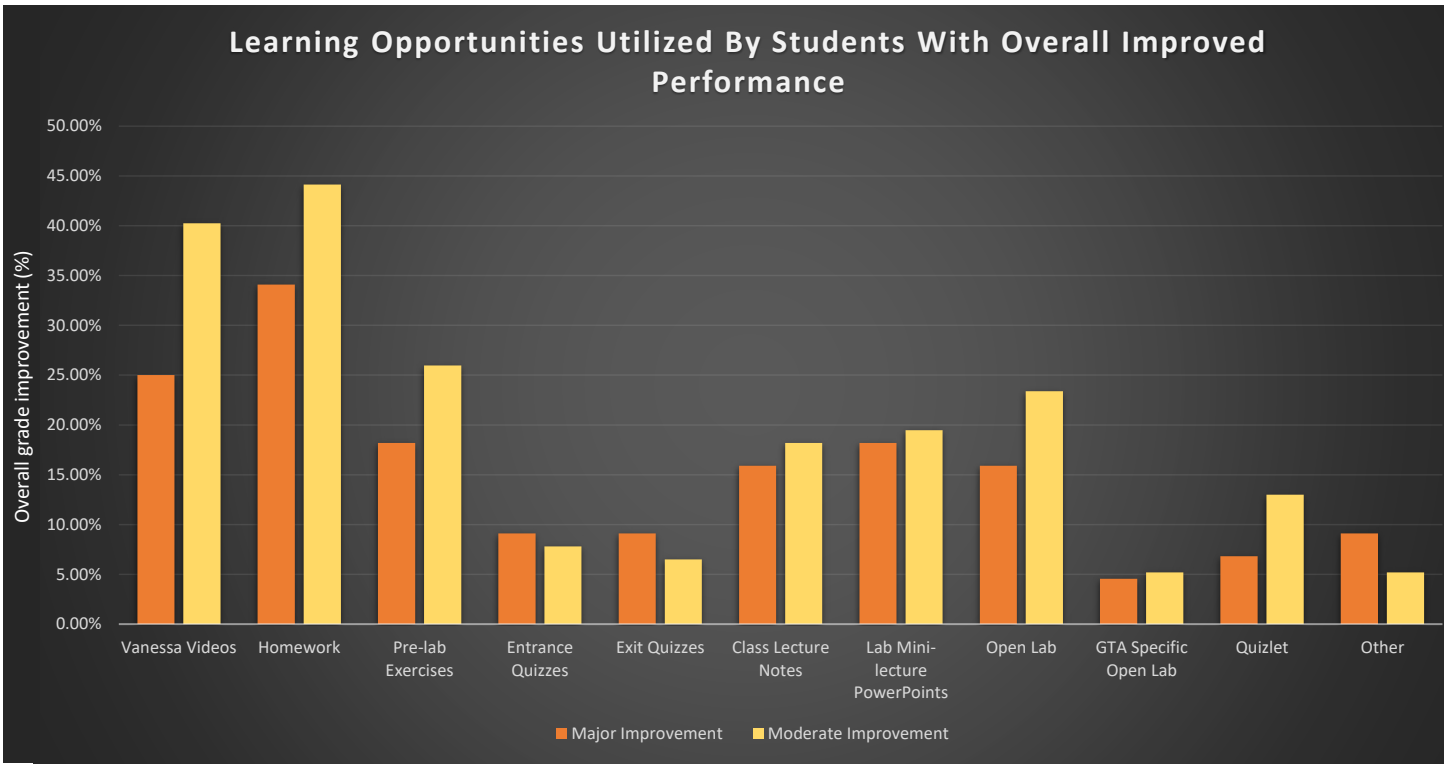

Figure 2. *Comparisons of major (orange; >5% overall grade improvement) and moderate (yellow; 1-4.99% overall grade improvement) grade improvement with the learning opportunities students reported using as a study tool.*

continued on next page

We found some learning opportunities may affect grade improvement. For example, Homework, Mini-lecture PowerPoints, Other, and Class Lecture Notes all had positive relationships ($R^2 = 0.929$, 0.854, 0.789 and 0.288 respectively; Table 2). Vanessa Videos and Open Lab with GTA both show negative relationships ($R^2 = -0.65$ and -0.32 respectively; Table 2). We attribute this to the high number of A students reporting use of both learning opportunities. We calculated that 81% of students who earned A's were found in the two lowest "grade gain" categories.

In addition to investigating which learning opportunities students reported using and their relationships to grades, in the open ended question responses students reported using Homework PowerPoints, the term sheet, and Other most often on Survey 1 (25.5%, 25.2% and 21.0%, respectively. On Survey 2, students reported highest use of Homework PowerPoints, term sheets and Vanessa videos (31.8%, 21.5%, and 17.0%, respectively). On Survey 1, students reported PowerPoint Homework, term sheet, and other as most helpful (29.9%, 25.7%, and 16.4%, respectively). Survey 2 results indicted a change with students reporting Vanessa videos, Homework PowerPoints, and Other as most helpful (26.5%, 25.7%, and 16.9%, respectively.) Most students reported studying frequency as one to two times per week, closely followed by three or four times per week, and then vague responses that we were not able to quantify (27.9%, 27.3%, and 19.2%, respectively). Examples of answers counted as vague include "A lot" and "Weekly." Survey 2 showed slight changes in their answers but still reporting one to two times per week as most common, followed by three to four times per week and vague (28.1%, 25.8%, and 27.0% respectively).

Survey 2 also asked if respondents had changed how they studied during the semester. Fifty-four percent said they had and 45.5% said they had not changed how they studied. Explanations for how they changed included "I have used them more often and harder" (Participant 224). Students who did not report a change stated comments like, "These methods have worked well for me (Participant 031). The most common suggestions for improvement (25 % completing Survey 1 and 15% completing Survey 2) commented that exit quizzes were stressful and they would prefer that they were removed. For example, "Remove exit quizzes" (Participant 251) and "The exit quizzes make me feel worse about myself" (Participant 44). Tables 3 and 4 report results for all learning opportunities and study frequencies.

Table 3 *Summary of learning opportunities used most frequently and perceived as most helpful by participants. Homework PowerPoints created by the students during lab were used most often and were considered most helpful on both surveys. Use of Vanessa videos increased between Survey 1 and 2, as well as their rating of helpfulness. Students reported meeting with their GTA as a learning opportunity least often used and least helpful.*

Code	Survey 1 most often (%)	Survey 2 most often (%)	Survey 1 most helpful (%)	Survey 2 most helpful (%)
Homework ppts	25.5	31.9	29.9	25.7
terms sheet	25.2	17.0	25.7	9.6
Other	21.0	16.3	16.4	16.9
Vanessa videos	12.6	21.5	10.7	26.5
class lecture notes	3.8	0.0	2.8	0.7
open lab, any	3.5	7.4	6.2	12.5
Mini lecture ppts	2.1	1.5	0.8	0.0
pre-labs	1.7	2.2	2.3	1.5
exit quizzes	1.4	0.7	1.1	2.2
open lab, with GTA	1.0	1.5	1.1	0.0
practice practical	1.0	0.0	1.7	0.7
entrance quizzes	0.7	0.7	1.1	1.5
Meet with GTA	0.3	0.0	0.0	2.2

continued on next page

Table 4 *Reported study frequencies per week. Most students reported studying once or twice per week on both Survey 1 and Survey 2. Other than those who claimed not to study, the lowest category of frequency was five or six times per week.*

Frequency per week	Survey 1 (%)	Survey 2 (%)
0	0.0	1.1
1 or 2	27.9	28.1
3 or 4	27.3	25.8
5 or 6	9.9	6.7
daily	15.7	11.2
vague	19.2	27.0

Discussion

This study began with the question, "Are the changes introduced to the lab helpful?" Similar to Sternberg (1990), we presented students with a variety of learning opportunities to try to provide methods that work for many different learning styles. It appears that our changes are helping our students. We found that having a multitude of learning opportunities both during and outside of scheduled lab time provided the greatest benefit to our students. However, we found that not all learning opportunities are equally helpful to all students. We concluded that the answer to this question depends on which students and which learning opportunities are considered. For example, students who demonstrated the largest improvements (5% or more) between the first practical compared to their final grades reported that they used the Vanessa Videos most. Students who earned A's and B's reported high use of Vanessa Videos and Homework PowerPoints. However, the A-students did not demonstrate the largest grade improvements, which we attribute to the fact that high scores do not have the potential for large improvements. In addition, the A and B-students, along with D-students, often reported using "Other" learning opportunities. Our hypothesis is that high achieving students are advanced enough in their understanding of how to seek additional sources of information independently. Our hypothesis for D-students who sought "Other" opportunities is that they are grasping for any potential opportunity to either maintain or improve their grades, thereby to avoid having to re-take the course. . These results do not imply that the learning opportunities are solely responsible for improvement in overall lab grades. We do however believe that students who utilize more resources have the potential to perform better as measured by proficiency in lab assessments.

We used data collected from Survey 2 to try to determine whether students changed the learning opportunities they used over the semester and why. We found that 54.5% (n = 90) reported changing the learning opportunities they used during the semester. However, not all students answered this question, nor did participants provide clear reasons as to why they made changes. Instead, many students simply replied with what they changed to. For example, participant 046 answered, "added Quizlet to study," and participant 134 stated, "I went to open lab and when I started to go and ask questions, I learned faster." Participant 068 answered "Yes, I did not know about the videos until the second practical." Students who answered "no" when asked if they changed study techniques generally agreed with participant 219's sentiment, "No, because they've worked well for me."

Additionally, as Armbruster *et al.* (2009) described in their article about active learning and student-centered pedagogy, we changed how we present and teach lab material and provoke students into taking initiative for their own learning through learning opportunities that encourage learning before, during, and after labs. For example, we have specifically designed the pre-labs so students begin familiarizing themselves with some of the terms for the week before coming to lab. Therefore, if they put effort into completion of the pre-labs, they have been exposed to a certain proportion of the terms before coming to lab.

Due to the negative perception and stress students reported with the "Exit Quizzes," we have reframed them as an "extra" learning opportunity (and have renamed them "Extra Credit"). This reframing strategy was implemented in the summer 2018 session. Spring 2018 semester study participants indicated high stress related to "Exit Quizzes" as they felt they only had 2.5 hours of hands-on lab time to study. We disagree with this perception of limited study time because students have access to all lab materials for the whole semester from the first day of class. Moreover, students perceive a score of 2 out of 5 points on the exit quizzes as a failing grade, even though these points earned make up only 7% of the overall final lab grade. After changing the verbiage of the assessment and denoting the outcome as extra credit worth up to five points, students' comments became positive. They saw the assessment as a chance to stockpile extra points to counteract deficiencies on practical grades, which is the largest point component of the lab.

Finally, as flipped classrooms become more common (Awidi and Paynter 2018), we have incorporated some of those methods into the way we teach our labs and those seem to be the most beneficial for our students as indicated by results tied to the Vanessa videos. Given the relationship of some of these learning opportunities to grade improvement, we plan to integrate them into curriculum (such as online quizzes over the Vanessa video's as a pre-test before lab and to get the students familiar with the material instead of pre-labs (Table 1 and 2). We can take this data and build on the items that students have used and which appear to help them learn and retain information. However, this is not what we tested. We

continued on next page

just wanted to know if providing these materials really does help students; based on these results, it seems to.

Limitations of this study include low return rates on survey 2. The highest return rates for the second surveys were obtained from the three most experienced GTAs' labs. Also, the lead investigator in the study acknowledges that she should have administered surveys to all lab sections in person, which may have resulted in a higher completion rate for both surveys. Survey questions that were not specific enough or perhaps were too long, such as the two-part question asking if and why students changed the learning opportunities they used, could have been broken into two separate questions and perhaps would have resulted in clearer explanations from students. We acknowledge that making inferences about effectiveness of our techniques using information gathered from only 29.36% of students taking the class during the survey period is not ideal. Further monitoring in future semesters is proposed, particularly to include demographic information. Our university has a large first-generation college student population, and variables such as work schedules, family commitments, and socioeconomic status could all affect how students access the available materials and which they choose to use. This is an introductory anatomy and physiology class and there are no prerequisite course requirements, nor is there a GPA minimum requirement. We did not collect data from the students regarding overall GPA, nor did we ask what their past science experiences were.

Conclusion

As researchers and instructors, we feel that this research has provided feedback to improve how we will teach anatomy and physiology in the future. We have gained insight into the methods of learning that students find most effective. While our results often indicated low relationships between grade improvements and learning opportunities used, we are encouraged in our efforts by any grade improvements.

In conclusion, we hope the three articles on the processes involved in reformulating a course in the interest of student success has been interesting and helpful. We will continue adjusting and improving our teaching methods as indicated by student feedback and practical results. Integration of viewing the Vanessa videos (described in parts one and two of this series) as a pre-lab, followed by an online quiz is being considered. Using the best student homework pictures as entrance quizzes seems to encourage pride in doing their homework effectively so we will investigate the effects of that. We now conduct all of our office hours in the lab and have had positive feedback toward location as a beneficial learning opportunity. Overall, we are encouraged by the results we have seen in students' lab grades and plan to continue incorporating creative approaches that include best practices, multiple learning styles, and use of technology while also promoting hands-on learning experiences.

About the Authors

Heather Rudolph, PhD, is a community college faculty member who is passionate about teaching Anatomy and Physiology. She draws from both active and applied learning techniques in order to connect the formal classroom environment to real life experiences.

Anna Schwabe is a biology education doctoral candidate and certified scientific botanical illustrator whose teaching expertise lies in maximizing student learning while fostering a teaching environment conducive to novice student instructors in anatomy and physiology labs.

Nastaran SoleimaniBarzi Mues is a PhD candidate specializing in genetics and stem cells. While working on her PhD, she developed an interest in biology education and its implications in helping students.

Literature cited

Avraamidou L, Osborne J (2009) The role of narrative in communicating science. *International Journal of Science Education,* 31(12): 1683-1707. 10.1080/09500690802380695

Bresler L (Ed.) (2004) *Knowing bodies, moving minds: Toward embodied teaching and learning.* Boston: Kluwer Academic Publishers.

Gopalan C (2016) The impact of rapid change in educational technology on teaching in higher education. *HAPS Educator* 20(4): 85-90.

McDaniel K and Daday J (2017) An undergraduate anatomy lab revision success story. *HAPS Educator* 21(2): 8-18.

Pulvermüller F (2005) Brain mechanisms linking language and action. *Nature Reviews. Neuroscience*, 6(7): 576-582.

Rawson KA and Dunlosky J (2012) When is practice testing most effective for improving the durability and efficiency of student learning? *Educational Psychology Review* 24: 419-435.

Rudolph HA and Schwabe A (2017) "Yes! Use your cell phones!" Active learning with technology in an anatomy and physiology lab. *HAPS Educator* 21(3): 75-79. doi: 10.21692/haps.2017.057

Rudolph HA, Schwabe A and Johnson VJ (2018) "It started because of a snow day!" Making online videos as customized learning tools. *HAPS Educator* 22(1): 69-72. doi: 10.21692/haps.2018.008

continued on next page

Appendix A

Survey 1 Bio 245

Name: _____ Date: _____
Lab section:_____
Please check all of the following learning opportunities you used to prepare for the practical.

☐	Terms Sheet	☐	Class Lecture Notes
☐	Vanessa YouTube videos	☐	Lab Mini- Lecture PowerPoints
☐	Homework PowerPoints	☐	Open Lab (any)
☐	Pre-Labs	☐	Open Lab (with your GTA)
☐	Entrance Quizzes	☐	Practice Practical
☐	Exit Quizzes	☐	Meeting with GTA
☐	Other (Explain.)		

Which of the above learning opportunities did you use most often?

If you used more than one learning opportunity, which one(s) did you find most helpful?

How often did you use the learning opportunity(ies) that you found most helpful?

What suggestions do you have for learning opportunities we could add to the lab or remove from the lab to help you be better prepared for the next practical?

_____

Survey 2 Bio 245

Name: _____ Date: _____
Lab section:_____
Please circle or highlight all of the following learning opportunities you have used to prepare for this practical.

☐	Terms Sheet	☐	Class Lecture Notes
☐	Vanessa YouTube videos	☐	Lab Mini- Lecture PowerPoints
☐	Homework PowerPoints	☐	Open Lab (any)
☐	Pre-Labs	☐	Open Lab (with your GTA)
☐	Entrance Quizzes	☐	Practice Practical
☐	Exit Quizzes	☐	Meeting with GTA
☐	Other (Explain.)		

Which of the above learning opportunities did you use most often?

If you used more than one learning opportunity, which one(s) did you find most helpful?

How often did you use the learning opportunity(ies) that you found most helpful?

Have you changed the learning opportunities you used since the first practical? Please explain why or why not.

What suggestions do you have for learning opportunities we could add to the lab or remove from the lab to help you be better prepared for the next practical?

Back to TOC

Kids in the Gross Anatomy Lab: How an Outreach Program in Anatomy Educates High School and Undergraduate Students about Health Care

Edgar R. Meyer, MAT, MS[1,2], Sutton Williams, PhD, Marianne Conway, MD[1,2,3]; Andrew Notebaert, PhD[1,2,3]
[1]School of Graduate Studies in the Health Sciences, University of Mississippi Medical Center
[2]Department of Neurobiology & Anatomical Sciences, Clinical Anatomy Division, University of Mississippi Medical Center
[3]School of Medicine, University of Mississippi Medical Center
Corresponding Author: Edgar R. Meyer, Department of Neurobiology & Anatomical Sciences, Clinical Anatomy Graduate Program, University of Mississippi Medical Center, 2500 N. State St., Jackson, MS 39216
emeyer@umc.edu

Abstract
The growing movement in STEM education is popularizing active learning, which has benefits for high school and undergraduate students, especially those in rural areas with limited resources. In Mississippi, daunting health disparities in incidence of stroke, cardiovascular disease, type II diabetes, and obesity warrant the implementation of outreach programs that expose high school and graduate students to the health sciences, such as anatomy, and to health literacy. The outreach program at the University of Mississippi Medical Center (UMMC) incorporates informal learning sessions whereby high school and undergraduate students rotate in the gross anatomy lab through various stations featuring organ systems. These sessions benefit the students by engaging their critical thinking skills, motivating them to lead healthier lifestyles, and incentivizing them to pursue careers in the health sciences. https://doi.org/10.21692/haps.2018.031

Key words: anatomy education, anatomy outreach, STEM education, active learning, health disparities

Introduction
In April 2013, a nationwide push began for the adoption of the Next Generation Science Standards (NGSS) developed by Achieve, Inc., a non-profit organization with a mission dedicated to educational reform (Pruitt 2015). Due to the growing movement in STEM education in secondary schools in response to NGSS, there is a need to incorporate STEM concepts into classroom learning experiences.

Since many schools in rural areas may have limited resources or limited access to advanced science materials or technology, academic medical centers or other institutions of higher learning, specifically in the sciences, have an opportunity to respond. In this respect, these institutions can lend their own resources in designing and implementing lessons that serve as community outreach efforts for local and regional undergraduate and secondary education students. These lessons can revolve around health-related topics and integrate basic science knowledge with pathological conditions in order to foster critical thinking skills and problem solving practices between undergraduate and high school students.

The Science Teaching Excites Medical Interest (STEMI) project at the University of Mississippi Medical Center (UMMC) advocates STEM education in high schools. It also brings basic scientists, clinical anatomy students, and high school teachers together with the goal of training teachers in the development and implementation of flipped classroom modules that incorporate active learning strategies and healthcare disparity content (Notebaert *et al.* 2018). A similar outreach program at the University of Arkansas for Medical Sciences (UAMS) also provides high school science teachers with training and resources for teaching content in the health sciences (Burns 2002). Such outreach programs hold the potential for improving STEM education and could greatly benefit STEM education in Mississippi.

Why Outreach?
According to the Centers for Disease Control and Prevention, Mississippi has among the highest levels in the nation of obesity (CDC 2017b), type II diabetes (CDC 2017a), cardiovascular disease (CDC 2018), and stroke (CDC 2018) in which major disparities of incidence, severity, and mortality exist. The high number of healthcare disparities in Mississippi creates a compelling need to better educate students and the general public in areas such as risk factors for disease, social determinants of health, and preventative measures for addressing such discrepancies.

This need is especially apparent in Mississippi where students have average scores below the national mean in every Advanced Placement (AP) science and mathematics course, except AP Physics C: Electricity & Magnetism (College Board 2017). This need is also apparent due to the fact that Mississippi high school graduates had the second-lowest mathematics, science, and composite ACT score averages

continued on next page

in the nation in 2017 (ACT 2017). The application of clinical relevance to biological concepts allows Mississippi students to see the influence that knowledge can have in encouraging lifestyle changes that improve their own health and the health of other Mississippians.

Program Description

The University of Mississippi Medical Center (UMMC) hosts an informal outreach program that receives undergraduate students from one of the local colleges and students from several high schools in the state. Faculty and students in the Clinical Anatomy program introduce students to the medical center's gross anatomy laboratory and guide them through activities.

The Institution

UMMC is the only academic medical center in the state of Mississippi. Located in Jackson, the state's capital and most populated city, the institution has the responsibility to serve as a hub for learning in the health sciences at all levels of education. Given the fact that Mississippi has some of the highest health disparities in the country, the medical center has the responsibility to serve as a beacon for health literacy, especially for individuals who have limited access to advanced scientific resources. Through the Clinical Anatomy program in the Department of Neurobiology and Anatomical Sciences, UMMC offers a free program for educating students about these health disparities, human anatomy, and the various health professions whose training includes anatomy.

Formal Anatomy Instruction

Of the seven different health professional schools on UMMC's campus, four of them incorporate formal anatomy instruction into their curricula. The first-year students in the School of Medicine, the School of Dentistry, and the Clinical Anatomy Ph.D. program in the School of Graduate Studies in the Health Sciences are required to learn content in macroscopic anatomy (gross anatomy), microscopic anatomy (histology), developmental anatomy (embryology), and neuroanatomy. In addition, the first-year students in the Occupational Therapy and Physical Therapy programs within the School of Health-Related Professions and students in the Master's in Biomedical Sciences (BMS) program within the School of Graduate Studies in the Health Sciences are required to learn content in gross anatomy and neuroanatomy. During their first year, the Ph.D. students in Clinical Anatomy also take two pedagogy courses to understand the nature of learning in general and to learn about various learning theories and teaching methods. In their second year, the anatomy graduate students take a pedagogy course in which they learn how to design curricula specific to anatomy. Moreover, in their second, third, and fourth years of the program, the anatomy graduate students gain training in teaching by serving as teaching assistants in at least one of the content areas in medical or dental anatomy each semester. Furthermore, anatomy graduate students have opportunities to teach undergraduate and high school students anatomical content during multiple visits scheduled throughout the year.

The gross anatomy lab is one of the traditional learning environments used by the on-campus health professional students and visiting undergraduate and high school students. While formal instructional lessons are conducted in the lab for the health professional students who take gross anatomy, the undergraduate and high school students participate in informal learning experiences with cadaveric material. Since anatomy courses are ongoing throughout the year, cadaveric material is available for informal teaching opportunities with the undergraduate and high school students. First-year medical students use the lab during the fall semester, first-year dental and BMS students use the lab during the spring semester, and first-year occupational and physical therapy students use the lab during the summer semester. The health professional students dissect donors during their lab sessions, affording undergraduate and high school students the chance to learn human anatomy from cadaveric specimens to which they would have otherwise not had exposure at their own institutions.

Informal Anatomy Instruction

During the experiential learning sessions, the undergraduate and high school students and their instructors become oriented to the gross anatomy lab. The students receive an introduction to the components of the laboratory space, the courses taught in the space, the Clinical Anatomy program, and the Body Donation Program through which the donors are acquired. The visitors are informed of the utmost respect given to the bodies of the donors, and they are prohibited from taking any photos of the donors or any other cadaveric material. The visitors are then provided with the appropriate protective attire, including white lab coats, latex gloves, and safety glasses. Afterwards, the visitors are given an overview of the lab stations through which they will rotate.

Each lab station is assigned an anatomy faculty member or an anatomy graduate student as a facilitator, and the time for each lab station is apportioned equally within the allotted timeframe for the lab visit. Since most undergraduate and high school students who visit the gross anatomy lab are enrolled in a human anatomy and physiology course, the topics of the lab stations revolve around individual organ systems. While human anatomy and physiology courses are taught using a systemic approach, gross anatomy is traditionally taught using a regional approach. In the informal learning sessions, the anatomy facilitators blend the regional approach of gross anatomy with the continuity of the systemic approach of human anatomy and physiology. This combination of approaches allows the students to understand entire organ systems in the context of specific body regions and these organ systems' relationships to the body as a whole.

continued on next page

The instructional process typically begins with a brief introduction initiated by the facilitator and an invitation for the students to share their intended career goals. Most students have aspirations of pursuing careers in healthcare fields such as medicine, nursing, dentistry, and occupational and physical therapy. However, some students have plans to earn degrees in other disciplines such as psychology, music, and art. Other students have not made a decision about their future career. Regardless of their decided or undecided career plans, the students convey a genuine interest in the ensuing learning sessions.

All of the lab stations usually incorporate four primary elements, including 1) the engagement of the students' prior knowledge, 2) the practices of probing, or deep and guided questioning, 3) inquiry and experiential learning, and 4) emphasis on the clinical relevance of the content discussed. First, the students are asked to explain what they already know about the organ system featured in the station, including its functions. The students' responses inform the progression of the discussion that follows. This discussion usually connects the functions that the students mention to the compositional structure conveyed by the various organs comprising the organ system under investigation. The interaction gradually evolves into an exploratory journey through each component of the organ system. This anatomical adventure is punctuated with a series of guided questions that lead the students through the system. The students are encouraged to name each consecutive organ within the system and to explain its respective function. An interrogation with deep questions compels the students to delve further into the rationale underlying the structure-function relationships. The students are encouraged to think more critically with "why" and "how" questions. The students are immersed in a condition of inquiry whereby they are allowed and encouraged to ask additional questions. They are invited to experience the preserved, prosected specimens pertaining to the respective organ system by holding them in their hands and feeling their textures.

Several specimens in the lab exhibit abnormalities due to diseases or health conditions of the donors to whom they belong. For example, when discussing the cardiovascular system, the students are shown hearts that exhibit stitches denoting by-pass surgery on one or more of the coronary arteries. Many of these hearts are retrieved from donors whose sterna have been sutured closed with wire—evidence suggestive of open-heart surgery. Such clinical correlations allow the students to extend their thinking to see the relevance of anatomy to heath care. By comparing both normal and abnormal anatomy, students can understand just how crucial the practices of heath care professionals are in preserving the lives of patients by restoring their anatomy to normal or as close to normal as possible. Overall, the lab experience begins as a discussion and evolves into something much more. The exchange of comments and replies unfolds as a conversation that engages the students in active intellectual participation with the facilitator and each other rather than subjecting them to a one-way didactic lecture in which they serve merely as passive learners.

An Example of Lab Stations on the Cardiovascular System

1) The facilitator (ERM) began with an overview of the cardiovascular system and its components (heart and blood vessels, e.g. arteries, veins, and capillaries), and its basic functions. The names of the components were posed to the students as questions to gauge what the students already knew prior to the session.

<<The students moved to station with heart and lungs prosection.>>

2) The facilitator and the students discussed the circulation of blood beginning with the heart, allowing the students to describe the flow of blood from one structure to the next in both pulmonary and systemic circulation. (This was a situation where the facilitator asked them where the blood goes next, etc.)

After this discussion, the facilitator allowed the students time to ask questions that they might have concerning the circulation. Upon answering their questions, the facilitator moved to the next station.

<<The students moved to station with dissected heart.>>

3) The facilitator then showed the students the exterior and inferior surfaces of the heart, emphasizing the main features that contribute to the function of the heart (e.g. valves, muscular walls).

The facilitator also briefly discussed the histological significance of cardiac muscle and its function as a syncytium via the sinoatrial (SA) node, the pacemaker of the heart. The facilitator asked if they knew anyone who has ever had a pacemaker and if they could explain the function of pacemaker.

<<At this point, the facilitator showed the students an example of a donor who had a pacemaker.>>

The facilitator also asked students: "If the heart does all the pumping for blood, does it too need a blood supply?" The facilitator then listened to how the conversation ensued and proceeded with a discussion of the coronary arteries, asking if the students knew anyone who had had bypass surgery.

continued on next page

<<The students moved to station with prosected cadaver.>>

At this station, the facilitator showed the students an example of a bypass surgery.

Once again, the facilitator allowed time for any questions the students might have.

4) At the same cadaver, the facilitator then led the students in a discussion of the major vessels in the body, pointing out some in the cadaver, but allowing them to first try to describe them (e.g. arteries: aorta and its branches, pulmonary artery and its branches, bifurcation of the aorta, etc., and veins: superior and inferior vena cava and some of their major tributaries). The portal system was left for another clinical anatomy graduate student facilitator (SW).

As the facilitator (ERM) talked about the bifurcation of the aorta, he showed the students an example of an abdominal aortic aneurysm. This demonstration catapulted the conversation into an inquiry about the causes of aneurysms, their dangers, and their treatments.

After, discussing the major vessels, the facilitator briefly discussed the histological differences between arteries and veins, allowing students to tell him some of the differences and similarities that they might have already known. While the students knew that arteries carried blood away from the heart (*a* for *away*) and that veins carried blood back toward the heart, they did not consider their different textures and histological compositions. Thus, the students were encouraged to touch various smaller arteries and veins so that they could compare the more elastic, thick-walled, opaque arteries to the flatter, thin-walled, transparent veins.

And finally, the facilitator initiated a transition to discussing the general idea of collateral circulation by pointing out the multiple arteries in places like the forearm and elbow. Students contemplated the rationale of extra sources of blood flow. The students eventually arrived at the importance of such additional blood circulation in the event that one source might be occluded, severed, or surgically removed.

A final two-part question the facilitator posed at the very end of the series of lab stations on the cardiovascular system was, "What part of the body receives the greatest amount of blood supply. And why?" The fact that the students spent a prolonged period of time contemplating the questions and pondering their responses further emphasized the answer since the brain is in fact the part of the body that requires most of the body's blood flow due to its high demand for glucose as a readily available energy source.

<<When they finished viewing the cardiovascular system stations, the students moved to stations with prosected cadavers pertaining to the digestive system with facilitator (SW).>>

Discussion

Despite the informality of these learning sessions in the gross anatomy lab, there is a rationale behind the techniques used in the guided questioning process. In addition, there are inherent benefits and challenges to implementing such a program that are worth noting.

Rationale

The guided questions which were provided in the aforementioned station descriptions were addressed to the students by the clinical anatomy student and faculty facilitators as a means to provide a just-in-time teaching experience (Marrs and Novak 2004) whereby the information the students needed to arrive at the answers was only given at the appropriate time. In this way, there was never a time that the facilitators simply gave the students the answers without first compelling them to think critically to arrive at their responses. These questions served as a formative assessment to appeal to the students' "logical thinking," a concept within the common set of values of scientific inquiry that is the nature of science. This concept is outlined under the disciplinary core idea of "Science and Engineering Practices" within the "Structure and Function" competency of the NGSS framework (NGSS Lead States 2013). In their reasoning, the students relied on their powers of observation to discover characteristic features within displayed images of photomicrographs of both normal and pathological tissues and organs. They also learned to relate structure to function.

One study showed that active learning promotes high school students' intrinsic motivation, their intent to enroll in advanced placement (AP) courses, and their interest in professions in scientific fields (Bryan *et al.* 2011). Therefore, programs that include learning sessions similar to the one used in the gross anatomy lab can encourage the incorporation of active learning modules regarding healthcare disparities into high school classrooms to encourage students to pursue careers in fields such as the health sciences.

Benefits

There are a number of benefits for students who participate in the lab visits. The students from the local college do not have access to cadavers due to the cost of maintaining a sustainable body donation program; therefore, field trips to the gross anatomy lab at the medical center nearby allows the students to gain experience with cadavers, even though the exposure may only be once per semester. Since their experience is free, access to human donors, though temporary, provides them with a cost-effective learning opportunity that

continued on next page

is simultaneously realistic and enriching. Similarly, the high school students also receive the same learning benefits. The students also receive authentic learning experiences during which they understand the importance of health literacy for themselves, their families, and the citizens within their state. Such motivational learning has the potential to encourage students to consider health care fields more intently if they were already on the health care track or to consider them as an option if they were undecided.

These lab sessions also benefit the facilitators as they provide them with training opportunities to gain more experience in teaching anatomy. However, these opportunities provide the facilitators not only with another traditional instruction experience to add to their curricula vitae but also with low-stakes training to enhance their skills in facilitating small-group discussions and active learning sessions. These teaching methods are becoming more common in primary and secondary education. Evidence-based research suggests that active learning improves students' average scores on examinations by about 6% (Freeman *et al.* 2014). In addition, these types of instructional techniques are becoming more common in undergraduate and graduate education. Graduate students in Clinical Anatomy also learned ways in which they could assist in classroom lessons regarding anatomical content and its relevance to adverse health conditions. Just as university science students mentored elementary students and engaged them in scientific activities in one study (Bruce *et al.* 1997), so too can graduate university students challenge high school students.

There is also a mutual benefit for both the students and the facilitators as the dynamics of their interactions during the lab sessions are fully dependent on the engagement and activity of each individual member of the group during the interchange of ideas. The queries and explanations posed by each participant have the power to transform the discussion and prolong the life of the rich conversation to which each participant becomes a vital member; thus, no one is more important than the other. Such experiences teach young students that their thoughts are highly valued and have the potential to instill within them the confidence to voice their thoughts with resolution.

Challenges

Along with the benefits, there are a number of challenges to organizing these learning sessions. The first challenge is the obvious challenge of time. Most faculty and students within the Clinical Anatomy program are extremely busy with course work and other obligations; thus, they have limited time to devote to outreach. Professors at undergraduate institutions as well as teachers in secondary education are also busy with their own teaching schedules, so the planning of visits to the gross anatomy lab often requires considerable efforts. Moreover, while the program and department do not charge visitors for learning in the gross anatomy lab, there are associated costs with travel to the institution from high schools as far away as two and a half hours from the campus. Furthermore, even local high schools may not always receive district approval to use school transportation to visit UMMC, and some districts, even local ones, may not even have the funds to support such visits, even as educationally beneficial as they may be for the students. With the health disparities in Mississippi being so high, the efforts this program exerts in making a difference for the health care of the entire state may seem minimal. Nevertheless, the efforts of this program in concert with the efforts of the multiple outreach endeavors initiated by students and faculty at UMMC together have the capacity to engender small changes that, over time and with consistent implementation, can make a substantial impact on the wellbeing of Mississippi citizens.

Future Directions

Desired directions for the future of the program include expanding invitations for more high schools and undergraduate institutions to include more of those with large underserved and minority populations. In addition, the rural parts of the state with high schools that have limited resources and no means of providing their students with a mode for traveling to the medical center may require visits from representatives of the medical center. Instead of the students traveling to UMMC, faculty and students in the Clinical Anatomy program could make site visits to rural schools, especially those in dire need of exposure to health literacy content. Although UMMC faculty and students could not transport cadaveric material, they could bring anatomical expertise and skills in active learning to secondary students while engaging with existing embedded health sciences content. Further directions include exploring the existence of grants to fund such endeavors.

Conclusions

While the program described in this paper is a small, informal venture, it provides an enjoyable and informative experience for both students and facilitators. The learning sessions are also cost-effective ways for participating students to gain exposure to authentic human anatomy, real-world exposure to the health disparities evident in the pathologies of donors, and empowering discussions that foster higher order thinking and validation. These provisions serve as the ultimate incentive for students to become invested in the mission of improving the health care of Mississippians, either directly as health care providers, or indirectly as advocates for health care practices and the dissemination of health literacy.

continued on next page

About the Authors

Edgar R. Meyer, MAT, MS**,** is a fourth-year graduate student in the PhD program in Clinical Anatomy in the Department of Neurobiology and Anatomical Sciences at the University of Mississippi Medical Center (UMMC). He has served in formal teaching assistant roles in medical gross anatomy, medical histology, and medical neuroscience and behavior courses as well as volunteer teaching assistant roles in medical developmental anatomy, dental microscopic and developmental anatomy, and dental gross anatomy. His research interests include the evaluation of virtual three-dimensional (3D) stereoscopic anatomical models on health professional students' short- and long-term retention.

Sutton Williams, PhD, is a fifth-year graduate student in the PhD program in Clinical Anatomy in the Department of Neurobiology and Anatomical Sciences at UMMC. He has served in formal teaching assistant roles in medical gross anatomy, medical histology, medical neurobiology, graduate anatomy, and graduate histology. He recently defended his dissertation exploring comparisons between prosections and dissections and their impact on first-year medical students self-efficacy.

Marriane Conway, MD**,** is an Assistant Professor in the Clinical Anatomy Division in the Department of Neurobiology and Anatomical Sciences at UMMC. She serves as the course director for medical developmental anatomy, and as an instructor in the medical gross anatomy, medical histology, and dental gross anatomy courses. Her research interests include team-based learning (TBL) and body donation.

Andrew Notebaert, PhD**,** is an Assistant Professor in the Clinical Anatomy Division in the Department of Neurobiology and Anatomical Sciences at UMMC. He serves as the Program Director for the PhD program in Clinical Anatomy, and he serves as the course director for the health sciences pedagogy courses. He also serves as an instructor in medical gross anatomy. He has an interest in qualitative research of various forms of survey responses from health sciences students.

Literature cited

ACT, Inc. Average ACT scores by state: Graduating class 2017. Retrieved from https://www.act.org/content/dam/act/unsecured/documents/cccr2017/ACT_2017-Average_Scores_by_State.pdf

Bruce BC, Bruce SP, Conrad RL, Huang HJ (1997) University science students as curriculum planners, teachers, and role models in elementary school classrooms. *J. Res. Sci. Teach.* 34: 69-88. doi: 10.1002/(SICI)1098-2736(199701)34:1<69::AID-TEA6>3.0.CO;2-M

Bryan RR, Glynn SM, Kittleson JM (2011) Motivation, achievement, and advanced placement intent of high school students learning science. *Sci. Educ. 95*: 1049-1065. doi: 10.1002/sce.20462

Burns ER (2002) Anatomy of a successful K-12 educational outreach program in the health sciences: Eleven years experience at one medical sciences campus. *Anat. Rec. (New Anat.)* 269: 181-193. doi: 10.1002/ar.10136

Centers for Disease Control and Prevention (2018) Heart disease mortality by state. Retrieved from https://www.cdc.gov/nchs/pressroom/sosmap/heart_disease_mortality/heart_disease.htm

Centers for Disease Control and Prevention (2017) Maps and trends in diagnosed diabetes. Retrieved from https://www.cdc.gov/diabetes/statistics/slides/maps_diabetes_trends.pdf

Centers for Disease Control and Prevention (2017) Prevalence of self-reported obesity among U.S. adults by state and territory. Retrieved from https://www.cdc.gov/obesity/downloads/data/overall-obesity-prevalence-map2016.pdf

Centers for Disease Control and Prevention (2018) Stroke mortality by state: 2016. Retrieved from https://www.cdc.gov/nchs/pressroom/sosmap/stroke_mortality/stroke.htm

College Board (2017) AP program participation and performance data 2017. Retrieved from https://research.collegeboard.org/programs/ap/data/participation/ap-2017

Freeman S, Eddy SL, McDonough M, Smith MK, Okoroafor N, Jordt H, Wenderoth MP (2014) Active learning increases student performance in science, engineering, and mathematics. *P. Natl. Acad. Sci. U.S.A.* 111: 8410-8415. doi: 10.1073/pnas.131903111

Marrs KA, Novak G (2004) Just-in-time teaching in biology: Creating an active learner classroom using the Internet. *Cell Biology Education* 3: 49-61. doi: 10.1187/cbe.03-11-0022

NGSS Lead States (2013) Next Generation Science Standards: For States, By States (HS. Structure and Function). Washington, DC: The National Academies Press. Retrieved from https://www.nextgenscience.org/sites/default/files/topic-arrangement/HS.SFIP06212017WithFooter.pdf

Notebaert A, Barnard M, Meyer E, Dehon E, Compretta C, Allen D, III, Stray S, Taylor J, Sullivan D, Rockhold R (2018) Science Teaching Excites Medical Interest: A teacher professional development program in Mississippi. Journal of STEM Outreach 1. Retrieved from https://www.jstemoutreach.org/article/5066-science-teaching-excites-medical-interest-a-teacher-professional-development-program-in-mississippi

Pruitt SL (2015) The next generation science standards: Where are we now and what have we learned? *Science and Children*. Summer.

Back to TOC

Making It Real: Case-Study Exam Model

Disa Smee PhD and Julie Cooke PhD
Faculty of Health, University of Canberra, Canberra A.C.T. 2601 Australia
Disa.smee@canberra.edu.au; Julie.cooke@canberra.edu.au

Abstract
Anatomy and Physiology at the University of Canberra, Australia, as in many institutions, is taught to large cohorts at the foundational level. There is a requirement for students to not only retain information regarding physiological systems and their associated anatomy, but also the integration and relationship between systems. Furthermore, students are often employed in clinical settings post-graduation and the need for real-world understanding is paramount. Assessing this type of understanding is challenging and academics must move beyond tradition exam models. Using case studies as the basis for exam questions is an exciting format for eliciting student understanding of physiological systems which, at the University of Canberra, has resulted in 7% higher grades compared to traditional short answer exams. The benefits go beyond grades, with students appearing more confident in answering questions and better able to integrate information.
https://doi.org/10.21692/haps.2018.029

Key words: teaching, undergraduate, anatomy and physiology

Introduction
Anatomy and Physiology provides a necessary foundation for students studying in the health science disciplines and students may be employed in a variety of clinical settings following graduation. Work in clinical settings requires not only the recall of information but also the ability to integrate that information for treatment and diagnostic purposes. Cliff and Wright (1996) stated that "Mere accumulation of a massive body of facts can never be the sole goal for this type of course." Furthermore, they add that students "only truly know the information once they can apply that content to real-world scenarios." Successful engagement with Anatomy and Physiology content is fundamental to the retention of the information (Cliff and Wright 1996) and the ability to apply that knowledge to critical thinking is essential for graduates to be effective in their chosen careers.

Though the use of case-studies as a learning tool has been an option for a number of years, it has increased in popularity over the past two decades (Herreid *et al.* 2011). Case-studies are an effective strategy for improving critical thinking and problem-solving skills of students (Dori and Herscovitz 1999). Case studies focus on the "real world" using constructivist learning theory (Hein 1991) allowing students to actively engage with content and understand theoretical concepts in a practical way (Herreid *et al.* 2011). Appropriately designed case studies can create an effective learning approach to guide students, especially if the case studies are specifically designed for a particular class. This ensures that students are engaging at a level that is appropriate to both themselves and the content. Furthermore, case-study learning has been shown to improve exam performance in Anatomy and Physiology (Cliff and Wright 1996).

While exams are generally not considered cutting-edge in the current teaching and learning environment they do have value within certain subject areas, particularly at the foundation level (Butler and Roediger III 2007). There is some evidence that studying for (McDaniel *et al.* 2007) and taking exams (Butler and Roediger III 2007) can actually promote learning, and that exams are not just an evaluation method. "Exams don't just provide a targeted, fit-for-purpose opportunity for students to demonstrate what they know: they also have the power to enhance what students know" (Van Bergen and Lane 2014). Furthermore, a variety of assessment items are necessary allowing students multiple options within varying contexts to demonstrate what they know. Exams, when purposeful and effectively designed are a viable assessment tool that provides a deeper level of understanding as opposed to the mere superficial level learning of the content.

At the University of Canberra, Anatomy and Physiology is taught in first year. It is considered a foundational, or service subject, accommodating up to 600 students who are enrolled in a range of professional degrees such as physical therapy, exercise science (kinesiology), nursing, and nutrition. The cohort is extremely diverse in terms of previous educational background and the degrees they are undertaking. As with many similar courses the concepts introduced at the introductory level are explored in greater detail in higher-level courses.

Over the last few years it became evident that students were not particularly competent in the skill of information integration that requires higher order thinking. Students were focused on remembering rather than understanding or applying information. Physiological systems are inherently related and it is therefore vitally important that students

continued on next page

are able to understand and explain basic physiological interactions and extrapolate ideas rather than just learn stand-alone information. Case studies were introduced into weekly lecture content and final examination assessment as a way to help students develop the skills of information integration.

Appropriately designed case-studies can create an effective learning approach to guide students (Bonney 2015). The case studies were structured to build on the information, providing a scaffolding-learning environment for students progressing through the topic. Case studies were introduced to the lectures so that students might become familiar with the style of question development prior to their final exam. Furthermore, this fit-for-purpose design of the case studies for lectures and final exam in Anatomy and Physiology utilises authentic learning and aims to encapsulate not only what the students know but also how they apply and integrate that knowledge. This new assessment format requires higher level learning that includes understanding the content and having the ability to apply and analyse the material (Krathwohl 2002).

Guidelines for design

Case studies should be designed as fit-for-purpose vehicles and as such, they should include the following:

- Real-world scenarios that students can associate with and therefore understand more easily.
- Limited extraneous information so as to not dilute what is being asked.
- Clear learning objectives with simple language.
- Scaffolded questions to guide the progression of learning.
- Questions should seek clarification of how anatomy and physiology are integrated in achieving homeostasis.
- An answer scheme that is determined prior to student exposure to ensure that the questions provide adequate information to illicit an appropriate answer.

Incorporating the Human Anatomy and Physiology Society (HAPS) learning outcomes (HAPS 2013) into case-study questions facilitates higher levels of learning such as analysis and synthesis. The example provided below incorporates topics from the HAPS guidelines and learning goals (LG) (identified in bold).

Example Question and Answer

Cardiovascular System (9 points)

Josh undergoes an echocardiogram (a picture of the heart) and everything looks normal until they see his left ventricle. The myocardium here is significantly thicker than normal. Josh's blood pressure is 160-110. (Note: add a picture of normal versus hypertrophy and corresponding ECG trace)

A. Why would doctors be concerned about a thickening of the myocardium (hypertrophy)? (1 point)
   HAPS LG 1, 2, 4, 6
   Answer: Stroke volume would be reduced (1 point) OR less blood would be pumped (0.5 point)

B. Why would the left ventricle be more of a concern than the other chambers of the heart? (1 point)
   HAPS LG 1, 2, 4, 5, 6
   Answer: It pumps systemically – to whole body (1 point)

C. How would Josh's heart attempt to maintain homeostasis given this thickening. Provide a physiological explanation (2 points).
   HAPS LG 1, 3, 6
   Answer: CO = SV X HR (point) therefore if SV decreased HR increased (1 point).

D. The doctor diagnosed Josh's left ventricular hypertrophy after noting an increased R wave amplitude (peak of QRS complex) on his electrocardiogram. Why was this section of the electrocardiogram affected by hypertrophy in the left ventricle? (3 points)
   HAPS LG 1, 4, 5, 6, 8
   Answer: This represents ventricular depolarization (1 point). The more "muscle" there is, the greater the amplitude of the QRS complex (1 point), which means there is more muscle to be electrically excited (1 point).

E. From the information provided, what other condition does Josh have? (1 point) And how might it relate to his left ventricular hypertrophy diagnosis? (1 point)
   HAPS LG 1, 6, 8
   Answer: Josh has hypertension (1 point). Individuals with elevated blood pressure have greater peripheral resistance. Thus the heart must work harder (0.5 point) and as with any muscle the result is an increase in hypertrophy (0.5 point).

continued on next page

Student outcomes/instructor observations

The results have been promising. The average percent for the traditional short answer exam questions on the final exam in 2014 (n = 450) and 2015 (n =484) was 46.01%. Upon introduction of the case-study model in 2016 (n = 490) and 2017 (n= 534) the short answer component of the final exam increased to 51.39% and 53.18% ($p<0.05$) respectively.

Student comments:
"*The case studies are great in assisting me to think critically about the content learned and not just learning it.*"

"It gives you something to relate to in the real world and see why it is relevant."

"The case studies in the exam allowed me to show how much I knew about each topic"

From our observations in face-to-face laboratory classes, students are able to breakdown and piece back together the physiological systems with greater understanding of the processes. They are better able to apply basic physiological information to an array of scenarios, not just the ones that are presented in lecture. Students appeared to be more confident in their knowledge and were more willing to engage in class discussions. In the final examination, students were more assured in attempting questions and incorrect answers were more academically logical.

Anecdotally, colleagues teaching upper level units have commented that students display improved ability to integrate information. Furthermore, students have greater understanding of the physiological foundations and can practically apply that information more successfully to a variety of contexts.

Discussion

Traditionally, learning in an Anatomy and Physiology setting used a surface approach where facts were memorized and then regurgitated back to answer questions (Pandey and Zimitat 2007). However, academics from a range of diverse settings, agree that educational success, regardless of the discipline, depends on learning concepts rather than memorizing facts (Hughes 2011). Deeper level learning, the process of learning for transfer, permits students to take knowledge from one situation and apply it to another. Thus, this approach seeks for students to find meaning within the material and be able to interpret it and fit it with what they already know (Pandey and Zimitat 2007). In Anatomy and Physiology being able to move to a deeper level of learning "may require a preliminary stage of rote learning that is difficult to distinguish from a surface approach" (Enwistle and Ramsden 1983). As such this surface level or rote learning approach, specifically in the development of the necessary complex anatomy and physiology vocabulary/terminology, is required to permit an understanding at the deeper level (Laurillard 1984). For students, being able to develop deeper approaches to learning is essential as it is believed this enhances students engagement with the content and ultimately results in improved analytical and critical thinking skills (Hall *et al.* 2004).

Different approaches to learning are also important when developing assessment items including examinations. For example, multiple choice question (MCQ) examinations are perceived to assess knowledge-based (lower levels of) academic processing with better performance associated with the employment of surface learning strategies (Scouller 1998). In contrast, students are more likely to be successful in tasks which are perceived as assessing higher levels of cognitive processing, if deeper strategies are used (Scouller 1998). Case studies use higher levels of cognitive processing and thus use deeper level learning. Case studies, purposefully designed for exams, allow students to demonstrate their understanding and ability to interpret the information as they apply that information to real-world scenarios (Cliff and Wright 1996). Case studies also permit active engagement with the content.

Conclusion

The use of case studies as the basis for exam questions at the University of Canberra has been a successful tool and student engagement with this format has resulted in better retention of the information and improved interdisciplinary application. The use of case studies provides students with exposure to scenarios they could encounter in clinical practice. This demonstrates the relevance of the foundational unit/course content long-term. Since this format has only recently been introduced, follow up with students and their supervisors during clinical placements or work-integrated learning internships in future years would be beneficial. Further research into student perception and engagement with this format is also warranted.

About the authors

Disa Smee, PhD, is an assistant professor of anatomy and physiology at the University of Canberra, Australia. Dr. Smee completed her BMed Sci (Hons) and her MCom at The University of Sydney and earned her PhD in physiology (University of Canberra) in 2015 but has been teaching at the tertiary level for 20 years. She teaches a variety of anatomy and physiology service units to freshman students and is passionate about ensuring their success.

Julie Cooke, PhD, is a senior lecturer in anatomy and physiology at the University of Canberra and is the Program Director of the Australian Biology Olympiad Program. Julie commenced her teaching career in 1998 at the Copenhagen International School. She completed a BSc (Hons) at Flinders University, South Australia, and received her PhD from the University of Adelaide in 2000. She teaches anatomy and

continued on next page

physiology to freshman students and Sports Medicine at the junior and senior level. Julie has a passion for teaching and enjoys engaging students so that learning is easier and more enjoyable.

Literature cited

Bonney KM (2015) Case study teaching method improves student performance and perceptions of learning gains. *Journal of Microbiology & Biology Education* 16(1): 21.

Butler AC and HL Roediger III (2007) Testing improves long-term retention in a simulated classroom setting. *European Journal of Cognitive Psychology* 19(4-5): 514-527.

Cliff WH and AW Wright (1996) Directed case study method for teaching human anatomy and physiology. *Advances in Physiology Education* 270(6): S19.

Dori YJ and O Herscovitz (1999) Question-posing capability as an alternative evaluation method: Analysis of an environmental case study. *Journal of Research in Science Teaching* 36(4): 411-430.

Enwistle N and P Ramsden (1983) Understanding student learning (pp. 169–194). *London: Croom Helm*

Hall M, A Ramsay and J Raven (2004) Changing the learning environment to promote deep learning approaches in first-year accounting students. *Accounting Education* 13(4): 489-505.

Human Anatomy and Physiology Society (2013) "Learning Outcome Documents For Anatomy & Physiology I." Retrieved 18 October 2018, from https://www.hapsweb.org/page/AP1_Outcomes.

Hein G (1991) Constructivist learning theory. *Institute for Inquiry. Available at:/http://www. exploratorium. edu/ifi/resources/constructivistlearning. htmlS*

Herreid CF, NA Schiller, KF Herreid and C Wright (2011) In case you are interested: results of a survey of case study teachers. *Journal of College Science Teaching* 40(4): 76.

Hughes KS (2011) Peer-assisted learning strategies in human anatomy & physiology. *The american biology Teacher* 73(3): 144-147.

Krathwohl DR (2002) A revision of Bloom's taxonomy: An overview. *Theory into practice* 41(4): 212-218.

Laurillard D (1984) Learning from problem solving in Marton, F. Hounsell, D and Entwhistle, N (Eds) The Experience of Learning, Edinburgh, Scottish Academic Press.

McDaniel MA, JL Anderson, MH Derbish and N Morrisette (2007) Testing the testing effect in the classroom. *European Journal of Cognitive Psychology* 19(4-5): 494-513.

Pandey P and C Zimitat (2007) Medical students' learning of anatomy: memorisation, understanding and visualisation. *Medical education* 41(1): 7-14.

Scouller K (1998) The influence of assessment method on students' learning approaches: Multiple choice question examination versus assignment essay. *Higher Education* 35(4): 453-472.

Van Bergen P and R Lane (2014) "Exams might be stressful, but they improve learning." *The Conversation* Retrieved 6 December 2017 from https://theconversation.com/exams-might-be-stressful-but-they-improve-learning-35614.

■

Back to TOC

Make Plans Now!

THE HAPS 33RD ANNUAL CONFERENCE

WILL BE HELD IN PORTLAND, OR

MAY 22-26, 2019

Workshops hosted by
The University of Portland

CLICK HERE for more details!

Current HAPS-I Course Offerings

CLICK HERE to find out more

Spring 2019

Lab Course in Cardiac Physiology
(1 credit) January 8 - January 29, 2019
Patrick Eggena, M.D.
Novateur Medmedia LLC

In this course participants review basic concepts in cardiovascular physiology by conducting laboratory experiments on themselves or supervising experiments performed by their students. Upon completion of these experiments, HAPS Professors email their observations to Dr. Eggena who organizes them in preparation for weekly discussions held on Tuesdays from 3-4 pm EST in Google Hangouts internet classroom. HAPS Professors may send only one set of experimental results by email to Dr. Eggena each week, and the HAPS Professor will be the spokesperson for her/his students during the weekly meetings. In these weekly meetings the lab exercises are related to clinical case scenarios from Dr. Eggena's iBook series, "Cardiac Physiology as a Country Doc". This course will prepare HAPS Professors to present a laboratory course in Cardiac Physiology at their institutions independent of Dr. Eggena.

HAPS COMMITTEES AND BOARDS

Standing Committees:

2019 ANNUAL HOST COMMITTEE

Jacqueline Van Hoomissen

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.

CADAVER USE

Cindy Wingert

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.

COMMUNICATION

Wendy Riggs

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.

CONFERENCE

Tom Lehman

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 AND INSTRUCTION

Christie Canady

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

EXECUTIVE DIRECTOR

Peter English

FUNDRAISING

Jon Jackson

GRANTS & SCHOLARSHIPS

Carol Veil

MEMBERSHIP

Jon Wisco

This committee is charged with expanding our membership base to include all Human Anatomy and Physiology educators or those individuals, institutions and corporations crucial to the HAPS mission statement of "Promoting Excellence in the Teaching of Human Anatomy and Physiology."

SAFETY / ANIMAL USE

Richard Simons

This committee develops standards for laboratory safety. The committee maintains a variety of safety documents available for download.

STEERING

Kyla Ross, Chair

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

Click here to visit the HAPS committees webpage.

Special Committees:

EXECUTIVE

Judi Nath

Composed of the HAPS President, President-Elect,

Past President, Treasurer and Secretary

NOMINATING

Mark Nielsen

This committee recruits nominees for HAPS elected offices.

PRESIDENTS EMERITI ADVISORY COMMITTEE

Terry Thompson

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

HAPS EDUCATOR

Kerry Hull, Editor-in-Chief

Sarah Cooper, Managing Editor

This committee is responsible for publishing a quarterly edition 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.

TESTING

Jennifer Burgoon, Co-Chair

Valerie O'Loughlin, Co-Chair

This committee has completed, tested and approved the HAPS Comprehensive Exam for Human A&P and is developing an on-line version of the exam.