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

Polly R. Husmann*

*Corresponding Author: phusmann@indiana.edu

HAPS Educator. Vol 22 (3), pp. 187-198. Published December 2018.
doi: 10.21692/haps.2018.024

Husmann P.R. (2018). Medical Student Study Strategies in Relation to Class Size and Course Length. *HAPS Educator* 22 (3): 187-198. doi: 10.21692/haps.2018.024

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-NSPLjRI.

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