The Effects of Repeated Lectures and Class Time on Learning Outcomes and Teaching Evaluations

Libo Sun
California State Polytechnic University, Pomona, Pomona, California, USA

Xiaohui Xu*
California State Polytechnic University, Pomona, Pomona, California, USA

In higher education institutions, classes are offered according to a block schedule. Courses are fitted into different time blocks based on classrooms and faculty availability. We examine the effects of class time on student learning outcomes and student evaluations of teaching (SET). To control for non-relevant factors, we collected data from pairs of sections of the same business courses taught by the same instructors in the same term at a western public university. The main findings are that time blocks do not impact learning outcomes and SET; the most popular time blocks do not lead to superior or inferior outcomes as well. However, when a section is a repeated lecture on the same day, students usually have better academic performance but rate their teachers adversely.

* Corresponding Author. E-mail address: xxu@cpp.edu

I. INTRODUCTION

In many higher education institutions, instructors may teach the same course in multiple sections in an academic term (either a quarter or a semester). The teaching schedule for an instructor of a multi-section course is usually arranged on the same days of the week. For example, an instructor may teach a course on Tuesday and Thursday with one section at 8 a.m. and the other at 10 a.m.

There is little research on the learning outcomes from repeated lectures. One can argue that instructors may get bored when have to retell the exact content multiple times. It is also true that the physical fatigue can prevent the faculty from achieving the ideal delivery of the content. However, many faculty state the opposite that the later sections are usually better taught. One reason is that the instructors have a chance to adjust the lesson plan in the subsequent section after observing earlier audience reaction. If something does not work, it can be left off and replaced with a different approach. If students hang up on a topic in the first section, the experienced instructors will lighten the lecture by slowing down, providing examples and metaphors, or experimenting hands-on exercises. Due to these multiple factors, it can be argued either way if a later section is executed better than the earlier one. In this paper, we compared the student learning outcome measurements and instructor evaluation results between earlier and later sections of the same course taught on the same days of the week by the same instructors.

Often instructors claim students are more focused and alert during certain time of the day. We frequently heard the comments that classes at these "better learning times" receive better teaching evaluations from the students. We are curious if this is true and set off to test these hypotheses. We also wonder if the better teaching evaluation is associated with better learning outcomes of students. An article (Shellenbarger, 2018) from *Wall Street Journal* claims that most adults perform best in the late morning, and easily get distracted from noon till 4 p.m., then the focus and efficiency start to rise again in the late afternoon. If this is the case, then classes scheduled in late morning or late afternoon may deliver better learning outcomes. Along the same line, the students may give higher teaching evaluations to classes taught in late morning or late afternoon. We call late morning and late afternoon the "peak time" in this paper and we are interested in comparing sections in peak time with ones offered in non-peak time.

University course scheduling is often processed at the department or college level, based on the perceived student needs, faculty availability, and classroom constraints. At the higher education institution the authors work at, most classes are offered between 10 a.m. and 4 p.m. The concentrated course offering time may be attributed to student choice and faculty preference. However, it is important to examine if this has any effect on student learning outcomes and student satisfaction. In this research, we name 10 a.m. to 4 p.m. time block as "crowded time" and investigate if there is a significant difference in learning outcomes between crowded time and non-crowded time sections.

Finally, we also have a chance to examine the relationship between the exam scores and teaching evaluation results across sections of the same course taught by the same instructor. It is a well-received theory that easy grades and better instructor evaluations are highly correlated. Students tend to give favorable ratings to instructors who assign more As and Bs to students.

All in all, we have collected a sample of pairwise sections. Each pair contains two sections offered at different times on the same day(s), in the same term, which were taught by the same instructor. A systematic approach is utilized to examine and compare the learning outcomes and instructor evaluations to verify statistical relationships among class time and student / instructor performance.

The rest of the paper is organized as followings. Section 2 summarizes related literature. Section 3 explains how data are collected and processed. In section 4, hypotheses are presented and examined. In section 5, we conclude the paper and discussed future research topics.

II. LITERATURE REVIEW

2.1. Data Collection

A learning curve is a visual presentation illustrating the fact that better outcomes are associated with repeated practice or training. The concept was first introduced in studies of production efficiency. For example, Wright (1936) demonstrated the marginal cost of aircraft manufacturing decreases with output volume. Henderson (1968) introduced the Power Law into the learning curve and explained that the later half of the learning takes much longer time than the first half. Many contemporary literature used an S-curve or Sigmoid function to capture the relationship between the learning outcomes and efforts spent ((e.g. (Speicher, Nussbaum, White et al., 2014) and (Ward, Mohammed, Walt et al., 2014)). In an S-curved learning experience, we gain progress very slowly in the beginning, then advance at a much faster pace, and eventually slow down as learning activities approach the limit. However, there is little discussion in the literature about how the learning curve applies to the teaching skills. Our research contributes to the stream of research in learning curve by examining if there is statistically significant improvements in teaching effectiveness when instructors repeat the same lecture on the same day.

Learning outcomes of individual students may be affected by their chronotypes. Early birds and night owls perform quite differently in an early morning class. Breus (2016) popularized the theory of four chronotypes of human beings and Arcady, Olga, and Evgeniy (2015) confirmed similar findings. According to this stream of research, most people are productive during late morning and late afternoon. In this paper, we call these more efficient time of the day the "peak time" (Shellenbarger, 2018). It has shown that the chronotypes of human beings are the results of gene-environment interaction. Research shows that the clock gene plays a major role as an activator of many physical body functions leading to circadian rhythm (Dunlap, 1999; Nusinow, Helfer, Hamilton, King et al., 2012). Important environmental zeitgebers include light, feeding, social behavior, and work and school schedules all contribute to body's internal bio-clock. Little has been done on the effects of the scheduled class meeting time on academic performance. After an extensive search, we only found a conference proceeding by Akpom and Huller (1994) that is relevant to the topic. Their work compared evening class and daytime class and found no significant differences in student learning outcomes. Our research is built upon recent findings in human chronotypes and classifies classes between those offered at "peak time" and at "non-peak time."

Our research also relates to the stream of literature on factors that contributes to the student evaluations of teaching (SET). Student ratings of teaching are widely used by higher education institutions to measure teaching quality (Cashin, 1999; Clayson, 2009; Davis, 2009; Seldin, 1999). It is also a common practice to include student teaching evaluation results into decisions for faculty retention, tenure, and promotion. Interestingly, a lot of research argues that students' feedbacks are not effective ways to measure teaching competency (e.g. Hornstein, 2017; Beecham, 2009; Braga, Paccagnella, and Pellizzari, 2014; Spooren, Brockx, and Mortelmans, 2013). There are evidences to support that physical environment, grade expectations, and students' gender bias are highly related to teaching evaluations (Lizzio, Wilson, and Simons, 2002; Braga, Paccagnella, and Pellizzari, 2014; Boring, Ottoboni, and Stark, 2016)). In this paper, we examine if class meeting times contribute to the differences in teaching ratings.

III. DATA AND METHODOLOGY

To examine the effect of repeated lectures and class meeting time on learning outcomes and teaching ratings, we pair courses taught by the same instructors during the same academic terms (quarters). We collect measurements from the pairwise sections. The data allow us to limit the impacts of many factors (such as class size, instruction modes, instructor personality, among many others) on the teaching outcomes. Course materials in repeated sections are usually the same during the same quarter. The final cleaned sample consists of 39 paired sections from 2007 to 2018. The class starting time varies from 8 a.m. to 6 p.m.

The courses used in this study are required of all business majors at a public higher education institution in California. The co-authors (full time tenure track faculty members) have been teaching these courses with relatively mature and consistent content for more than 10 years. There are around 5,000 students majored in business on campus as of fall 2017. Students usually take these courses at their junior or senior years after finishing prerequisites on mathematic courses. The focus of the instruction is on critical thinking and analytical problem solving skills. Due to the large demand, multiple sections had been offered every quarter with the average of the class sizes from 35 to 41 per section. PowerPoint slides, whiteboards, projectors are used to facilitate face-to-face lecture delivering; course materials, grades, announcements, and other course related content is available on course Blackboard sites; assignments, quizzes, and additional practices are conducted through online learning management platforms. Exams consist of a mixture of 40% conceptual and 60% computational questions. These courses are challenging for many students. The average of final exams in our sample is in the range of 59.8% - 64.9%; the non-passing rate - the percentage of students with grades of Ds and Fs - ranges from 11.6% to 22.2% in our samples.

Two sets of data were collected to assess student performance. The first one is the reported end-of-quarter course letter grades retrieved from the registrar's office. For the first data set (course letter grades), we converted them to course GPAs according to the following scales:

Letter Grade	GPA	Letter Grade	GPA
A	4.00	C	2.00
A-	3.67	C-	1.67
B+	3.33	D+	1.33
B	3.00	D	1.00
B-	2.67	D-	0.67
C+	2.33	F	0.00

Based on student GPAs, we calculated class *success rate* - the percentage of students received letter grades of C- or better. This is a measure widely used in the literature and by administrators to assess the students' passing rate.

Accordingly, we calculated failure rate for a class as the percentage of students received letter grades of Ds, Fs, or unauthorized withdrawals. Students received those grades may repeat the courses to boost GPAs to avoid the disqualification risk.

The second measurement of student learning outcomes is the untreated final exam scores. Final exams usually are in a format of closed book, closed note with a time limit. The advantage of these data is that they are untreated and reflect student performance across sections; while course letter grades can be modified (with curve, extra credit, etc.) and have reduced variances due to combining a number of different assessment tools: participation, homework, quizzes, midterm and final exams. Although homework is a less effective indicator of academic performance, we still collected assignment scores and converted them to percentages. As students have access to the course materials and unlimited time to do the homework, homework percentage is a weaker indicator for academic performance but more of a measure of efforts. As a result, the average of homework tends to be high (86.6% to 93.1%), and variation across sections is usually low.

The same SET (student evaluations of teaching) form and process have been in effect in all sections over the whole time span of the study. The SET form includes 16 questions (see Table 2). The rating is based on a reversed five point Likert Scale with 1 being "Strongly Agree." For each evaluation question, three measures are constructed: the average evaluation rating scores (the lower the better), the percentage of students who responded with "strongly agree" or "agree", and the percentage of students who responded with "strongly disagree" or "disagree". SET has a summery item (Q16), in which students are asked the likelihood of recommending this instructor for the same course to other students. For this summary item, we collected the scores, the percentage of "agree / strongly agree" responses, and the percentage of "disagree / strongly disagree" responses.

IV. RESULTS AND DISCUSSION

4.1. The Effect of Immediate Repeating in Face-to-face Lectures

Many faculty members have talked about the improvement in content delivery when teaching the same lecture the second time during the day. Instructors may re-organize the content, change the examples, jokes, or media, and focus more on rough spots after observing students' response in an earlier section. In our paired sections, we compare the earlier and later sections in terms of learning outcomes and student evaluations of teaching (SET). The adaption to early audience responses may help instructors to teach more effectively and get better teaching ratings in the repeated sections. However, one could also argue that after teaching the same courses several times in a number of years, an instructor's teaching effectiveness could approach the limits and does not vary much over the time as the learning curve is usually S shaped. In addition, physical fatigue or a lack of excitement may prevent an instructor from doing his or her best in a later section.

Our first research question can be stated as follows: do later repeated sections receive better student learning outcomes? Accordingly, the first null hypothesis is:

H1o: There are no significant differences in the student learning outcomes between early and later repeated sections.

We use a dummy variable "repeated" and it takes value of "1" if a section was a later repeated lecture and "0" if a section was taught earlier during the day. A series of difference-in-mean T-tests were conducted to compare course letter grades, failure rate, success rate, and untreated final averages across paired sections. The results are reported in Table 1.

TABLE 1. EARLIER SECTION VS. LATER REPEATED SECTION DIFFERENCE-IN-MEAN T-TEST

Means	**Results**				
	Earlier Sections	**Later Repeated Sections**	**Difference in Means**	**p-value**	***N***
GPA	2.468	2.582	0.114	0.138	39
Student Success Rate	80.580	84.163	3.583	0.137	39
Student Failure Rate	15.743	14.992	-0.751	0.528	39
Final Exam	61.40%	63.60%	0.022	0.042	24
Homework	89.50%	90.50%	0.011	0.342	24
SET - Summary Item (Q16) Scores (lower is better)	1.679	1.777	0.098	0.094	32
SET- % Agree to Summary Item	82.0%	77.9%	-4.1%	0.039	32
SET - % Disagree to Summary Item	6.8%	9.1%	2.4%	0.109	32

From Table 1, it is clear that the later repeated sections are associated with higher untreated final exam scores, and the result is statistically significant at 5% level. The untreated final exam results indicate students' comprehensive proficiency in the content area. This finding moves us to reject the null hypothesis of H1o and accept there is a significant difference in student performance between earlier and later sections. Students in later sections, on average, also have higher letter GPA, higher successful rate, and lower failure rate, though the differences are not statistically significant.

The result manifests the fact that even the experienced professional instructors benefit from the same day repetition. To some extent, teaching is similar to performing in a show. There is never an identical performance even with extensive training and rehearsals. Experiments, forgotten lines, random incidents, and new findings are always part of teaching. Although the learning curve theory states that with experience, we obtain decreasing gains from repeating. The result here shows that teaching profession is actually a sophisticated art that takes a life time and a lot of repetition to be effective.

With better statistical significant student learning outcomes in the later repeated sections, one would expect a better teaching ratings for the later classes. We then test the following null hypothesis:

H2o: Student evaluations of teaching (SET) in later repeated sections are significantly better rated than those in earlier sections.

For SET, we first use the summary question (Q16) in the evaluation form: the likelihood of a student recommending the instructor to other students. We compared the average score for Q16, the percentage of responses that were "strongly agree" or "agree" (% Agree) to Q16, and the percentage of responses being "disagree" or "strongly disagree" (% Disagree) to Q16. The difference-in-mean T test results are presented in Table 1.

We are surprised to see that our null hypothesis H2o that later repeated sections have better teaching ratings does not hold. The later repeated sections, according to our data, achieved statistically significantly higher final exam scores, but received worse student teaching evaluations. The possible explanation is that when instructors teach better in class, naturally more efforts are demanded from students. These efforts could be comprehending a difficult theory that an instructor successfully introduced or mastering a complex algorithm when a professor kept students engaged long enough. Most students dislike this forced exertion of efforts and thus would be less likely to recommend this instructor. The summery item (Q16) is designed to measure overall teaching competency and to captures students' impression of instructors' efforts. However, in this particular study, we see that human beings' "energy saving" preference leads to unfair ratings to later repeated sections. Braga, Paccagnella, and Pellizzari (2014) confirmed similar observation in a different setup. They reported that lower SET scores lead to better student academic performance and suggested that comprehensive learning assessment at the end of the term is a better measurement for teaching effectiveness.

More interestingly, when we examined students' responses to all SET questions, 16 in total, we found out the later section SET scores are worse in a few items that are not relevant or explainable by the repetition of the same lecture. For example, the earlier and later sections in a pair shared the same office hours in a quarter. But the students of later sections rated instructors lower on availability during office hours (Table 2, Q10). Similarly, students rated later sections worse in terms of returning graded exams/assignments in a timely fashion (Table 2, Q9). In practice, however, the exams and assignments are collected and returned on exact same dates for sections in a pair. This means that the extra efforts spent in later sections by students contributes to unfair ratings to unrelated aspects of teaching. Our study is not the first one reporting biased SET ratings in non-relevant items. Boring, Ottoboni, and Stark (2016) found in a big data analysis that gender bias leads to lower SET for female instructors, not only in summary item, but also in random items that are considered non-relevant.

TABLE 2. DIFFERENCE-IN-MEAN T-TEST FOR 16 SET QUESTIONS.

		Earlier Sections	Later Sections	Difference-in-Means	p-value	N
Q1. Instructor encouraged critical thinking.	**Score** (Lower =Better)	1.660	1.724	0.065	0.258	32
	% Agree	0.847	0.836	-0.011	0.604	32
	% Disagree	0.042	0.050	0.008	0.586	32
Q2. Instructor helped me understand concepts.	**Score** (Lower =Better)	1.800	1.872	0.071	0.232	32
	% Agree	0.776	0.768	-0.008	0.730	32
	% Disagree	0.064	0.088	0.024	0.144	32
Q3. Instructor was genuinely interested in teaching.	**Score** (Lower =Better)	1.513	1.572	0.073	0.184	31
	% Agree	0.882	0.874	-0.011	0.471	31
	% Disagree	0.036	0.035	0.001	0.938	31
Q4. Instructor presented material clearly.	**Score** (Lower =Better)	1.852	1.938	0.086	0.162	32
	% Agree	0.757	0.748	-0.009	0.718	32
	% Disagree	0.085	0.106	0.021	0.142	32
Q5. Instructor responded well to student questions.	**Score** (Lower =Better)	1.648	1.761	0.113	0.054	32
	% Agree	0.818	0.803	-0.016	0.398	32
	% Disagree	0.053	0.071	0.018	0.095	32
Q6. Practical examples were used.	**Score** (Lower =Better)	1.597	1.696	0.099	0.106	32
	% Agree	0.855	0.804	-0.051	0.028	32
	% Disagree	0.045	0.061	0.016	0.189	32

Q7. Instructor was well prepared.	**Score** (Lower =Better)	1.467	1.496	0.028	0.607	32
	% Agree	0.900	0.901	0.001	0.978	32
	% Disagree	0.036	0.036	0.000	0.978	32
Q8. Grading system was clearly explained.	**Score** (Lower =Better)	1.479	1.537	0.058	0.324	32
	% Agree	0.901	0.882	-0.019	0.301	32
	% Disagree	0.031	0.045	0.014	0.194	32
Q9. Exams/assignments were returned within a reasonable period of time.	**Score** (Lower =Better)	1.422	1.497	0.075	0.121	32
	% Agree	0.923	0.900	-0.023	0.066	32
	% Disagree	0.028	0.040	0.011	0.095	32
Q10. Instructor was available during scheduled office hours.	**Score** (Lower =Better)	1.374	1.493	0.119	0.014	32
	% Agree	0.897	0.873	-0.023	0.124	32
	% Disagree	0.021	0.039	0.019	0.010	32
Q11. Class participation was encouraged.	**Score** (Lower =Better)	1.553	1.603	0.057	0.205	31
	% Agree	0.860	0.847	-0.015	0.336	31
	% Disagree	0.042	0.051	0.009	0.348	31
Q12. Instructor treated students with respect.	**Score** (Lower =Better)	1.357	1.432	0.075	0.123	32
	% Agree	0.933	0.913	-0.020	0.087	32
	% Disagree	0.030	0.038	0.008	0.258	32
Q13. Course objectives were clearly specified.	**Score** (Lower =Better)	1.509	1.598	0.088	0.094	32
	% Agree	0.893	0.869	-0.024	0.21	32
	% Disagree	0.034	0.044	0.011	0.259	32
	Score (Lower =Better)	1.435	1.496	0.062	0.271	32

Q14. Course topics described in the syllabus were covered.	**% Agree**	0.920	0.900	-0.02	0.161	32
	% Disagree	0.027	0.038	0.011	0.223	32
Q15. Exams/assignments were related to the material.	**Score (Lower =Better)**	1.497	1.554	0.057	0.283	32
	% Agree	0.894	0.867	-0.027	0.112	32
	% Disagree	0.045	0.047	0.002	0.814	32
Q16. I would recommend this instructor for this course to other students.	**Score (Lower =Better)**	1.679	1.777	0.098	0.094	32
	% Agree	0.820	0.779	-0.041	0.039	32
	% Disagree	0.068	0.091	0.024	0.109	32

4.2. The Effect of "Peak Time"

According to recent research (Nusinow, 2012), most adults' focus, efficiency, and performance peak in the late morning and late afternoon. The peak time arguments apply to both the instructors and students.

Our second research question asks if there is a "peak-time" effect in student learning outcomes and student evaluations of teaching (SET). A dummy variable "peak time" was created to differentiate classes scheduled at peak time and those at non-peak time. "Peak time" variable takes the value of "1" if a course starts between 9:30 to 11:00 a.m., or between 2:30 to 4:00 p.m., "0" otherwise. Accordingly, the next two null hypotheses are:

H3o: There are no significant differences in the student learning outcomes between "peak time" and "non-peak time" sections.

H4o: There are no significant differences in student evaluations of teaching (SET) between "peak time" and "non-peak time" sections.

Table 3 reports the paired T-test results on sections in peak and non-peak class times. There is no statistical difference in students' GPA or final exam scores. Instructors received similar student evaluation ratings in these classes. Based on these findings, we cannot reject null hypotheses H3o and H4o. No peak-time effect is detected in courses scheduled in the late morning and late afternoon. Note that Akpom and Huller (1994) reported no statistically significant differences in academic performance between day-time class and evening class when students were working professionals.

It is possible that students' out of class efforts may mitigate the peak time effects in learning. At the higher education level, students are expected to manage their learning, especially for juniors and seniors. If a student takes a course during non-peak time and feels extra effort is needed, she or he may spend proportionally more time after class for this course. Similarly, if a student gets more information in a peak time class, she or he may choose to allocate less study time on the subject. College students, on average, have better time management and learning capability than younger learners. As a result, the peak time effects in higher education is not quite visible when measured towards the end of the course. The implication of this result is that when scheduling classes, we do not need to limit courses with high failure rate to peak time blocks.

To confirm our findings, we also conducted similar tests in subsamples. Four pairwise subsamples were created. One subsample has all the pairs of sections with one starting at 8 a.m. The second one has al pairs of sections with one starting at 10 a.m. The third one contains pairs with one starting at 1p.m. These starting times are most frequent ones in our aggregate sample. The fourth subsample has pairwise classes with one taught in the late afternoon or evening (4 p.m. or later). Note that we expect more noticeable pairwise difference in the fourth subsample because the late afternoon sections are both peak time sections and repeated later classes.

Results in Table 4 echo the findings in Table 3: no academic performance or teaching ratings differences were found in 8 a.m. or 1 p.m. paired section comparisons. Students in 10 a.m. sections actually has a lower success rate to pass the course. The only exception is late afternoon classes, which have a better learning outcomes. This finding has to be interpreted with caution as it is hard to tell if it is from the peak time effects or repeated later section effects or both. Overall, the peak-time effect is not supported by our data. Therefore, we believe that the peak time effect may not be enough to affect overall learning outcomes, or are balanced out with the adjustment in students' learning efforts outside of the face-to-face lecture.

TABLE 3. PEAK TIME VS. NON-PEAK TIME DIFFERENCE-IN-MEAN T-TEST.

Means	**Results**				
	Non-peak Time	**Peak Time**	**Difference in Means**	**p-value**	***N***
GPA	**2.485**	**2.558**	**0.073**	**0.343**	**39**
Student Success Rate	**81.655**	**82.742**	**1.088**	**0.656**	**39**
Student Failure Rate	**14.581**	**16.5**	**1.918**	**0.104**	**39**
Final Exam	**62.30%**	**62.60%**	**0.002**	**0.845**	**32**
Homework	**90.50%**	**89.60%**	**-0.009**	**0.424**	**32**
SET - Summary Item Scores (Lower is better)	**1.738**	**1.748**	**0.011**	**0.864**	**24**
SET-Agree to Summary Item (%)	**80.5%**	**78.2%**	**-2.3%**	**0.286**	**24**
SET - Disagree to Summary Item (%)	**8.5%**	**7.9%**	**-0.6%**	**0.689**	**24**

TABLE 4. PEAK TIME SUBSAMPLES DIFFERENCE-IN-MEAN T-TEST.

		Subsample Difference-in-Mean T-test for			
		8 a.m. Pairs	**10 a.m. Pairs**	**1 a.m. Pairs**	**After 4 p.m. Pairs**
	No. of Pairs	10	18	11	11
GPA	**Difference in Means**	-0.375	-0.064	0.000	0.068
	p-value	*0.208*	*0.1610*	*0.999*	*0.046*
Student Success Rate	**Difference in Means**	-8.932	-3.123	-0.571	3.721
	p-value	*0.323*	*0.0750*	*0.672*	*0.085*
Student Failure Rate	**Difference in Means**	-1.362	3.212	0.313	-3.735
	p-value	*0.563*	*0.1110*	*0.799*	*0.134*
Final Exam	**Difference in Means**	-0.033	-0.005	-0.012	0.033
	p-value	*0.43*	*0.8250*	*0.456*	*0.010*
Homework	**Difference in Means**	-0.018	-0.005	0.011	0.007
	p-value	*0.621*	*0.8170*	*0.458*	*0.696*
SET - Summary Item (Q16) Scores	**Difference in Means**	-0.093	-0.016	0.016	0.087
	p-value	*0.323*	*0.8100*	*0.908*	*0.464*
SET- % Agree to Summary Item	**Difference in Means**	0.055	-0.014	0.031	-0.031
	p-value	*0.144*	*0.4640*	*0.468*	*0.311*
SET - % Disagree to Summary Item	**Difference in Means**	-0.013	-0.025	0.005	0.026
	p-value	*0.468*	*0.2110*	*0.862*	*0.363*

4.3. The Effectiveness of "Crowded Time"

At our institution under a quarter system, the class runs between 8 a.m. and 10 p.m. However, the most popular class time is between 10 a.m. and 4 p.m., a time block that is convenient to both students and instructors. We call this popular class time window "crowded time." Coming to class during the crowded time may allow students and instructors more time for sleep and meal, but it also means the hassle of parking and crowded classrooms. A dummy variable "crowded time" was created that takes the value of "1" if a class is taught during the crowded time, and "0" otherwise.

Our third research question aims to answer: do sections in "crowded time" receive better student learning outcomes and student evaluations of teaching (SET)? We test the following two null hypotheses:

H5o: There are no significant differences in the student learning outcomes between "crowded time" and "non-crowded time" sections.

H6o: There are no significant differences in student evaluations of teaching (SET) between "crowded time" and "non-crowded time" sections.

Results in Table 5 suggest students enrolled in classes at popular time block on average earned higher GPA and higher success rate, and also rated instructors more favorably; though the difference is not statistically significant, meaning we cannot reject the null hypotheses of H5o and H6o. This finding partially supports the common practice of scheduling classes during convenient hours.

TABLE 5. CROWDED TIME VS. NON-CROWDED TIME DIFFERENCE-IN-MEAN T-TEST.

Means	**Results**				
	Non-Crowded Time	**Crowded Time**	**Difference in Means**	**p-value**	***N***
GPA	2.455	2.544	0.089	0.417	27
Student Success Rate	80.994	82.857	1.863	0.588	27
Student Failure Rate	14.481	16.318	1.837	0.201	27
Final Exam	61.40%	61.40%	0.000	0.995	22
Homework	90.00%	90.80%	0.009	0.490	22
SET - Summary Item Scores (Lower is better)	1.796	1.738	-0.058	0.408	17
SET-Agree to Summary Item (%)	77.3%	78.8%	1.5%	0.460	17
SET - Disagree to Summary Item (%)	9.8%	7.2%	-2.6%	0.165	17

V. CONCLUSION

Instructors who have been teaching multiple sections of the same courses have been pondering at teaching experience: students in 10 a.m. sections seem to be smarter than those in 8 a.m. early morning classes; everyone is asleep at 1 p.m. and teaching ratings suffer for these sections. Prior literature does not provide answers to these questions. This study examines the effects of class time and repeated lectures on the student learning outcomes. In addition, we explore if student teaching evaluations change with class time. Using pairwise data sets, we conducted difference-in-mean T-tests to examine several related questions.

First of all, we confirmed a noticeable and statistically significant learning effects of instructors when they repeat the same lectures on the same day. Students in the later repeated sections received higher grades, higher passing rate, and significantly better final exam score. This finding confirms the hypothesis that teaching skills can be improved when lecture is repeated within a day.

However, better student learning outcomes in the repeated sections do not lead to better teaching evaluations. Quite to the opposite, instructors received worse evaluations in these sections. The areas of teaching received worse ratings could be somewhat arbitrary such as instructor's availability during office hours and promptness in grading and returning tests. We argue that when instructors teach better in the later section, which leads to higher academic achievements, students actually are forced to exert more efforts and thus rate instructors unfavorably. The finding is in line with Kornell and Hausman (2016) that the teachers who had received relatively lower teaching ratings appeared to have been most effective.

Our study finds little evidence on the effect of class time blocks on the learning outcomes or teaching evaluations. We examine if students in late morning and afternoon classes performed better due to a higher level of focus and efficiency. Our data do not support chronotype effect hypothesis. Students in peak-time classes performed similarly to their peers in non-peak time ones. Subsample results also show that students' performance and SET do not vary with class schedule.

Though courses are offered throughout the day, classes during certain crowded time are usually filled up first during the registration. These class time blocks (10 a.m. to 4 p.m.) are preferred by both students and instructors. We examine and find no differences in student learning outcomes and SET for classes in "crowded time" versus the ones in "non-crowded time."

Our study sheds valuable insights on the effect of repeated lectures and class time on learning outcomes and teaching evaluations. We contribute to the existing literature with the new multi-perspective understandings of the class time effects. The findings would be of value for students, faculty, and administers in making better decisions in registration, teaching preparation, and scheduling.

The paired sections in our sample are all undergraduate business core courses with a focus on problem solving skills. More research is recommended to verify the results in a variety of courses. In addition, we want to examine when students are younger or lack of learning capability, would the peak time effects lead to significant differences in the academic performance. The study of time of the day effects on learning outcomes and student satisfaction may also be applied to workshops and seminars. Therefore data could be collected through workshops and seminars as well. Finally, the days of the week effect may be examined and tested to see if there is a peak day in a week for learning purpose.

REFERENCES

Akpom, U.N. and Huller, I.S., "Class Attendance and Student Performance in Principles of Economics Course", *Midsouth Academy of Economics and Finance Papers and Proceedings*, 1994, 319-328.

Beecham, R., "Teaching Quality and Student Satisfaction: Nexus or Simulacrum?" *London Review of Education*, 7(2), 2009, 135–146.

Boring, A., Ottoboni, K. and Stark, P. B., "Student Evaluations of Teaching (Mostly) Do Not Measure Teaching Effectiveness", *Science Open Research*, 2016, doi:10.14293/S2199-1006.1.SOR-EDU.AETBZC.v1.

Braga, M., Paccagnella, M. and Pellizzari, M., "Evaluating Students' Evaluations of Professors", *Economics of Education Review,* 41, 2014, 71–88.

Breus, M., *The Power of When*, Little, Brown and Company, 2016.

Cashin, W. E., "Student Ratings of Teaching: Uses and Misuses", In P. Seldin (ed.), *Current practices in evaluating teaching: A practical guide to improved faculty performance and promotion/tenure decisions*. Bolton, MA: Anker, 1999, 25–44.

Clayson, D. E., "Student Evaluations of Teaching: Are They Related to What Students Learn? A Meta-Analysis and Review of the Literature", *Journal of Marketing Education*, 31(1), 2009, 16–30.

Davis, B. G., *Tools for Teaching* (2nd ed.), San Francisco, CA: Jossey-Bass, 2009.

Dunlap, J. C., "Molecular Bases for Circadian Clocks", *Cell*, 96 (2), 1999, 271–290.

Henderson B., "The Experience Curve", *Perspectives*, Boston Consulting Group, 1968. https://www.bcg.com/en-us/publications/1968/business-unit-strategy-growth-experience-curve.aspx (accessed October 13, 2018).

Hornstein H. A., "Student Evaluations of Teaching Are an Inadequate Assessment Tool for Evaluating Faculty Performance", *Cogent Education*, 4(1), 2017, doi: 10.1080/2331186X.2017.1304016.

Kornell, N. and Hausman H.,"Do the Best Teachers Get the Best Ratings?" *Frontiers in Psychology,* 7, 2016, 570.

Lizzio, A., Wilson, K. and Simons, R., "University Students' Perceptions of the Learning Environment and Academic Outcomes: Implications for Theory and Practice", *Studies in Higher Education*, 27(1), 2002, 27–52.

Nusinow, D.A., Helfer, A., Hamilton, E.E., King, J.J., Imaizumi, T., Schultz, T.F., Farré, E.M. and Kay, S.A., "The ELF4-ELF3-LUX Complex Links the Circadian Clock to Diurnal Control of Hypocotyl Growth", *Nature,* 475, 2012, 398-402.

Putilov, A. A., Donskaya, O. G. and Verevkin, E.G., "How Many Diurnal Types Are There? A Search for Two Further 'Bird Species'", *Personality and Individual Difference*, 72, 2015, 12-17.

Seldin, P., "Building Successful Teaching Evaluation Program", In P. Seldin (ed.), *Changing practices in evaluating teaching: A practical guide to improved faculty performance and promotion/tenure decisions,* Boston, MA: Anker. 1999, 213–242.

Shellenbarger, Sue., "The peak time for everything," *Wall Street Journal*, Updated September 26, 2018, https://www.wsj.com/articles/SB10000872396390444180004578018294057070544. (Accessed October 8, 2018).

Speicher, P.J., Nussbaum, D.P., White, R.R. et al., "Defining the Learning Curve for Team-Based Laparoscopic Pancreaticoduodenectomy", *Annals of Surgical Oncology,* 21(12), 2014, 4014-

4019, https://doi.org/10.1245/s10434-014-3839-7 (Accessed October 13, 2018)

Spooren, P., Brockx, B. and Mortelmans, D., "On the Validity of Student Evaluation of Teaching: The State of the Art", *Review of Educational Research*, 83(4), 2013, 598–642.

Wright, T.P., "Factors Affecting the Cost of Airplanes", *Journal of Aeronautical Science*, 3(4), 1936, 122-128.

Ward, S.T., Mohammed, M.A., Walt, R., Valori, R., Ismail,T. and Dunckley, P., "An Analysis of the Learning Curve to Achieve Competency at Colonoscopy Using the JETS Database", *Gut*, 63,2014, 1746-1754.