4th Time Around: Do Classes Get Better with Instructor Repetition?

Dr. Joshua L. Hertz, Northeastern University

Dr. Hertz earned a B.S. in Ceramic Engineering from Alfred University in 1999 and then a Ph.D. in Materials Science and Engineering from the Massachusetts Institute of Technology in 2006. Following this, he worked at the National Institute of Standards and Technology as a National Research Council postdoctoral fellow. He joined the Department of Mechanical Engineering at the University of Delaware as an Assistant Professor in September 2008, leading a lab that researched the effects of composition and nanostructure on ionic conduction and surface exchange in ceramic materials. In 2014, he moved to Northeastern University to focus on teaching and developing curriculum in the First Year Engineering program.

Dr. Richard Whalen, Northeastern University

Dr. Richard Whalen is a Teaching Professor at Northeastern University in Boston, MA and is Director of First-year Engineering. The mission of the First-year Engineering team is to provide a reliable, wide-ranging, and constructive educational experience that endorses the student-centered and professionally-oriented mission of the University. He also teaches specialty courses in the Department of Mechanical and Industrial Engineering at Northeastern and has published and presented papers on approaches and techniques in engineering education. He has won multiple Outstanding Teaching Awards at Northeastern and numerous Best Paper and Best Presentation Awards with fellow First-year faculty coauthors at ASEE.

Prof. Constantine Mukasa, Northeastern University

Constantine Mukasa received a B.S. degree in Computer Engineering from Bethune-Cookman University, Daytona Beach, Florida, USA in 2007, and his M.Sc. and Ph.D. degrees in Electrical Engineering from Florida Atlantic University, Boca Raton, Florida, USA, in 2013 and 2017, respectively. He is currently an Assistant Professor at Northeastern University in Boston, MA. His research interests include Engineering Education, Wireless Communications, satellite and mobile communication Systems, vehicular networks, wireless network connectivity, and interference modeling.

Dr. John Sangster P.E., Northeastern University

Dr. Sangster is an Assistant Teaching Professor in the First Year Engineering program at Northeastern University. Prior to joining Northeastern in 2018, he served for three years as an Assistant Professor in the Department of Civil Engineering at the University of Nebraska. He received his Ph.D. in 2015 from Virginia Tech in Civil Engineering with a focus on Transportation.

Fourth Time Around:
Do Classes Get Better with Instructor Repetition?

Abstract

This complete, evidence-based practice paper presents a study that compares student learning outcomes and perceptions of class experience for students enrolled in different sections of a course taught by the same instructor in the same semester. For a variety of reasons, instructors in first-year engineering programs are particularly likely to lead multiple sections of a course each semester. A commonly held perception among instructors is that student grades and student course evaluations improve, on average, in sections that are taught during later repetitions within their daily or weekly schedule. Relatedly, instructors (and students) commonly think of an early morning class as less effective than a class later in the day; however, the cause may be—at least in part—due to the fact that an earlier class is more likely to be the first time this semester that the instructor is teaching that day's content. In this paper, we describe our findings in a study that compiles data from multiple instructors and multiple years. Final semester grades were used as a metric of student learning outcomes, and responses to quantitative questions in the end-of-semester student course evaluation surveys were used as a metric of student perceptions of their class experience. Contrary to instructor perceptions, only a slight improvement of a few percentage points in the student perceptions was found for the third section in the instructor's week. This increase was found across multiple metrics and, though small, is found to be statistically significant. A case study of a new instructor teaching new content is included, examining the confounding factors that may be masking the expected impact of improvement with repetition.

Introduction

Those of us who teach in the first-year community are often tasked with teaching multiple sections of the same course in each semester. Recently, faculty at Northeastern University began to discuss the perceptions we all felt in that we seem to have a better class experience later in the sequence of classes, when presenting the same material a second, third, or even fourth time. The authors decided it was time to see if these perceptions were accurate and if students felt the same.

While these students do receive a nominally identical course experience, instructors who present the same course content multiple times within a day or week will inevitably modify their approach each time. For example, instructors are better able to predict the areas that will confuse students after witnessing the response to course delivery to an earlier section, and they are likely to respond by increasing the time or modifying the approach devoted to that topic in later lectures that day or that week. Do students in these later sections respond more favorably to the adjustments? Are the faculty perceptions correct? Does the data show these perceptions are more likely to be accurate for new faculty and less likely for more seasoned faulty? What are the other factors that might also be in play, such as student GPA and time of day for the class?

This paper will answer these as well as other questions by comparing student learning outcomes and perceptions of class experience for students enrolled in different sections of a first-year cornerstone of engineering course taught by the same instructor in the same semester. The data is compiled from the end-of-term course evaluation surveys administered anonymously by the university over a period of several years. Final semester grades are used as a metric of student learning outcomes, and responses to quantitative questions in the end-of-semester student course evaluations are used as a metric of student perceptions of their class experience.

Literature Review

Many studies have looked into characteristics that impact student evaluations of teaching (SET), including student, instructor, and environmental aspects. It has become part of the lore of academia that many kinds of external factors impact SET, with instructors confidently stating that many factors are the cause of their evaluations, both positive and negative. The literature is riddled with studies focusing on student evaluations from a handful of classes in a single department, examining student factors such as gender, maturity, and academic level, reason for taking the course, and anticipated grade; course-related factors such as grading, class size, course content, and timing of the course; and instructor related factors such as gender, age, experience, rank, research or teaching focus, and attitude and influencing tactics [1]–[4]. A few studies seek to explore nuance with long lists of questions about faculty behavior, looking to correlate specific actions with SET [5]. Ultimately, meta-studies are employed, such as the oft-cited article by Pounder which systematically assesses the various forms of bias in SET and espouses a move toward capturing the "richness and diversity of what happens in a typical classroom" a concept not encompassed by SET [6].

An often examined question is whether time of day impacts the teaching evaluation, and many studies have looked at various aspects of this question. A number of studies that looked at time of day found it to not be significant, instead finding that other factors had a larger effect, such as class size [1], [2], [4]. Conversely, the meta-study by Pounder also mentions prior research that implies that the course evaluations themselves may be impacted by the time of day and day of the week on which they are conducted [6].

One interesting angle on the time-of-day question came from Heckert, et al.; instead of examining the time of completion of the SET itself, they looked at the impact of a student's satisfaction with the time of day of their scheduled course, and found this factor to be significantly related to all dimensions of teaching performance [3]. This question of timing keeps coming back into focus not because the research is greatly divided on the topic, but because it makes logical sense to both faculty and students that the time of day at which a class is held will impact the quality of the course evaluations. In a survey of students and faculty by Hinkin, only 30% of faculty and 60% of students believed that time of day would *not* impact teaching evaluations; their evaluation data found that time of day had no effect on SET despite these fairly widespread expectations that it would be there [7].

The relationship between time of day and quality of teaching, or at least student *perceptions* of the quality of teaching, becomes a critical question for administrators in the context of first-year engineering programs. The growing body of research in the area of retention and matriculation of students has consistently suggested a move away from department-specific introductory seminar courses, and toward separate unified programs in first-year engineering [8]. Although the current percentage of programs adopting the first-year engineering model is unknown, a study from 2005 by Brannan and Wankat found that 17 out of 99 universities responding to their survey used this model, with an assumption by the authors of this report that this percentage has increased significantly in the subsequent 15 years [9]. Along with this move to first-year programs come courses where faculty teach multiple sections per week of the same course. The time-of-day versus teaching effectiveness question can then be examined for the content of a single course, controlling for many of the external factors that lead to mixed results and confounded data in the studies referenced above.

The First-year Engineering Program

The Northeastern University College of Engineering, following a successful pilot in 2014, decided to adopt a "Cornerstone to Capstone" curriculum design for all incoming first-year engineering students. The Cornerstone course incorporates hands-on, project-based design work with computer programming. Previously taught in two separate first-year courses, the new Cornerstone course model blends programming and design in a way that demonstrates the intertwined nature of the two skills. The project-based Cornerstone includes occasional incongruent learning of course content. By highlighting the fact that problem-solving in engineering brings together groups of competencies in a networked rather than a linear fashion, the program improves instruction by demonstrating that this incongruence is acceptable. Specifically, the emphasis is on the Cornerstone as an example of how engineering can develop practical problem-solving applications.

At Northeastern University, the first year is common for all engineering majors. The Cornerstone course was carefully designed to help first-year students achieve success in the program regardless of the specific engineering major they select in their second year.

Therefore, the course includes themes centered on several design-and-build projects with the following program objectives:

1. Provide students with the opportunity to experience engineering as an evolving, creative, and interdisciplinary career that impacts global society and daily life.

2. Provide students with the opportunity to develop process-driven problem-solving skills that recognize multiple alternatives and apply critical thinking to identify an effective solution.

3. Provide students with the opportunity to integrate math & science in an engineering context.

4. Create motivated and passionate engineering students by challenging them with authentic engineering problems across multiple disciplines.

5. Instill in students the professional, personal and academic behaviors and common competencies needed to move to the next stage of their development.

The Cornerstone courses are taught in two forms to accommodate the varying needs of our first-year students. There is a 4 credit two-semester *split* version where all course content is divided between the Fall and Spring semester and an 8 credit one-semester *full* version where all the content is presented in either the Fall or Spring semester. The need to create the 8-credit *full* version was driven by the fact that at the university first-year students entering with a large number of AP credits would be able to jump into a discipline-specific course in the Spring semester of their First-year instead of the Fall semester of their Sophomore year.

In a typical semester, there are 5-7 8-credit *Full* Cornerstone sections offered and 18-22 4-credit *Split* Cornerstone sections. Enrollment in each section is capped at 32 students. Each *Full* Cornerstone section is taught by one faculty member in a semester with the same students. Professors teaching the Split Cornerstone sequence almost always will have the same students enrolled in the Fall and Spring semesters. On occasion, due to scheduling conflicts a student who was in a Fall semester section with one professor may be moved out and placed in another professor's section (<2% of the total enrollment of ~600 students per term in a *split* section). Faculty full loads are 12-credit hours per semester. Those teaching a *Full* Cornerstone section will usually be assigned another 4-credit course of a different topic to make a full load or be assigned a second *Full* Cornerstone as a 4-credit overload. For those teaching the Split Cornerstone section, they are assigned 3 sections of Cornerstone in both the Fall and Spring term. Assigning courses this way minimizes the number of course preps a faculty member has to do while maximizing the opportunity to repeat a lecture to refine and perfect it.

Methods

At the end of each semester, Northeastern University conducts a university-wide Teacher Rating and Course Evaluation (TRACE) survey that collects students' responses to specific questions for each course they were enrolled in during that semester. TRACE was inaugurated in the Fall of 2011 and the responses are collected electronically before final grades are posted. The data is compiled for each course or section and sent to the instructor after final grades have been posted. Typical response rates at the university are 60-65%; data included in this study had a minimum of 55% and a maximum of 100%.

For this study, we were interested in investigating the perception that an instructor teaching several sections of the same course in a given semester gets better as they present the material from one section to the next. With this in mind, we requested first-year engineering faculty who teach several sections of the same course each semester to provide specific data from their TRACE evaluations. All classes considered in this study were first-year engineering courses with the same content. Both Full and Split Cornerstone sections were included in the data. Honors sections were removed and their data was not included in this study because grades and other data were significant outliers from the non-Honors sections (of which there were many more). After this elimination, we had a total of 82 class sections split among 7 instructors and 9 semesters of data to use for the study. Students were assigned to sections at random by their first-year academic advisors, with no student ability to select their instructor or class meeting time.

Data was not keyed to an individual instructor by name, only by an anonymous code. The instructors had a range of experience as college instructors leading this course from one to six years. In order to remove the instructor as a confounding factor, a normalization procedure (described later in this section) was performed on the data.

The data collected for this study included the section's registrar scheduling block (which can be correlated to a class start time), the academic year and semester the course was taught, and the section's average end-of-semester grade. Based on the set of scheduling blocks present for an instructor within a particular semester, we were able to trace the order in which the instructor presented the course material that semester, that is, what was the first section the instructor taught in any given week of that semester, what section followed, and so on. The assumption here was that the instructor always presented new material in the first section taught. For example, if an instructor has sections that meet at 8 AM on Monday and Wednesday, 11:50 AM on Monday and Thursday, and 8 AM on Tuesday and Friday, those would be encoded as order = 1, 2, and 3, respectively.

A cross-tabulated listing of the number of sections studied for each combination of start time and order in the instructor's week is given in **Table 1**. An order value of 1.5 corresponds to one of two possible scenarios: 1) a section's weekly meeting schedule made it the first one led by the instructor for one of the weekly meetings but the second one for another of the weekly meetings, or 2) the length of the sections' meeting times were different such that the instructor presented new content in the first section and then presented the same content in the next section but with extra new content due to a longer class period.

Table 1. Crosstab of sections for which data is used in this study. A total of 7 instructors and 9 semesters of data are included.

		Start time								
		8:00	9:15	9:50	10:30	11:45	13:35	14:50	16:35	Total
Order of class within the instructor's week	1	2	8		4	6				20
	1.5					11		11		22
	2		1	3	7	1	6			18
	3					3	13		5	21
	4	1								1
	Total	3	9	3	11	21	19	11	5	82

For this study, we were most interested in TRACE questions that focused on the instructor's apparent level of preparation, class time organization, and ability to communicate. Specifically, instructors were requested to include data corresponding to the following questions:

1. The lectures helped me to learn.
2. The in-class discussions and activities helped me to learn.
3. I learned a lot in this course.
4. The instructor clearly communicated ideas and information.
5. The instructor came to class prepared to teach.
6. The instructor used class time effectively.
7. The instructor displayed enthusiasm for the course.
8. What is your overall rating of this instructor's teaching effectiveness?

TRACE asks students to provide numerical ratings to these questions on a 5-point scale, with only integer responses allowed. For this study, each section's average response value was normalized by instructor and semester to remove both the instructor-to-instructor and semester-to-semester variations. A normalized value of 1 thus represents an average section for *that* instructor *that* semester. Values above 1 indicate that the section had a more positive response or higher course grade relative to that instructor's other sections that semester, and values below 1 indicate that the section had a more negative response or lower course grade relative to that instructor's other sections that semester. Similar normalization was performed with the sections' average end-of-semester course grade such that values above 1 indicate that the section had a higher final average course grade relative to that instructor's other sections that semester.

Results

First, we look at the overall results submitted by the faculty in this study. The values in this section represent the non-normalized, section-wide average final semester grade and TRACE values. In general, the results reported by the instructors showed students achieving high average grades in the courses as well as having positive student perceptions of learning and of the instructor. The average end-of-semester grade across all sections considered in this study is 91.7, corresponding generally to an A-, with a minimum of 85.1 and a maximum of 98.0. As an example of the student perception scores, the average overall rating of instructor effectiveness was 4.60, with a minimum of 3.3 and a maximum of 5.0. For comparison, over the last few semesters, the university-wide average score on this question has been 4.4.

Contrary to what some may expect, **Figure 1** shows that there is no meaningful correlation in our data between the section's average course grade and the students' perceptions on either the overall rating of instructor effectiveness or the amount learned in the course. Trendlines show, in fact, a *negative* slope, but the R^2 values are both extremely close to 0. It should be noted that there is the possibility of sampling bias, since the choice of whether a student submits a TRACE survey may correlate with the student's expected final grade.

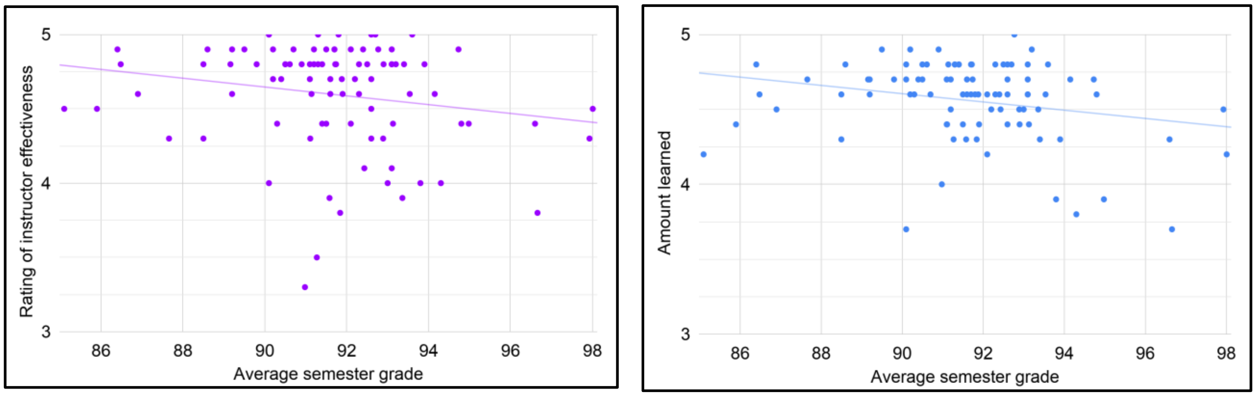

Figure 1. For all sections included in the study, there is no correlation between the average grade received by the section and that section's students' average response to (*left*) "What is your overall rating of this instructor's effectiveness?" or (*right*) "I learned a lot in this course." Student responses are encoded as 5 = "almost always effective" / "strongly agree", 4 = "usually effective" / "agree", 3 = "sometimes effective" / "neutral", 2 = "rarely effective" / "disagree", 1 = "never effective" / "strongly disagree". The trendlines have R^2 values of 0.036 and 0.056, respectively.

Student Perceptions

We now take a closer look at the effect of instructor repetition upon the students' perceptions. As described in the Methods section, the data from here have been normalized such that both the instructor-to-instructor and semester-to-semester variations have been removed. When averaging over the order of the class within the instructor's week, we find that the student perceptions collected in this data often have a total range of less than ±1%.

The four notable exceptions are given in **Figure 2**. In each of these cases, the instructor's third class meeting within the week has a better-than-average value. For three of the questions, all sections before the third one in the instructor's week have a very nearly average student perception score. These three TRACE questions refer to the instructor's clear communication of ideas and concepts, the instructor's effective use of class time, and the overall rating of instructor effectiveness. In the case of the fourth question, which referred to student perceptions on whether the lectures helped them to learn, both the first and third sections in the instructor's week exhib higher-than-average student perceptions. Though the overall change in average student perception is relatively small, these results correlate with the commonly held perception among the authors and their colleagues that our use of class time and our ability to clearly communicate improves with repetition.

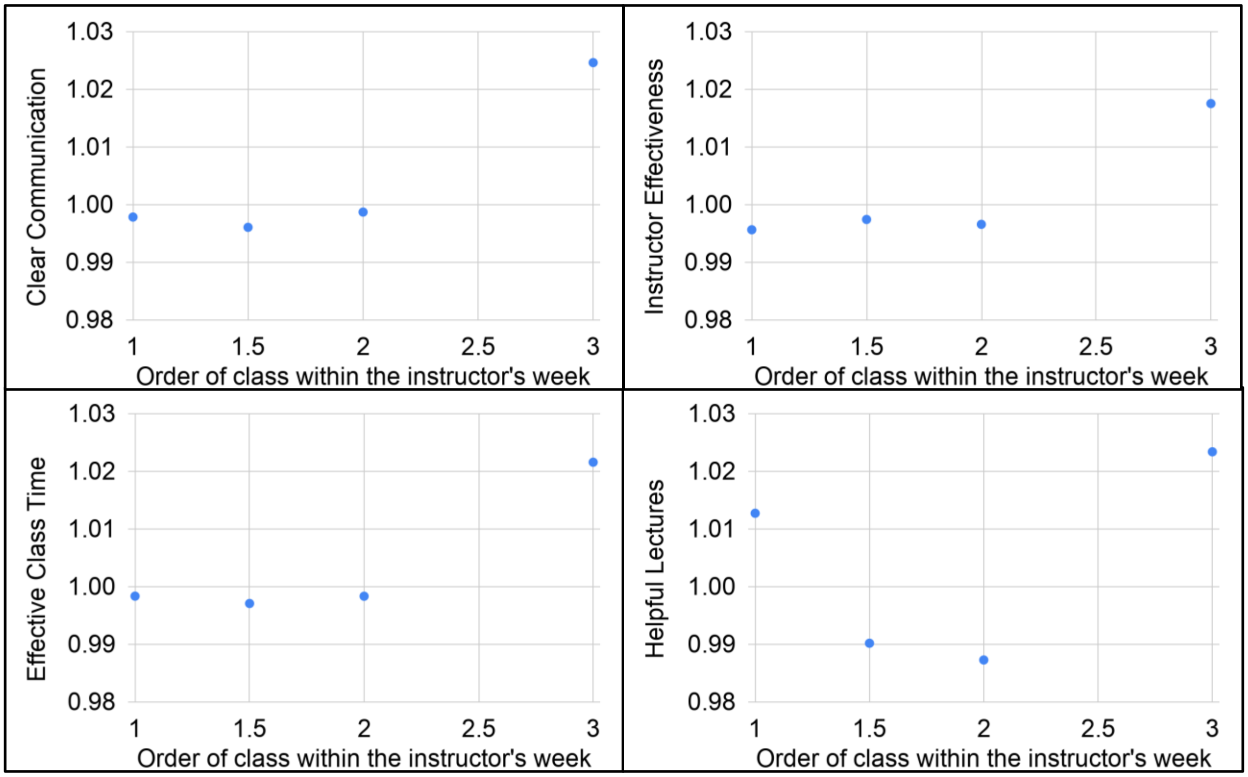

Figure 2. The average normalized student perception score for the questions (*clockwise from top-left*) 1. "The instructor clearly communicated ideas and information.", 2. "What is your overall rating of this instructor's teaching effectiveness?", 3. "The lectures helped me to learn.", and 4. "The instructor used class time effectively." In these four areas, student perception values differed by 1% or more from average for different class meetings within the instructor's week.

Since we observed that there appeared to be a difference in the values for the instructor's third section relative to the rest, we further investigated the data to find evidence of the statistical significance of this difference. We first performed the Analysis of Variance (ANOVA) to test the null-hypothesis that there is no significant differences in the sample means among the sections. We reject the null-hypothesis alternate based on the *p*-values presented in **Table 2**, which implies that there are significant differences between at least two of the means. We further investigate using a *t*-test to examine whether the third section is significantly different from the others (i.e., compared to the combination of all sections before the third). The results of our analysis are presented in **Table 3**, below. Since all *p*-values are less than 0.05 and 0.016 (using Bonferroni's Method), we conclude that the results from the instructor's third section were statistically different from the other sections. Thus, we conclude that the sample mean for the third section is significantly different from the rest of the sections.

Table 2. Summary of the ANOVA statistics between the sections.

Question of interest concerning instructor	F-value	*df*	*p*-value
Clear communication of ideas and concepts	3.796	3(B),75(W)	0.014
Effective use of class time	3.204	3(B),73(W)	0.028
Overall rating of effectiveness	3.119	3(B),73(W)	0.031
Lectures helped me learn	3.085	3(B),73(W)	0.032

Table 3. Summary of t-test statistics between the third and the rest of the sections.

Question of interest concerning instructor	*p*-value	*t*-statistic	*df*
Clear communication of ideas and concepts	0.001	3.399	38
Effective use of class time	0.008	2.840	28
Overall rating of effectiveness	0.010	2.760	28
Lectures helped me learn	0.004	2.995	54

When comparing student perceptions across the time of day that the section meets, we find for all questions that the middle-of-the-day meeting times are all highly similar. We do find a lower-than-average response for 8:00 AM classes and a higher-than-average response for 4:35 PM classes, but, notably, very few sections with these meeting times were included in this study. As a typical example, Figure 3 shows the normalized student perception scores for the question "The instructor came to class prepared to teach" plotted against the time of day that the class started. Data for each section included in this study are shown along with the average value for that time of day.

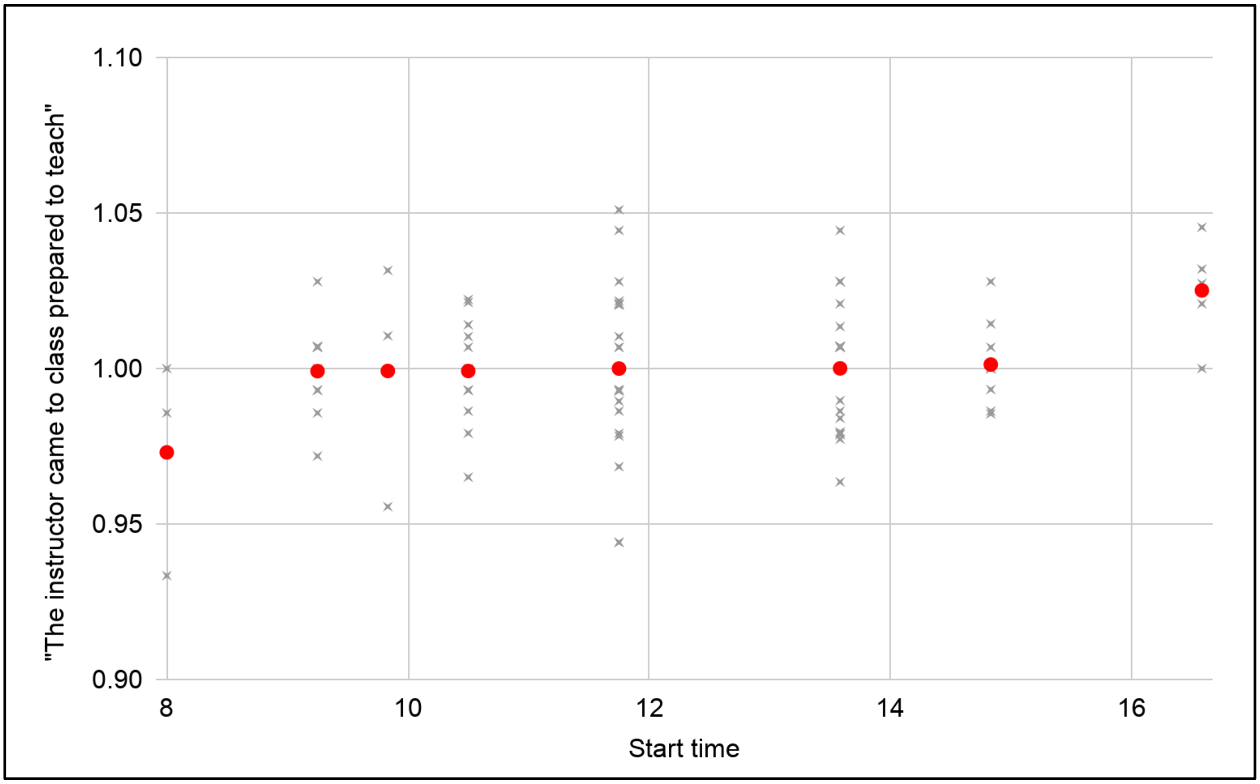

Figure 3. The normalized student perception score for the question "The instructor came to class prepared to teach" versus the hour start time of the class (24 hr clock). Every section included in this study is represented with a gray x, and the average value within each start time is represented with a red circle.

Learning Outcomes

In **Figure 4**, we observe that the final semester grade has no consistent correlation with the time of day that the class is held. On the other hand, in **Figure 5**, we observe that there is a positive trend between the final semester grade and the order of the section within the instructor's week. More specifically, the first section meeting in the instructor's week appears to be negatively affected. In fact, of the 20 sections that represent an instructor's first section meeting of the week, only one had a better-than-average final semester grade.

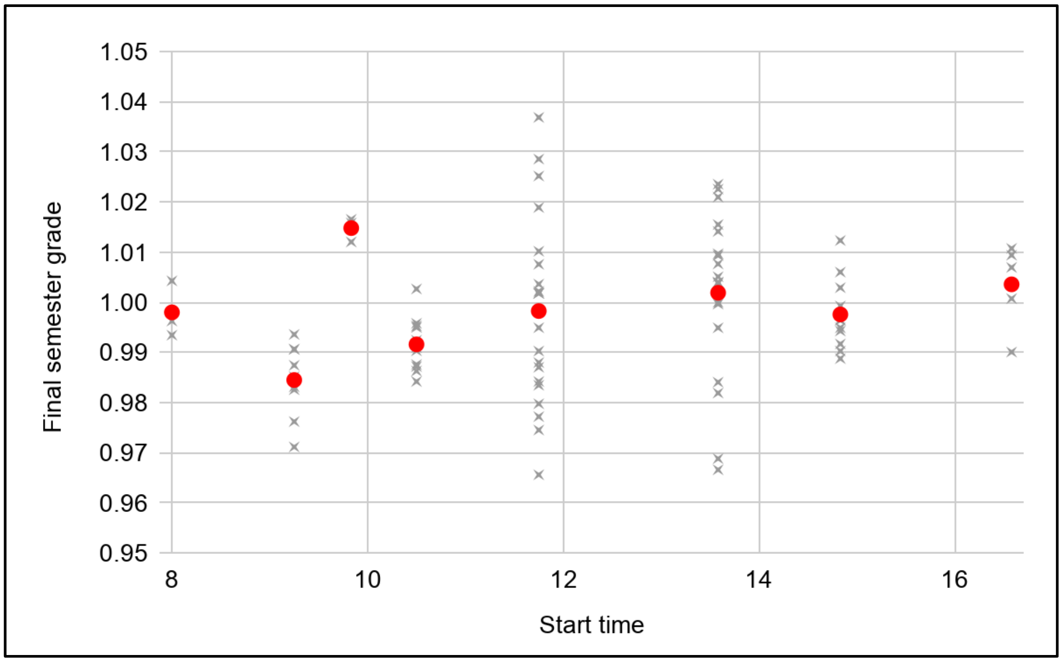

Figure 4. The normalized average student course grade versus the hour start time of the class (24 hr Clock). Every section included in this study is represented with a gray x, and the average value within each start time is represented with a red circle.

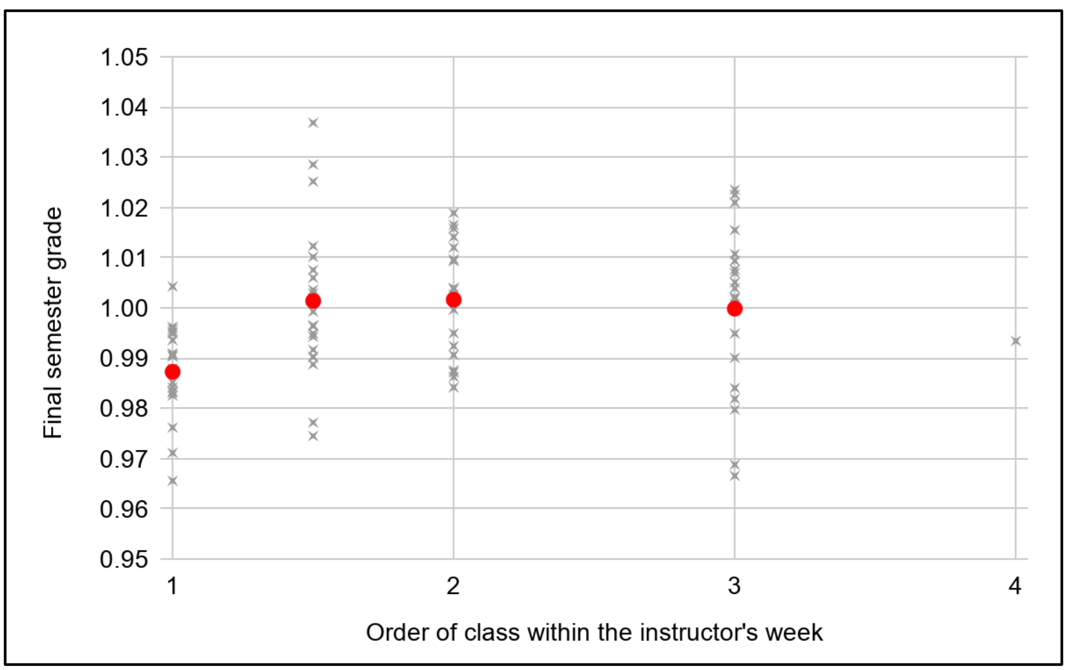

Figure 5. The normalized average student course grade versus the order of the class within the instructor's week. Every section included in this study is represented with a gray x, and the average value within each start time is represented with a red circle.

Case Study: New Faculty with New Content

A parallel case study was conducted in the previous year at Northeastern University by a faculty member new to the department. A previous study, presented in [10], was not focused on improvement over the day, but instead on the impact of the student prior experience and perceived need in the various content areas of the class, and how these prior experiences impacted time spent, grades, and ultimately retention. However, because of the detailed nature of the data collected for the prior study, it serves as an important case study supplement to the data presented above. In the data presented above, the authors did not observe the expected clear trends of improvement, given multiple deliveries of the same content throughout the week. What the inclusion of the case study helps to show is the great number of confounding factors present in each classroom, with the heterogeneity of classes greatly overshadowing the marginal proficiency increases in content delivery. It is difficult to discern trends from the data presented in the case study because the individual students in each instance of the class impart a great effect on the collective students' experience of the class.

A single faculty member for which data was collected in the prior study had a wide range of experience levels with the content of the course, including extensive workplace experience as a consulting engineer utilizing AutoCAD and Excel, but minimal experience with MATLAB, and no experience within the previous 20 years in either C++ or SolidWorks. As a result, the faculty member was routinely learning new material throughout the semester before teaching the same material to the classes. With three sections of the same course taught by the instructor, this case study presents the worst-case scenario for a faculty needing to improve with experience, but the best-case scenario to verify the hypothesis that teaching improves with repetition, particularly for faculty teaching a new course or new content.

The data below was collected as part of the previous study, and comes from three sections of the first-half of the year-long engineering cornerstone courses at Northeastern University. The pattern of responses to the TRACE evaluation questions that was found from the department-wide analysis presented earlier in the paper is consistent with the experience of the first-time faculty, where improvement is seen specifically in the areas of "the instructor clearly communicated ideas and information" and "the instructor came prepared to teach" as the day progressed. However, these improvements in instruction were not seen to result in improvements to grades overall. This data is shown below, in **Table 4**.

Table 4. New-faculty TRACE evaluation results by section.

Block	Time	Size of Dataset	Grade Overall	The lectures helped me to learn.	The in-class discussions and activities helped me to learn.	I learned a lot in this course.	The instructor clearly communicated ideas and information.	The instructor came prepared to teach.	The instructor used class time effectively.	The instructor displayed enthusiasm for the course.	What is your overall rating of this instructor's effectiveness?
1	8:00-9:05	29	92.6	3.5	3.7	3.4	3.8	3.7	3.6	4.4	3.7
3	10:30-11:35	26	94.5	3.1	3.6	4.0	3.8	3.6	3.6	4.2	3.7
4	1:35-2:40	30	92.1	3.5	4.1	3.9	4.1	4.3	3.8	4.6	4.0

A long-standing complaint with SET is that it fails to provide context for the nuance of the classroom situation, assuming homogeneous populations in every classroom and using this assumption to allow for comparison of instructors against an assumed standard metric. The data collected in the fall of 2018 for the case study previously performed captured additional information that is pertinent to the current analysis, and may shed some light on the confounding factors that may be limiting the size of the improvement being measured with multiple iterations of a lecture/class. Additional data collected from the previous study includes objective and subjective measures of student performance in the course, with detailed information regarding their performance on the C++ material, having been identified by the faculty as their area of least expertise going into the semester. Student perceptions of their prior experience with programming, as well as their assumed need for programming in their future degree program/career was collected in the first week of the semester. Student time on homework was reported with each assignment, and the aggregated total time spent on C++ homework is provided along with the aggregate grade earned on the C++ homework, as well as the C++ exam. At the end of the semester, students were asked to report their perception of the difficulty level of the C++ content in the course, as well as additional questions about the faculty performance such as "I was respected by the instructor," "my success was important to the instructor," and "I was made to feel like I belonged in the class." The aggregated results of this portion of the data are provided below in **Table 5**, with analysis following.

Table 5. Additional contextual information for student success by section.

Block	Time	Size of Dataset	Grade Overall	Grade on C++ Exam	Grade on C++ Homework	Time on C++ Homework	Prior Experience with Programming	Assumed Need for Programming	Percieved Difficulty of C++ Content	"I was respected by the instructor."	"My success was important to the instructor."	"I was made to feel like I belonged in the class."
1	8:00-9:05	29	92.6	85.3	93.7	22.1	2.11	3.93	2.93	4.79	4.61	4.57
3	10:30-11:35	26	94.5	88.6	95.1	27.3	1.65	3.50	3.16	4.84	4.72	4.44
4	1:35-2:40	30	92.1	82.9	92.8	22.0	2.25	4.10	2.86	4.90	4.79	4.79

The picture that develops from the additional data begins to show the stark lack of homogeneity between class sections. The middle section of the day had significantly less prior experience with programming and perceived less need for programming in their future degree/career, relative to the earlier and later classes; subsequently, they spent much more time completing their homework assignments, ultimately earning better grades on homework, exams, and overall in the course. They perceived the difficulty of the C++ content to be higher than their peers who had more prior knowledge coming in, but their lack of experience led to better outcomes on the course-designed metrics than their more-experienced peers achieved. There were marginal gains through the day in terms of student perceptions of being respected and belonging, but these perceptions are not tied back to performance in the class either of grade earned or perception of "I learned a lot in this course." Examining the TRACE evaluations against student performance on homework, the metrics of lecture and discussion quality match up with time spent on the homework, with students spending less time to complete their work if they believed that they got more information from lecture and discussion, but this seems to a function of prior knowledge in the topic area, and does not track at all with the iteration of the course being taught.

There are a great many ways in which the data above can be sliced and recombined that may help to better understand the dynamics of class performance and how it changes with multiple iterations of course content through the day/week, but the key take-away is that these sections are far from homogeneous, and that the relatively small upward trends in performance may be swallowed up by confounding factors in the data. The authors of this paper hold a belief based on logic that our instruction improves with multiple iterations of the content, but that the data ultimately may fail to show this because of the many extraneous factors that impact the learning outcomes for each class.

Conclusion

In this paper, we have shown that, after normalizing to remove semester-to-semester and instructor-to-instructor variation, the effects of instructor repetition of course content within the week upon student perceptions of teaching and final course grades are statistically significant yet small, on the order of a few percentage points. The effects may be larger for faculty who are new to teaching or lack background knowledge in particular concepts. Still, natural variation in student population between sections—even when these populations are created at random—can easily mask these effects. Future research is recommended to look more specifically at disciplinary versus general engineering design topics, as well as student perception and grades on particular content areas as opposed to the overall course.

References

[1] M. J. Seiler, V. L. Seiler, and D. Chiang, "Professor, Student, and Course Attributes that Contribute to Successful Teaching Evaluations," *Financ. Pract. Educ.*, 1999.

[2] R. B. Bausell and C. R. Bausell, "Student Ratings and Various Instructional Variables from a Within-Instructor Perspective," *Res. High. Educ.*, vol. 11, no. 2, pp. 167–177, 1979.

[3] T. M. Heckert, A. Latier, A. Ringwald, and B. Silvey, "Relation of Course, Instructor, and Student Characteristics to Dimensions of Student Ratings of Teaching Effectiveness," *Coll. Stud. J.*, 2006.

[4] S. Liaw and K.-L. Goh, "Evidence and control of biases in student evaluations of teaching," *Int. J. Educ. Manag.*, pp. 37–43, 2003.

[5] C. Kim, E. Damewood, and N. Hodge, "Professor Attitude: Its Effect on Teaching Evaluations," *J. Manag. Educ.*, vol. 24, no. 4, pp. 458–473, 2000.

[6] J. S. Pounder, "Is student evaluation of teaching worthwhile? An analytical framework for answering the question," *Qual. Assur. Educ.*, 2007.

[7] T. Hinkin, "The Effects of Time of Day on Student Teaching Evaluations: Perception versus Reality," *J. Mangement Educ.*, 1991.

[8] M. W. Ohland, S. D. Sheppard, G. Lichtenstein, O. Eris, D. Chachra, and R. A. Layton, “Persistence , Engagement , and Migration in Engineering Programs,” *J. Eng. Educ.*, 2008

[9] K. P. Brannan and P. C. Wankat, “Survey of First-Year Programs,” in *2005 ASEE Annual Conference and Exposition* (2005).

[10] J. Sangster, “Work in Progress: Privilege and Diversity as Determiners of Engineering Identity and Success,” in *2019 ASEE Annual Conference and Exposition*, paper ID# 26153 (2019).