Teaching Undergraduate Artificial Intelligence Classes: An Experiment with an Attendance Requirement

Sven Koenig* **Tansel Uras** **Liron Cohen**
Computer Science Department
University of Southern California
941 Bloom Walk
Los Angeles, CA 90089-0781, USA
{skoenig,turas,lironcoh}@usc.edu

Abstract

We report on an experiment that we performed when we taught the undergraduate artificial intelligence class at the University of Southern California. We taught it – under very similar conditions – once with and once without an attendance requirement. The attendance requirement substantially increased the attendance of the students. It did not substantially affect their performance but decreased their course ratings across all categories in the official course evaluation, whose results happened to be biased toward the opinions of the students attending the lectures. For example, the overall rating of the instructor was 0.89 lower (on a 1-5 scale) with the attendance requirement and the overall rating of the class was 0.85 lower. Thus, the attendance requirement, combined with the policy for administering the course evaluation, had a large impact on the course ratings, which is a problem if the course ratings influence decisions on promotions, tenure, and salary increments for the instructors but also demonstrates the potential for the manipulation of course ratings.

We performed an experiment with the undergraduate "Introduction to Artificial Intelligence (AI)" class (CS360) at the University of Southern California (USC) in Fall 2018 since class attendance had quickly decreased during the semester in the past (to an average of about 50-60 percent in the lectures, as our experimental results show) despite the instructor receiving good course ratings, which is consistent with reports for other classes (Romer 1993; Marburger 2001; Friedman, Rodriguez, and McComb 2001). CS360 is required for a B.S. degree in computer science and taken by some other undergraduate students as well, resulting in a high enrollment. The university therefore often offers more than one section of CS360 per semester. The first author was the instructor of both sections of CS360 in Fall 2018. The other two authors were the two teaching assistants. With permission of the computer science department, we evaluated the advantages and disadvantages of imposing an attendance requirement. The approval was necessary since the attendance requirement was imposed selectively on students, thus treating some students differently than others, which could potentially affect their grades. We were not sure about the impact of the attendance requirement. An attendance requirement limits the freedom of the students and might be inconvenient for them. Students that comprehend the material without attending the lectures or, perhaps more likely, exercises might not improve their performance with the attendance requirement and feel that they waste their time. Students that struggle with the material might or might not improve their performance with the attendance requirement and, depending on the results, might or might not feel that they waste their time.

In the following, first, we describe CS360 in more detail to provide the readers with the necessary context. Second, we describe how we imposed the attendance requirement to provide the readers with information on the methodology of our experiment. Third, we discuss some issues that arose during our experiment and how we resolved them, thus showcasing both some unforeseen problems caused by the attendance requirement and some non-ideal experimental conditions with this first experiment in a real-world education setting. Finally, we discuss our experimental results about the attendance of the students and how it affected their performance and course ratings. We conclude by discussing the similarities and differences of our findings and the findings of others as well as future work that could address some of the shortcomings of our experiment.

Context: The Class

CS360 is a standard undergraduate "Introduction to AI" class that uses the popular textbook "AI: A Modern Approach" (third edition) by Russell and Norvig (Russell and Norvig 2009). It is centered around knowledge in AI, namely how to represent it, how to reason with it, how to learn it, and how to use it to achieve given objectives. The first part of the class discusses these issues in the given order in a deterministic setting, and the second part discusses them in a probabilistic setting. Successes of AI in computer games and ethical issues are also discussed. Figure 1 shows the

*This study was supported by the National Science Foundation (NSF) under grant numbers 1724392, 1409987, 1817189 and 1837779. Swantje Willms compiled the statistics based on the questionnaires, and Tansel Uras compiled the statistics based on all other data. The USC Institutional Review Board determined that this study did not require its approval.
Copyright © 2020, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

Week 1	Mon, Aug 20	Intelligent Agents [1-2]
	Tue, Aug21	No Exercise
	Wed, Aug 22	Review: Propositional Logic [7.1-7.5]
Week 2	Mon, Aug 27	Reasoning with Propositional Logic [7.1-7.5]
	Tue, Aug 28	Exercise
	Wed, Aug 29	First-Order Logic [8 and 9.5]
Week 3	Mon, Sep 03	No Class (Labor Day)
	Tue, Sep 04	Exercise
	Wed, Sep 05	Rule-Based Systems [9.3-9.4]
	Fri, Sep 07	Add/Drop Deadline 1 (without any W and tuition expense)
Week 4	Mon, Sep 10	Ontologies and Semantic Networks [12.5.1]
	Tue, Sep 11	Exercise
	Wed, Sep 12	Decision Tree Learning [18.3]
	Thu, Sep 13	Project 1 Out (10% of the class score)
Week 5	Mon, Sep 17	Perceptron Learning [18.7.1-18.7.4]
	Tue, Sep 18	Exercise
	Wed, Sep 19	Neural Network Learning [18.7.1-18.7.4]
Week 6	Mon, Sep 24	Function Optimization with Local Search [4.1-4.2]
	Tue, Sep 25	Exercise
	Wed, Sep 26	Review: Probabilities [13]
	Thu, Sep 27	Project 1 Due
Week 7	Mon, Oct 01	Bayesian Networks [14.1-14.4]
	Tue, Oct 02	Exercise
	Wed, Oct 03	Bayesian Networks continued [14.1-14.4]
	Fri, Oct 05	Drop Deadline 2 (without a W on the transcript)
Week 8	Mon, Oct 08	Midterm Review
	Tue, Oct 09	Midterm (35% of the class score)
	Wed, Oct 10	Naive Bayesian Learning [13.5.2 and 20.2.2]
	Thu, Oct 11	Project 2 Out (10% of the class score)
Week 9	Mon, Oct 15	Planning Agents [2.4 and 3.1-3.2] and STRIPS [10.1]
	Tue, Oct 16	Exercise
	Wed, Oct 17	Motivating Guest Lecture: AI for Space Exploration
Week 10	Mon, Oct 22	SAT-Based Planning [10.4.1]
	Tue, Oct 23	Exercise
	Wed, Oct 24	Uninformed Search [10.2.1 and 3.3-3.4]
	Thu, Oct 25	Project 2 Due
Week 11	Mon, Oct 29	Constraint Satisfaction [6]
	Tue, Oct 30	Exercise
	Wed, Oct 31	Heuristic Search [3.5-3.6]
	Thu, Nov 01	Project 3 Out (10% of the class score)
Week 12	Mon, Nov 05	Heuristic Search continued [3.5-3.6]
	Tue, Nov 06	Exercise
	Wed, Nov 07	Search-Based Planning [10.2.3]
	Fri, Nov 09	Drop Deadline 3 (with a W for "withdrawn")
Week 13	Mon, Nov 12	Decision Theory [16.1-16.3 and 16.5]
	Tue, Nov 13	Exercise
	Wed, Nov 14	Markov Decision Processes [17.1-17.2]
	Thu, Nov 15	Project 3 Due
Week 14	Mon, Nov 19	Adversarial Search [5.1-5.3]
	Tue, Nov 20	Exercise
	Wed, Nov 21	No Class (Thanksgiving)
Week 15	Mon, Nov 26	AI Ethics
	Tue, Nov 27	Exercise
	Wed, Nov 28	Wrap-Up, Final Review, and Course Evaluation
Exam Days	Wed, Dec 05	Final (35% of the class score)

Figure 1: Schedule of CS360 in Fall 2018. Blue text highlights the dates when we took attendance. Red text highlights the projects and exams that determined the class score. Numbers in brackets are the assigned chapters or sections from the textbook.

schedule of CS360 in Fall 2018. The slides and assignments can be found at the following URL: http://idm-lab.org/intro-to-ai.html.

CS360 is meant to be taken by junior students, so that they have the option to take more advanced AI classes in their senior year. Some students consider CS360 to be easy but others struggle with it as it is one of their first classes where they need to utilize their knowledge of the prerequisites (such as algorithms and data structures and programming in C/C++) in the context of new material that focuses not only on gaining an understanding of a large number of AI algorithms but also on acquiring an abstract computational thinking ability in general.

The class personnel consisted of the instructor, two teaching assistants (who were both Ph.D. students of the instructor), and seven course producers (who were undergraduate students whose job was to help the teaching assistants with grading and the students with their programming problems during the projects).

CS360 consists of two 80-minute lectures per week given by the instructor and one 50-minute exercise per week given by one of the teaching assistants in presence of the instructor, where the teaching assistant goes over textbook-style problems (often from the assignments) and their solutions. There were two sections in Fall 2018. The lectures of one section were given on Mondays and Wednesdays from 3:30pm to 4:50pm, and the lectures for the other section were given on Mondays and Wednesdays from 5:30pm to 6:50pm (that is, immediately afterward) in the same classroom. To achieve the best possible experimental conditions, the lectures for both sections were taught completely identically (for example, using the same slides, which were posted before each lecture), and the exercises were given to both sections simultaneously on Tuesdays from 6:00pm to 6:50pm in a larger classroom.

All class personnel offered office hours to help the students. In addition, we used Campuswire (campuswire.com) as a common discussion forum for all students. The teaching assistants monitored it and answered questions in case none of the students answered them first.

There were three graded two-week projects that needed to be solved individually and contained both theoretical problems and hands-on problems, most of which involved C/C++ programming. The projects were intended to 1) give the students a deeper experience with some of those AI techniques taught in class that they can likely use after graduation; and 2) demonstrate the power of those AI techniques to them in an otherwise more abstractly taught class.

There were one midterm and one final, both 110 minutes long. To help the students prepare for these exams, we posted textbook style assignments on Wednesdays (covering the material taught in the current week). We encouraged the students to form study groups to come up with solutions but did not collect or grade their solutions. We posted fully worked-out solutions to the posted assignments four days later. We posted about 100 assignments and solutions over the course of the class. Both exams were open textbook ("AI - A Modern Approach" only) and open printed or hand-written (but not electronic) notes (including, but not

limited to, the assignments and their solutions), although the number of problems on the exams typically prevents students from studying the textbook or notes extensively during the exams. All exams were comprehensive but the final focused on material not yet tested in the midterm. To achieve the best possible experimental conditions, both sections wrote the same midterm simultaneously during one of the exercises. Similarly, both section wrote the same final simultaneously on the date given in the official final examinations schedule for the section without the attendance requirement. With permission of the department, we obeyed the final examinations schedule, where both sections were assigned different dates for the final, by allowing students during the first three weeks of the lecture period to opt for writing their final during those dates instead, which eventually affected only students with the attendance requirement. Nine students with the attendance requirement made use of that option and wrote a different final. Three additional students missed the joint final (for example, due to sickness) and joined these nine students in writing their final, and one student did not write the final and received an incomplete. Three of these four students were subject to the attendance requirement.

We re-used some parts of older projects and exams from previous years (although typically not from the year before) but also added a substantial amount of new content. We used Gradescope (gradescope.com) as the grading environment for all projects and exams. The solutions of the students were scanned, and a list of common mistakes was uploaded together with their penalties. Grading then proceeded in parallel by checking the mistakes made in each part of a solution, after which its score was computed automatically. All students could then view their solutions online together with the grading key and their scores and, if necessary, appeal their scores with the push of a button.

Figure 1 includes the project and exam weights. We did not provide opportunities for extra credit except that students could earn a bonus of up to 1.5% for frequently helping other students on the discussion forum. We did not grade on a curve. Instead, we transformed the weighted class scores (in form of percentages) into the class grades using the linear transformation (if the university would recognized an A+, which it does not) shown in Figure 2.

A	90%-100%	B-	70%-75%	D+	50%-55%
A-	85%-90%	C+	65%-70%	D	45%-50%
B+	80%-85%	C	60%-65%	D-	40%-45%
B	75%-80%	C-	55%-60%	F	0%-40%

Figure 2: Mapping class scores to class grades.

Methodology: The Attendance Requirement

Until Add/Drop Deadline 1 (three weeks into the lecture period), the university allows students to add classes, drop classes, and change sections of classes freely, which often results in substantial changes to the class roster. Afterward, students can still drop classes, but this typically happens for CS360 only in case of severe performance problems since the class is mandatory for almost all students. We thus told the students up front that there would be an attendance requirement for approximately half of the students. All students were subject to the attendance requirement (which covered both the lectures of the sections they were registered for and the exercises) but, for approximately half of the students, that attendance requirement would be lifted directly after Add/Drop Deadline 1. We would announce only at that time for which students the attendance requirement would be lifted. This announcement ensured that all students knew that they needed to be available for all lectures of the sections they were registered for and all exercises, which was important because the university imposes capacities for the sections of CS360 to distribute the students over the different sections (for example, due to limited capacity of the lecture room) but the students had come to expect that they could register for a section with available seats but then attend any section convenient for them (which in the past had resulted in popular sections being overcrowded). In Fall 2018, as in previous semesters, the section of CS360 with the earlier lectures was by far more popular than the one with the later lectures.

Students with the attendance requirement could miss up to five meetings (lectures or exercises) without an excuse. We required excuses for all longer absences of three consecutive meetings. We announced that we would subtract 5% percent from their class score for every absence without a valid excuse beyond the five-meeting absence allowance, which corresponded to one ”plus” or ”minus” in the class grade. However, when determining the class grades at the end of the class, we decided to subtract only 0.5% for every such absence. Four students received such penalties, three of them for missing one meeting over the absence allowance (which did not change their grades) and one of them for missing three meetings over the absence allowance (which decreased their grade by one ”plus” or ”minus”). We required excuses for absences from exams from all students.

The instructor and teaching assistants discussed several ways of determining which students should be subject to the attendance requirement. For example, we could have selected the students randomly from both sections. However, we eventually decided to impose the attendance requirement exactly on all students registered for the section with the earlier lectures, for the following reason: When two sections are taught by different instructors, they typically have different class policies since the university allows the instructors to implement their own class policies, so our experiment did not set a precedence. More importantly, class attendance had quickly decreased during the semester in the past. We are not sure about the reasons for this trend but there could be several of them. For example, we supply both detailed lecture slides and assignments with solutions for exam preparation, so that some students can skip the lectures, learn from the textbook and the supplied class material only, and still do well in the exams. Of course, other students might skip the lectures for other reasons, for example, because they get frustrated by the wealth of material and the speed with which it is presented. Thus, imposing the attendance requirement on a whole section can be expected to increase attendance

and change the learning environment, for example, because the students are now more anonymous and each receive less attention from the instructor, which might result in them paying less attention. This effect gets lost unless a whole section is subject to the attendance requirement, similar to how an attendance requirement would be imposed outside of an experiment. We imposed the attendance requirement on the students registered for the more popular section because it would have caused more inconvenience for the other students since many of those students had registered for their section only because the more popular section had reached capacity. We also discussed whether the attendance requirement should be imposed for only the lectures, only the exercises, or both. We eventually decided to impose it for both the lectures and exercises because we might get stronger and thus easier-to-measure effects of the attendance requirement this way.

Experimental Conditions: Issues

We encountered the following issues during our experiment:

First, several students – both those with and without the attendance requirement – pointed out during the semester that they could not attend the lecture of their section on a particular day, often due to special events, and asked whether they could attend the lecture of the other section on the same day. Since it would have been an unnecessary hardship for the students not to be allowed to do that, we granted these requests and the word spread (for example, because the instructor now asked after each headcount during a lecture how many of the student were from the other section), in effect changing the policy to allowing all students to attend either one of the lectures on a day. Many more students without the attendance requirement made use of this informally relaxed requirement to attend the other (more popular) section than vice versa, which increased attendance of the earlier lectures even more, perhaps resulting in stronger and thus easier-to-measure effects of the attendance requirement.

Second, five students pointed out during the semester that they could not attend either lecture on several days due to onsite interviews for jobs or internships. We eventually decided not to count these days toward their absence allowance but we still count these and other excused absences as absences in our experimental results.

Third, the room used for the exercises was very large and equipped with bad sound and projection equipment. Many of the students who wanted to participate in the exercises sat in front, while the students who were not interested in participating typically sat in the back and either worked on material not related to the CS360 or entertained themselves by surfing the web, watching videos, or playing games. The attendance requirement for the exercises was likely counterproductive for these students.

Fourth, we used the Arkaive app and website (arkaive.com) to track the attendance of the students. Initially, the computer science department had provided us with one scanner that, when connected to a laptop, could be used to scan the student IDs of the attending students but this process was slow, error-prone, and thus unusable for our purposes. On suggestion of a student, we then switched to Arkaive since it was used by other classes of the university as well. For most of the lectures, the students needed to check in on Arkaive at any time from about 20 minutes after the beginning of the lectures to about 20 minutes before their end. The instructor reminded the students to check in once this time period had started and, at the same time, performed a headcount and asked how many of the students were from the other section. This was done for both sections. Toward the end of the class, when we noticed a larger discrepancy of the Arkaive data and the headcount data, we used a very short, randomly picked time period during which the students had to check in. For most of the exercises, the students could check in on Arkaive at any time during the exercises. We could not enforce, neither for the lectures nor the exercises, that the students without the attendance requirement indeed checked in since not doing so did not result in any disadvantages for them.

Fifth, we used the process mandated by the university for letting the students fill out the official course evaluation. We decided to let the students fill them out online during the last lecture. While the students did not need to attend the lecture to be able to fill out and submit the course evaluation, the time window heavily biased the course evaluation toward the opinions of those students who attended the last lecture. The resulting return rate was small since the students without the attendance requirement did not need to attend the lecture and all but 14 students with the attendance requirement still had not used up their absence allowances and thus did not need to attend the lecture either. The students often make use of that freedom to skip class during the busy time at the end of the lecture period, especially in this case where some students might have considered this lecture not to be relevant for the final and it was given by one of the teaching assistants since the instructor was on unexpected jury duty.

The above issues demonstrate the noise inherent in a first experiment in a real-world education setting, where the experimental conditions are not ideal and might need to be adapted slightly during the experiment to account for unforeseen situations and ensure that some students do not have unfair disadvantages compared to other students. Additional noise is due a small number of students who were granted special exam conditions (for example, due to disabilities) or wrote a different ”makeup” final than the other students (for example, due to sickness), and similar factors.

Experimental Results

At Add/Drop Deadline 1 (three weeks into the lecture period), the section with the attendance requirement had 76 registered students, and the section without the attendance requirement had 63 students. At the end of the semester, this had changed to 74 and 62 students, respectively. We have excluded the data of the three students who dropped the class from the experimental results. In the figures, we refer to the section with the attendance requirement as Section 1 and the section without the attendance requirement as Section 2.

Figure 3 shows the results of an anonymous questionnaire that we asked the students to fill out during the lectures on September 10 (directly before we informed them for

whom the attendance requirement would be lifted) to provide us with their opinions on the attendance requirement. We allowed for free-form answers that we then classified by hand into categories. The students also suggested a variety of softer versions of our attendance requirement, such as making attendance a percentage of the class score, giving extra credit for attendance, having a larger allowance for missed meetings, being more lenient with excuses for missed meetings, writing in-class quizzes, enforcing attendance only for students with poor performance, or letting each student decide themselves whether they would like to be subject to some form of an attendance requirement.

We asked the students in an anonymous questionnaire during the next-to-last lectures on November 26 again for free-form "thoughts about attendance (for example, whether it is helpful, whether it should be enforced in some form, and so on)." 6.9% (44.8%) of 29 students with the attendance requirement and 12.5% (37.5%) of 16 students without the attendance requirement explicitly stated that the attendance requirement is (is not) helpful, while additional students voiced a variety of concerns about the attendance requirement, including that

- students should be responsible for their own learning;
- students might have other obligations during class time;
- students do not pay attention if forced to attend; and
- headcounts disrupt lectures.

Student Attendance

We measured attendance using both Arkaive (where the students needed to check in during the lectures and exercises) and headcounts (where the instructor counted the number of students during the lectures). The Arkaive data was collected for (almost) all lectures and exercises but the headcount data was collected less regularly. The Arkaive data might have errors due to students not checking in during class or by checking in without being in class (which is possible, for example, since the location-based check-in uses a large location radius). The headcount data might have errors due to miscounting and students not correctly reporting the sections they were registered for.

Figure 4a shows the percentage of students, as a function of the section they were registered for, that attended lectures on particular dates, according to both the Arkaive and headcount data. The lecture on October 17 is excluded from the data since the lecture was replaced with a motivational guest lecture (by Steve Chien from NASA JPL on the growing role of AI in space exploration) at a different time on that date. The headcount for the section with the attendance requirement on September 17 is not shown in the figure since it is greater than the number of registered students. Figure 4b shows the Arkaive data for the exercises. The exercise on October 9 is excluded from the data since the exercise was replaced with the midterm on that date. Figure 4c shows the average attendance during the semester.

The attendance requirement resulted in a substantial increase in the percentage of students who attended the lectures and exercises, as expected. For example, it increased from 56.3% on average for the lectures to 77.03% according to the headcount data. The Arkaive and headcount data differ, in a very systematic way: For students with the attendance requirement, the attendance in the Arkaive data is greater than the attendance in the headcount data. We suspect that this difference is due to some students exploiting a loophole in the Arkaive app, where students can check in from the vicinity of the classroom rather than the classroom itself. For students without the attendance requirement, the attendance in the Arkaive data is smaller than the attendance in the headcount data. We suspect that this difference is due to some students not bothering to check in with the Arkaive app since doing so did not result in any disadvantages. Overall, we expect the headcount data to be much more accurate than the Arkaive data.

Student Performance

Figure 5a shows the average project and exam scores of the two sections (and the standard deviations). The scores of the two sections are similar but the students without the attendance requirement had slightly higher exam scores on both the midterm and final and performed slightly better overall than those with the attendance requirement. The difference in the total scores (calculated as a weighted average of the project and exam scores without the penalties for missed class attendances of students with the attendance requirement) is only about 1.64%, which is roughly one third of the 5.00% that corresponds to one "plus" or "minus" in the class grade. The difference in student scores across the two sections is not statistically significant, with z-values of 0.99, 1.20, and 0.78 for the midterm, final, and total scores, respectively.

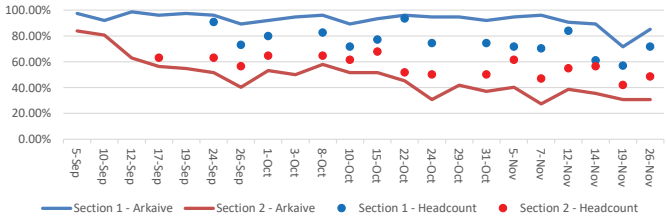
We asked the students in an anonymous questionnaire during the next-to-last lectures on November 26 (before the final was written) to rate the difficulty of CS360 on a scale from 0 (= easy) to 10 (= difficult). The averages over 25 and 17 responses by students with and without, respectively, the attendance requirement were both 7.3. Thus, the attendance requirement did not affect the perceived difficulty of CS360 (despite the selection bias described in the upcoming footnote), which is consistent with the attendance requirement not substantially affecting the total scores either. Figure 6 shows the results of an exit poll at the finals (including the makeup final) which did not differentiate between students with and without the attendance requirement and asked them to provide a self-assessment of the difficulty and length of the final (via sheets taped at the doors of the exam rooms).

Figure 5b shows the (Pearson) correlation of attendance in the lectures, the exercises, and both with the total scores. The correlation is positive but does not sheds light on causation, namely whether students who attend more classes perform better or whether students who perform better attend more classes. The correlations of the section without the attendance requirement are larger than the correlations of the section with the attendance requirement since all students in the latter section attended most lectures and exercises, so the small differences in their attendance do not explain their performance differences well.

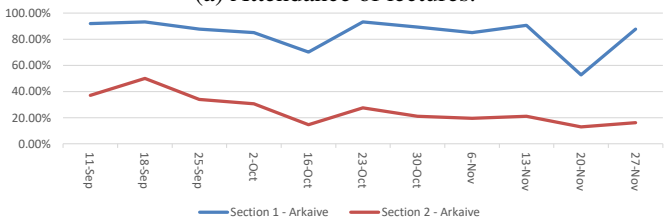
Do you like the idea of being subject to an attendance requirement?				Responses
yes, lectures and exercises	yes, lectures only	no	indifferent	
11.6%	3.3%	76.9%	8.3%	121

Will our attendance requirement be beneficial for students?			Responses
yes	no	maybe or I don't know	
29.4%	51.3%	19.3%	119

Figure 3: Results of an early questionnaire.



(a) Attendance of lectures.

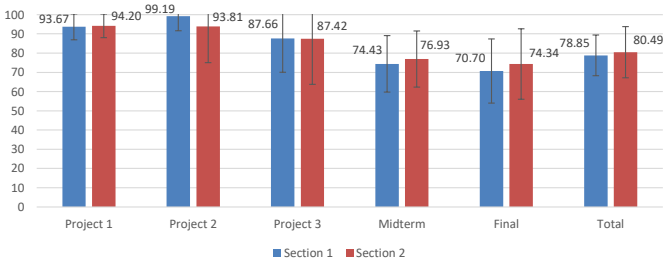


(b) Attendance of exercises.

	Lectures			Exercises
	Arkaive	Headcounts	Difference	Arkaive
Section 1	92.51%	77.03%	15.48%	84.28%
Section 2	47.87%	56.35%	-8.48%	25.81%

(c) Average attendance during the semester.

Figure 4: Student attendance.



(a) Project and exam scores.

	Attendance in		
	lectures	exercises	both
Section 1	0.2685	0.2302	0.3009
Section 2	0.3694	0.3864	0.4095

(b) Correlation of attendance with total score.

Figure 5: Student performance.

The final was...					Responses
too easy	easy	ok	hard	too hard	
0.0%	7.4%	44.4%	40.7%	7.4%	81
too short	short	ok	long	too long	
0.0%	5.0%	27.5%	45.0%	22.5%	80

Figure 6: Results of the exit poll at the final.

Student Satisfaction

The official course evaluation asks the students to provide course ratings on a 1-4 scale (where 1 = strongly disagree, 2 = disagree, 3 = agree, and 4 = strongly agree) for 17 statements, which are grouped into five categories, namely course design (3 statements), instructional practices (3), inclusion practices (4), assessment practices (4), and course impact (3). Two supplemental questions ask the students to rate the instructor and the course on a 1-5 scale (where 1 = poor, 2 = below average, 3 = average, 4 = above average, and 5 = excellent). The students can also provide free-form comments for the most valuable aspect of the course, for the least valuable aspect of the course, for a change that would improve a lacking aspect of the course, and for additional feedback in addition to providing data for their engagement in three different aspects of the course.

37 out of 74 students (that is, 50.0% of the students) with the attendance requirement and 10 out of 62 students (that is, 16.1% of the students) without the attendance requirement filled out the course evaluation since more students with than without the attendance requirement attended the last lecture (during which time was set aside for the students to fill out the course evaluation). Figure 8 shows the average course ratings for the five categories and the two supplemental questions as reported by the university (and the standard deviations). The ratings differ in a very systematic way: All ratings are lower for the section with the attendance requirement, which is also true for 16 of the 17 statements that are grouped into the categories (the exception being C3 in Figure 7). The mean difference over all 17 statements is 0.34 on the 1-4 scale, with the largest difference for any statement being 0.82 (for statement C2) and the largest difference for any category being 0.43 (for the category on assessment practices). The overall rating of the instructor was 0.89 lower and the overall rating of the course was 0.85 lower in the 1-5 scale. These two ratings have traditionally been taken into account for decisions on promotions, tenure, and salary increments. The average number of free-form comments per student (added up over the four opportunities) is 2.4 for students with the attendance requirement and only 1.9 for students without it.

The way how the course evaluation was administered could have resulted in a selection bias, namely in case a larger percentage of students happy with the instructor and

A: Course Design

1. The course objectives were well explained.
2. The course assignments were related to the course objectives.
3. I understand what was expected of me in this course.

B: Instructional Practices

1. The instructor carefully explained difficult concepts, methods, and subject matter.
2. The instructor encouraged questioning and discussion of course topics from the students.
3. The instructor encouraged me to do my best work.

C: Inclusion Practices

1. The course materials included diverse perspectives OR applications to diverse populations.
2. The instructor used a variety of teaching approaches to meet the needs of all students.
3. The instructor was receptive to the expression of diverse student viewpoints.
4. The instructor demonstrated sensitivity to students' needs and diverse life experiences.

D: Assessment Practices

1. The assessments/assignments reflected what was covered in the course.
2. The grades I have received thus far reflect the QUALITY of my performance in the course.
3. The criteria for good performance on the assignments or assessments were clearly communicated.
4. The instructor's evaluation of my performances was constructive.

E: Course Impact

1. I learned a lot from this course. I learned perspectives, principles, or practices from this course that I expect to apply to new situations.
2. This course challenged me to think critically and communicate clearly about the subject.
3. This course provided me with information that may be directly applicable to my career or academic goals.

Supplemental Questions: Overall Assessment

1. Overall, how would you rate the instructor?
2. Overall, how would you rate this course?

Figure 7: Course evaluation statements.

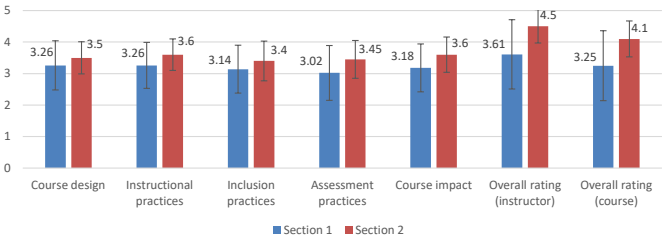


Figure 8: Student satisfaction. The five categories (on the left) use a 1-4 scale, while the two supplemental questions (on the right) use a 1-5 scale. Larger scores are indicative of higher student satisfaction.

class attended the last lectures on November 28 (during which time was set aside for the students to fill out the course evaluation) without the attendance requirement than with the attendance requirement, which is a plausible assumption.[1] It is also important to take into account that the number of students who filled out the course evaluation in the section without the attendance requirement is small, which limits the significance of the experimental results. It was therefore surprising to us that the standard deviation of the course ratings was higher for the section with the attendance requirement for 16 of the 17 statements, all five categories, and both supplemental questions.

[1] Such a selection bias also seems to be supported by the results of an anonymous questionnaire that we asked the students to fill out during the next-to-last lectures on November 26 to self-report rough estimates of the percentages of the lectures and exercises of CS360 that they had attended during the semester. The average over 29 responses by students with the attendance requirement was 94.3% of the lectures and 88.8% of the exercises, while the average over 16 responses by students without the attendance requirement was 90.2% of the lectures and 74.5% of the exercises.

Discussion

A recent review of prior research on the effect of attendance policies agrees with some of our findings: There is typically a positive correlation between attendance and performance when using observational methods. On the other hand, imposing an attendance requirement typically increases attendance but does not always result in increased performance (Golding 2011). Different experiments have resulted in different outcomes (St. Clair 1999; Golding 2011), which resulted in the insight that the "empirical research evidence on the relationship between attendance rates or attendance policy and academic achievement is inconclusive" (St. Clair 1999) and the later insight that "[m]andatory attendance policies appear to have a small positive impact on average grades" (Credé, Roch, and Kieszczynka 2010). In an experiment by Golding in 2009, imposing an attendance policy lowered the exam scores (Golding 2011), similar to what we have observed. However, the review differs from our findings in one crucial aspect: In the experiment by Golding, imposing an attendance policy did not have a negative impact on the rating of the instructor (Golding 2011), different from what we have observed. The availability of course videos and other study information on the internet has made many studies outdated that were performed before the internet was a main source of information for students.

Future Work

It is future work to repeat the experiment under cleaner experimental conditions, with a course evaluation whose results are not biased toward the opinions of the students attending the lectures, and with an attendance requirement that only applies to the lectures but not the exercises, so that students who understand the lecture material well are not forced to attend the exercises (especially since we provide assignments with solutions that the students can use to practice outside of the classroom and check their solutions).

References

Credé, M.; Roch, S.; and Kieszczynka, U. 2010. Class attendance in college: A meta-analytic review of the relationship of class attendance with grades and student characteristics. *Review of Educational Research* 80(2):272–295.

Friedman, P.; Rodriguez, F.; and McComb, J. 2001. Why students do and do not attend class. *College Teaching* 49:124–133.

Golding, J. 2011. The role of attendance in lecture classes: You can lead a horse to water... *Teaching of Psychology* 38(1):40–42.

Marburger, D. 2001. Absenteeism and undergraduate exam performance. *Journal of Economic Education* 32:99–110.

Romer, D. 1993. Do students go to class? Should they? *Journal of Economic Perspectives* 7:167–174.

Russell, S., and Norvig, P. 2009. *Artificial Intelligence: A Modern Approach*. Pearson, third edition.

St. Clair, K. 1999. A case against compulsory class attendance policies in higher education. *Innovative Higher Education* 23(3):171–180.