Leveraging Space:
Testing the Flipped Classroom

1. Objectives

Classroom space is at a premium on many college campuses. Large class sizes are common at many universities and tight budgets, which often block new construction, make efficient use of space a high priority.

Digital technology provides a variety of possible solutions. Fully online courses have the potential to yield learning outcomes comparable to those students achieve in face-to-face courses (Means at al., 2010), but there is a large amount of variability in the outcomes reported by different studies (Bernard et al., 2004; Zhao et al., 2005) and the dropout rates in fully online courses are notoriously high (Xu & Jaggars, 2011).

Students in face-to-face courses supplemented by online technology tend to outperform their counterparts in unsupplemented courses (Means et al., 2010), but this blended approach does not address the issue of high demand for physical classroom space.

This investigation tests a new, technology-based solution to this conundrum, namely an instructional model with three main components:

1. A "flipped" classroom in which the more didactic components of the class are moved online, using recorded video presentations and Web-based activities, leaving face-to-face time free for collaborative problem-solving and instructor tutorials;
2. Holding face-to-face class meetings in newly configured, technologically-equipped active learning spaces that enhance collaborative problem-solving and instructor tutorials; and
3. A reduction in face-to-face class meeting time by 67% (from 3 nominal hours per week to 1 hour).

We find that such an approach takes advantage of the learning gains afforded by blended approaches (Brown et al., 2007; Means et al., 2010; Vignare, 2007) and active learning classrooms (Brooks, 2011; Walker, Brooks, Baepler, 2011) while accommodating the institutional demand for large enrollment introductory courses.

2. Theoretical framework

In general, reducing instructor-student contact should hinder student learning, because quantity of substantive interaction with faculty members is positively associated with the learning outcomes students achieve (Pascarella & Terenzini, 2005). So while reducing face-to-face meeting time may address the issue of limited classroom space, one would expect this change to lead to diminished student learning outcomes, compared to a control condition.

However, the recent wave of interest in active learning classroom design points to a possible solution. Beginning with MIT and North Carolina State University, in recent years numerous colleges and universities have begun to reconfigure existing classrooms or to construct new classroom spaces (Oblinger, 2006). Although the details of design vary, in general the new spaces involve a move away from stadium-style seating, placing the instructor in the middle of the student seating area and incorporating display technologies designed to facilitate collaborative teaching-learning activities.

Recent research into the effectiveness of these active learning spaces shows that with proper use, they can not only improve student affect and engagement, but also help students to outperform expectations in learning outcomes when compared with traditional classrooms (Beichner et al., 2007; Brooks, 2011; Dori & Belcher, 2005; Walker, Brooks & Baepler, 2011). The spaces examined in this study seat students at round, nine-person tables with microphones, movable chairs, and switchable laptop connections which permit students to project what their laptops are showing onto flat-panel displays associated with each table. From an instructor station, faculty members can control what is shown on large projection screens as well as the smaller flat panel displays.

This study was designed to test whether the use of active learning spaces in conjunction with a flipped classroom model could overcome the reduction in student learning outcomes theoretically expected from reduced face-to-face instructional time.

3. Methods

Researchers partnered with a professor of chemistry to implement a post-test-only quasi-experimental design that compared student outcomes from one section of introductory chemistry taught in spring 2012 (the control) with the outcomes from a second section taught in fall 2012 (the treatment section). The study was then replicated, comparing the outcomes from another section of introductory chemistry taught in spring 2013 to the control outcomes.

In this study, the same professor taught all three of the classes using the same materials, readings, assignments, and schedule of assessments. The factors that were allowed to vary were the classroom in which the classes met; the number of contact hours between instructor and students; and the teaching-learning activities that took place during face-to-face meetings (Table 1). In the control class, students met for 3 nominal hours in a traditional lecture hall with a capacity of approximately 400 students; the teaching in that class centered on lecture and demonstration. In the treatment classes, students met for only 1 nominal hour per week in an active learning space with a capacity of 128 students; teaching in these classes centered on team-based learning activities, with the professor circulating among the tables providing guidance.

Table 1: Research design

	Spring 2012 (control)	**Fall 2012 (treatment #1)**	**Spring 2013 (treatment #2)**
Classroom	Traditional lecture hall	Active learning space	Active learning space

Contact hours	3 hours per week	1 hour per week	1 hour per week
Learning activities	Lecture-based, with occasional group work	Collaborative, student-centered activities	Collaborative, student-centered activities
N	383	345	372

The total number of student participants in this study was 1100 across all three sections. Most of the participants were in the early stages of their academic careers (freshmen or sophomores), and all were undergraduate students. In terms of ethnicity, the sample was heavily white (73.6%), which is consistent with the student population at the university. The sample included students from a wide variety of academic majors, and because this class satisfies an important academic requirement, it is likely that the students in the sample are representative of the broader population at the university.

It was not possible to randomly assign students to treatment and control sections in this study; instead, students enrolled as they normally would. Both of the treatment sections were described as follows in the official course guide:

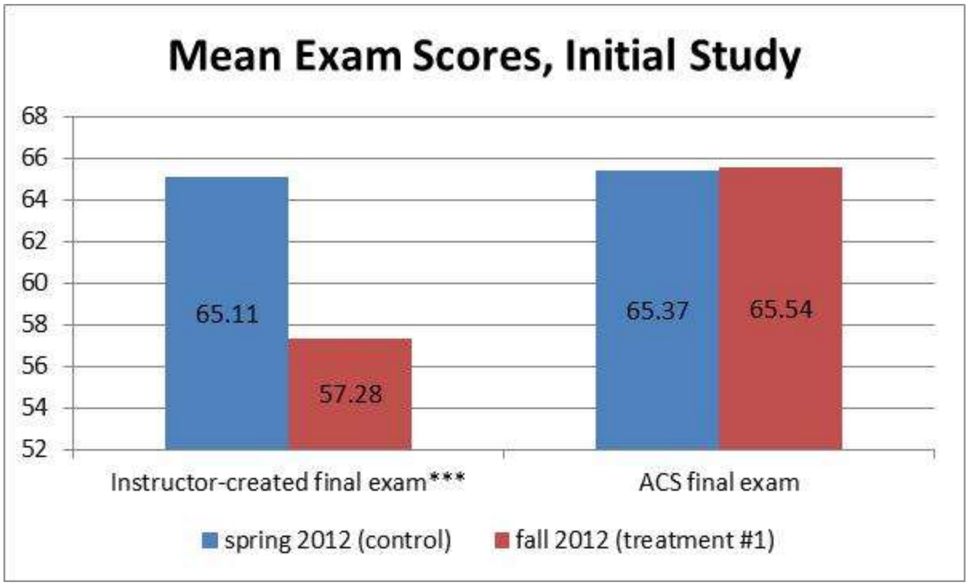
> "This is a HYBRID section of CHEM 1061. Traditional lecture materials will be accessed by students outside of class time, as videos posted online, in the class Moodle site. The online videos can be accessed at any time. The weekly 50-minute class period will be a required problem-solving session with the instructor."

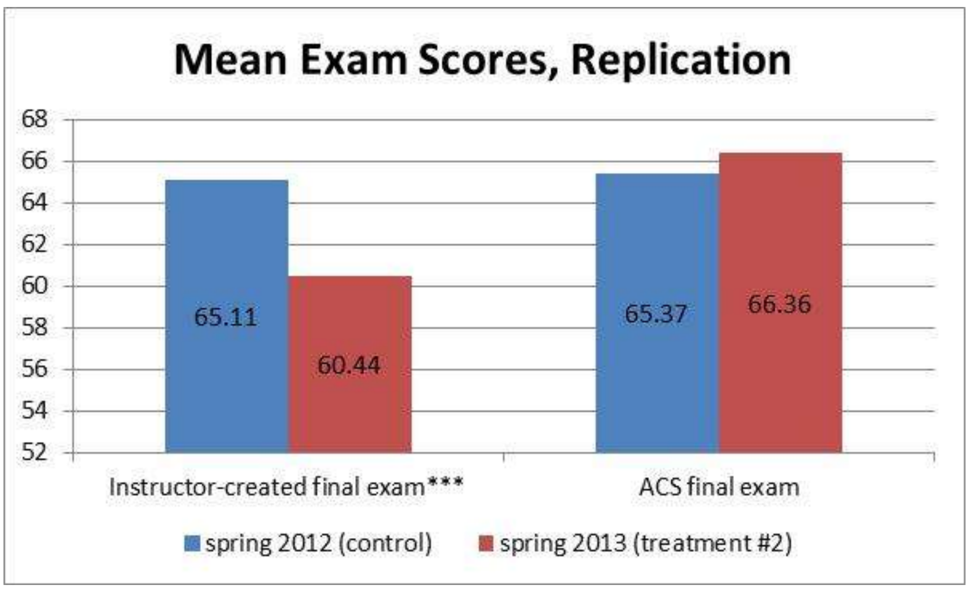
It is possible that students self-selected into the treatment sections of this class in a way that could affect the outcomes of this study. Therefore special care was taken to determine, using all available data, that the students in each of the treatment sections were comparable to the students in the control section. Using measures of central tendency, it was found that the treatment samples were statistically identical to the control sample on all exogenous variables, including age, year in school, ethnicity, major, and sex (results not shown).

4. Data

The main outcome variables of interest in this study were student performance on a final exam created each semester by the professor, and on a standardized introductory chemistry exam created by the American Chemical Society (ACS). (Data bearing on a secondary set of outcomes having to do with important aspects of the student classroom learning experience are currently being analyzed.)

Initial independent-samples t-tests showed a clear pattern: in both the initial study and the replication, students in the control section significantly outperformed students in the treatment section on the instructor-created final exam, but achieved statistically identical scores on the ACS final exam (Charts 1 & 2; ***$p < .001$). Further analyses were conducted in order to explore this pattern more fully.





4.1. Initial study

Two multivariate ordinary least squares (OLS) regression models were constructed for the initial study (models 1.1 and 1.2, spring 2012 compared to fall 2012), and these were used to predict students' performance on the two final exams. The main predictor variable of interest in each model was a dichotomous measure representing the semester in which students had taken the course (Spring 2012 = 0; Fall 2012 = 1).

The regression coefficients for these models indicate that students in the control section of the class outperformed students in the treatment section by about 6.5 percentage points on the instructor-created final exam, while the reverse pattern held for the standardized ACS final exam, with treatment students outperforming the control students by about 2 percentage points.

Pre-model tests revealed a significant association between certain exogenous demographic variables and the outcome variables. In particular, being male, being white, and being a science major were all positively associated with scores on the instructor-created exam and on the ACS exam, and these were included in the models. Age was negatively associated with exam scores and was included in the models as well. Finally, students' composite ACT scores and cumulative GPAs were included in the models due to their likely association with exam performance.

Table 2: OLS Regression of Exam Performance on Semester

	Model 1.1: Instructor-created final exam	Model 1.2: ACS standardized final exam
Semester (spring 2012 vs fall 2012)	-6.663*** (1.274)	2.029* (0.954)
ACT	0.540* (0.223)	1.518*** (0.167)
GPA	17.605*** (0.997)	7.402*** (0.842)
Age	0.367 (0.483)	1.187** (0.366)
Sex	3.311* (1.306)	7.090*** (0.983)
Science major	2.299 (1.633)	1.619 (1.221)
White	-1.842 (1.513)	1.882 (1.142)
Constant	-5.250 (12.203)	-27.874 (9.290)
N	512	499
Adjusted R^2	.449	.404
F Test	60.532***	49.318***

NOTE: Cell entries are unstandardized OLS coefficients with standard errors in parentheses.
*$p < .05$, **$p < .01$, ***$p < .001$

Each of the first two models was highly significant in predicting student performance on each of the two final exams in the course as described by the F test (Table 2). Moreover, the models account for a moderate amount of the variation in the dependent variables with adjusted R^2 coefficients between .40 and .50.

4.2. Replication

Similar OLS models were created for the replication of the initial study (models 2.1 and 2.2, spring 2012 compared to spring 2013).

Table 3: OLS Regression of Exam Performance on Semester		
	Model 2.1: Instructor-created final exam	Model 2.2: ACS standardized final exam
Semester (spring 2012 vs spring 2013)	-4.876*** (0.527)	-0.762 (0.451)
ACT	0.715*** (0.193)	1.581*** (0.166)
GPA	17.096*** (0.986)	8.721*** (0.872)
Age	0.154 (0.334)	1.118*** (0.285)
Sex	1.919 (1.127)	6.656*** (0.964)
Science major	1.173 (1.408)	1.886 (1.203)
White	-0.978 (1.320)	2.645* (1.131)
Constant	-8.239 (9.074)	-30.326 (7.787)
N	562	560
Adjusted R^2	.404	.416
F Test	55.356***	57.925***
NOTE: Cell entries are unstandardized OLS coefficients with standard errors in parentheses. *p < .05, **p < .01, ***p < .001		

Each of the two replication models was highly significant in predicting student performance on each of the two final exams in the course as described by the F test (Table 3). The models account for a moderate amount of the variation in the dependent variables with adjusted R^2 coefficients around .40.

The regression coefficients for these models indicate that, as in the initial study, students in the control section of the class outperformed students in the treatment section on the instructor-created final exam, in this case by nearly 5 percentage points. For the standardized ACS final exam, treatment students and control students achieved statistically identical results.

5. Results and significance

At first glance, the findings from these two studies send a mixed message regarding whether the classroom can be successfully "flipped" by moving the didactic and information-acquisition components of a class online and reserving the reduced face-to-face time for interaction and problem-solving activities. If we focus on the instructor-created final exam, it seems that student performance is impaired by the reduced instructor contact hours. But if we focus on the standardized ACS final exam, on which treatment students performed at least as well as control students, it seems that contact hours can successfully be reduced while leveraging the pedagogical strength of active learning spaces.

Analysis of model results, exam content, and qualitative data is ongoing, but an initial explanation of this mixed message may be provided by data from a series of interviews with the instructor. Those interviews indicate that the content of the instructor-created final exam changed from semester to semester, with the instructor selecting questions based on her face-to-face interactions with students. She attempted, in other words, to choose final exam questions that corresponded to material she had explicitly discussed with students during class. We hypothesize that this process resulted in final exams that were more proximal to the exact content of the traditional, control class taught in spring 2012 than to either of the treatment classes, and that this explains the control students' strong performance on this exam.

By contrast, the ACS final exam was consistent, covering standardized content that all students should have mastered when they complete an introductory chemistry course. The questions on this exam did not vary from semester to semester, and for that reason it seems to be the better of the two outcome measures used in this study.

We conclude that it is possible to cut face-to-face student-faculty contact hours by two-thirds without loss of student learning, if the remaining face-to-face class time is devoted to team-based learning activities conducted in an active learning space. The strong pedagogical effects of newly designed learning spaces in combination with more student-centered teaching techniques appear sufficient to overcome the expected reduction in student performance which is to be expected when contact hours are decreased. This points the way toward a more efficient use of scarce classroom space, when approximately 360 students can succeed in an introductory chemistry course that is taught in a space that accommodates about 1/3 of that number at one time.

J.D. Walker, Ph.D.
jdwalker@umn.edu

Paul Baepler, Ph.D.
baepl001@umn.edu

University of Minnesota
Information Technology
212 Walter Library
117 Pleasant Street, S.E.
Minneapolis, MN 55455
http://www.oit.umn.edu/research-evaluation/

References

Beichner, R., Saul, J., Abbott, D., Morse, J., Deardorff, D., Allain, R., Bonham, S., Dancy, M., & Risley, J. (2007). Student-Centered Activities for Large Enrollment Undergraduate Programs (SCALE-UP) project. In E. Redish & P. Cooney (Eds), *Research-based reform of university physics*. College Park, MD: American Association of Physics Teachers, 1-42.

Bernard, R., Abrami, P., Lou, Y., Borokhovski, E., Wade, A., Wozney, L., Wallet, P.A., Fiset, M., & Huang, B. (2004). How does distance education compare with classroom instruction? A meta-analysis of the empirical literature. *Review of Educational Research* 74, 379-439.

Brooks, D.C. (2011). Space matters: The impact of formal learning environments on student learning. *British Journal of Educational Technology* 42, 719–726.

Brown, G., Smith, T., and Henderson, T. (2007). Student perceptions of assessment efficacy in online and blended classes. In Picciano, A. and Dzuiban, C., Eds. *Blended learning: Research perspectives*. Needham, Massachusetts: SLOAN-C Series, 145-160.

Dori, Y. & Belcher, J. (2005). How does technology-enabled active learning affect undergraduate students' understanding of electromagnetism concepts? *Journal of the Learning Sciences* 14, 243–279.

Means, B., Toyama, Y., Murphy, R., Bakia, M., & Jones, K. (2010). Evaluation of evidence-based practices in online learning: A meta-analysis and review of online learning studies. Washington, DC: Center for Technology in Learning, U.S. Department of Education.

Oblinger, D., Ed. (2006). *Learning spaces*. EDUCAUSE: Boulder, Co.

Pascarella, E., & Terenzini, P. (2005). *How college affects students: Volume 2, A third decade of research*. San Francisco: Jossey-Bass.

Vignare, K. (2007). Review of literature: Blended learning: Using ALN to change the classroom - will it work? In Picciano, A. and Dzuiban, C., Eds. *Blended learning: Research perspectives*. Needham, Massachusetts: SLOAN-C Series, 37-63.

Walker, J.D., Brooks, D.C., & Baepler, P. (2011). Pedagogy and space: Empirical research on new learning environments. *EDUCAUSE Quarterly* 34, no pagination.

Xu, Di, & Jaggars, S. (2011). Online and hybrid course enrollment and performance in Washington State community and technical colleges. Community College Research Center Working Paper No. 31. Online at http://ccrc.tc.columbia.edu/media/k2/attachments/online-hybrid-performance-washington.pdf.

Zhao, Y., Lei, J., Yan, B., Lai, C., & Tan, H.S. (2005). What makes the difference? A practical analysis of research on the effectiveness of distance education. *The Teachers College Record* 107, 1836-1884.