

## **2006-2076: ENHANCING LEARNING OF LOW ABILITY STUDENTS IN MULTI-SECTION FRESHMAN LECTURE/LABORATORY CLASSES**

### **Jon Sticklen, Michigan State University**

Jon Sticklen is an Associate Professor in the Department of Computer Science and Engineering at Michigan State University. He has a strong research record in knowledge-based systems. His main contributions have been in the theory and application of task specific approaches and in model-based reasoning. Dr. Sticklen has led the effort to rejuvenate the MSU College of Engineering freshman gateway course in computational tools.

### **Mark Urban-Lurain, Michigan State University**

Mark Urban-Lurain is Director of Instructional Technology Research and Development in the Division of Science and Mathematics Education at Michigan State University. He is responsible for providing vision, direction, planning and implementation for using technology mathematics and science education and developed several introductory computer science courses for non-computer science students serving 2000 students per semester.

## **Enhancing Learning of Low Performance Students in Multi-section Freshman Lecture/Laboratory Classes**

### **Abstract**

Because of a scheduling “glitch,” in fall semester, 2004, our large enrollment introductory computer science was offered in two lecture sections at opposite ends of the week. One lecture section met on Monday nights at 7:00 p.m., with associated labs meeting later in the week. A second lecture section met on Friday mornings at 10:20 a.m., with associated labs meeting before lecture during that week. Rather than moving the lab sections for the Friday lectures to the following week, we decided to compare the results of using the Monday lecture for preparation for the upcoming labs and the Friday lecture as a “wrap up” of that week’s labs. Analysis of the 2004 data shows that there was no statistically significant difference in student outcomes between the two lecture days. In 2005, the same scheduling occurred. However, this year we added the use of a Personal Response System (PRS or “clickers”) to encourage student preparation and participation. For middle and high performing students, there are no differences in outcomes. However, low performing students in the Friday section performed significantly better than the low ability students in the Monday section in total points, midterm exams, and individual quizzes. Comparing women to men showed no differences in the Monday section, but women performed significantly better than men in the Friday section. The results of this study have implications for improving outcomes and retention for at-risk populations in engineering.

### **Introduction**

This is the follow-on paper to research reported at ASEE 2005<sup>1</sup>. In the earlier work, we began an examination of the effect of the timing of lecture in relation to laboratory, for high enrollment freshman engineering courses. The research question we addressed was “Does lecture-before-lab versus lecture-after-lab affect objective student performance?” In the work reported here, we extended our earlier work. The major finding in this study is that lecture-after-lab, when compared to lecture-before-lab, has a statistically significant effect in raising the performance of the *lowest performance students*.

A common instructional model for freshman engineering is the lecture/laboratory model. In this model, students usually spend two to four hours per week in a large lecture section typically of one hundred or more students, and three to six hours per week in small laboratory (or recitation) sections typically of twenty or fewer students.

Although not universal, the most common implementation of this instructional model is that lecture introduces material of a given “unit” before laboratory (or recitation) sections give students an opportunity to provide hands on, detailed experience with applying knowledge introduced in assigned readings and lecture. The assumption is that students need a framework for understanding before they can apply material of a given unit, and that such a framework is

best developed by students reading assigned material then hearing a professor go over the same material to emphasize important points.

One critical flaw in the standard lecture-before-lab implementation is that it depends on students reading assigned material before lecture. If they do not, then either (a) lecture is unintelligible to students who have not done the assigned readings if the instructor only addresses the key concepts of a unit, or (b) lecture becomes a substitute for the assigned readings. Neither of these two results is desirable, and neither places the freshman learner in a position to actively engage in her own learning.

An alternative implementation of the “large lecture/lab” instructional model would reverse the order of lecture and lab (or recitation). Students would be expected to read material, attend laboratory sections emphasizing hands on work, then at the end of the “cycle” students would attend lecture. The lecture in this implementation would play the role of “wrap up” for the unit students have completed - including making generalizations from the specifics students have learned, and demonstrating any common mistakes students make when applying the material of the unit. In this implementation, students are more or less obligated to read assigned material before their first unit meetings (labs/recitations) because in lab they must “perform” using what they have learned from the assigned readings.

The bottom line would be that it’s easy for an unprepared student to “hide” in lecture, but not in small lab sections. The pedagogically larger picture is that “lecture as wrap up” should require students to take more responsibility for their own learning, and in the end be more actively engaged in their learning than the more common “lecture as introduction” path.

In fall semester, 2004, we undertook a study to compare the “lecture as wrap up” implementation to the more standard “lecture as introduction” implementation. For a number of reasons, our initial study was based on internal clustering methods<sup>1</sup>, although we cautioned that clustering based on external performance metrics was desirable.

At the end of fall semester, 2005, we reanalyzed our 2004 data using the more standard external metric clustering. We also went further and did a preliminary external cluster-based analysis on data obtained in fall semester, 2005. It is these two new studies that we describe here. Results suggest that for large, freshman engineering lecture/laboratory classes a combination of “lecture as wrap up” and the use of personal response systems (PRS) may provide a means to enhance the performance of lower-performance students.

## **Background and initial hypothesis**

One of the current bedrocks of pedagogy is active learning and its importance in transforming the educational enterprise from a view of the student as a vessel into which the professor pours “knowledge” to one in which the learner is actively engaged in her own construction of knowledge.<sup>2</sup> One example of the introduction of principles of active learning into engineering studies can be found in<sup>3</sup>. The goal of establishing active learning has become wide spread in

computer science and engineering to the extent of enabling students to set the term grade they desire, then work towards that goal. <sup>4</sup>

A specific area of active investigation with the goal to enhance active learning/student engagement lies in the large and growing work on the use of the “personal response system” (PRS). PRS units are typically small, student held units that resemble a TV remote, and are typically called “clickers” by faculty and students alike. PRS units are used in modes ranging from giving in-lecture quizzes for grade, to use as student-to-faculty feedback devices. As feedback devices, they enable instructors to tailor lecture material on the fly to help meet, for example, the needs of “just in time teaching” (JiTT) <sup>5</sup>.

Enhancement of active learning is one of the backdrops for research reported here. As noted above, we believe that requiring students to do assigned reading before any class (lecture or lab) that is dependent on the assigned reading could have the effect of actively engaging students in their learning process, certainly more so than the standard lecture situation in which lecture material closely mirrors assigned reading.

Specifically, our hypothesis remains that students who participate in *lecture as wrap up* will perform better than those who participate in *lecture as introduction*. Other than our work reported at ASEE 2005, other past studies on this specific issue were difficult to find, and in fact, we found no other directly relevant literature.

There is a body of literature that on the surface seems to address similar issues to those that we address: the body of work on inductive learning/teaching. Inductive learning focuses on presenting special cases (problems) before presenting to students the theories which can provide conceptual underpinning to explain the specific examples. Examples of studies involving inductive learning/teaching have appeared in ASEE in recent years. <sup>6-8</sup>

Inductive teaching/learning is usually contrasted to deductive teaching/learning. In deductive teaching/learning general principles are presented first in order, followed by display of specific examples. The issue of ordering is shared with our research. However, in our work, the significant point is not between whether examples or principles are given to students first. Rather our focus is on whether students have a (relatively) passive lecture period before or after they have a (relatively) active laboratory/recitation period. Note in our work, if we were forced to consider the approach inductive or deductive, then we would be forced to choose *deductive* since in our model, students are required to do outside reading as a first step, and that outside reading is largely principle driven. The point here is really that the deductive/inductive dimension does not help to illuminate our findings.

### **Results from preliminary study (ASEE 2005 report)**

As reported in the ASEE 2005 Proceedings <sup>1</sup>, our preliminary work involved an exploratory clustering of student performance by using internal metrics only. Based on our exploratory method using total course points for the students in CSE 131 in fall, 2004, we reported

preliminary results that seemed to imply a positive benefit for the *lecture as wrap up* approach, especially for the higher achieving students. The result was tentative.

We noted however, that further work was required to confirm the putative effect. In particular, we noted the need for future research to reanalyze the data using a clustering metric(s) external to the objective measures of performance in CSE 131.

### **Setting for studies in both Fall 2004 and Fall 2005 and variables**

Computer Science and Engineering (CSE) 131 is a high enrollment (approximately 250-300 students per term), multi-section (approximately 24 sections), freshman engineering course in technical problem solving with MATLAB. It is offered fall and spring semesters with an additional offering in summer term with a substantially lower enrollment. CSE 131 is a required gateway course for most majors in the College of Engineering, Michigan State University. The standard "Calculus 1" is a pre-requisite/co-requisite for CSE 131.

Because of a scheduling "glitch," in fall semester, 2004, CSE 131 was offered in two lecture sections at opposite ends of the week. One lecture section met on Monday nights at 7:00 p.m., with associated labs running on Tuesdays, Wednesdays, Thursdays, and Fridays. A second lecture section met on Friday mornings at 10:20 a.m., with associated labs running on Tuesdays, Wednesdays, Thursdays, and Fridays. Lecture section enrollments were approximately the same size in both years. Laboratory sections entailed enrollments of a maximum of 16 students, and were not mixed – that is laboratory sections included students either totally in the Monday lecture section, or totally in the Friday lecture section.

Each student in CSE 131 meets for one lecture session per week lasting one hour and twenty minutes, and meets twice per week in laboratory sessions twice per week with each lab meeting lasting one hour and twenty minutes. Thus the Monday lecture session students met in lecture *before* participating in any lab assignments, while Friday lecture session students met in lecture *after* participating in lab assignments for the week. Both lecture sections had identical reading assignments, and both sections had identical laboratory exercises, laboratory quizzes, midterms, term project and final examination.

An exception for the Friday lecture section was that several associated lab sessions met after the Friday lecture for the second of the two lab meetings each week. This exception covered 24 students of the total 109 students in the Friday lecture section.

Because the initial study and its results were preliminary, we continued the same schedule for CSE 131 in fall, 2005, to enable continuation of the research. The same faculty person was lecturer for all lecture sections, both in 2004 and in 2005. In fall, 2004, and fall, 2005, the two lecture sections (Monday and Friday) received different treatments:

Lecture material for the Monday lecture section consisted of the typical **introduction** of a unit. Lectures largely paralleled assigned readings for the unit. MATLAB problems were worked that

were drawn from examples in the assigned readings. Lecture material for the Friday lecture section consisted of **wrap-up** for a unit. Lectures focused largely on two areas: (a) demonstrating MATLAB points that beginning students are likely to misunderstand and (b) working MATLAB problems drawn from the exercise sets that students were assigned for lab sessions.

The two different lecture treatments were used to form a retrospective comparison of student performance under “lecture as introduction” versus “lecture as wrap-up.” In fall, 2004, the single metric selected for student performance was the total of “course points” earned by a student over the entire term (of a possible 100) plus the number of “extra credit points” earned (of a possible 4). This metric was also the basis for student term grades.

In Fall, 2005, the study was broadened to include the use of PRS units in both Monday and Friday lectures. PRS units were used in 13 of 15 lecture meetings in both lecture sections. PRS use was focused mainly on quizzing of students on material they were assigned for the lecture meeting. Students answered questions with their PRS where questions were spaced out through the lecture, with a rough average of 5 PRS questions per 80 minute lecture. Finally, the course grade weighting for PRS quizzes was purposefully set high – at 30% of the course grade dependent on PRS quizzes. This high weighting for lecture quizzes was set to strongly encourage students to attend lecture and to come to lecture prepared.

The fall 05 data gathering was extended for both fall04 and fall05 datasets to include the following:

- category totals for all categories of graded work,
- total term points (out of 100),
- percentile ranking based on total term points,
- ACT composite and area scores,
- cumulative GPA as of the end of Fall, 2005.

In synopsis, a top level view of our evolved study based on both the fall04 and fall05 datasets is that we have two major variables with two treatments each:

- **timing of lecture:** either lecture as introduction ... OR ... lecture as wrap up, and
- **use of PRS technology:** either PRS used ... OR ... PRS not used.

### **Comparing FS 04 to FS 05**

Does use of the PRS result in improved student outcomes? Since both the FS04 and FS05 were taught by the same instructor and used similar assignments and exam items, the primary

difference between the two years' offerings was the use of the PRS. There were no significant differences between the students in FS 04 and FS 05 as measured by GPA or ACT scores.

In FS05, the total point distributions were re-allocated to award credit for the PRS questions. However, the instructor believes that the individual mini tests, group mini tests, midterm exams and final exam items were comparable in terms of content and difficulty between semesters. We compared performance on these items between FS 04 and FS 05, adjusting the points so that they were proportionally the same for both years. Because the scores on these items were not normally distributed, we computed ranks and performed independent sample t-tests on the resulting ranked scores. The results are shown in Table 1

Table 1: Comparison of student performance between FS 04 and FS05

| Course Component     | FS 04 n = 203 |         | FS05 n = 246 |         | Significance two-tailed |
|----------------------|---------------|---------|--------------|---------|-------------------------|
|                      | Mean          | Std Dev | Mean         | Std Dev |                         |
| Individual mini test | 14.12         | 3.33    | 14.47        | 3.43    | .164                    |
| Group mini test      | 13.23         | 4.29    | 14.73        | 2.64    | < .001                  |
| Midterm exams        | 18.51         | 9.45    | 21.55        | 7.45    | < .001                  |
| Final                | 6.99          | 2.15    | 8.66         | 2.06    | < .001                  |

### Results of reanalysis of fall 04 dataset: external clustering method, no PRS

We used a two-dimensional clustering space based on composite ACT (as one clustering dimension) and cumulative GPA (as the second clustering dimension) to classify students into groups of low, medium and high academic performance. The K-means clustering algorithm in SPSS was used to form three performance clusters: Low (mean GPA = 2.60, N = 62); Medium (mean GPA = 2.89, N = 90); and High (mean GPA = 3.4, N = 51).

The three-cluster dataset was then divided into two groups: one for the Monday lecture section (*lecture as introduction*, N=112) and one for the Friday lecture section (*lecture as wrap up*, N=91). Its important to keep in mind that the entire fall04 dataset is in the *PRS not used* treatment.

The research question for the fall04 dataset then is “Was there any statistically significant difference between the Monday and Friday lecture sections on any performance measures: total term points, term percentile based on total term points, or any individual category of graded work?”

The somewhat unexpected answer was **NO**. This was contrary to our preliminary study, but given that our preliminary study used an exploratory method based on internal clustering, not a startling outcome since we are not using the much more stable and established method of performance clustering based on ACT/GPA data.

## **Results of analysis of fall 05 dataset: external clustering method, PRS in play**

The fall 05 data was clustered on the basis the external ACT/GPA data, again using as before the SPSS K-means clustering algorithm in SPSS, and again three performance clusters were formed for the entire fall05 dataset: : Low (mean GPA = 2.67, N = 53); Medium (mean GPA = 2.96, N = 120); and High (mean GPA = 3.34, N = 47).

As before, the three-cluster dataset was divided into two groups: one for the Monday lecture section (N = 123) and one for the Friday lecture section (N = 124). For this data, it is important to remember that the PRS units were in use for all students.

Because the distribution of students' scores on the various course components was not normal, we computed ranks for the scores. All analyses were performed on the ranked values. We examined differences between the Monday and the Friday FS05 students in the three clusters using the standard SPSS t-test for equality of means of the ranked scores. For the FS05 students, we found no statistically significant differences between the Monday and the Friday treatments the high performance cluster nor for the middle performance cluster.

However, we found statistically significant differences for the low performance cluster for the following course performance metrics: total term points (which determine final course grade); total points from the two hour exams; total scores on individual lab quizzes. Details of the results are shown below.

### **Total term points**

For total term points for the low ability group, the Monday lecture was lower (mean = 57.93 N = 27) than the Friday lecture (mean = 67.36, N = 26;  $p = .031$ ). The distributions of the three ability group scores are shown in the boxplot in Figure 1.

Boxplots summarize data that are not normally distributed<sup>9</sup>. Figure 1 shows the distribution of total points for by each student ability group in both lecture days. The boxes indicate the interquartile range, the data between the 25<sup>th</sup> and 75<sup>th</sup> percentiles. The heavy black line shows the median value for that group's data. The vertical lines extend to 1.5 times the interquartile range. Outliers beyond the vertical lines are indicated by circles. Note that the scores for the low ability group (blue boxes) in the Friday section are higher than those in the Monday section and are almost as high as the scores for the middle ability group in the Friday section.

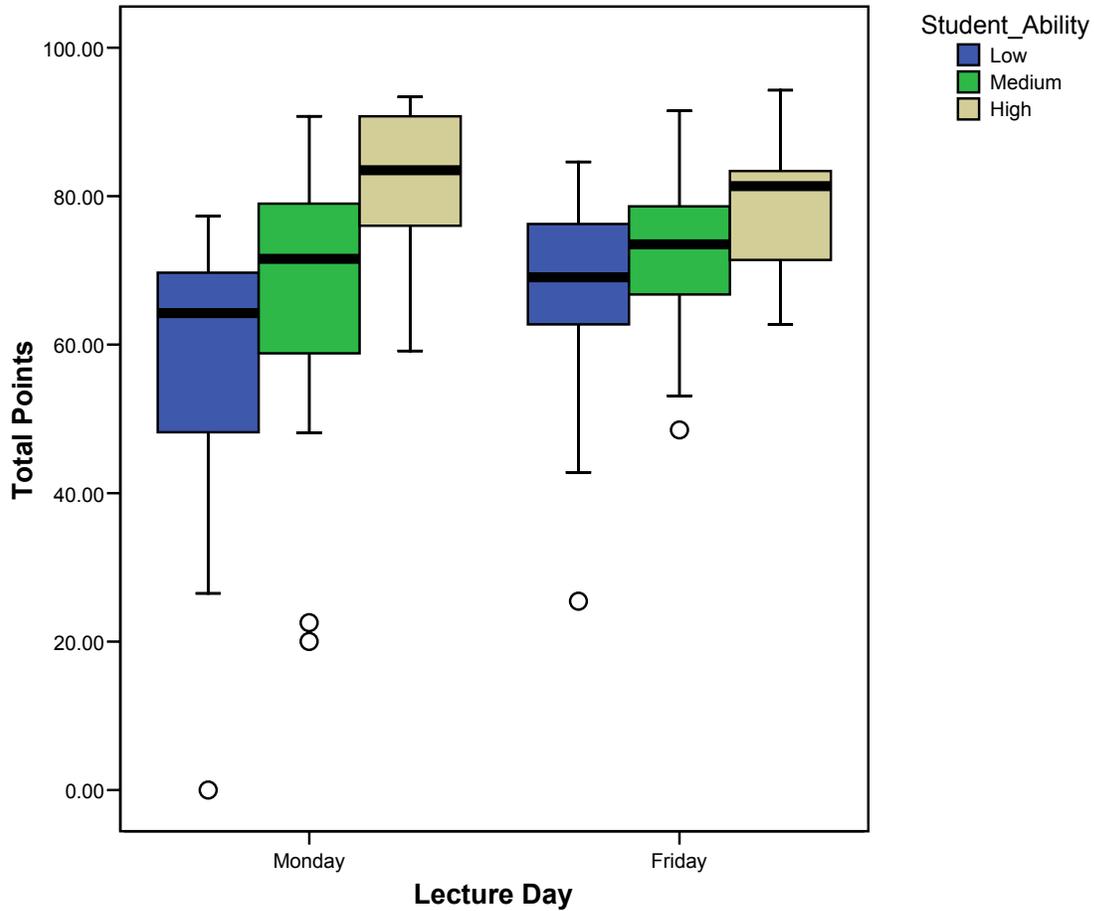


Figure 1: Fall 05 dataset, external clustering, comparison of Total Points in Monday and Friday treatments

### Total hour exam points

For total points from earned on the two hour exams in the low ability group, the students in the Monday lecture performed worse (mean = 8.35, n = 27) than the students in the Friday lecture (mean = 11.78, n = 26; p = .005). These distributions are summarized in Figure 2. Again note that not only are the scores for the low ability group higher on the Friday lecture, but they are almost as high as the scores for the middle ability group.

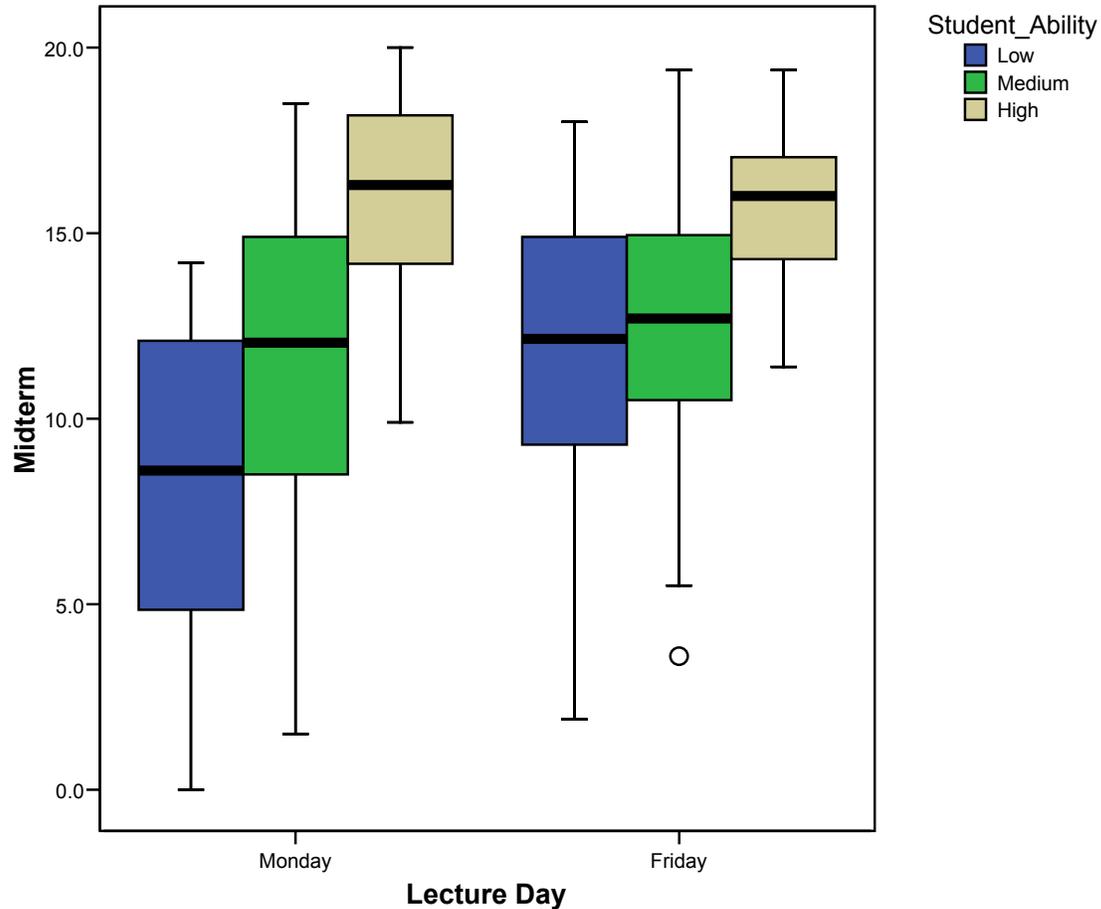


Figure 2 Fall 05 dataset, external clustering, comparison of Midterm Exams in Monday and Friday lectures

### Individual laboratory quizzes

For the individually taken laboratory quizzes, the low ability students in the Monday lecture (mean = 11.34, n = 27) had lower scores than those in the Friday lecture (mean = 13.65, n = 26; p = .009). These results are shown in Figure 3. While the low ability students in the Friday section did not do quite as well as the middle ability students in the Friday section, they again did better than their counterparts in the Monday section. This is particularly interesting as the individual quizzes are taken during the lecture sections and cover materials from the readings in addition to problems worked by the students during lecture.

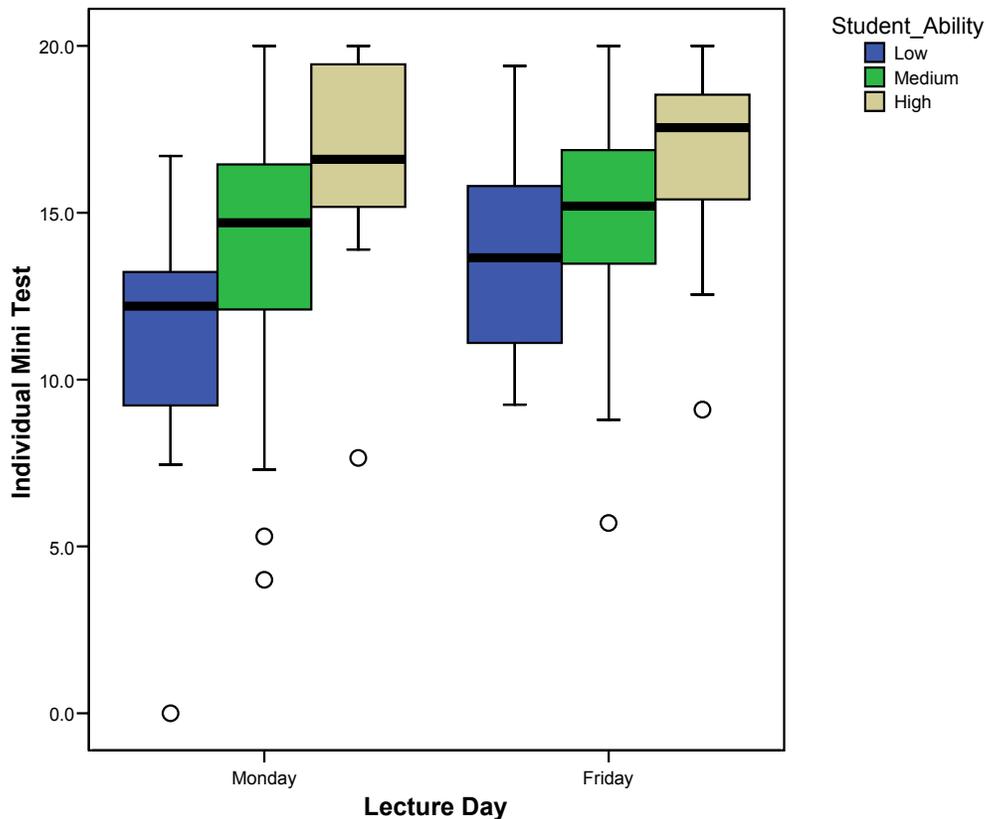


Figure 3: Fall 05 dataset, external clustering, comparison of Individual Lab Quiz scores in Monday and Friday lectures

The results show a very clear pattern. For the fall 05 students, the cluster of low performance students is significantly and systematically (across these three performance measures) “brought up” by being in the Friday lecture section – the lecture section that utilized *lecture as wrap up*.

It is important to note that the overall most important performance measure from a student viewpoint is total term points. However, from the viewpoint of the instructor for the class, the most important performance measure is the hour exam (midterm) data. The hour exams in CSE 131 are taken in the laboratory, and involve the students solving problems directly with MATLAB, some problems being novel to the student. Thus given that the goal for the course is to develop problem solving ability, and in particular to develop problem solving ability using MATLAB, the hour exam data is central.

## Analysis of gender differences

Recruitment and retention of women in engineering courses is a concern. Introductory courses are often seen as dissuading women from persisting in STEM disciplines<sup>10</sup>. We examined the outcomes from FS 05 to see if there were any gender differences. There were a total of 56 women and 190 men in the course (no gender information on one student.) Because of the sample sizes, we were not able to perform performance clustering across Monday/Friday lecture days. Overall, women (mean = 73.79) actually did better than men (mean = 69.30) in the course ( $p = .038$ ).

When we examined the data by lecture day, there were no gender differences for the Monday class. However, in the Friday class, women (mean = 74.69,  $n = 27$ ) did better than men (mean = 67.35,  $n = 95$ ,  $p = .014$ ). Although the overall outcome for women in the course was better than for men, these results suggest that for a population that is considered “at risk” in engineering, the *lecture as wrap up* may provide opportunity for more success.

## Discussion of results of combined analysis of 04 and 05 datasets and future research

On three of the four course components that were comparable between years, students in the FS05 class, which used the PRS, performed significantly better than did the students in FS04 across the board.

When analyzing the outcomes by student performance group, for the case where there was no use of the PRS units – and no credit given for coming to lecture – the *lecture as wrap up* versus *lecture as introduction* showed no statistically significant differences for any of the three performance clusters. But when the PRS units were used in both lecture timing treatments and students received credit for the questions answered during lecture, there was very clear advantage to the *lecture as wrap up* treatment.

Although it is premature to assert a causal explanation for these findings, we find it intriguing that the combination of PRS units for answering questions and the lecture as wrap up timing results in a **significant improvement in the outcomes of the lowest performance students**. One of the most important problems in undergraduate engineering education is to help struggling students “over the initial hump” that many students feel on entry to engineering colleges. Indeed retention of students in the first few freshman courses is becoming increasingly important to most engineering colleges.

While premature to suggest a causal understanding to explain our results, we are comfortable stating our hypothesis for future work. First, we believe that our results need to be replicated by others, and we would welcome discussion and collaboration with other researchers to further test the validity of our results presented here.

Second, we believe that the basic effect we have observed – raising the performance of the lowest performance students – can be understood as follows. The use of PRS systems encourages

student engagement in lecture, and our heavy weighting for in-lecture PRS-based quizzes strongly encourages students to come to lecture prepared. This helps to get past the tendency of some low performance students to depend solely on lecture to supply “what they need to know.” Further however, the timing of the lecture (as wrap up) and in particular the content of the wrap up helps students, particularly the lower performance students, to form a mental scaffold so that the material of the week’s work becomes an integrated whole.

Pedagogical researchers have for some time argued that a flaw in engineering education is that our students are expected to hold too many unrelated facts – as the students see it – without developing a mental structure to hang all those facts. This is particularly challenging for the lower performance students. Our use of PRS and lecture as wrap up demands that students engage and prepare themselves, and when they come to lecture, promotes the development of a knowledge framework.

## References

1. Sticklen J, Urban-Lurain M, Hinds T, Eskil T, Amey M. Multi-section Freshman Classes with Laboratories: Lecture as Intro vs. Lecture as Wrap-up. 2005; Portland, OR. American Society for Engineering Education. p paper #2005-1601.
2. Bransford JD, Brown AL, Cocking RR. How People Learn: Brain, Mind, Experience, and School. Washington, D.C.: Committee on Developments in the Science of Learning, Commission on Behavioral and Social Sciences and Education, National Academy Press; 1999.
3. Hall SR, Waitz I, Brodeur DR, Soderholm DH, Nasr R. Adoption of active learning in a lecture-based engineering class. In: Bjedov G, editor; 2002 November 7; Boston, MA. IEEE. p T2A\_9-15.
4. Urban-Lurain M, Weinshank DJ. "I Do and I Understand:" Mastery model learning for a large non-major course. Special Interest Group on Computer Science Education 1999;30:150-154.
5. Bailey T, Forbes J. Just-in-time teaching for CS0. 2005; St. Louis, Missouri, USA. ACM Press. p 366-370
6. Moor S, Piergiovanni P. Inductive Learning in Process Control. 2004. p session 2213
7. Hesketh R, Ferrell S, Slater CS. The Role of Experiments in Inductive Learning. 2002. p session 3613.
8. Dahm KD. Process Simulation and McCabe-Thiele Modeling: Specific Roles in the Learning Process. 2002. p session 3513.
9. Tukey JW. Exploratory data analysis. Reading, Mass.: Addison-Wesley Pub. Co.; 1977. xvi, 688 p. p.
10. Seymour E, Hewitt NM. Talking about leaving: why undergraduates leave the sciences. Boulder, Colo.: Westview Press; 1997. x, 429 p.