Comparison of Mastery Learning and Traditional Lecture-Exam Models in a Large Enrollment Physics Course

Dr. Barbara Masi, University of Rochester

Dr. Masi is the Director of Education Innovation and Assessment Initiatives in Arts, Sciences and Engineering at the University of Rochester.

Dan M. Watson, Department of Physics and Astronomy, University of Rochester

Dan Watson, Professor of Physics and Astronomy, is currently chair of the Department of Physics and Astronomy at the University of Rochester. His research centers on the formation of planets and stars, and on the development of detectors and instruments for infrared astronomy. His educational activities include mastery learning, tutorial learning, and the application of online resources and assessment in physics and astronomy courses, both for STEM students and non-majors.

Prof. Arie Bodek, Departmente of Physics and Astronomy, University of Rochester

Prof. Bodek received his B.S. in Physics (1968) from the Massachusetts Institute of Technology, and his Ph.D. in Physics (1972) also from MIT. He was a postdoctoral associate at MIT (1972-74) and a Robert E. Millikan Fellow at Caltech (1974-77). Prof. Bodek joined the University as an Assistant Professor of Physics in 1977. He was promoted to Associate Professor in 1980 and to Professor in 1987. Prof. Bodek was appointed as an Alfred P. Sloan Fellow (1979-81); NSF-JSPS Fellow, KEK, Japan (1986); and Fellow of the American Physical Society (1985). He served as a project director at the Department of Energy (1990-91), was Associate Chair (1995-98) and then Chair of the Department of Physics and Astronomy (1998-2007). He is on the editorial board of theEuropean Physics Journal C. Prof. Bodek was awarded the 2004 APS W.KH. Panofsky Prize in Experimental Particle Physics "for his broad, sustained, and insightful contributions to elucidating the structure of the nucleon, using a wide variety of probes, tools, and methods at many laboratories." In 2004, Prof. Bodek received the University of Rochester Award for Excellence in Graduate Teaching. His doctoral thesis provided some of the evidence of the quark's existence that was the basis for the 1990 Nobel Prize in Physics. Prof. Bodek's research interest is in the field of Experimental High Energy Physics.

Mr. Dev Ashish Khaitan

Dev Ashish Khaitan is a doctoral student in the University of Rochester Department of Physics.

Erik Garcell

Erik Garcell is a doctoral student in the University of Rochester Department of Physics.

Comparison of Mastery Learning and Traditional Lecture-Exam Models in a Large Enrollment Physics Course

Abstract

This study describes the impact of two pedagogical models, mastery-based, self-paced learning versus traditional lecture and exam, on student performance, study behavior and confidence in an introductory physics course. The "control" course was designed as a traditional lecture and recitation course with three major exams and final exam (LRE). The "experimental" course, which drew on Bloom's Learning for Mastery and Keller's Personalized System of Instruction models, was designed as mastery-based, self-paced instruction (MSP) where students were required to pass each of 17 module exams with a minimum score of 90% score and to complete a final exam. To compare groups, both received the same written, final exam. A quasi-experimental design was used. Final exam results showed that the MSP group mean was higher (M=67.4, SD=15.7, N=151) than that for the LRE group (M=60.6, SD=17.5, N=160) (t(309)=2.179, p<0.001). At-risk students in the MSP course also performed better than a comparative group in the LRE course. Pre- and post-course written surveys revealed that the MSP group's mean confidence in physics skills was higher than LRE group mean by the end of course. Observations and student and teaching assistant interviews about student study processes highlighted a greater prevalence of "deep learning" and "strategic learning" strategies used by the MSP group compared to the LRE group.

I. Introduction

The introductory mechanics course, required of most engineering majors, poses a challenge for many students, particularly those with weaker high school math backgrounds. This course was one of several STEM foundation courses identified that adversely affected retention in engineering majors. In recent years, an intensive tutoring program was added to the physics course to improve learning and performance for students who performed poorly on the first or second major exam. As engineering major enrollments grew, however, the tutoring program grew to a size that was hard to manage. In the final analysis, it was concluded that tutoring as part of the traditional course was an insufficient solution for two reasons. Students in need of tutoring were often identified only after the first or second major exam. Moreover, since tutors only had student performance on a major exam that generally covered one-third of the course, each tutor needed to take additional time in tutoring sessions to identify individual student weaknesses before actual tutoring could commence.

To address these issues and improve student performance, the traditional lecture, recitation and exam physics course was split into two separate courses: a traditional lecture, recitation and exam (LRE) course as "control" and an "experimental" mastery-based, self-paced (MSP) course that drew on elements of Bloom's Learning for Mastery (LFM) model and Keller's Personalized System of Instruction (PSI) model. The overall goal of this study was to determine if the promising results noted in the literature on LFM and PSI could be replicated in a large enrollment, introductory physics course. A meta-analysis of 108 studies that employed either

LFM or PSI found that the mastery group learning was able to raise the student scores by an average of 0.48 standard deviations or .48 σ (Kulik 1990). A survey of recent mastery learning implementations, whether LFM, PSI or a combination of the two, in medium to large enrollment undergraduate science and engineering courses points to positive differences when compared to the traditional group: college physics and other sciences (Diegelman-Parente 2011; Kulik et al. 1990; Leonard 2010; Zeilik 1981); engineering (Capaldi 2014; Sangelkar 2014; Wankat and Oreovicz, 1984, 1993, 2001). In each study, positive differences in mastery group performance were found compared to the traditional group. This study was also driven by the question of whether a large enrollment introductory physics course could be designed and delivered given the constraints noted by Eyre (2007) and Fuller and Winch (2005). In particular, could a course be designed for mastery learning given the time limitations of a 14-week semester, could test proctoring be staffed appropriately, and would student procrastination present a difficulty?

Given the prominence of both institutional and national goals of attracting and retaining more students in STEM majors, finding alternative pedagogies that are successful in supporting all students was an important rationale for this study. Another goal of this study was to compare the performance of similar groups of students considered "at risk" in the two different courses. These groups are defined on the basis of their performance coming into the course. This study explores how and by how much the MSP approach impacts a group's performance. Of particular interest is the performance on the tails of the course's normal distribution.

Bloom, Keller, and Bandura argued that students could only move to the next level of difficulty in their studies if they were able to master tasks through well-structured learning experiences and gain confidence through that effort (Bandura 1973, 1975). To explore student confidence, this study incorporates a student self-assessment based on a Bandura confidence scale to measure student confidence in physics skills at the beginning and end of the term.

Lai and Biggs (1994) observed that a mastery learning process of study, test and retest would, in fact, promote surface learning over deep or strategic learning approaches if test questions were lower order memorization questions rather than higher order problem solving questions. In order to examine student study behavior, this study incorporates an observation protocol and semi-structured interview protocol to investigate how students approached physics study and test-taking.

Drawing on the above models, this study hypothesizes that three factors would be important in determining a student's success. These factors are math and physics background, study habits, and test-taking strategies. Therefore, in addition to study protocols to examine study habits and test-taking strategies, the study also analyzes students' prior math and physics background including high school math and physics course background, SAT and ACT math scores, and calculus course background and grades.

II. Bloom and Keller Models

Bloom's LFM model is structured so that the course is broken down into small units, roughly two weeks of instruction, with each unit presenting clear learning objectives. The teacher

lectures on the unit to the entire group of students and gives a unit test. Students who perform poorly receive individualized tutoring before moving on to the next unit. Individualized tutoring that is related to a specific unit is key since a tutor would be able to more quickly identify the areas of student misunderstanding and support the student's mastery of the unit material. By the time students reach the end of the course, Bloom predicted that all students would have mastered the course material and should all receive final mastery test scores that only the top 5% of the distribution usually received. He also predicted that not only would students' cognitive entry characteristics (specific knowledge, abilities related to the course), but also their affective entry characteristics (attitude, confidence) would be improved. He noted that both characteristics were essential for academic progress and success (Bloom 1973, 1976).

In the personalized system of instruction (PSI) proposed by Keller, students could progress at their own pace; unlike in LFM where they had to keep pace with the lecturer. In the PSI, course material was presented in small, written compartmentalized units with clear objectives. The units would each contain less course material than LFM so that it could be completed in a week. Students would study based on these units and take a proctored unit test before moving on to the next unit. If the student did not pass, the proctor would discuss the test results and identify key areas of misunderstanding for the students. The student would repeat this process of studying and testing until they scored 100%. This meant that in some cases not all students completed all units. Unlike Bloom's LFM, mastery of each unit was considered a sufficient measure of mastery (Keller 1974, 1981).

III. Course Set-Up and Structure

The undergraduate physics course in mechanics was a typical first semester course of a two-course sequence for students planning to major in engineering. Topics addressed include motion in one and two dimensions, Newton's laws, work and energy, conservation of energy and momentum, systems of particles, rotations and oscillations and periodic systems. The course drew on student knowledge of geometry, algebra, trigonometry, vectors and calculus.

IIIA. Control: Traditional Lecture-Recitation-Exam (LRE) Format

The structure of the LRE format included three hours per week of face-to-face lectures delivered by faculty and two-hour recitations delivered by teaching assistants. The lectures were recorded and the videos were available through the course website. In the recitations, students completed a set of special physics problems developed by the faculty instructor that highlighted key weekly lecture concepts and methods. Weekly homework sets and reading assignments were based on the book, *Fundamentals of Physics Extended (10th Edition)* (Halliday 2013).

Students were also required to complete five laboratory experiments. This component accounted for 5% of their final grade. Recitation participation accounted for 5% of their grade and they also had fourteen homework sets; the scores for the best of twelve accounted for 10%. They were subjected to three midterm tests; the scores of the two best midterms accounted for 25% each. The common final exam for the course accounted for 30% of the final grade.

IIB. Experimental: Mastery-Based, Self-Paced (MSP) Format

The MSP format was offered concurrently with the LRE. Students in this section had no formal lecture sessions, but could attend the LRE lecture or view the recording on the course website. The website also hosted a variety of material, chiefly the concise units based on the PSI method (the same one shared with LRE). They also had access to the same reading assignments and homework sets from the LRE course. In addition, both courses provided access to past homework sets from previous iterations of the LRE course.

A self-paced study room was appropriated and staffed every weekday from 2pm to 10pm. Adjacent rooms were set up with individual tables and chairs for student testing and grading. Normally this area was staffed by three teaching assistants with greater deployment of TAs during peak hours between 3 and 6pm. Several copies of the textbook, module units and whiteboards were available for students to use. Students were free to use the study room at their convenience during the staffed hours and were encouraged to work either individually or in groups. When appropriate, the TA would attempt to form study groups for students to work together. When the students were stumped, they could ask questions and receive help from the staff.

The course material was divided into seventeen units. Each unit summarized the key concepts, background math and methods required to solve the problems, and about ten to fifteen physics problems to solve. Often reading assignments were also included on these units in addition to being posted on the course website.

Students were free to work through the unit(s) at their own pace and solve as many problems as they felt necessary. There was no requirement of completing these for a grade like the LRE. When a student felt prepared to demonstrate mastery of a module, the student completed a short oral quiz with a teaching assistant. The purpose of these quizzes was to determine whether the student had grasped the key concepts and to get them familiar talking the language of physics. After this, they could schedule a unit exam.

The unit tests were often three to four problems long. For each unit, there were at least four varieties of tests that were chosen at random for the student. Students were not permitted to take anything into the exam room; blank sheets of paper, pencil, eraser and a basic scientific calculator were provided. While the tests were designed to be completed in thirty minutes, students were given an hour to complete these tests. After which, they went over the test with a TA.

The tests were graded immediately after the student completed it. Their score was determined on the basis of determining the relevant concepts, using the appropriate method and finally arriving at the correct answer. A score of 90% was needed to pass the module.

After grading the exam, the TA went over the exam with the student and provided feedback. If they passed, they received a note stating that they passed. If they did not, the TA would go over the exam with the student, provide constructive feedback on how to fix the misconceptions and additional problems to work on. The key messages from this conversation were written on the form that the student would take with them while the TA retained the test.

If a student scored less than 75%, it was determined that significant studying needed to be done before they should take the test again. Thus, they were required to wait at least a day before taking another variety of the same unit test. If a student failed the same unit three times, they were required to talk to the head TA. This strategy was used as an opportunity to force the student to seek help through official channels rather than attempting to work on their problems on their own.

The 17 module quizzes were worth 65% of a student's total grade, and periodic labs and lab reports were worth 5%. At the end of the semester, students completed a written final exam, worth 30% of their final grade. In order to compare MSP and LRE student groups, they both received the same final exam jointly developed by the MSP and LRE professors

IIIC. Teaching Assistants

Staffing of the MSP course in the pilot phase drew on experience with the University of Rochester Workshop model of peer-led, group work whereby not only graduate students, but also undergraduates can be effectively utilized as teaching assistants (Roth et al. 2001). This experience was combined with projections of requirements for self-paced course operation. It was estimated that the population of students working in the course workshop would range between 15 and 30 at any given time during the 40 hours of weekly operation, with peak population in the 3 to 6pm time period. Based on this estimate, three teaching assistants were scheduled to be present in non-peak hours and five during peak hours. One or two teaching assistants were present to tutor on physics problems with individuals or student groups. In addition, there were specified times for pre-screening by an additional teaching assistant, and at least one available for administering and grading mastery quizzes, which also expanded to two during peak hours. In the inevitable crush of additional demand during the last two weeks of the semester, it was necessary to add an additional twenty hours of teaching assistant time to normal operating hours, during which additional teaching assistants were on hand only to administer and grade mastery quizzes.

In contrast, staffing for the LRE course required teaching assistants to run the weekly recitations and grade exams.

Given the above estimates of MSP course workload, five physics graduate teaching assistants and eleven physics undergraduate teaching assistants were assigned to the course. In comparison, 2 graduate TAs and 6 undergraduate TAs were assigned to the traditional course. In general, graduate TAs work 16-20 hours per week and undergraduate TAs work 5 hours per week. Utilizing a large number of undergraduates as teaching assistants also addressed potential difficulties with English language skills of some graduate students.

Optimization of the student/teaching assistant ratio was not attempted on the pilot offering of the self-paced course; the authors plan to track this issue over the course of several semesters. It can be noted, however, that the main difference from the parallel lecture/workshop course PHY 121 is the presence of additional instructors for prescreening, administering, and grading mastery quizzes, which exceeds the requirements for administering and grading midterm exams. It would seem that MSP courses would always be more expensive than the traditional format as a result.

Traditional and self-paced course teaching assistants completed the same two-day, basic TA training prior to the term's start. In this regimen, students learn about running peer-led workshops and offering constructive feedback. During the term, the MSP TA team shared information via email and informal meetings.

IIID. Student Course Choice Factors

The introductory physics course is a requirement for all of the university's engineering majors and several of the science majors, including geological sciences, biology, and physics. Given the large enrollments of the engineering majors relative to science majors, 95% of the 311 enrolled students registered in the course were engineering majors. Eighty-eight percent were freshmen.

The LRE and MSP courses were described to students in an email to all freshmen in majors that require the introductory physics course, and additional written information was provided to all students' academic advisers in the form of emails and presentations that could be used during the course registration period. It was intended that students should use this information to choose one of the courses. The study also tracked students who dropped the courses altogether, switched from one course to the other, or withdrew late in the term. Of the original 231 students who signed up for the LRE course, 30% dropped the course and did not sign up for the other course. Of the original 195 students who signed up for the MSP course, 18% dropped the course and did not sign up for the other course. Of the 160 students who completed the LRE course, 13% switched from the MSP course to the LRE course early in the term. Of the 151 students who completed the self-paced course, 27% had switched from the traditional course early in the term.

Since random assignment to the LRE and MSP courses was not possible, a quasi-experimental design was used. To check homogeneity of groups, student standardized math test scores were compared. In addition, student grades in previous calculus courses taken in the College, a necessary math background for the physics course, were compared. As additional pre-course measure, a special one-hour written, 19 question, basic math assessment (BMA) that included algebra and trigonometry methods needed for the physics course was given to both groups. Finally, a half-hour written physics concept exam, the Force Concept Inventory (FCI), was given at the beginning of the term to both groups, and scores used for comparison (Hestenes 1992).

Two-tailed t-tests were used to compare whether there was a statistical difference between any of the traditional or self-paced group scores (SAT math, ACT math, college calculus grade, pre-course basic math test, FCI test); no statistically significant difference between groups was found for any of the five test scores. The FCI test of mechanics and BMA test of algebra and trigonometry completed by students at the beginning of the course were the most important measures for comparing LRE and MSP groups. For the FCI test (max score=30), mean LRE score was M=16.2, SD=7.3, N=153 and mean MSP score was M=16.9, SD-6.9, N=138 and t-test result was (t(289)=.973, p<.332). For the BMA test (max score=30), mean LRE score was M=18.5, SD=4.8, N=145 and mean MSP score was M=18.1, SD=5.1, N=147 and t-test result was (t(290)=.756, p<.450).

Despite the written information provided about the MSP course, both students and advisers were unfamiliar with the pedagogy. As a result, the unfamiliarity appeared to provide some measure

of random assignment of students to each course. For example, to address student concerns about which course to choose, students were permitted to move from one course to the other during the first two weeks of the term. Once enrollment settled, an online survey was distributed to students. The survey asked students to rate the importance of factors that determined their final choice of physics course. Students were also given the opportunity to give an open-ended response. Table 1 presents numerical survey results.

Table 1. Student feedback on factors that determined MSP or LRE course choice

	Percent students reporting that factor was important or essential (Scale: 1-5 where 1=not important at all, 2=slightly important, 3=moderately important, 4=important, 5=essential)	
Choice Factor	**LRE Course (N=49, 30% response rate)**	**MSP Course (N=63, 42% response rate)**
Prefer traditionally structured course of scheduled lectures and exams	51%	18%
Prefer to set study pace and take tests when I am ready	18%	38%
Course is a better fit with my strong physics background	17%	23%

In open-ended comments, students noted other reasons for choosing LRE or MSP. Several LRE course students reported that they were afraid that they would procrastinate and preferred a strict schedule, while several MSP students reported that they felt they would learn physics in greater depth.

Given students' freedom of course choice, the numerical survey results were surprising. It was expected that a larger percentage of students would report strong preferences for one course pedagogy over another. However, students' open-ended comments suggest that the enrolled students continued to hold misconceptions of the two courses, wished they had chosen the other course, or made choices based on factors other than learning preferences, such as scheduling, instructor preference, or final grade concerns. Several student comments illustrate these points:

LRE course student comments:
"I didn't want to risk my grades with a self-paced course, because I had never heard of it before."
"I like the professor better for the regular-paced (traditional) course."
"(I was) Unsure if PHY 121P (self-paced) would be effective in teaching me what I need to learn. Now hearing other student's experiences, I wish I had taken the self-paced version."
"I have to take it for my major and I didn't really understand what self paced was."
"I'm not going to pay money to teach myself."

MSP course student comments:
"I wanted more control over my final grade."
"I needed a course that worked better with my packed schedule."

It was therefore concluded that students' lack of experience with MSP pedagogy meant that there was less self-selection bias in students' choice of course in the pilot offering of the self-paced course.

IV. Data

In order to compare the LRE and MSP courses, several direct and indirect assessment measures were designed and implemented. The methods included a final 3-hour physics exam, self-paced module test tries and completion rates, online pre-course and post-course surveys, classroom observations, and post-course interviews of students and teaching assistants.

IVA. Final Physics Exam

A three-hour written physics exam was jointly defined by the MSP and LRE faculty and given to both LRE and MSP course students at the end of the semester. The six-question test addressed representative physics concepts and problem solving methods from the course. In order to ensure that the test questions were fair to both student groups, the shared course workshop problems assigned to all were used as the basis for the exam questions.

IVB. Self-Paced Unit Test Completion Rates

A key difficulty of an MSP course completed during a 14 week college semester is that students do not have unlimited time to master course material. Students must continuously balance the mastery task of passing each module test and completing as many of the seventeen required module tests as they can in fourteen weeks. The study hypothesized that three factors would be important in determining how a student accomplished this task: math and physics background, study habits, and test-taking strategies. It was hypothesized that student patterns of module test completion would be affected by study habits, test taking strategies, high school math and physics background, SAT and ACT math scores, calculus background and grades, and difficulty of a given physics module (eg. rotational versus linear motion). To analyze these factors, several pieces of student data were gathered:

- Percent of the 17 module tests completed by the end of the term
- Number of module tests completed per module in order to pass with 90% score
- Total number of module tests taken during the entire term
- Calculus course type and math grade

This data was joined with qualitative student data on study habits and test-taking strategies gathered from student and teaching assistant interviews and surveys.

IVC. Online Surveys

Using scaled and open-ended questions, the surveys were designed to capture information on what course factors supporting student learning, student impressions and attitudes about their teaching and

learning experience, and confidence in physics and math abilities. The pre-course survey was designed to capture:

- Student confidence in physics and math abilities (scaled questions)

The post-course survey was designed to capture:

- Course factors, including face-to-face and online lectures, readings, workshops and workshop problems, tutoring, homework, and peer study groups that supported learning (scaled questions)
- Whether the course assessment processes supported learning (scaled questions)
- Student confidence in physics ability (scaled question)

IVD. MSP Physics Workroom Observations

The MSP physics workroom, a large room that seated 30 students at 5 large tables, was designed as a space for individual and peer group studying, work on physics problems alone or in groups, and to get help from teaching assistants. The room contained an area at the side for students to complete their pre-module test oral exams with a teaching assistant. A smaller room next to the workroom contained areas for students to complete their proctored module tests and to complete face-to-face test grading sessions with teaching assistants.

A semi-structured classroom observation protocol was designed to capture several types of student-student and student-teaching assistant interactions:

- Number/percent of students studying alone or in groups
- Types of student-student conversations in workroom
- Process for students to obtain tutoring help from a teaching assistant
- Pre-module test screening process

The observation protocol was implemented for two hours each week during the term.

IVE. Post-Course Interviews

A semi-structured one-hour interview protocol for students was developed to capture qualitative aspects of the teaching and learning experience in the LRE and MSP courses. Interview topics included:

- Reasons for choosing the MSP or LRE course
- Study process used to prepare for each major exam or module exam and whether this process changed during the term
- Student impressions of the pre-module test screening process and module test feedback process (self-paced only)
- Student impressions of the major exam or mastery test process and how each process supported learning of physics concepts

Ten students from each course were chosen for interviews. All students were invited to participate in the interviews; twenty students were chosen at random from the 43 respondents.

Five teaching assistants assigned to the self-paced course were also interviewed for one hour each using a semi-structured interview protocol. Interview topics included:

- Impressions of the physics workroom teaching and learning process
- Impressions of module test pre-screening process and module test feedback process

V. Results

VA. Final Physics Exam

Final exam results showed that the MSP group mean score was higher (Mean = 67.4, SD=15.7, N=151) than that for the LRE group (Mean = 60.6, SD=17.5, N=160) (t(309)=2.179, p<0.001).

It was found that the mastery learning experience raised physics final exam scores by delta (Δ) standard deviations (σ) and was calculated as (Kulik 1990):

$$\Delta = (S(MSP) - S(LRE)) / \sigma(LRE) = 0.39\sigma$$

where S(MSP) and S(LRE) are the mean final exam scores of the MSP and LRE groups, respectively, and σ(LRE) is the LRE exam score standard deviation.

At-risk students in the MSP course also performed statistically better than a comparative group in the LRE course. At-risk students were defined for this study as students whose overall background scores were lower than group average SAT Math, ACT math, or calculus course grades. Final exam results for this group showed that the MSP group mean score was higher (Mean = 56.8, SD=14.1, N=36) than that for the LRE group (Mean = 39.0, SD=14.9, N=27) (t(62)=-4.728, p<0.001).

VB. BMA Score

Since proficiency in math is important for performance in the course, the relationship between pre-course basic math assessment score (BMA) and final physics exam performance was examined. Figures 1 and 2 present the final exam score versus BMA score for the MSP and LRE groups, respectively.

Linear regression of the relationship between final exam score and BMA score alone is the predictor found a low to moderate correlation between the two. The results also show that, when BMA score is taken into account, the MSP group performed better than LRE group on the final exam. The linear regression calculations were able to account for a small amount of the variance for MSP and LRE, 27% and 24%, respectively (MSP group ($F_{(1,137)} = 51.230$, p<0.000, $R^2 = 0.27$ and LRE group: ($F_{(1,140)} = 44.202$, p<0.000), $R^2 = 0.24$). Based on the linear regression results, MSP and LRE group predicted physics final exam scores are calculated as:

MSP group: Predicted physics final exam score = 39.707 + 1.539*BMA score
LRE group: Predicted physics final exam score = 29.502 + 1.748*BMA score

The MSP and LRE linear models for predicted final exam score as a function of BMA score are also shown in Figures 1 and 2.

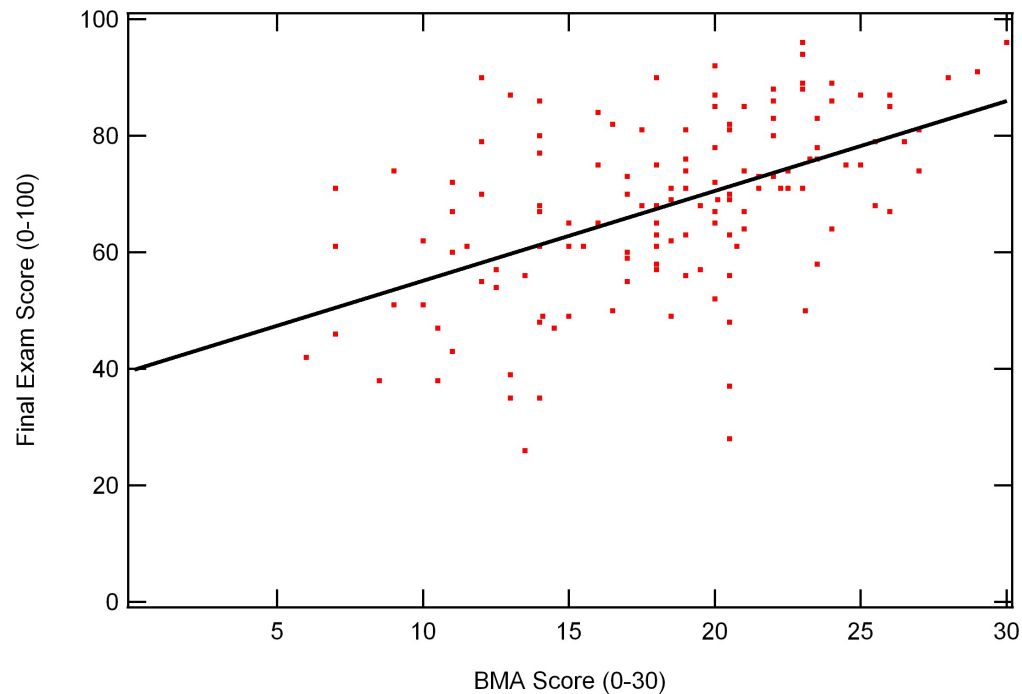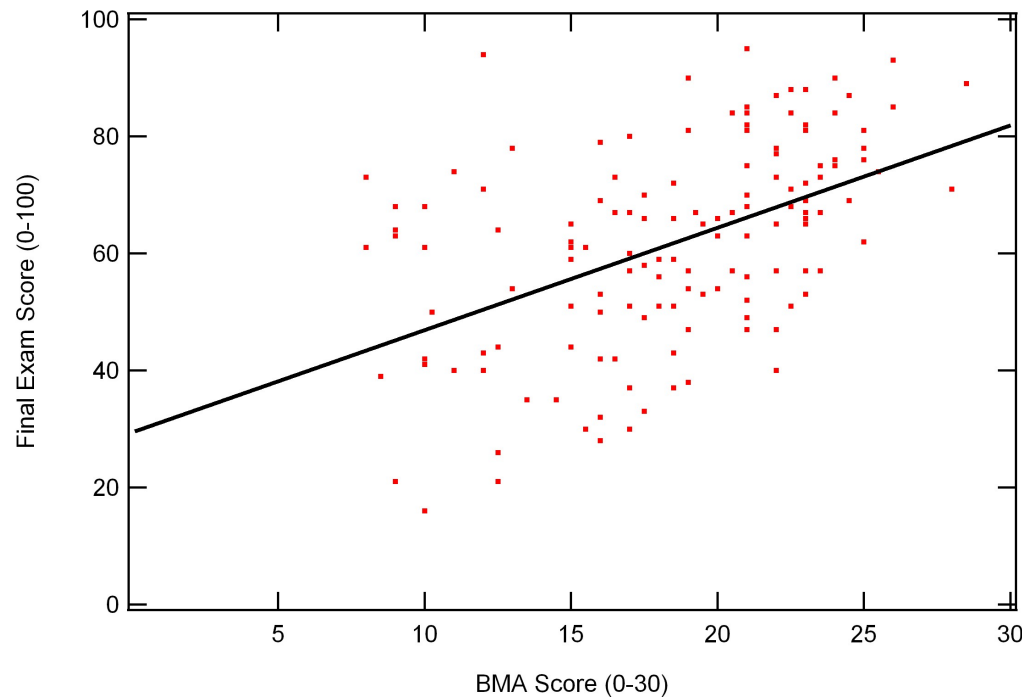**Figure 1. MSP group: Final exam score versus BMA score**

Figure 2. LRE group: Final exam score versus BMA score

VC. MSP Module Tests

The relationship between mastery test behavior and physics final exam score was also examined. The total number of mastery test modules completed was calculated for each student. It was found that 43% of the 151 enrolled students completed all 17 modules and 28% completed modules in the range of module 12 – 16. It is important to note that a total of 71% of the 151 students completed at least 12 modules. A count of the total module test questions answered correctly by the students in this group puts this result in perspective. Since each module test was comprised of 3 or 4 problems, students who successfully passed 12 module tests would have correctly answered from 36 to 48 test problems; students who completed 17 module tests would have correctly answered from 51 to 68 test problems. Twenty-nine percent completed 11 modules tests or less and were below the group average. Table 2 presents module completion data for the mastery course.

Table 2. MSP Course: Percent Students Completing Modules 1 - 17

Module Completion Groups	Percent Students Completing Modules (N=151)
Completed all 17 modules	43%
Completed modules 12 - 16	28%
Completed modules 6 – 11	23%
Completed modules 1 - 5	6%

The average number of tries per MSP module test was also calculated for each student. Overall, the average student passed each module test in just under 2 tries per module. However, students varied in the average number of tries from under 2 to over 3. Table 3 presents average number of module test tries by percent students in the course.

Table 3. MSP Course: Average number module test tries by percent students

Average number of test tries per module	Percent Students (N=151)
< 2	60%
$\leq 3 , \geq 2$	31%
> 3	9%

The number of module test tries also varied by the difficulty of the module material. While the average test tries per module averaged in the range from 1 to 2 for most modules, the average test tries was significantly higher for three of the 17 modules. The average tries per module for these three modules ranged between 2.3 to 2.5. The three modules included introductory mechanics topics that often challenge students:

Module 6: Applications of Newton's Laws (frictional forces in a moving system, circular motion)
Module 10: Collisions
Module 14: Oscillation and Simple Harmonic Motion

It was hypothesized that a possible "snowball effect" would limit the weakest students from progressing through the module tests and thus would be ill-prepared for the final exam. For

example, students who struggled with the course material would take more tries to pass the module tests. In turn, if students took more tries on average to complete module tests than others, they would be slower to progress through the 17 modules. Finally, students who did not complete all of the module tests would not have the opportunity to study all of the mechanics topics that would be included on the physics final exam.

Student math background was hypothesized to also be an important factor determining performance in the course. In this case, BMA score was used to measure student math skills at the entry point to the course. It was further hypothesized that students with low BMA scores would obtain lower final exam scores and have higher average test tries per module.

To test the hypotheses, Pearson R correlations were calculated for each relationship (Table 4).

Table 4. Pearson R Correlations for final exam score, module behavior and prior math grade

	Physics final exam score	Total module tests passed	Average test tries per module
Total module tests passed	R=.592** (N=151)		
Average test tries per module	R=-.663** (N=151)	R=-.338** (N=151)	
BMA test score	R=.522** (N=139)	R=.344** (N=142)	R=-.406 (N=142)

** Correlation is significant at the .001 level (2-tailed t-test)

The table shows that total module tests passed and average test tries per module are moderately related to physics final exam score (Pearson R correlation of .592 and -.663, respectively). BMA test score was also moderately related to physics final exam score (Pearson R correlation of .552).

Given the results of the Pearson R correlation calculations, it was hypothesized that module test-taking behavior, or average test tries per module, and math background, as measured by BMA score, would be better predictors of physics final exam grade. To examine this hypothesis, a linear regression was calculated with average test tries per module and BMA score were used as predictors of final exam score. The two predictor model was able to account for 49% of the variance (F (2,135) = 66.577, p<.000, R^2=0.49). MSP group predicted final exam score was calculated as:

MSP group: Predicted physics final exam score = 75.73 – (12.681*Average number test tries per module) + (0.883*BMA score)

The equation shows that physics exam score increases with each unit of BMA score, but is negatively affected by the average number of test tries per module.

Figure 3 presents this relationship for average module test tries of 1,2,3. The figure shows that as BMA score increases for students in the MSP group, final physics exam score increases. However, as average test tries per module increases from 1 to 3, the predicted physics final exam score decreases. The figure also shows that final exam scores for the MSP group are higher than

for LRE group when average module test tries equals 1. When average test tries equals 2, the MSP group final exam scores are higher up to a BMA score equal to 25. When average test tries rises to 3, the BMA score threshold is 10 for MSP and LRE final exam scores to be equal. Finally, the figure shows that the difference between MSP and LRE group performance is least where BMA score is highest.
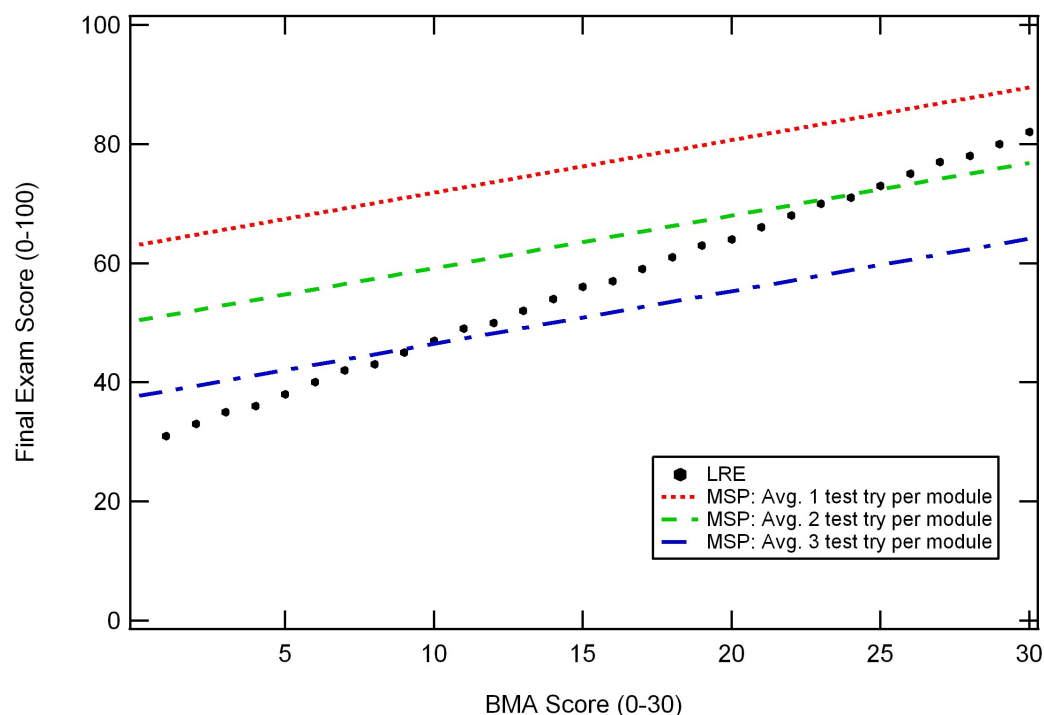

Figure 3. Predicted MSP Final exam score as function of BMA score and average number of test tries per module and predicted LRE final exam score as function of BMA score

VD. Online Pre- and Post Course Surveys

In the pre-course survey, all students were asked to report confidence in their math and physics abilities. The scaled questions used a Bandura-style confidence scale where 0% equals no confidence to complete a task and 100% equals very strong confidence to complete a task. One hundred and thirty of the 311 enrolled students responded to the online survey. Table 5 presents these results. The table shows that the majority of students were highly confident in their math and physics abilities.

Table 5. Pre-course survey: confidence in math and physics abilities at start of course (N=130)

Confidence	80% or higher	50-70%	40% or lower
Ability to apply math methods in problem solving	88%	10%	1%
Ability to apply physics concepts in problems solving	73%	20%	8%

The post-course survey was designed to compare MSP and LRE groups' use of course learning resources. Figures 4 and 5 present the survey results. Figure 4 shows that few MSP students used the

organized learning sessions including face-to-face lectures or online lectures. Rather students chose informal learning opportunities, including attendance in the MSP study room and help from friends, hat they could use when needed. A high percent of students also used textbook readings and problems to prepare for module tests. Finally MSP students completed many of the additional workshop problems on their own.
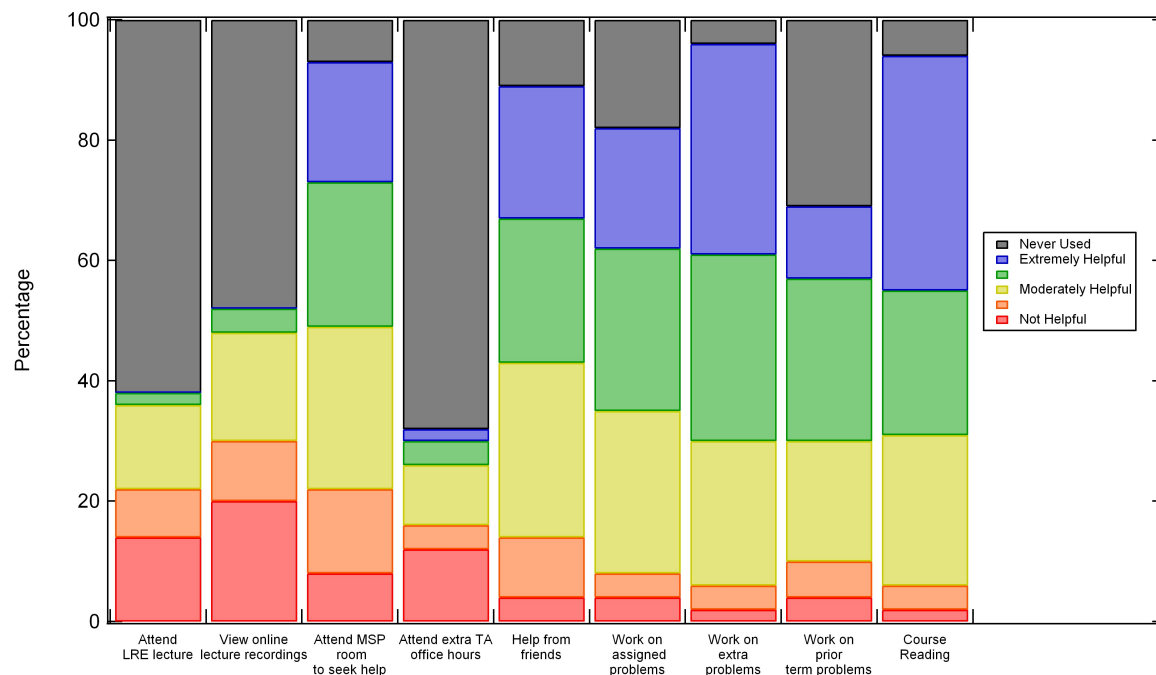

Figure 4. MSP Group: Course learning resource use by percent of students
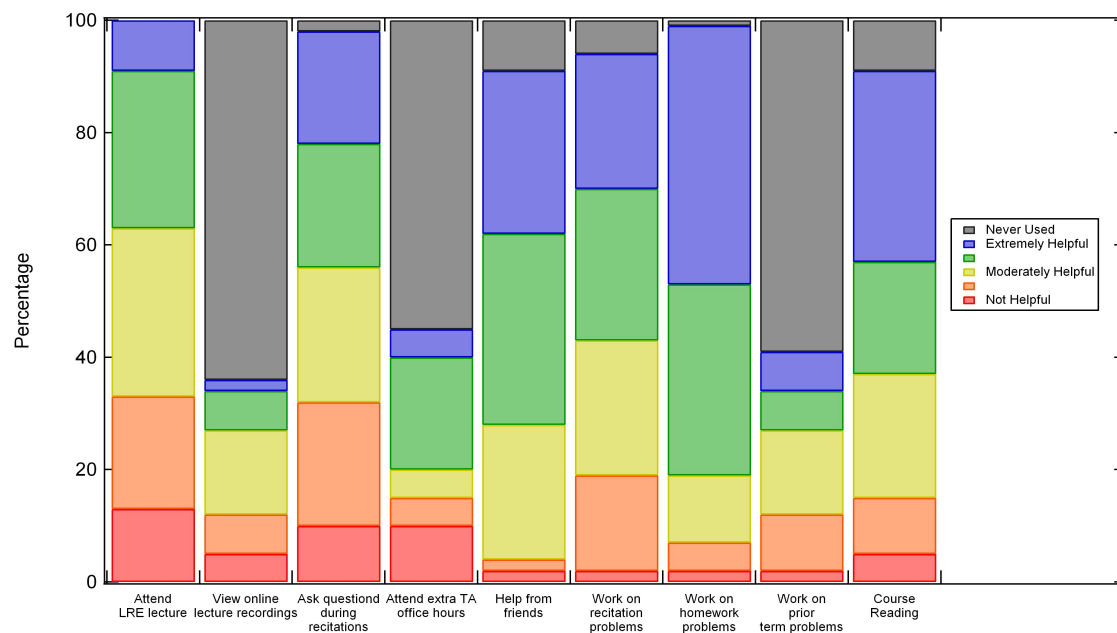

Figure 5. LRE Group: Course learning resource use by percent of students

LRE students depended more heavily on organized learning resources, including face-to-face lectures and recitations with teaching assistants, to learn course material. This group of students also felt that assigned homework problems helped them learn material. Like MSP students, LRE students used textbook readings to learn course material, but, unlike MSP students, few made use of the additional textbook problems or prior term homework problems.

LRE and MSP groups were asked about the testing processes in each course. The majority of both groups agreed or strongly agreed (70% or more) that the tests included physics and math methods and concepts that they were able to grasp, and that the test questions were fair. LRE students were asked whether test review sessions were helpful in preparing for the exams, however, only 20% agreed with this statement. MSP students did not have access to test review sessions. Rather they participated in test pre-screening oral exams to discern readiness for a module test; 67% agreed or strongly agreed that this method helped them prepare for module tests. In addition, MSP students participated in module test feedback sessions with teaching assistants as their tests were graded; 77% agreed or strongly agreed that this method helped them to learn physics concepts.

At the end of the course, students were asked to rate their sense of mastery of physics concepts. Using a scaled response, where 1= not mastered at all, 2=moderately and 3=strongly mastered, 61% of MSP group reported that they had strongly mastered material compared to 36% of LRE group (Table 6).

Table 6. MSP versus LRE groups: Self-assessment of physics concept mastery

Mastery of physics concepts	MSP Response Percent, N=52	LRE Response Percent, N=43
Not mastered	10%	16%
Moderately mastery	29%	48%
Strongly mastery	61%	36%

VE. Classroom Observations, Student Interviews, and Teaching Assistant Interviews

Classroom observations of the MSP workroom, a large room assigned solely for MSP students to work, found that the physics workroom promoted active, informal learning experiences through positive student-student and student-teaching assistant interactions. Several observation examples of this finding:

> On a given day, the physics workroom is filled with students and teaching assistants. Students sit alone or in groups and are involved in actively working on physics problems. Teaching assistants roam the room and join a student group to ask what they are working on. Students may use this prompt to ask a question or the teaching assistant may remain and observe students as they work on a problem. If students seem to be having difficulties progressing, the teaching assistant makes suggestions to move students forward.

In another section of the room, a teaching assistant is leading an impromptu mini-lecture on a physics topic related to one of the course modules. Several students sit and listen to the lecture, and ask questions when they do not understand. The group then moves back into the room to continue working on physics problems.

Several students are working alone on physics problems, however, they often look up and listen intently to discussions of physics problems around them. While working alone, they appear eager to learn from other students' approaches to problem solving.

In one corner, a teaching assistant is leading a student through her pre-module test screening process. In the prescreening, a student must respond to a series of semi-structured oral exam questions; a teaching assistant can probe student responses if he or she does not feel the student is explaining a concept or method with clarity. In one observation, a teaching assistant asked a student about a specific problem involving collisions of two objects. When the student was unable to explain how to approach the problem, the teaching assistant explained the concept and worked a sample problem. He then suggested several ideas and problems for the student to try before taking the module test. The student returned to the workroom to work on the suggested problems.

In interviews, MSP students often noted their initial frustration with the mastery learning approach. None of the enrolled students in the MSP group had ever taken a mastery course. Many students noted initial confused and feelings of regret concerning their course choice. After the first module exam, a simple test of unit calculations, most students felt energized to move forward through the first few equally straightforward modules on vectors and linear motion. Once students failed a module test, however, students' sense of frustration returned as they attempted to figure out what "mastery" meant. At that point, students noted several strategies for taking their next module test:

I just went back in and tried again without working on more physics problems. I failed again. I thought that maybe I'd just been given a more difficult (module) test, and maybe the next one would be easier. I failed again. At that point, the head TA (teaching assistant) talked to me about what I was doing wrong. We worked on a few problems and I realized that I wasn't really looking at what the problems were asking me. I needed to set up the problem (statement) more clearly. When I tried this on the next few (module) tests, I passed most of the (module) tests after one, sometimes two, tries.

I went to another teaching assistant with the notes I collected from the TA who reviewed my failed (module) test. The new TA worked on a few related problems with me and helped me figure out what I was doing wrong. I tried a few more problems before going back in to take another test. I passed the next time around.

A small percentage of MSP students, about 5% of those enrolled (based on percent of students with consistently high number test tries), did not change their "strategic" approach to test taking. Rather, these students appeared to maintain a "surface" approach to test taking. One student said:

I figured that if I took as many tests as possible that I would eventually see every kind of problem they could throw at me. It worked for me. I finished nearly all of the (module) tests.

This small percentage of students presented a frustration to teaching assistants. Some noted that the MSP course, while positive in terms of student learning, had its negative aspects. In particular, some noted that, in the early part of the semester, students tried to "game" the module test taking process. One TA noted:

Some students would not prepare for a (module) test, and just come in, look over the test, and leave. All the students want to know was what kind of problems would be on a (module) test, and thought this would help. It wasted my time and the students' time since I could have told (the students) that this strategy would not help in passing a module test.

To prevent students from trying quizzes repeatedly without trying to improve mastery, teaching assistants noted that students were strongly encouraged to actively work on all of the assigned workshop problems and ask questions to pass a module quiz the first time. Students were also encouraged to note the time necessary to take and retake quizzes in one module would create a sense of frustration rather than mastery. As a result, teaching assistants observed that most students moved smoothly from module study to quiz completion with only a few modules requiring multiple quizzes completions per module. If a student was unable to pass a module quiz after three tries, he or she was asked to meet with the head teaching assistant for special tutoring. The tutoring process was designed to target a student's key misunderstandings of concepts, problem solving and strategizing approach, or math.

The majority of MSP students realized that the mastery method was promoting mastery of physics compared to traditional courses. Several students noted:

I realized that I needed to work more problems than I used to do in a regular science course. The more problems I worked, the better I got at solving them. I have to spend more time on this course…., but I really have a solid hold on physics.

Students in the LRE lecture course revealed a greater prevalence of "surface" learning strategies. In interviews, students discussed their test taking strategies and study habits:

I only did the required work for the class. There was no reason to do more to do well in the class. That included the workshop problems and the homework. Then I'd just hope that the exam questions are like the workshop or homework problems.

I just have to get class average on the exams.

On one exam, the problems didn't look anything like the problems I'd done in the homework or workshops. I had no idea what to do. Class average was really low for that one.

VI. Discussion

The physics final exam results showed that MSP students outperformed students in the LRE course. Moreover, at-risk students in the MSP course also outperformed a similar group in the LRE course. The calculated effect size of 0.39σ is consistent with findings of other studies where student performance in MSP and LRE courses is compared (Kulik 1990). Figure 6 presents a graph prepared by Kulik of their meta-analysis of 67 of the mastery/traditional studies they reviewed (including 35 college science courses and 32 college social science courses). The graph shows the range of effect size values for the 67 studies, where the average value is 0.48σ.
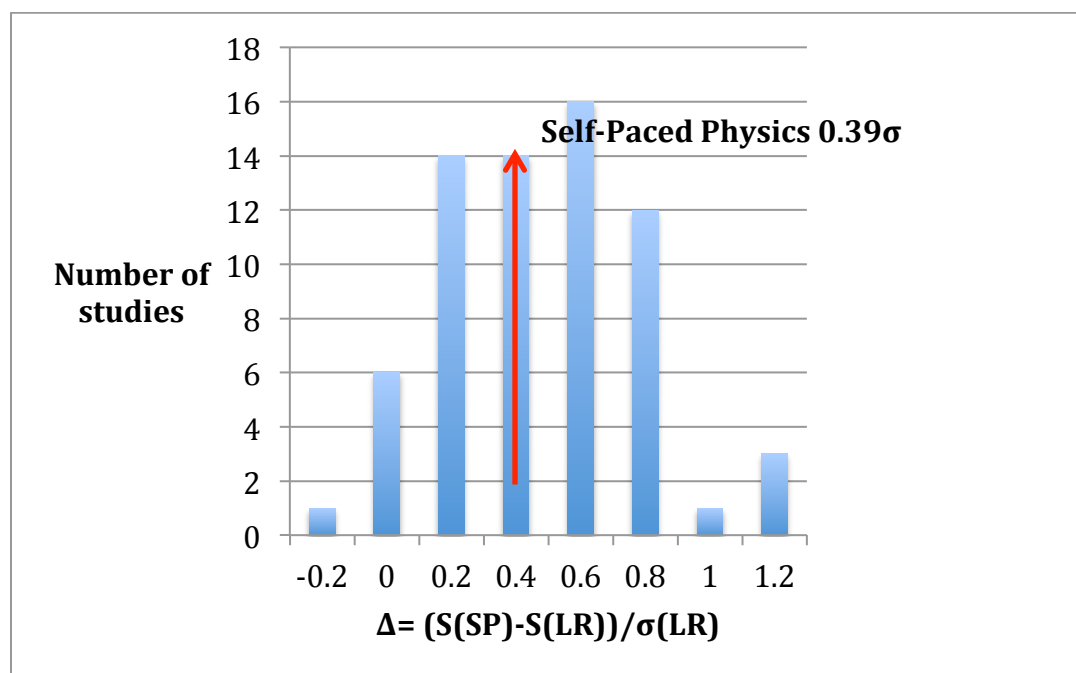

Figure 6. Kulik (1990) Meta-Analysis of Mastery vs. Lecture Course Effect Size and Comparison with Effect Size for this Physics Self-Paced/ Lecture Study

In this study, it was found that student study habits and test-taking strategy needed to incorporate both deep learning and strategic learning strategies in order to master the physics material and to demonstrate mastery on the module tests. If a student did not take the time to figure out what they did not understand, and tried to just take a module test over and over, they would fail over and over. Most students in the course quickly learned how to deeply learn. Students reported that they would work on assigned problems; when they did not understand a problem, they would seek out additional, non-assigned problems to try. They would also discuss their questions with TAs or peers until they felt ready to take a mastery test. However, when a particularly challenging module topic appeared, such as circular motion or oscillation or simple harmonic motion, students struggled anew. Again, however, students realized that deep and strategic learning approaches eventually led them to pass even these more difficult module tests.

This study also found that math background, particularly calculus course background and performance, was an important variable determining progress and performance in a challenging

math-based mechanics course. As student grades in a prior calculus course increased, performance on the physics final exam increased. A key finding was that the MSP course experience helped students with weaker math backgrounds learn physics and demonstrate mastery. In fact, MSP students outperformed students with similar math backgrounds from the LRE group. In fact, as Bloom predicted, students with the weakest backgrounds gained the most from the mastery approach.

It was also found that performance on the physics final exam was significantly leavened by students' study habits and test taking strategies. If students did not try to deeply understand module material before taking a module test in the MSP course, or after failing a module test, their average test tries per module increased which, in turn, led them into frustrating and unproductive repeats of module tests. Most students realized that readiness for module tests meant changing their study behavior in order to avoid unproductive test attempts.

Some students did not change their study or test taking behaviors, and continued throughout the term with consistently high number of test tries. More attempts meant more time testing rather than spending that time deeply learning the material. In fact, our model showed that physics final exam scores declined with increasing number of average module test tries.

This study has shown that it is possible to design a large enrollment course that utilizes a mastery learning approach. Staffing of the physics workroom with experienced physics students who were able to respond to student questions in an informal setting was an essential feature of the course. Providing individualized feedback as part of the module test prescreening and after each module test were essential features of the mastery approach. While teaching assistants and students alike were new to the seeming chaos of the mastery approach, a calmer process quickly emerged. Changing staff hours to meet the "crunch periods" around 4 p.m. to 6 p.m. that most students seemed to favor is an example. In addition, clarity of message to students that they must spend more time on working problems, and ask questions when they do not understand is another essential feature. In short, mastery learning does take more time than traditional course learning. While this aspect was a surprise to students, they also gained in confidence as they mastered the challenging physics material.

Student procrastination was another issue that arose during the semester. While most students jumped into the course and completed 3 or 4 of the 17 modules, a lull in the middle of the term occurred where the pace of module completion slowed. In the semester's last few weeks, students again picked up the pace of module completion. In future offerings, the faculty plan to design a somewhat stricter policy of module completion that does not sacrifice the primary goals of the mastery method.

Due to the intensive process of TA-student interactions needed for workshop Q&A, pre-module test oral exams, and individualized feedback on module tests, a higher number of teaching assistants was needed to teach the MSP course than the LRE course. In the next year, as the faculty gain more experience in managing such courses, it is expected that that some process efficiencies will be identified. However, while the MSP student:TA ratio may be lower than the LRE course, the benefits clearly outweigh the costs.

VII. Conclusion

This study confirms previous studies that found that the mastery learning approach, including Bloom's mastery model and Keller's personalized system of instruction (PSI) (Bloom 1973, 1976, 1980; Keller 1974, 1980), improves student performance, attitude and confidence in physics and other sciences (Diegelman-Parente 2011; Kulik et al. 1990; Leonard 2010; Zeilik 1981) and in engineering (Capaldi 2014; Sangelkar 2014; Wankat and Oreovicz, 1984, 1993, 2001). This study extends previous works by providing a full description of implementation in a large enrollment, one-term course, often considered a challenge for mastery learning use. This study extends previous works by establishing a stronger connection between mastery learning and deep learning models. This study recommends that the mastery approach, popular in the 1970s, should be revived given its positive impact on at-risk student performance.

Bibliography

Bandura, A. (1973). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*, 191-215.

Bandura, A. (1995). *Self-efficacy: The exercise of control*. New York: Freeman.

Bloom, B. (1976). *Human characteristics and school learning*. New York: McGraw Hill.

Bloom, B. (1980). The new direction in educational research: Alterable variables. *Phi Delta Kappan*, *62*, 382-385.

Bloom, B., T. Hastings and G. Madaus (1973). *Learning for mastery*. National Laboratory for Higher Education, 1973.

Capaldi, F. (2014). Mastery Learning in Statics Using the STEMSI Online Learning Environment. *Proceedings of 2014 Zone I Conference of the ASEE*, Feb. 14, 2014.

Diegelman-Parente, A. (2011). The Use of Mastery Learning with Competency-Based Grading in an Organic Chemistry Course. *Journal of College Science Teaching, 40(5),* 2011, 50-58.

Eyre, H. (2007). Keller's Personalized System of Instruction: Was it a Fleeting Fancy or is there a Revival on the Horizon? *The Behavior Analyst Today, 8(3),* 317-324.

Fuller, R. and D. Winch (2005). The Personalized System of Instruction (PSI), or Keller Plan Thirty Years Later. *American Association of Physics Teachers, AAPT Summer Meeting, 2005*.

Halliday, D., R. Resnick, and J. Walker (2013). *Fundamentals of Physics Extended, 10th Edition*. New York: Wiley, 2013.

Hestenes, D., M. Wells, and G. Swackhamer (1992). Force Concept Inventory. *The Physics Teacher. Vol. 30*, March 1992, 141-158.

Keller, F. (1981). PSI and educational reform. *Journal of College Science Teaching*, *11(1),* 37-38.

Keller, F., J.G. Sherman and C. Martuscelli Bori (1974). *PSI, the Keller Plan Handbook: Essays on a personalized system of instruction*. Menlo Park, Calif.: WA Benjamin, 1974.

Kulik, C. C., J.A. Kulik and R.L. Bangert-Drowns (1990). Effectiveness of mastery learning programs: A meta-analysis. *Review of Educational Research*, *60(2),* 265-299.

Lai, P. and J. Biggs (1994). Who Benefits from Mastery Learning? *Contemporary Educational Psychology*, *19(1)*, 13-23.

Leonard, W. J. (2010). Work in progress — Application of Mastery Learning in an introductory mathematical physics course. *ASEE/IEEE Frontiers in Engineering Education Annual Conference 2010*," Arlington, VA, Oct. 27-30, 2010.

Roth, Vicki, E. Goldstein and G. Marcus. *Peer-Led Team Learning: A Handbook for Team Leaders*. Upper Saddle River, NJ: Prentice Hall, 2001.

Sangelkar, S., O. Ashour, R. Warley and O. Onipede (2014). Mastery Learning in Engineering: A Case Study in Statics. *ASEE Annual Conference 2014*, Indianapolis, IN, June 15-18, 2014.

Wankat, P. and F. Oreovicz (1993). *Teaching Engineering*. McGraw-Hill, New York, 1993.

Wankat, P. and F. Oreovicz (1984). Teaching Prospective Faculty Members About Teaching: A Graduate Engineering Course. *Engr. Education*, *75(2),* Nov. 1984, pp. 85-85.

Wankat, P. and F. Oreovicz (2001). Mastery Learning. *ASEE Prism 11(3),* 2001, pp. 35.

Zeilik, M. (1981). Flexible, Mastery-Oriented Astrophysics. *American Journal of Physics, 49*, 827-829.