The "RIPL" Effect on Learning Gains in Lecture

Patricia E. Allen and John E. Cockman

Physics & Astronomy, Appalachian State University, Boone, NC 28608-2106

Abstract. The main goal of the Redesigned Introductory Physics Lab (RIPL) project at Appalachian State is to improve student performance and attitudes in the algebra-based sequence. Modifications of the student lab experience were examined in terms of their impact on performance in the lecture portion, independent of the lecture instructor's pedagogical approach. Preliminary results for one lecture section, based on Modeling Instruction, indicate a large positive difference in all course measures for students in the redesigned lab compared to those in the more traditional lab offered by the department. On the other hand, FCI and other diagnostic scores show little difference between the two groups. While these measures indicate a discrepancy in the redesigned lab impact, an item-by-item analysis of the diagnostics reveals a rich story, one that can be used to improve both lecture and lab activities. In this paper, we examine some of the factors that strongly affect student performance, as well as the implications for the redesign process.

Keywords: Physics education, undergraduate education, introductory laboratory.

INTRODUCTION

Introductory physics laboratories at Appalachian State are being redesigned to move from structured, step-by-step verification activities to guided inquiries culminating in open-ended experimentation. The goal of the Redesigned Introductory Physics Lab (RIPL) project is not only to improve the undergraduate lab experience, but just as importantly, to improve student conceptual gains in the lecture portion of the course. In this paper, we report on the preliminary results and implications for both lecture and lab instruction.

PROJECT DESIGN

The algebra-based sequence (PHY 1103-1104) was selected as the starting point for making changes to the lab curriculum for many reasons. Analysis of FCI[1] and in-house diagnostics have shown that the mathematical analysis and reasoning skills of these students lag behind their peers in the calculus-based course. Since most of the students are not physics majors, their interest and motivation tend more to survival than learning science. Departmental student evaluations indicate current dissatisfaction with the lab component of the course that could be reduced by engaging the students with activities that encourage students to explore phenomenon with minimal guidance.

Second, modifications made in the algebra-sequence can be easily tailored for the conceptual-based and calculus-based courses offered by the department. The mathematical and experimental complexities of an activity/experiment can be ramped up or down depending on learning goals of the course.

Another reason for selecting the algebra-based sequence is that our program traditionally has low student scores. The average pre-test FCI score is 29% while the post-test average is 39%, well below Hake[2] scores for traditional instruction. Additionally, none of the item responses were above the Newtonian entry threshold of 60% indicative of a student's ability to recognize the agents of force. These low scores motivated one of the authors (Allen) to adapt the Modeling Instruction[3] (MI) approach to the course lecture. In the MI "lecture" section, in-class worksheets, table-top experiments, and student presentations help students develop conceptual models that are then applied to appropriate situations.

While minor increases have been accomplished with MI, the conceptual gains have been undermined by the rigid, cookbook-oriented lab activities. For instance, buoyancy was explored using plastic water bottles. Students were soon able to analyze various buoyant force situations and problems correctly. After lab, however, students abandoned the broader class-developed model in favor of one "magic equation" that was verified in the lab activities, despite admonitions that this equation was valid only for one specific situation. Similar phenomena occurred for 1-D collisions, calorimetry, and circuits.

In response to these issues, as well as the need to modify the lab experience department-wide, the authors began to redesign the lab activities in Fall

2008. The full details of the redesign process are part of Cockman's (co-author) dissertation and will be reported elsewhere. The key elements include:

- Replacing the departmental lab manual with a series of QuickGuides (brief, online documents that provide an overview of a concept, equipment, or policy, and two to four activities to aid students in gaining practical experience);
- Dividing the curriculum into four to five themes, each related to major topics in the course; and
- Adding an end-of-the-semester practicum activity.

Redesigned curriculum impact was examined in two experimental sections, while the remaining control sections used the in-house lab manual. Since the students self-select the lab and lecture section times that best fit their schedules, there were different students in experimental sections in the Spring course (1104) than in the Fall (1103). Students with at least one semester of the redesigned labs were labeled the experimental lab (ELab) group while those with no experience with the redesigned labs were labeled the traditional lab (TLab) group. Of the 74 students who took both the pre- and post-tests, 33 students were in the MI lecture section while 41 students were in a traditional lecture (TL) section.

The impact of the RIPL approach on student attitudes and performance were examined in a variety of ways: pre- and post-testing using an in-house conceptual diagnostic, final exams, student grades, student evaluations and interviews, and CLASS (Colorado Learning Attitudes about Science Survey).[4]

EVALUATION METHODS

While multiple evaluation techniques were used to assess the effectiveness of the redesigned labs, only the content-based methods will be discussed in this paper. The department "Diagnostic Tool" (DT) is a series of 34 conceptual questions. Nearly two-thirds of DT are based on the FCI. Although DT is not as well validated as other instruments used in the PER and SER communities, Allen has studied the responses for some time. Student responses have been used to formulate a plan for curriculum improvements and to establish a mechanism for timely feedback on the effectiveness of instruction. The database of responses has been well correlated with more "standard" forms of assessment (full FCI, FCME[5], etc.) As such, the subset of FCI questions can be compared with results from times when the full FCI has been administered. Diagnostic post-test scores were compared with pre-test scores for the MI and TL sections.

Based on the results from the diagnostic tests, twenty-nine multiple-choice questions from the MI Fall and Spring final exams were also examined. Responses from the 2008-09 academic year were compared with the database of responses from previous years. Finally, student performance in various graded components of the MI section were analyzed, including averages for class work, three exams, the final exam, laboratory, and overall course grade. Although these are not standardized forms of assessment, they provide invaluable insights into the transferability of conceptual-based and performance-based skills from the laboratory into the classroom, the main objective of this study.

RESULTS AND DISCUSSION

Student Grades

At the end of the Fall 2008 course (1103), student grades were examined for all lecture sections (three) as a quick test of the impact of the lab experience on student performance in the course. In the MI section, there was an 8-15 point (out of 100) improvement in all measurable categories for the Elab students. (Lab scores were normalized for the course.) The difference in grades provided encouragement to continue the RIPL project and to search for factors contributing to this difference. Course grades were then examined after the Spring 2009 course (1104) with similar results, as shown in Table 1 for the MI section. Fall and Spring grades for students within each TL section showed no statistical differences between the experimental and control lab groups.

TABLE 1. Modeling Instruction (MI) student grade breakdown by lab experience for PHY1104, Spring 2009

Grade Category	MI: ELab	MI: TLab
Classwork (46%)	83	73
Exams (27%)	81	69
Lab (15%)	88	87
Final Exam (12%)	68	59
Final Grade (100%)	81	72

These results are surprising for several reasons. Although the Elab students had similar diagnostic pre-test scores (Table 3 below) and were academically similar to the traditional group, it is apparent that the combination of MI and ELab resulted in higher scores in the MI section. Table 2 highlights the breakdown of factors affecting the students' grades.

Students who practice physics regularly tend to perform better on all exams and other course measures.[6] Several factors indicate that ELab students in the MI section performed more assigned and elective work than their TLab counterparts. Elab students attended more classes, submitted more assignments, and were more likely to complete extra credit problems and assignments.

ELab students also spent more time, on average, in the lab room than the TLab students to complete assigned tasks. While they covered the same content material, they had fewer full lab reports due over the course of the semester (6 versus 12). Even with fewer assignments, CLASS and other survey results[7] indicate that ELab students felt that their work was more relevant to the course than did the TLab students. This may have contributed to ELab student "buy in" into the course that was then built upon in the MI section since the work environments were similar.

TABLE 2. MI student classwork performance by lab experience for PHY1104, Spring 2009.

Grade Sub-Category	MI: ELab	MI: TLab
Homework	89	70
Attendance	94	88
In-Class Quizzes	86	76
Online Quizzes	87	76
Overall Assignments	94	85

Our goal was to examine how changes in the laboratory affect student learning in the lecture. Grade improvement is suggestive that the inquiry-based, hands-on lab approach is having a positive impact on student lecture performance. Future research will examine how the combination of ELab activities with inquiry-based lecture instruction, like MI, achieves this goal. Although grades have improved, an examination of the diagnostic scores reveals a different, more interesting story.

Diagnostic Tests

Pre- and post-test results for the Diagnostic Tool (DT) and the subset of FCI questions for the various sub-groups (MI, TL, ELab, or TLab) are shown in Table 3. The pre-test scores for all sub-groups were similar to each other and to previous groups of students. Hake scores for the various groups averaged to +0.13 for the DT and +0.10 for the FCI. This is consistent with historical program data in the traditional lecture, but lower than what is usually achieved in the MI section. While these values are surprisingly low (especially when student grades are considered), a detailed analysis of each diagnostic question reveals some of the subtleties of what the lab experience contributed to the lecture and vice versa.

Item Analysis of Diagnostic Tests

There are examples that demonstrate an increase in student understanding simply because it was supported by inquiry-based lab experiments. However, select questions were chosen for this paper to highlight some interesting features that arose when each item was examined instead of focusing solely on the Hake scores. For instance, two FCI questions that deal with the path an object takes after an impulse ('kick/hit") or a constant force (thrust) show some differences between the experimental and traditional students. In both cases, the TLab students outperformed the ELab students, with the TL students scoring higher than the MI students. By examining how the two lecture sections differ, the explanation becomes clear.

During lab, the ELab students studied a collision between two KickDis® pucks on the floor of the lab using video analysis of their collision while the TLab students received carefully selected photocopies of spark tracks of a puck collision on an air table. For the ELab group, the pucks were on an uneven floor and tended to rotate resulting in curved paths for many of the student collisions. When this activity was repeated in the MI section, similar results were observed. As such, the ELab students tended to select the reality-based responses of curved paths for all path-related diagnostic questions instead of including the role of zero or constant net forces.

This phenomenon was also observed with the DT question in which one in a series of three light bulbs is short-circuited. During ELab, many students incorrectly built circuits involving batteries and light bulbs, while the TLab students conducted an unrelated activity in electricity. This error was discovered only when additional circuits were built during MI class meetings. The groups that improved their circuit-building skills increased their diagnostic scores. Those who struggled with building circuits responded according to their observations of an incorrect circuit, instead of considering the underlying principle of

TABLE 3. Scores for the Diagnostic Tool and subset of FCI questions broken down by lecture and lab instructional approach.

	Diagnostic Tool (DT)		FCI Questions from DT	
	Pre-Test (%)	**Post-Test (%)**	**Pre-Test (%)**	**Post-Test (%)**
All Students (historical program data)	34	45	29	39
All 2008-09 students - matched (74)	33	42	28	34
All MI section (33)	31	44	26	35
ELab (19)	32	45	26	34
TLab (14)	30	44	25	36
All TL section (41)	35	40	29	34
ELab (12)	33	40	25	32
TLab (29)	35	40	31	35

resistance on current. While becoming more observant is an important learning goal in science, our focus now is to link observations with appropriate concepts.

In the previous example, MI activities helped to minimize misunderstandings generated during lab. However, MI activities can introduce new issues. For instance, the FCI question dealing with what happens to the velocity of a box when a pushing force is removed led to a surprising result. Significant gains were obtained by the ELab students in the TL section while students in the MI section indicated the box would *immediately* stop instead of *slowing* to a stop. While the ELab helped students observe the motion of objects, MI students also examined the effect of pulling blocks at constant velocity to measure coefficients of friction. From the student's perspective once the pulling force was removed, the blocks stopped over a very short distance that could be interpreted as "immediately" stopping.

The previous examples provide some insights into the effect of lab activities when combined with in-class activities on student understanding. Some of the largest Hake scores, both positive and negative, occurred as a consequence of the combination of ELab and MI. The various combinations of lab and lecture activities are currently being explored to determine which factors most positively affect Hake scores.

In addition to the above issues, some situations still require a full explanation. During the rotational equilibrium lab, both ELab and TLab students performed similar activities. The sole difference between the two groups was that ELab students generated the equation to solve for unknowns (lever arm, balancing mass, etc.) while TLab students applied a given, generic equation to different situations. Both ELab groups (MI and TL) showed significant improvement in the diagnostic score for a rotation question in comparison to the TLab groups, this in spite of the fact that the topic of rotation was not covered in the MI section due to time constraints.

Similar improvement occurred with a question regarding the cause for rising smoke. Both ELab student groups outperformed TLab groups, even when the topic was covered in lecture, but not in lab. These examples appear to demonstrate an improvement of cognitive skills that helped ELab students to reason out the answer. When an item-by-item analysis is repeated for the multiple-choice questions on the final exams, similar results support observations made with the diagnostics. The full results of this analysis will be presented in a future paper.

Implications for Future Work

Findings from the DT may now be used to plan future lab activities effectively and provide some suggestions for lecture instructors. Improvements in the collision, force, and circuits labs are now in place for implementation in all lab sections beginning Fall 2009. Workshops in non-lecture-based activities are now scheduled for all instructors starting this Fall. In addition, the RIPL curriculum will be tested with two of six lab sections in the calculus-based sequence. Some of the issues raised from the previous item-by-item analysis will be re-examined for this audience.

CONCLUSION

The RIPL project has demonstrated that the "ripple" effect goes both ways. Significant changes in student learning, both positive and negative, can be affected by the combination of lab and lecture activities. To be most effective, both lecture and lab instructors need to be aware of the impact of lab on lecture learning and vice versa. It appears that our redesign of the laboratory portion of the course improved student understanding in some areas while impeding understanding in others. This resulted in low Hake scores for both FCI and DT. For diagnostic scores to improve, greater care must be made to ensure that some of the shortcomings of real-world situations in the laboratory are identified and accounted for. However, ELab students did give a greater effort and achieve higher grades in the MI section, and CLASS and other survey results indicate that their attitudes about science exceeded their TLab counterparts.

REFERENCES

1. D. Hestenes, M. Wells, and G. Swackhamer, *Physics Teacher* **30**, 141 (1992)
2. R. R. Hake, *Am. J. Phys.* **66**, 64 (1998).
3. D. Hestenes, *Am. J. Phys.* **55**, 440 (1987)
4. W. K. Adams et al., *Phys. Rev. ST Phys. Educ. Res.***2**, 010101 (2006).
5. R. Susan, "The Force and Motion Concept Evaluation," Paper presented at the Annual Meeting of the Mid-Western Educational Research Association (Columbus, Ohio, October 16-19, 2002).
6. G. Gibbs and C. Simpson, "Conditions Under Which Assessment Supports Student Learning," Learning and Teaching in Higher Education, V. 1, pp. 3-31, (2004)
7. John Cockman, PhD dissertation, Appalachian State University, Boon, NC, to be published Fall of 2009.