Interim Report for the RealTime Physics and Interactive Lecture Demonstration Dissemination Project

Michael C. Wittmann, Chris van Breen
Department of Physics
University of Maryland
College Park MD 20742-4111
wittmann@physics.umd.edu
tel: 301 – 405 – 6184

I. Introduction

The evaluation of the Real Time Physics/Interactive Lecture Demonstrations (RTP/ILD) dissemination project looks at how the RTP and ILD curricula have been implemented at a variety of schools, ranging from community colleges to research universities.[1] To measure the manner in which research-based curriculum materials provide effective instruction to students whose physics understanding is shown to be sorely lacking when they leave the traditional classroom, we focus primarily on student learning. For a curriculum to be considered effective and relevant to the greater community, it must first prove to be effective and relevant for the students who occupy our classrooms. Though the RTP/ILD curricula has shown great promise at its primary, developmental site,[2,3] the open question remains whether the secondary implementation can be as effective and whether student learning shows gains as striking as at the primary site.

In this report, we show results that indicate that the RTP/ILD curriculum materials are effective at helping students gain a deeper understanding of the conceptual issues in learning introductory physics. Student learning gains are large in comparison to traditionally taught schools at the same university, though not as large as at the primary implementation sites. While conceptual learning is greatly improved, student attitudes and expectations toward the course are commensurate with those in traditionally taught classes. These results are to be expected due to the nature of the curriculum redesign that is asked for when using RTP/ILD materials.

In the sections below, we describe the data in more detail. We briefly summarize how the data were collected. More detailed descriptions of this process have been given to FIPSE previously (Nov. 99 report, available upon request). We then describe the data from the different test sites, with special emphasis on those schools and classes that do not match the general trends observed in the overall data. Finally, we describe which elements of the dissemination need further work, and how these additional elements will be incorporated into a final evaluation report.

II. Data collection

At this stage of evaluation, the primary emphasis lies on data gathered through the use of five conceptual tests and one attitudinal survey. These data are interpreted based on instructor comments supplied during the semester and in outside conversations.

The primary goal of the conceptual tests is to measure student performance in the classroom. It is the assumption of this evaluation project that the primary measure of effective dissemination is that students have a better understanding of and attitude toward physics when they leave the course than a traditionally taught course. A secondary assumption is that instructors will feel motivated and able to use these materials. This consideration is secondary, because a course that does not lead to improved learning should not be promoted to instructors, regardless of their enthusiasm for it.

The conceptual tests measure student learning in the following areas:

- Force and motion
- Electric circuits
- Heat and temperature
- Light and optics

Two of the conceptual tests deal with force and motion. To give an example of student learning as done in the evaluation project, this report will describe results on conceptual learning from the FMCE (Force and Motion Conceptual Exam) and the ECCE (Electric Circuits Conceptual Exam). Previous research[2,3] has shown that curricula based on physics education research can help students to a higher gain in scores when compared to traditional instruction.[4]

We evaluate student attitudes toward the course with the Maryland Physics Expectations Survey (MPEX). The study of expectations allows insight into students' attitudes about the hidden curriculum, the implicit expert description of the community of physics. We use the term expectations to describe the attitudes, beliefs about the classroom, and epistemologies that students bring to the classroom. Research has shown that students approach the classroom with a set of expectations that are unfavorable to properly learning physics in a qualitative, coherent manner that is linked to the real world around them.[5,6] Most of the data gathered for the MPEX have shown a decrease in student expectations due to instruction. These results have been found in the case of both traditional and modified curricula. Only students in modified curricula that drastically change the course format away from lecture/recitation/laboratory (for example, into a lab only setting, as in Workshop Physics) have consistently shown improvements in their MPEX scores.

Data have been gathered at the test sites using two methods. The first is the administration of pre- and post-instruction conceptual tests either in lecture or laboratory. The MPEX has also been asked in this format, but a second method has also been used for the MPEX. It has been asked on-line, using the Harvard Galileo site at Harvard University. Students are required to log in from a world wide web site created by the class instructor, and results are mailed from the Galileo team to the instructors and the evaluators. In all cases, only data that have been matched (i.e. students took both the pre- and post-test) are included in the evaluations.

III. Data Analysis

In this section, we will describe the conceptual test data and the expectations test data from the first two years of implementation. These data will be complemented by third year which we are in the process of gathering. It is expected that the third year data will be similar to (or improved from) the second year data as the instructors involved in the project become more familiar with the materials and are able to teach with them more effectively.

Conceptual Test Data: FMCE

Several consistent results have been found in the evaluation of student understanding of force and motion. These are:

- Students in RTP/ILD classes consistently perform better than students in traditional instruction classrooms with gains from 0.3 to 0.55 as compared to gains of 0.1 to 0.3,

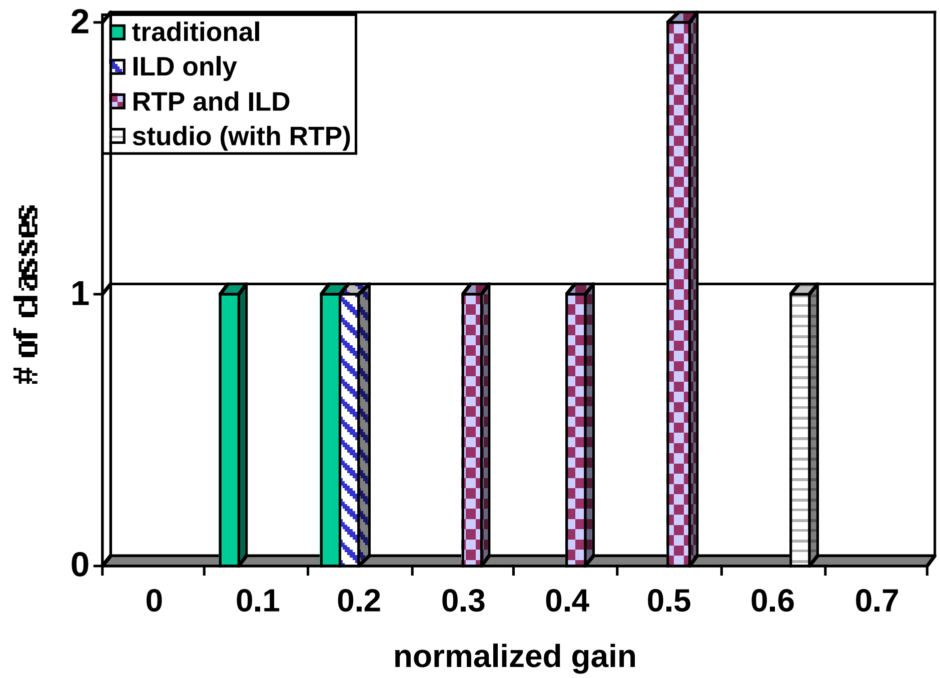

Figure 1: FMCE data from Year 1 of RTP/ILD implementation. Note that these are overall data using the full 47 question test in the analysis.

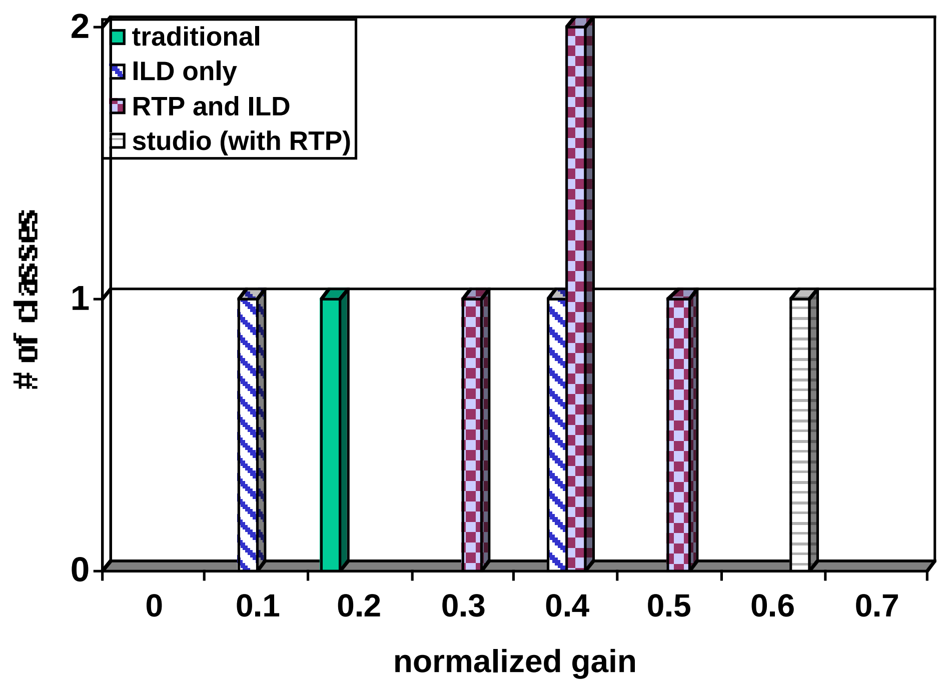

Figure 2: FMCE data from Year 2 of RTP/ILD implementation. Note that these are overall data using the modified analysis rubric which uses only 35 questions in the overall score.

- Students in ILD-only classes can perform as well as RTP-only and RTP/ILD classes, but these classes are far more (negatively) sensitive to implementation issues,
- Classes taught in the studio format (but using RTP materials) score higher than any other form of instruction on conceptual tests.

The charts and graphs shown above present the data in summary form. Figure 1 gives the results from year 1 of implementation, while Figure 2 gives the results from year 2 of implementation. Though the results are very similar, one important point should be noted. The year 2 data are analyzed using a rubric provided by Ron Thornton for proper analysis of the full 47 question FMCE test. The rubric calls for only 35 of the questions to be analyzed. Of the others, 4 are on energy topics that are not part of an overall understanding of force and motion, and the rest are "priming" questions designed to get students prone to thinking a certain way to actually think that way on subsequent questions. In addition, some questions are analyzed in clusters, meaning that students only get full credit (weighted to two questions) on a topic when they answer three questions of a set correctly. The data are very similar in each situation, showing that overall little changes, but the more careful and appropriate analysis has been used since year 1 of the evaluation.[7]

Of note in the comparison of year 1 and year 2 data are the ILD-only classes. The year 1 ILD data came from a course that was taught in a three week winter term. We would not expect students to achieve any major gains in such a short course. In addition, effects of "test memory," meaning improvement in scores due to recall of questions from the previous time the test was taken, should play a role here. Still, it is encouraging to see that the ILD results in such a setting can be at the top end of the traditional spectrum. More important to the evaluation project are the data from year 2, where two ILD-only classes at the same institution reported very different scores. The instructor who carried out the ILDs indicated that the class with low scores had several difficulties, including low class attendance, very bad attitudes toward the materials, and problematic implementation of the materials. In the semester with high scores, special attention was paid to attendance and implementation of the materials in the classroom. This comparison indicates the importance of proper implementation of specifically the ILD materials, which only rarely enter into the classroom and are not a part of the weekly classroom design.

To illustrate that comparisons across schools while using different curricula are at the very least plausible, see Figure 3. The results in this figure are taken from one school, where three different forms of instruction (traditional lecture and lab, traditional lecture with RTP lab, and studio physics) were used in a single semester. The students entering the three courses scored roughly evenly on the pre-instruction FMCE test (with scores between 26% and 33%). The great different in their scores comes on the post-instruction results. Here, we find that students in the studio class had a gain of 0.65, while students in the traditional only class had a gain of only 0.21. The class that used RTP labs in addition to the otherwise traditional lectures scored in between, with a gain of 0.42. These results are roughly consistent with what was found in other studio class settings.[8]

This test site is now able to make more informed choices about its instructional methods through a cost/benefit analysis of the two more effective instructional methods. With a studio class, they need more computing equipment, a redesigned laboratory space, and other large capital costs. With the RTP class, they keep the lecture format, while still having conceptual learning gains far above that of the traditional class. As part of the evaluation efforts, such results can have an effect on individual schools, in addition to helping evaluate the overall effectiveness of the materials. Note, also, that the studio class used RTP materials in its classes.

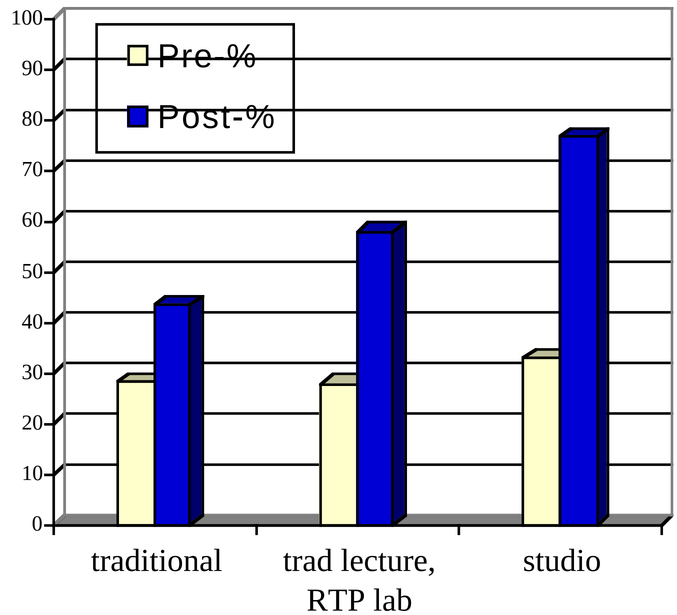

Figure 3: Comparison of FMCE data at one test site. Data are from three classes taught in the same term (gains are 0.21, 0.42, and 0.65 in each case).

Conceptual Test Data: ECCE

The data from the Electric Circuits Conceptual Exam are less clear than the FMCE data in terms of the a clear distinction between traditional instruction and RTP lab instruction. For an overview of the data, see Figure 4 and Figure 5. Note that fewer schools teach electric circuits and give their students the ECCE, so that the conclusions that can be drawn are based on fewer data than in the case of student understanding of force and motion as measured by the FMCE. Several additional elements of the evaluation are worth noting.

First, the results from one test site which used both RTP and ILD materials show that making a clear effort to help students in many ways (both in lab and in lecture, as modified with ILDs) can have a strong effect on student learning. The test site has reported that they put great effort into making these materials helpful for the students. One should note, though, that the ECCE and the RTP/ILD materials are often very close in design and style, such that students engaged in more than only RTP activities are more likely to perform well on the test. The very high gains (0.6 and 0.68 in years 1 and 2,. respectively) indicate that additional factors should be at work, though.

Second, two studio physics classes were also evaluated. One of these is not an official part of the RTP/ILD dissemination project, but data from the class were supplied by the test site as a comparison for the other classes being taught at that school. This class was taught using Workshop Physics materials which are closely related to the RTP materials, while the other studio class used materials modified from the original RTP labs. The data indicate that students in the studio classes perform on average better in the studio classes than in the RTP lab classes.

Third, the low-scoring RTP classes in each year are from the same school. This school had serious implementation issues, with many classes taught in many different fashions. The level of implementation in each of these classes was difficult to determine, some of the materials

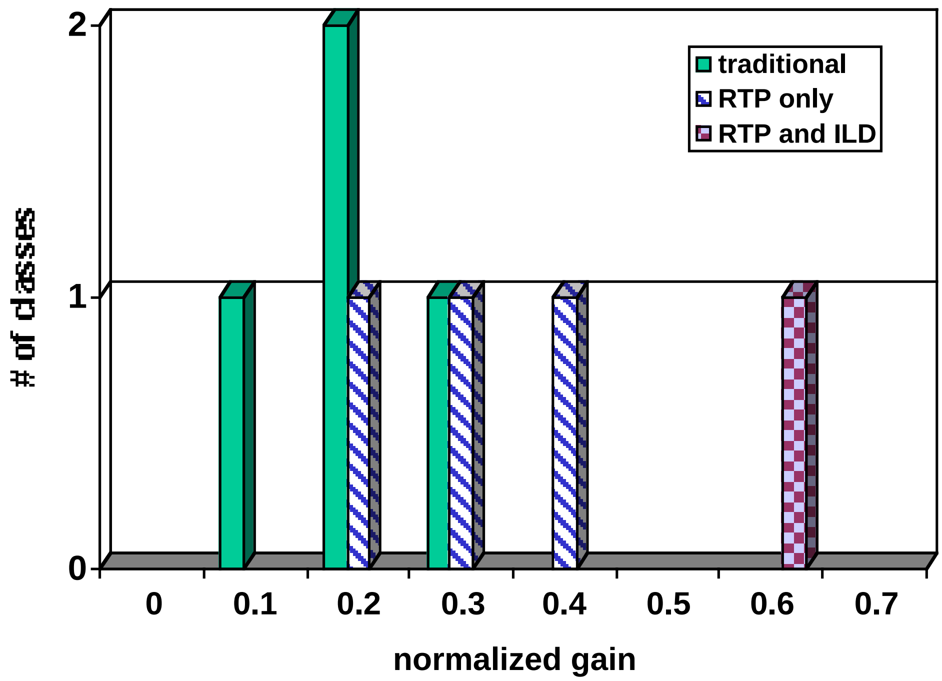

Figure 4: ECCE data from Year 1 of RTP/ILD implementation. Note that in year 1, no classes used solely ILDs on this topic.

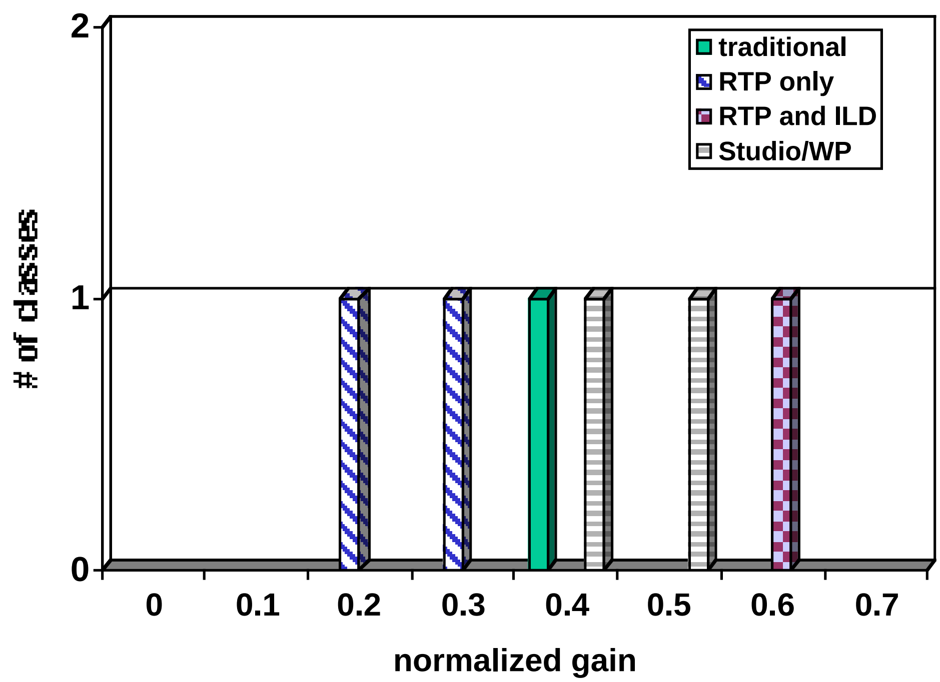

Figure 5: ECCE data from Year 2 of RTP/ILD implementation. Note that in year 2, no classes used solely ILDs on this topic, but two classes were taught in a studio format.

were rewritten at the test site, and in some cases it is possible that students were only doing a few of the RTP labs as originally designed. Thus, we view this school as an example of the difficulties caused by incomplete implementation of the materials. (Note, also, that this test site had the lowest scoring RTP/ILD classes on the FMCE for similar reasons, though there the results were better than on the ECCE.)

Finally, the traditional instruction classes in both year 1 and year 2 fall into a broad range of scores. Specifically, the year 2 traditionally taught course (with a gain of 0.40) scored much higher than expected. The test site has not indicated how this might have come about, and we are seeking further information to ascertain how this class scored so highly when other traditionally taught classes usually have gains between 0.1 and 0.3.

Overall, we find that the ECCE data indicate success in the RTP labs (especially when using ILDs in addition). The data are less compelling than in the case of FMCE data, though. We believe that the ECCE data indicate sensitivity to implementation in the same fashion that the FMCE indicated sensitivity to implementation issues in the ILD-only schools.

Expectations Survey

As with the FMCE and ECCE data, the MPEX data are analyzed using only students who answered both pre- and post-instruction questions. The responses on the MPEX are analyzed according to whether they are favorable or unfavorable (as compared to the responses given by experts in physics instruction). The percent of student favorable or unfavorable responses before and after instruction are compared to see the effect of instruction on student expectations. We had expected that students in the RTP/ILD project would show results similar to those found in traditional instruction classes because the general format of the class (lecture/lab/recitation) was not conducive to students rethinking their general approach to learning physics. This expectation is clearly seen in the data, where we see overall degradation in student scores over the course of instruction.

When presenting the data from the MPEX, we have left out the year 1 results. These will be included in the final report, but are not included here. There are several reasons for this. First, the year 1 data used a shortened set of MPEX questions (designed in order to save time in test-taking at the test sites). Comparing year 1 and year 2 data is therefore problematic, since the year 2 data contain more information. Second, the year 1 data are incomplete, with many schools having test-implementation issues. At some schools, the number of students taking both pre- and post-tests (or answering enough questions to be considered a fair representation of their expectations in the course) was so low that the data are very difficult to interpret. These shortcomings will most likely be addressed by the time of the final report, but at the moment, we focus on year 2 data. These give a clear representation of student expectations in and attitudes toward the course, allowing discussion of the RTP/ILD dissemination project.

In the data shown below, two representations are used. In Figure 6 and Figure 7, two charts are given. Chart (a) in each case shows the actual average student scores, indicating both pre- and post-instruction scores. Pre-instruction scores are indicated by open icons, post-instruction scores by filled icons. Chart (b) shows the movement between the pre- and post-instruction scores. All pre-instruction scores are placed at the origin, in effect normalizing all pre-instruction scores to a single value. A vector has been drawn on the figure to indicate the direction of movement in each class. Note that the classes are not labeled by name but rather by type of instruction in each situation. Icons in chart (b) match those in chart (a). In order to save space, only those quadrants which show data points are shown in Figure 6(b).

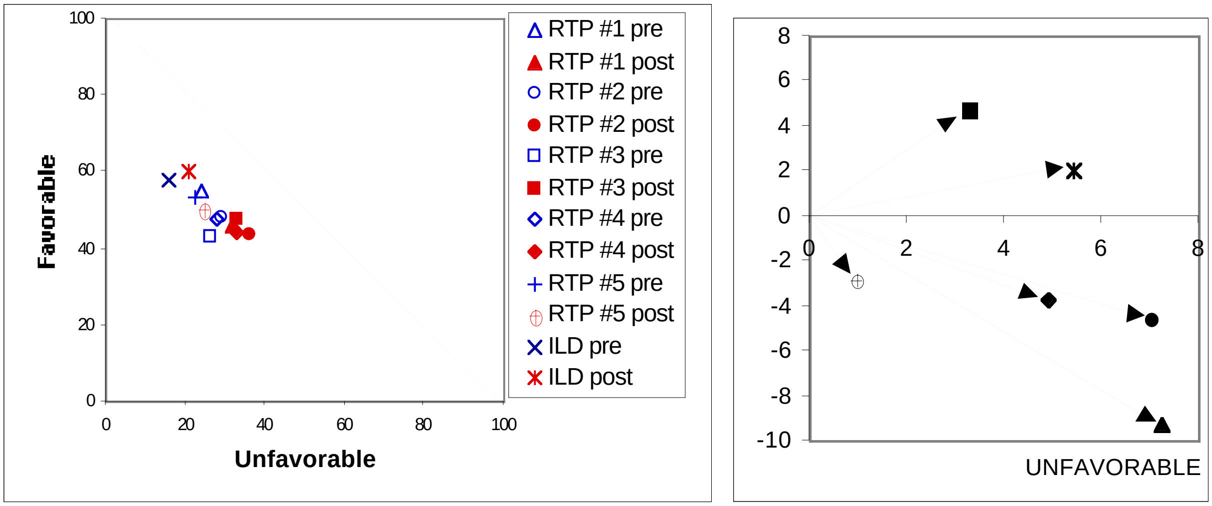

Figure 6: MPEX data from Year 2 of RTP/ILD implementation, mechanics instruction. a) Pre- and post-instruction scores from XX classes teaching first semester physics, including force and motion. b) The change in scores of classes from their pre-instruction scores. The origin is taken as the pre-instruction score for each school (i.e. normalized across schools).

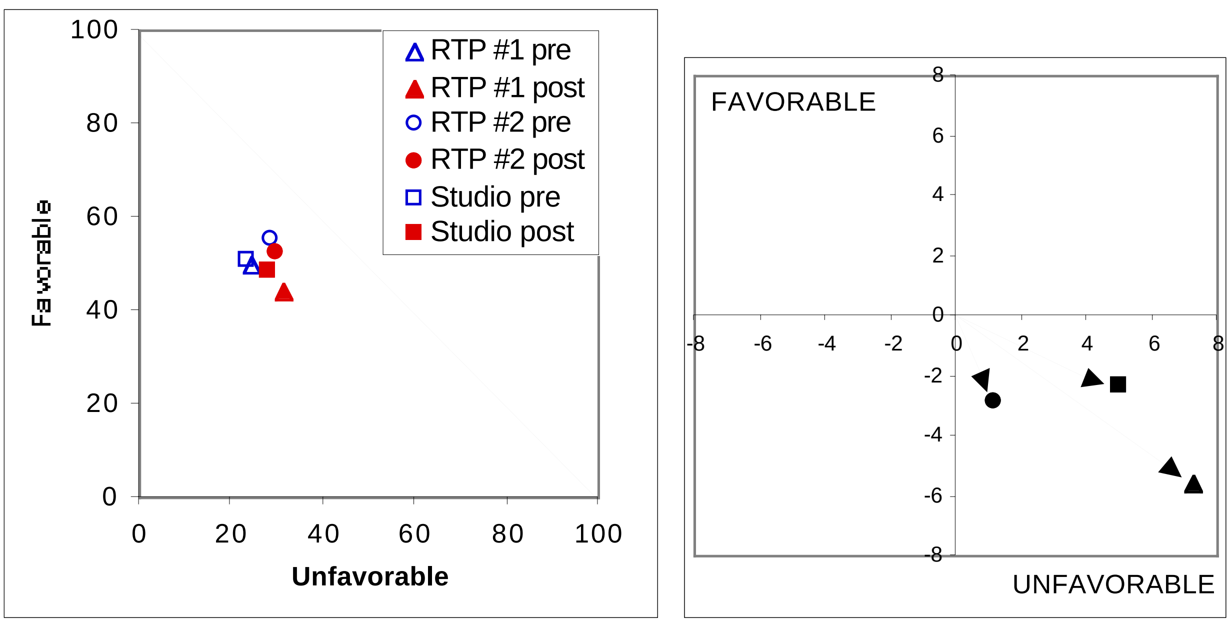

Figure 7: MPEX data from Year 2 of RTP/ILD implementation, electric circuits and heat and temperature instruction. a) Pre- and post-instruction scores from 3 classes teaching second semester physics, including electric circuits and heat and temperature. b) The change in scores of classes from their pre-instruction scores. The origin is taken as the pre-instruction score for each school (i.e. normalized across schools).

Two major trends are obvious when discussing the MPEX data. On the one hand, students begin instruction with expectations that are only slightly favorable. Thought the scores are spread out over a large range, students are generally between 40 and 60% favorable and 20 to 40% unfavorable as they enter the course. This obviously creates a problem for instructors wishing to help their students develop a deeper understanding of the physics. Weak expectations when entering the course can lead to bad study habits and an incomplete picture of what is meant by physics.

On the other hand, the movement of the classes is in all but two cases (discussed below) in the unfavorable direction. Students end the semester with less favorable scores than when they entered the class. This leads to two different conclusions. First, students are being lead to believe or follow unfavorable study habits, expectations of learning, and attitudes toward the course. This is a negative result, and one that is highly problematic. It is consistent with results from other lecture oriented classes, indicating that the RTP/ILD materials lead to improved conceptual learning (as shown above), but do not lead to improved expectations in the course. Second, students are able to gain a conceptual understanding of the material while having increasingly unfavorable expectations of the material. This indicates that students may not be aware of their learning while they are learning. This question of productive student epistemologies toward learning lies outside the scope of this paper and evaluation report, but is an on-going research question that needs to be addressed in more detail.

The two MPEX results that show neither favorable nor unfavorable motion (see Figure 6) have anomalies in the test-taking methods and should not necessarily be interpreted as neutral results. In both cases, the MPEX was given as a take-home test rather than an online test. Previous research[9] has shown that the MPEX is extremely sensitive to the setting in which it is given. In one class, both the pre- and post-test were given in take-home format, but very few students completed both. This indicates that possibly only those motivated to take the test and turn it in are included in the data, and this level of motivation may play a large role in the final result. Similarly, in the other class, the post-test was given as a take-home (the pre-test had been given online, but with serious problems in implementation), and possibly only the highly motivated students returned the tests. As a result, we include the data, but give little weight to the results.

IV. Conclusion

As has been shown by Hake[10] and Saul,[9] interactive engagement classrooms provide a better setting for student learning than traditional instruction classrooms do. But, as Karen Cummings et al.[11] have pointed out, the interactive engagement activities must be implemented correctly, otherwise large expense leads to little result. The results in this report indicate that, for the most part, the RTP/ILD results fit into the picture provided by previous research. Students using these materials in a properly implemented fashion leave the course with improved understanding of the concepts of force, motion, and electric circuits. We have preliminary results that indicate the same for student understanding and learning of heat and temperature. These results give strong evidence that the RTP/ILD materials are effective and of great value to the real consumers in the classroom, namely the students.

As was predicted at the outset of this project, the MPEX results on student expectations in the course are less satisfying. The data show that students leave the modified RTP/ILD classes with roughly the same attitudinal shift as students in the traditional lecture/lab classes that have been studied previously. It seems that the format of instruction, with a heavy reliance on choices made by the instructor and questions heavily guided by the instructor, is the dominating problem in is situation. Only those classes with very high teacher involvement and an overall departmental atmosphere of student interaction show no loss in the expectations score (as measured by a degradation of scores toward the more unfavorable).

What remains to be done in the evaluation report is to combine these conceptual and expectations test data with the reports from the individual instructors. Additional surveys of the instructors are needed, together with a more detailed view of what occurred in some of the classrooms which are reporting results outside the average. By understanding both the positive and negative elements that help or hinder appropriate implementation, a more complete picture of the curriculum can be made for future adopters of the materials. It is worth emphasizing, though, that the average results for the modified RTP/ILD materials are better than the average for traditional instruction, supporting the idea that instructors can stray from the norm in the RTP/ILD setting and still see improved learning in their students.

In summary, the evaluation of the RTP/ILD dissemination project shows that the goal of student learning is being met successfully. We strongly support the continuation of this project.

---

[1] The schools involved are the University of Massachusetts, Dartmouth; California Polytechnic State University, San Luis Obispo; Salt Lake City Community College; the U.S. Naval Academy; Pacific University; and Hunter College of the City University of New York.

[2] As an example, see R.K. Thornton and D.R. Sokoloff, “RealTime Physics: Active Learning Laboratory,” AIP Conf. Proc. 399, 1101-1118 and references cited therein.

[3] See R.K. Thornton and D.R. Sokoloff, “Assessing student learning of Newton’s laws: The Force and Motion Conceptual Evaluation and the Evaluation of Active Learning Laboratory and Lecture Curricula,” Am. J. Phys. 66, 338-352 (1998) and references cited therein.

[4] Note that gain is measured in the percent of possible movement toward a perfect score. Thus, students with a pre-instruction score of 10 and post-instruction score of 55 (with an improvement of 45 out of 90 possible points) have the same gain as students with a pre score of 90 and post score of 95 (5 out of possible 10). In both cases the gain is .50 or 50%, even though the pre- and post-instruction scores vary widely. For more information on the normalized gain and its value in comparing different schools and courses, see R.R. Hake, “Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses,” Am. J. Phys. 66, 64-74 (1998).

[5] H. Halloun, “Views about science and physics achievement: The VASS story,” AIP Conf. Proc. 399, 605-613 (1997).

[6] E.F. Redish, J.M. Saul, and R.N. Steinberg, “Student expectations in introductory physics,” Am. J. Phys. 66, 212-224 (1998).

[7] Note that by the end of the dissemination project, we expect to return to the year 1 data and carry out the analysis using the more appropriate methods.

[8] See J.M. Saul and E.F. Redish, “An Evaluation of the Workshop Physics Dissemination Project,” FIPSE Grant #P116P50026 (1998) for more details.

[9] J.M. Saul, “Beyond Problem Solving: Evaluating Introductory Curricula Through the Hidden Curriculum,” dissertation, University of Maryland, College Park, 1998.

[10] R.R. Hake, “Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses,” Am. J. Phys. 66, 64-74 (1998).

[11] K.C. Cummings, J. Marx, R.K. Thornton, and D. Kuhl, “Evaluating innovation in studio physics,” Am. J. Physics, 67 (supplement 1 to no.7), S38-S44 (1999).