No Single Cause: Learning Gains, Student Attitudes, and the Impacts of Multiple Effective Reforms

Steven J. Pollock

Department of Physics, University of Colorado, Boulder.

Abstract. We examine the effects of, and interplay among, several proven research-based reforms implemented in an introductory large enrollment (500+) calculus-based physics course. These interventions included Peer Instruction with student response system in lecture, Tutorials with trained undergrad learning assistants in recitations, and personalized computer assignments. We collected extensive informal online survey data along with validated pre/post content- and attitude-surveys, and long answer pre/post content questions designed to assess learning gains and near transfer, to investigate complementary effects of these multiple reforms, and to begin to understand which features are necessary and effective for high fidelity replication. Our average [median] normalized gain was 0.62 [0.67] on the FCI, 0.66 [0.77] on the FMCE, yet we find we cannot uniquely associate these gains with any individual isolated course components. We also investigate the population of students with low final conceptual scores, with an emphasis on the roles played by demographics, preparation, and self-reported attitudes and beliefs about learning.

INTRODUCTION

Physics education research has documented the need for, and value of, interactive engagement in physics instruction[1,2]. However, the effect of implementing multiple, layered, and coordinated activities is somewhat less well understood. To help facilitate sustainable, structural integration of such activities we investigate necessary features of individual reforms and the dynamics between them. We present here preliminary data from a case study in a large reformed mechanics course. Our initial research questions center on two issues. First, which elements of the course reforms are most closely associated with student learning gains? Second, even with such high average gains, there remains a population of students who fail to develop satisfactory conceptual understanding. Can we identify and characterize such "at risk" students, with a future goal of better targeting and accommodating them?

COURSE STRUCTURE

Physics 1110 is a large introductory calculus-based mechanics course at the University of Colorado, Boulder (CU) with 500+ students in three 50-minute lectures/wk, plus a smaller 50-min. recitation section (with no lab). Reforms implemented included the following: we used Peer Instruction[3] in lecture, with individual personal response systems. Roughly 30-50% of class time was spent on ConceptTests. Recitations were converted to engage in Tutorials[4], with grad TA's and undergrad learning assistants. Tutorial homework[4] was assigned, as well as traditional problems graded online with CAPA[5]. There were two texts: traditional and multimedia[6]. A unique feature of the lectures was an increased emphasis on epistemology and metacognition, making explicit what it means to learn physics, the interconnectness of topics, the value of collaboration, and the nature of science. Exams were 75% multiple choice, with a strong conceptual focus, and 25% long answer based closely on the Tutorial instructor's manual[4]. Exams accounted for 60% of the grade, CAPA for 15%, Tutorial homework 15%, Tutorial attendance 5%, and online participation questions 5%. The final grade distribution matched departmental norms: 18% A's, 37% B's, 34% C's, and 11% D/F's.

DATA COLLECTION

We collected extensive student data for two terms, including a pre/post conceptual survey (FCI in Fa'03, FMCE Sp'04), and weekly Tutorial pretests (provided
online by the U. Washington PEG), with post-tests
given as collaborative in-class questions, and later
variants on exams. We gave pre/post demographic
surveys, along with questions on attitudes and beliefs
about learning, and the nature of science (CLASS[7])
We gave frequent informal online surveys probing
students’ attitudes toward class components. (No data
were anonymous except for a SALG[8] survey, the
results of which closely matched our informal
surveys.) Students in engineering Calculus
(approximately half the class) took a math pretest.

RESULTS

We now have an extensive database including
individual results on pre/post tests, assignments,
exams, and surveys. We first summarize some key
initial findings related to the guiding questions of this
work. Both semesters, we measured strong learning
gains on the FCI/FMCE. (See Fig 1 for results from Sp
04. The pretest average on the FCI was higher but
normalized learning gains were similar[9].) These
gains lie near the high end of the distribution shown in
Hake[2] for interactive engagement courses, and are
particularly significant for a large class with no
laboratory and limited financial resources. Learning
gains for males (who comprise 3/4 of the class) and
females are comparable, although females started with
a statistically significant lower average pretest score
(over 10% lower). Their final course grades are
statistically identical.

The "Posttest in-class" column refers to clicker
questions in lecture, with peer collaboration allowed.
Typical semester to semester variance of data is ±5%.

<table>
<thead>
<tr>
<th>Topic</th>
<th>Pretest online</th>
<th>Posttest in-class</th>
<th>Posttest exams</th>
</tr>
</thead>
<tbody>
<tr>
<td>a at top of ramp</td>
<td>20</td>
<td>70</td>
<td>75</td>
</tr>
<tr>
<td>Newton II: tension in string/block series</td>
<td>30</td>
<td>85</td>
<td>75</td>
</tr>
<tr>
<td>Galilean relativity</td>
<td>35</td>
<td>35</td>
<td>75</td>
</tr>
</tbody>
</table>

The results in the first row of Table 1 are typical of
many topics covered in both lecture and Tutorials,
showing a strong, rapid learning gain from one lecture
to the next with just one Tutorial in between,
indicating the significant impact of the Tutorial. The
2nd row shows another common feature: individual
scores from exams on difficult topics are often slightly
lower than in-class results, where peer collaboration is
encouraged. This has also been verified on isolated
"work alone" in-class questions. The last row shows a
less common but interesting result, in which there is
no apparent immediate learning gain from the Tutorial,
but later class/peer discussion and/or homeworks
appear to have a measurable impact. This particular
topic was covered only in Tutorial (and later in
Tutorial homework), but not in lecture or reading until
the in-class question.

Comparison with U. Washington on common exam
questions [10] is shown in Table 2. These questions
(and others, not shown) demonstrate remarkable
similarity between primary (UW) and secondary (CU)
implementation of Tutorials. By this measure, our
implementation appears to be of high fidelity.

<table>
<thead>
<tr>
<th>Topic</th>
<th>UW no Tutorial</th>
<th>UW with Tutorial</th>
<th>CU with Tutorial</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atwood: tension before/after release</td>
<td>25</td>
<td>50</td>
<td>55</td>
</tr>
<tr>
<td>Atwood: describe constrained motion</td>
<td>45</td>
<td>70</td>
<td>45</td>
</tr>
<tr>
<td>Coupled objects: force diagrams</td>
<td>30</td>
<td>90</td>
<td>95</td>
</tr>
<tr>
<td>Identify N-III partners</td>
<td>15</td>
<td>70</td>
<td>70</td>
</tr>
</tbody>
</table>

Correlations

It has been well documented elsewhere that Peer
Instruction and Tutorials positively affect student
conceptual learning. One of our central research
We see in Fig. 2 that there is a strong correlation (Pearson r = .73) between Tutorial homework score and course grade (excluding Tutorial homework). Since this course grade includes traditional homework, participation, and different types of exam questions (including long answer/explanation) on a broad set of topics extending beyond tutorial materials, Fig. 2 demonstrates the strong coordination of the entire variety of course elements. This is again suggestive that no single element is associated with measured learning gains, but rather the gains arise from the coupled aspects of course components, and perhaps the framing of the course itself.

### Characterizing Students With Low Final Conceptual Exam scores

Despite strong average FMCE gains for the class, a considerable fraction of students do not achieve final scores above 50%, indicating a lack of Newtonian conceptual understanding[9]. This is seen in the long, flat "tail" of the post-test distribution (Fig 1), which includes roughly 15-20% of the class. These students are not generally the ones failing the course, but are more typically C students. (Many D/F students missed one or both of the classes in which the test was given.)

What is most significant about this tail is that their average homework scores, online participation, and attendance (effort based measures) are almost the same as students getting high FMCE posttests. That is, these students are apparently working as hard, participating at nearly the same level, and doing almost as well on homeworks, but are failing to master the conceptual topics measured by the FMCE. Further investigation shows little significant difference in physics backgrounds, or self-reported use of resources (e.g. reading the texts). The tail population has a marginally lower pre-class math diagnostic score, and self-reports a marginally higher amount of time spent on physics outside of class (5.7 hrs/week vs. 5.3 hrs/week).

Demographically, women are somewhat more represented in this tail (34% of the tail, vs. 25% for the overall class), and undeclared majors make up a larger fraction of this population (26% of the tail vs. 16% overall). Their average physics pre-test score is lower (13% vs. 32%), but this alone does not characterize the population, because many other students with low pretests do very well - indeed, students in the bottom quartile of pretests overall have a normalized gain of nearly 50%. The normalized gain of the "tail" group is 18% (vs. 78% for those not in the tail). We are naturally very interested in this population, because they are the students for whom our efforts at reforming the class are apparently not working well, yet they are not "slackers", and perhaps could be better helped in some other way. Further research is clearly essential.
Student Attitudes and Beliefs

Student attitudes and beliefs about the nature of learning, and of science, are interesting for a variety of reasons. Part of our goal is to help shift students' attitudes from "novice-like" to "expert-like", and attitudes may play a role in learning, performance, and appreciation of the course[11]. We use the CLASS survey as a rough measure of self-reported attitudes and beliefs (AB's). We find no regression in AB's over the course of a semester[7,12], but do find statistically significant differences in AB's between lower and higher performing students (See ref [7] for more detailed discussion). We make no claim regarding cause and effect, but the correlation is itself of interest.

E.g., the tail population described above with low final FMCE score began the course with an average overall CLASS score of 60±2% compared to 71±1% for high FMCE scorers. This can also be seen in Fig 3, where students with low FMCE learning gain entered the class with lower CLASS pre-scores, and ended even lower. Students with higher learning gains tended to start with more "expert-like" attitudes, and became more expert-like after the course.

FIGURE 3. Average pre/post overall scores on the CLASS exam[7] which provides some measure of student attitudes and beliefs about learning, and the nature of science. Bins refer to average normalized gain (g) on the FMCE.

CONCLUSIONS

As the physics teaching community becomes increasingly sophisticated at using and integrating multiple proven reforms, questions of how various course components interact and impact students become correspondingly richer and more significant. We have implemented three distinct reforms in a large lecture course: peer instruction, Tutorials, and automated homeworks, and have begun an analysis of how they contribute to student learning and student attitudes. The data shown here focus first on the direct results of the reforms, and on the relationship of course elements to learning gains. We find that no single element alone appears to account for the measured successes, but the combination of course elements has a large impact. We have also identified a population of students for whom this course structure is not effective, and have begun the work of identifying and characterizing them, with the goal of further improving the mix of classroom approaches. No single demographic identifies this population in advance. We do find that student attitudes and beliefs are associated with both grades and learning gains, and believe that further study of e.g. explicit attention to epistemology is warranted.

ACKNOWLEDGMENTS

The author thanks the Physics Education Group at U. Washington for their support as we implemented Tutorials, particularly Peter Shaffer for his exceptional help with data collection, pretests, and exams. Thanks also to the Pew/CASTL program, the NSF STEM-teacher prep grant which supports our learning assistants, CU Physics for its support in implementing Tutorials and Peer Instruction, and all the members of the Physics Education Research at Colorado group.

REFERENCES

5. CAPA homework system: see www.lon-capa.org
9. SALG, the Student Assessment of Learning Gains: www.wcer.wisc.edu/salgains/instructor/