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 .62 [0.67] on the FCI, .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.
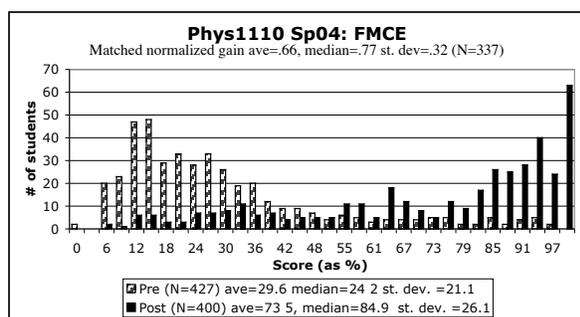
FIGURE 1. Pre/Post results for FMCE exam (Sp 04). Scoring of this exam follows Thornton [9]. All students who took the exam are shown in the histogram, but the quoted normalized gain includes only matched pre/post scores. The most frequent posttest score was 100%.

Table 1 shows learning gains as measured by questions targeting isolated, selected topics beyond those covered on the FCI or FMCE. The table shows a limited sample of our detailed data on topics matching the 12 Tutorials covered. Scores are averaged over both semesters, each with multiple related questions.

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 1. Average scores on selected conceptual topics targeted by Tutorials, rounded to nearest 5%

Topic	Pretest online	Posttest in-class	Posttest exams
a at top of ramp	20	70	75
Newton II: tension in string/block series	30	85	75
Galilean relativity	35	35	75

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 2. Comparison of specific long answer *exam* questions with primary implementation results from U. Washington [10] (Results rounded to nearest 5%)

| Topic	UW *no* Tutorial	UW *with* Tutorial	CU *with* Tutorial
Atwood: tension before/after release	25	50	55
Atwood: describe constrained motion	45	70	45 final: 75
Coupled objects: force diagrams	30	90	95
Identify N-III partners	15	70	70

Correlations

It has been well documented elsewhere that Peer Instruction and Tutorials positively affect student conceptual learning. One of our central research

questions is: what is the impact of layering these two methods? To this end, we have measured correlations between various course elements and learning gains on standardized (FMCE) exams, shown in Table 3.

TABLE 3. Pearson's correlation coefficients for normalized learning gain (g) on FMCE with specific course and curricular components (Data for Sp '04, results for Fa '03 (FCI exam) are similar)

Correlating gain with:	Correlation coefficient
FMCE pretest	.31
Math pretest (selected students)	.25
CLASS (attitude survey) overall score	.20
Tutorial homework	.22
Conventional homework	.14
Tutorial attendance	.13
In class (clicker) score	.11
Lecture attendance	.02
Average exam score	.63

Although there is a positive correlation of normalized learning gain with each of the specific course components, no single one dominates, nor are the correlation coefficients particularly high. It appears that one cannot isolate any single element (Tutorials, homeworks, attendance, or Peer Instruction) as strongly, individually coupled to conceptual learning gains as measured by the FMCE. The high correlation with average exam score may be a reflection of the conceptual focus of our exam questions, although the exams do cover a much broader range of topics (traditional introductory mechanics) than the FMCE does. But no single measurable course reform element, in isolation, appears to be uniquely associated with the high average learning gains in this course.

Correlating rest of course grade to tut hw
Correlation coefficient r=0.73 (N=562)
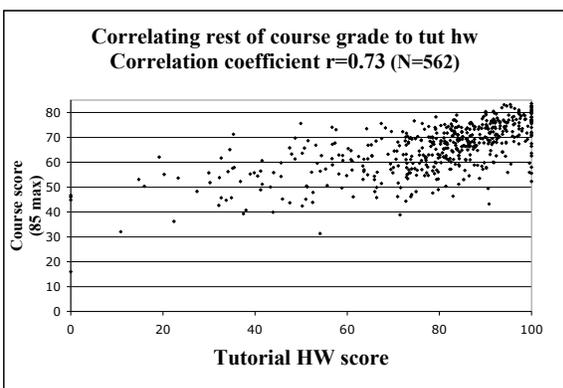
FIGURE 2. Relation of course grade (exams, conventional homework, attendance and participation) to Tutorial homework scores (Fa '03 data, Sp '04 is similar).

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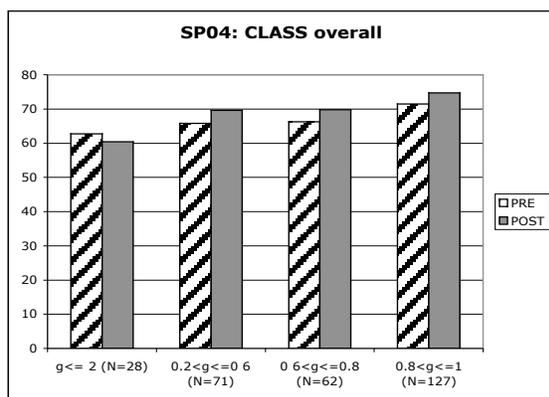
FIGURE 3. Average pre/post overall scores on the CLASS exam[7] which provides some measure of student attitudes and beliefs about learning, and the nature of science. Bins refer to average normalized gain (g) on the FMCE.

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

REFERENCES

1. McDermott, L., Redish, E., *American Journal of Physics*, **67**(9), 755-767, 1999, Redish, E., *Teaching Physics with the Physics Suite*, Wiley 2003
2. Hake, R., *American Journal of Physics* **66**, 64-74, (1998) Hake, R., Conservation Ecology **5**(2):28, (2002)
3. Mazur, E., *Peer Instruction,* Prentice Hall 1997
4. McDermott, L., Shaffer, P., and the PEG, *"Tutorials in Introductory Physics",* Prentice Hall 2002.
5. CAPA homework system: see www.lon-capa.org
6. Halliday, Resnick, Walker, *Fundamentals of Physics* Wiley 2004, and *Thinkwell Physics*: www.thinkwell.com
7. http://cosmos.colorado.edu/phet/survey/fall2003.html. W. Adams et al., "*Design and Validation of the Colorado Learning About Science Survey*", and K. Perkins et al., "*Correlating student attitudes with student learning using the CLASS*", both in review for Proceedings of the PERC 2004, Sacramento, CA.
8. SALG, the Student Assessment of Learning Gains: www.wcer.wisc.edu/salgains/instructor/
9. Thornton, R., Sokoloff, D., *Am. J. Phys* **B66**(4), 338-352, 1998, also Thornton, R., Kuhl, D., Cummings, K., Marx J., "*Comparing the FMCE and the FCI*", unpublished.
10. McDermott, L., Shaffer, P., Somers, M, *Am. J. Phys.* **62**(1), 46-55 (1994).
11. Elby, A. *Am. J. Phys* **69**(7), (2001)
12. Redish, Saul, Steinberg *Am. J. Phys.* **66** 212-224, (1998)