

*FTEP teaching improvement project '04. S. Pollock, N. Finkelstein, and M. Dubson.*

**Summary:** The FTEP summer teaching improvement project provided support for three faculty members to devote significant time (summer '04) to further develop and study a novel student-centered replacement for conventional recitations, called "Tutorials". Tutorials are based on materials developed over the last decade by the Univ. of Washington's physics education research group, modified to suit local conditions at CU. They are among the best investigated teaching interventions in introductory physics. This project coupled with several local education projects, including STEM-TP and our nascent PhysTEC program. This summer, we analyzed data collected over the first year of Phys 1110 Tutorials, created the infrastructure for implementing Tutorials in 1120, gathered/constructed materials needed for future faculty members to be able to implement tutorials at CU, and developed a new course Physics 4810/7810: Teaching and Learning Physics which, in part, supports the reforms in 1120 and 1140.

All three of us worked on all aspects, but we each took the lead in different areas. Prof. Dubson purchased, built, and tested equipment for the upcoming year of Phys 1120 Tutorials (which is more experimentally focused than 1110 was). Prof Finkelstein, in parallel, developed a new course on the Teaching and Learning of Physics, (running this semester, Fa '04) which provides a bidirectional link between education research and the practice of teaching in the Tutorials. Professor Pollock analyzed the course reforms in Phys 1110 and designed the implementation strategy for Tutorials in Phys 1120. The bulk of the work done by Prof. Dubson is best summarized by opening the cabinets in the basement tutorial room. Prof. Finkelstein's work is summarized at his course web page [www.colorado.edu/physics/phys4810/phys4810\\_fa04](http://www.colorado.edu/physics/phys4810/phys4810_fa04) and discussed below in the Appendix. Here, we focus on 1110 evaluation and 1110/20 material collection, as well as how these reforms are coupled.

Implementing Tutorials requires developing integrated activities, experiments, and simulations, as well as generating 14 week's worth of materials for training grad students and learning assistants. This development began summer/fall '03, and SJP has used the 1110 materials for two terms. These materials are now in place and available for anyone desiring to run 1110 tutorials in the future. The summer work allowed us to implement Tutorials in Phys 1120 for the first time this fall, and Prof. Dubson plans to continue this into the Spring of '05. We have begun assessing the 1110 tutorials in a variety of ways, including pre and post testing on the Force Concept Inventory (and a variant, the FMCE), on specific content areas, and in surveys of student attitudes. The 1120 sequence will be pre/posted tested with the BEMA ("Brief Electricity and Magnetism Evaluation") developed at NC State. The first implementation(s) of 1110 Tutorials went quite smoothly, and our normalized test gains rank among the highest in the nation. It is our ultimate goal that this work will help support and sustain longer term adoption of Tutorials (or some local variation) at CU.

**Results of this summer project: (See also the "Primer on Tutorials for CU Faculty" )**

**A) Available materials developed:** To run tutorials, the following are required:

1) A faculty member in charge who understands the structure, implementation and basis of the Tutorials. (UW runs workshops at AAPT. We could now do something roughly equivalent here. It is not trivial to understand what Tutorials are about and how to run them by just reading a simple, brief reference) The faculty member must run a *weekly* mandatory training session for the TA's, to prepare them for the week's tutorials. (This task is roughly equivalent, in workload for a faculty member, to a single recitation assignment) I am now prepared to provide training for this role, and am doing so with Prof. Dubson so that he will be able to run/use Tutorials next spring.

2) Physical space suited to small group work. We constructed two temporary spaces (G2B75 and 77) in the 1140 lab area this summer and last summer, with tables for 7 groups of 4 students. (1110 tutorials can be run in traditional classrooms, but it is awkward and undesirable.) 1120 has experiments which mandate this kind of space. We doubled the space this summer, adding temporary walls, whiteboards, storage cabinets, and student desks.

3) Materials for "hands on experiments" in several 1110 tutorials, and about half of the 1120 tutorials. These materials are gathered in storage cabinets in the temporary Tutorial area. (We have almost completed the 1120 materials, but some work is still in progress here)

4) Pretests. These are done by students every week before tutorial. We have collected word and/or pdf files of all the pretests used so far. We have used the UW server to let students do these online, and have established a relationship so this could (probably) be repeated in the future. If we institutionalize tutorials at CU, we may want to create our own online system to serve these. At the moment, these pretests are all available electronically from me, by request. Since the materials are copyrighted I am not putting them on the web.

5) Student pretest results. We now have pdf files containing student responses to the pretests for Phys 1110, and we are collecting them for Phys 1120. These are an essential part of the TA preparation meeting: TA's look over student responses to get a sense about the preconceptions coming into section. (See comments at end of #4 above)

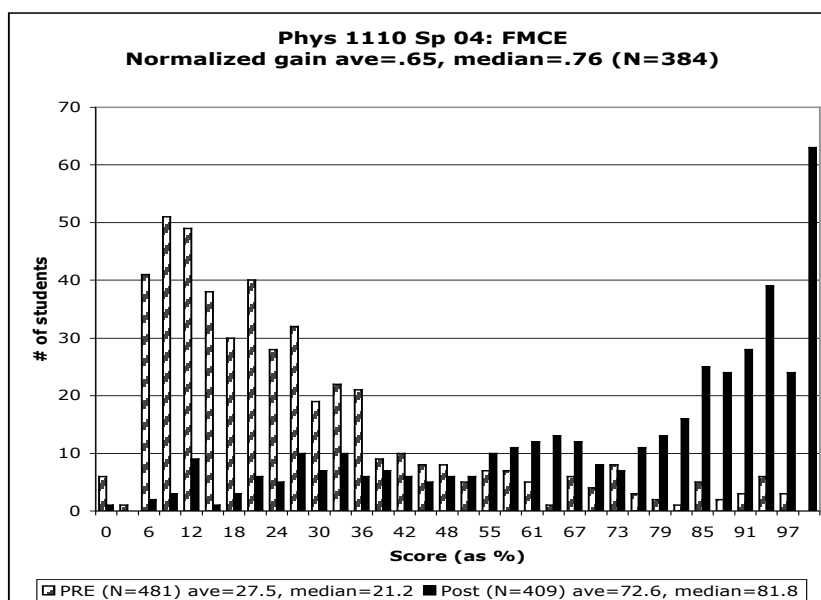
6) TA meeting notes. We have word docs containing notes (Introduction, general instructions, and then week by week homework assignments) This is complete for 1110, and is under construction for 1120. (Ditto on comments above)

7) Learning Assistants. (LA's) Tutorials are student centered, and require a smaller student/teacher ratio than our usual 28:1. We solved this by hiring undergrad LA's to work with the usual grad TA's. They have been funded by the STEM-TP project, which supports innovative pedagogy and preparation of future teachers. They provide financial support and training. The LA's attend weekly prep meetings with the TA's. We have established strong ties with the STEM project, and can provide LA's for 1-2 courses each semester for the next 2 years. Noah Finkelstein's Phys 4810 course has, among its goals, the intent of providing a "pool" of LA's which can supplement and ultimately perhaps replace the STEM-TP project to allow a longer term sustainable supply of prepared Learning Assistants. This is work-in-progress, and will be an essential ongoing result (whether it proves to be possible, or not!)

**B) Assessment:** Here are some of the early research results from 1110 tutorials. (Much of the material below has been presented at the AAPT/PERC meeting, Aug 2004)

1) Data collection and primary results: We collected extensive student data for two terms, and continue to do so this fall ('04). Data includes a pre/post conceptual survey (FCI in Fa'03, FMCE Sp'04, BEMA Fa'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 (this is the CLASS survey developed by the PER@CU group) We gave frequent informal online surveys probing students' attitudes toward class components. (No data were anonymous except for a SALG survey, the results of which closely matched our informal surveys.) Students in engineering Calculus (approximately half the class) took a math pretest.

We now have an extensive database including individual results on pre/post tests, assignments, exams, and surveys. Both semesters, we measured strong learning gains on the FCI/FMCE. (See Fig 1 for results from Sp 04. The pretest and posttest averages on the FCI were higher but normalized learning gains were similar.) These gains lie near the high end of the distribution shown in Hake's AJP article 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 significant lower average pretest score. (Their final course grades are statistically identical.)



**FIGURE 1.** Pre/Post results for FMCE exam (Sp 04). Scoring of this exam follows Thornton's (unpublished) work. All students who took the exam are shown in the histogram, but the normalized gain includes only matched pre/post scores.

Table 1 below shows learning gains as measured by questions targeting isolated, selected topics beyond those covered on the FCI or FMCE. The table shows a very limited sample of our detailed data on topics matching the 12 tutorials covered. Scores are averaged over both semesters, each with multiple related questions. Typical semester to semester variance is  $\pm 5\%$ .

**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

These are just samples - we have many more results of this type, which I can provide in a separate file. (There are roughly 3-6 conceptual entries for *each* tutorial, for a dozen tutorials, so the complete data tables span many pages) The results in the first row of Table 1 are typical of many topics covered in *both* class 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 allowed). 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 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%)**

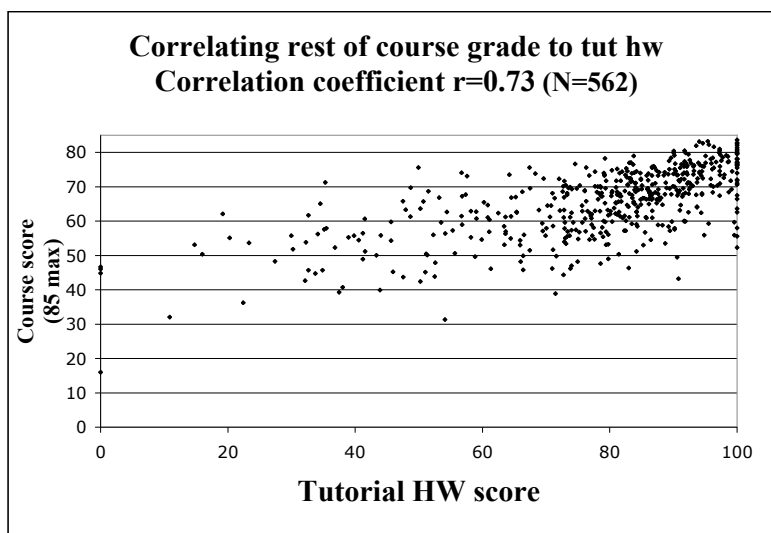
Topic	UW <i>no</i> Tutorial	UW <i>with</i> Tutorial	CU <i>with</i> Tutorial
Atwood: tension before/after release	25	50	55
Atwood: describing constrained motion	45	70	45 (75 on final)
Coupled objects: force diagrams	30	90	95
Identify N-III partners	15	70	70

2) Correlations: It has been well documented elsewhere that Peer Instruction and Tutorials positively affect student 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 measured learning gains on standardized (FMCE) exams, shown in Table 3.

**TABLE 3. 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.



**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 ( $r=.73$ ) 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

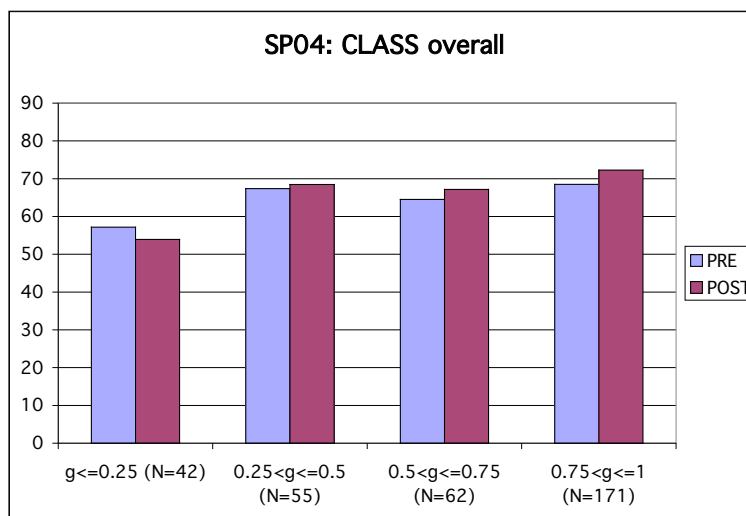
3) Focus on low performing students: Despite strong average FMCE gains for the class, a considerable fraction of students do not achieve final scores above 50%, indicating a lack of Newtonian conceptual understanding. This is seen in the long, flat "tail" of the post-test distribution (Fig 1), which 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 students with such low pretests do very well - indeed, low pretest students 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. In this regard, we are working on an assessment of "at risk factors" that might help identify this population at the start by combining physics pretest, math test, and CLASS attitude results.

4) 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. 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 (which is itself an unusual result, published data from similar surveys tend to show a decline in students' "expert-like" responses after an introductory course) We do find statistically significant differences in AB's between lower and higher performing

students. 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  $60 \pm 2\%$  compared to  $71 \pm 1\%$  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.

**C) Summary:** This summer FTEP project was a key component in a larger effort involving Physics 1110/20, STEM-TP and PhysTEC. We are attempting to implement and study Tutorials in introductory physics, creating a sustainable improvement in the structure of our large intro calculus based courses that could be taken over by other faculty members. This FTEP summer support provided us resources to expand our effort to second semester Phys 1120, and take time to analyze some of the data we have been collecting. We were able to build tools - physical (the room, the equipment), assessment (the pre/post tests, the TA prep notes) and curricular (the Phys 4810 materials) to help develop and support future use of the Tutorials. The data shown above strongly supports the conclusion (established elsewhere) that Tutorials are pedagogically effective, and we are encouraged to continue development. Based on survey data (not reported above: it hasn't been compiled in a simple form yet) student reactions are mixed, some are unhappy about these reforms. On the basis of this (low level, but not negligible) student discontent, coupled with the large amount of time and effort still required on the part of the professor teaching a reformed course using Tutorials, we do not feel that we are ready to "institutionalize" Tutorials here at CU yet. But, to a large extent due to the effort this summer, clear progress is being made, and deeper integration into the departmental "norms" seem possible in the future. We would like to create a robust, reproducible local Tutorial infrastructure, and at that time effectively expand the practice to faculty who have considerably less time and background knowledge in education, which is ultimately what would be required for these reforms to take hold at CU.

Some useful references:

Summary of PER results for intro physics:

1. McDermott, L., Redish, E., *American Journal of Physics*, **67**(9), 755-767, 1999,  
Redish, E., *Teaching Physics with the Physics Suite*, Wiley 2003

Statistical demonstration of effectiveness of active engagement:

2. Hake, R., *American Journal of Physics* **66**, 64-74, (1998)  
Hake, R., *Conservation Ecology* **5**(2):28, (2002)

The Tutorial materials themselves:

3. McDermott, L., Shaffer, P., and the PEG, *"Tutorials in Introductory Physics"*, Prentice Hall 2002.

Evaluation of the Tutorial materials:

4. McDermott, L., Shaffer, P., Somers, M, *Am. J. Phys.* **62**(1), 46-55 (1994).

The CLASS instrument:

5. <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.

The FMCE Instrument:

6. 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.

## APPENDIX: Physics 4810/7810 Teaching and Learning Physics.

The new course syllabus can be found on the Phys4810 web page. It describes the course goals and approach. Preliminary evaluation of the course suggests that it is already demonstrating success for increasing student involvement in the Phys 1120 program (as LA's), increasing student interest and understanding of teaching and education in physics and physics education research, and suggestive that students are learning more physics content. The course has been modified from its original design to couple strongly with Physics 1120. The physics content and pace parallels that of Phys 1120 and students enrolled in Phys 4810/7810 participate in the weekly homework assignments (on LON-CAPA) that are assigned to the Phys 1120 students.

The course currently has 23 students enrolled with additional faculty / staff sitting in as participants. The class is split roughly 50/50 between undergraduates and graduate students. All students engage in fieldwork opportunities. Two thirds of the class (about 15) are engaged in teaching opportunities on campus (7 in Phys 1120, others in 1020, APS and the Fiske planetarium). About 8 students are engaged in fieldwork opportunities in local community environments (Boulder High School, Casey Middle School, Stanley Lake High School, and elsewhere). All students enrolled in the course participated in both the BEMA and CLASS evaluations discussed above. Results will be collected and analyzed at the end of term.