Tutorials in Introductory Physics: The Pain and the Gain

Émerson Cruz, Brian O'Shea, Werner Schaffenberger, Steven Wolf, and Gerd Kortemeyer, Michigan State University, East Lansing, MI

In an introductory physics sequence with a large enrollment of premedical students, traditional recitation sessions were replaced by *Tutorials in Introductory Physics,* developed by the Physics Education Group at the University of Washington. Initially, summative test scores (as well as FCI scores) dramatically increased, but so did student complaints and workload. Both effects decreased over time. The paper discusses issues that instructors should consider when contemplating implementation of the tutorials.

The series is composed of a workbook with in-class exercises,[1] a book with homework assignments,[2] and an instructor's guide[3] that includes pre-tests and suggested exam problems (it is not immediately obvious that this guide is available; it needs to be ordered directly from the publisher). Some of the in-class exercises use simple equipment, such as cardboard boxes, balls, mirrors, and light bulbs. The in-class work is not graded, but is designed to begin the students' Socratic learning process. The tutorial homework is very much in the style of the in-class tutorials, but meant to be solved individually.

Tutorials are meant to be a supplement to traditional algebra- or calculus-based courses, not a full curriculum or textbook. A unique feature of the tutorials is that they are based on extensive research regarding students' understanding of physics concepts (e.g., Refs. 4 and 5). Yet, at the same time, the tutorials are not "out there"—the presented physics is solid and exactly what most physics instructors desire their students to master.

The course

Rather than trying to make general statements about the tutorials, we describe our experience in a particular course with a particular student population and staff structure. Our 250-student introductory calculus-based physics sequence is run on a tight weekly schedule, with the standard mechanics topics plus sound and thermodynamics in the first semester, and electricity and magnetism plus selected topics in relativity and quantum mechanics in the second semester.

The course has three lecture hours plus one recitation hour per week. Course materials and homework are made available online using LON-CAPA,[6] and students are expected to read the material before the first lecture on a given topic. Through embedded content-level questions and attached discussion boards, Just-In-Time Teaching is implemented.[7] During lecture, peer teaching and "clickers" are used.[8] After each lecture topic, homework questions are due online.[9] The course offers several help room hours over the course of the week, staffed by undergraduate learning assistants (LAs) and one or two graduate teaching assistants (TAs). Some of the LAs are physics majors, though most simply took our non-majors course in previous years and received high grades.

Instead of large-scale midterm exams, the first semester of the sequence had six tests, typically covering two topics (e.g., energy and momentum conservation). The second semester had three traditional midterm exams. At the end of both semesters, a traditional cumulative final exam was given. Due to the large size of the course and limited grader availability, all tests and exams were multiple choice, with roughly half of the questions being numerical and the other half conceptual ("what if?" questions, ranking, graph interpretation, etc). The grade is based on 40% midterms, 20% final, 20% homework, 15% recitation, and 5% in-class "clicker" questions. A co-requisite laboratory course accompanies the lecture course.

The students

The typical student in our course is a sophomore or junior life science major who plans to go on to medical, veterinary, or dental school. The students' background knowledge in physics is low, and typical Force Concept Inventory (FCI)[10] pre-test scores are 35%. Also, their percentage of favorable answers on the Maryland Physics Expectations survey (MPEX)[11] is low (48%), with a particularly low score in the independence cluster (38%), which indicates that their approach to physics is very different from that of expert physicists. The independence cluster answers show that, on average, students value information given by authority (instructor and materials) over learning independently. Typical statements that the students react strongly to from this cluster are: "In this course, I do not expect to understand equations in an intuitive sense; they must just be taken as givens," and "My grade in this course is primarily determined by how familiar I am with the material. Insight or creativity has little to do with it." This pattern has been consistent over at least the past three years.[12] Experiences with other student populations and class settings have been reported in Refs. 13-15 and can vary greatly.

Previous recitation format

Previously, recitation was composed of instructor-selected exercises related to the topic of the week as well as small quizzes. The sessions were led by LAs, and students turned in their quizzes and exercises immediately at the end of each session.

This format was largely unsuccessful: when putting a recitation problem on a test, success rates were frequently low, even though the recitation grades were high. In one case, only 20% of the students successfully solved a recitation problem when it appeared on a test, even though the vast majority had full points on the original recitation quiz. Students frequently described recitations as "useless" or "a joke" in course evaluations.

Introducing the tutorials

Tutorials were introduced to replace the previous recitation materials. The nine tutorial sessions, spread over the course of the day with up to 28 students each, were led by undergraduate LAs. Institutional policies do not allow LAs to lecture; this coincides with the philosophy of the tutorials that assistants in the room are there simply to facilitate and support collaborative work. All course staff participated in weekly preparation meetings, led by the TA.

Students were required to work through the tutorials in groups of three or four, get a receipt for their attendance in the session, and turn in the associated tutorial homework for grading at a later time. A key was prepared by a TA, applied to the homework by the LAs, and the graded homework was returned the following week.

The instructor's guide recommends using problems and scenarios from the tutorials on tests and exams so the students perceive them as relevant. It has always been our practice to include recitation scenarios on tests, and we continued to have a few tutorial-based (albeit multiple-choice) questions on every exam.

Purchasing and building the hands-on equipment introduces some cost in both time and funds, with the latter being approximately $1200 initially for this size of course for the tutorials we chose. We anticipate that there will be annual costs of a few hundred dollars a year to replace lost or damaged equipment and consumables. Rarely can a new curriculum be implemented at such a low cost for a course of this size.

The pain

Aligning the tutorials with the lecture topics was challenging. We go through basic kinematics and dynamics within the first four weeks of the course, yet eight out of the 20 tutorials deal with these introductory topics. In the second semester, topics such as RLC circuits and relativity are absent.

Using undergraduate non-physics major LAs to lead the sessions presented a significant challenge, as it soon became apparent that many of them were overwhelmed by the tutorials. We had to schedule additional weekly preparation time and also scheduled volunteer graduate students, post-docs, and faculty to participate in recitations as additional helpers. In the second semester, we modified teaching schedules so that we could have two LAs per session, which improved LA confidence and alleviated the need for the volunteers. The additional staff requirements could have been anticipated had we read Finkelstein and Pollock, for example.[13]

Very different from the lively discussions reported in Ref. 14, the sessions were often depressing, as many groups just sat quietly and waited for staff to come to their tables and help them out; little active learning occurred. As LAs followed directions and did not simply give away solutions, students complained to course faculty that "their peers" (i.e., other undergraduate students) intentionally withheld necessary information.

The tutorial materials do an excellent job of providing opportunities for students to check their thinking for internal consistency. Often the same scenario is approached from two different angles, and possible inconsistencies and logical errors are carefully exposed. Unfortunately, many groups simply work linearly through the material and hardly ever go back and check their previous statements.

Grading is work-intensive: some of the homework assignments took an average of eight minutes to grade. In a 250-student course, that is more than 30 hours of grading, and we needed to hire additional grading staff within the first two weeks of the first semester (getting "bailed out" by our dean). It is also rather unfortunate that the homework assignments are of varying lengths, ranging from 2 to 7 pages, which means very different grading workload between weeks. Some of the assignments are simply too long, for both graders and students. This was moderated in the second semester by assigning a subset of problems on larger assignments, though this decreased the ability of students to check their answers for consistency. Overall, the increase in workload appears to be typical.[15]

Recitation grades went down considerably: the average recitation score went from 92% in the previous year to 71% this year. Naturally, students complained, and some stated that *all* the tutorials would do is lower their GPA; in previous years, recitations in this course had the reputation of "diligence points."

Our student comments regarding the tutorials were different from the favorable evaluations reported in Ref. 14, the mixed reactions reported in Ref. 15, and even worse than the unfavorable evaluations in Ref. 13: the most positive comments were that the tutorials are "maybe ultimately" or "very, very slightly" helpful. In the evaluations, population and class standing effects seem to be very pronounced. The majority of the complaints correspond to the students' epistemology, as indicated by the MPEX; in fact, the expert-like tutorial philosophy apparently goes against the grain of what our students expect regarding teaching and learning. Several students stated that tutorials are not helpful because the correct answer is never given to them, and that their fellow students are just as "clueless" as they are, and thus of no help during the discussions. Students claim that, as they got things wrong during the in-class sessions, their mistakes propagated into their homework. They also asked for a key with the correct solutions after they turn in their homework, to help them prepare for the tests; this is not uncommon,[15] and once again shows that students are not taking advantage of the built-in "checkpoints," but instead strongly rely on authority. Other epistemological challenges become apparent in complaints that recitations are "pointless," because "we are not taught any equations and logic." Informal student interviews show that the tutorials have not changed the way that students think—the tutorials are

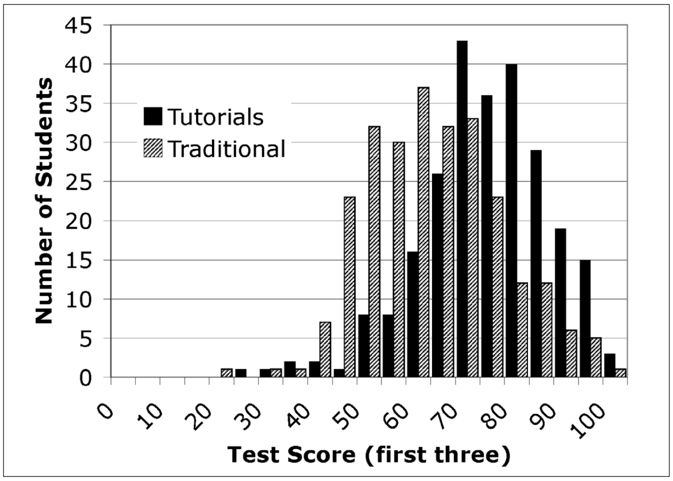

Fig. 1. Score distribution on the first three tests in a previous year and this year.

written by experts and attempt to mold students into an expert mode of thinking (understanding the underlying physics before the mathematics is added). As students got hung up on surface features of the presented scenarios, they often failed to see the connection and progression between them.

The number of students who dropped the course between the first day and the last day where it could be dropped without academic penalty (albeit with loss of tuition) more than doubled between last year (six students) and this year (14 students). Several students stated that they would switch to an alternative physics course between semesters, and cited tutorials as the reason for leaving the course. Indeed, 18 students who completed the first semester did not enroll in the second semester of our course, as opposed to six the previous year (a 7.8% intra-semester attrition rate as opposed to 2.4% in the previous year). In a subsequent year, less than 220 students enrolled in our course, which typically enrolls 250. More than 30 students apparently chose an alternative, highly traditional physics course.

The gain

The introduction of the tutorials was the only major change from the previous year implemented in the structure of the course sequence; most other aspects were the same, though a new faculty member taught the second semester.

Due to logistics and time constraints, no FCI pre-test could be given for the mechanics semester. A comparison of the score distribution of the FCI post-test between the previous year and this year showed a statistically strongly significant increase from 48 percent to 55% ($p < 10^{-5}$, $n = 244$ and 252, respectively). Also, the post-test "failure" percentage (less than 7 out of 30 points) strongly decreased from 6.3% to 1.2% (while we had increased attrition, drops happened after the data we are considering here was gathered). This effect on the FCI appears to be comparable to Ref. 13 and much stronger than reported in Ref. 16. The correlation between tutorial homework and FCI scores, though, is surprisingly weak ($R^2 = 0.15$), which may be due to a high degree of close collaboration and copying on the homework (Ref. 13 also reports low correlation of tutorial homework scores with other performance measures).

The increase in average test scores on the first three tests was even higher, from 60% last year to 72% this year ($p < 10^{-19}$, $n = 256$ and 250, respectively); Fig. 1 shows both distributions. The correlation between tutorial homework and test scores ($R^2 = 0.32$) is much stronger than the correlation with FCI scores. Disappointingly, later exam scores approached previous years' averages.

The course is graded on an absolute scale, without a curve. Due to the relative weighing of the recitation scores (which went down) and test scores (which went up), slightly higher overall grades were observed in the mechanics semester than in the previous year. Grades in the second semester were comparable to previous years.

Additional improvements in student learning, primarily in the area of reasoning, may have gone unnoticed, as neither the FCI nor multiple-choice exams would reveal them, and they might have gotten lost in the general and always-present decrease of MPEX "scores." Detailed student interviews would be needed to appropriately assess these.

A somewhat mixed blessing is how grading the tutorial homework reveals the persistence of physics "myths," such as the infamous "force of the velocity" that shows up in free-body diagrams. It is painful to the instructors to see that in spite of what they believe to be their best efforts, they were unable to debunk such notions.

A final benefit to the implementation of the tutorials in this course sequence is that, as also reported in Ref. 15, the course personnel, including instructors, deepened their understanding of basic physics during the preparation sessions. This was especially evident in the LA training sessions, where we regularly had lively discussions and an active learning process —even at 9:00 in the morning!

Discussion

It is rare to make any changes to an introductory physics course for premedical students that make a major difference in student learning, and almost unheard of that a single change between two years of the same course would have such a significant impact on student performance. In this course, the very same change also resulted in an astonishing increase in student complaints, apparent frustration, additional materials and staff cost, attrition, and increased workload on instructors and volunteers, only to see that the effect eventually decreased. This leaves us with a number of open questions:

- ***What was the reason for the initially strong performance increase?*** It could easily be understood if the recitation sessions were as active as our preparation sessions; alas, they were mostly far from it: many students did not engage in active collaborative learning. More often than not, they appeared to completely "miss the point." Still, the performance increase shows that somehow tutorials work in spite of many students not really taking advantage of them. Some side effects are increased time-on-task (some students reported spending three hours per week on tutorial homework) and presumably increased studying due to insecurity (or lack of a false sense of security).

- ***How can the initial success be sustained?*** Increased familiarity (and false comfort level) as well as copying of solutions may have subverted the active learning process. As one of the TAs put it, the students eventually "fixed the problem of tutorials" by mainstreaming them into their "efficient" way of surviving the course—the tutorials had "worn out." It is clear that the initial shock effect of the Tutorials is unsustainable, and that long-term success depends on a broad-based change in staff and student attitude toward acceptance of active learning.
- ***How can we foster active learning?*** In most cases, non-physics major undergraduate LAs should not be the facilitators of these sessions in the first place, as their grasp on physics is generally not firm enough for them to be effective Socratic teachers. More time needs to be spent training staff to raise their comfort level with both the material and the teaching method. Furthermore, all staff need to understand the point of the tutorials and take them seriously—otherwise, this will sabotage the learning process. In the future, we plan to have more intense initial training for our LAs on how to effectively facilitate a tutorial session, and to periodically revisit this during our weekly meetings.
- ***How can we get the students to "buy into" the active learning process?*** It may be difficult, if not impossible, to completely change the students' epistemology within our current course structure and limitations; however, we certainly could have improved the implementation of the tutorials. It is critical (particularly for premedical students) to demonstrate that doing the tutorials increases their understanding of the material and, as a result, their exam scores—we did this only once at the beginning of the second semester.
- ***How can this be achieved without ignoring student concerns and worries?*** The pre-tests and in-class introductions to the tutorials are an avenue that we should explore in the future, even at the expense of "clicker" questions and lecture content. The call for eventually giving out the "correct answers," though, cannot be accommodated: if the students expect to get them at the end of the session, they will not engage in any discussion with each other, and homework keys would be handed down to next year's students. To partly accommodate this desire, we started producing "Tutorial Guides," in which we tried to point out the deeper physics that may otherwise never be recognized by the students.

In summary, within our introductory course sequence the tutorials led to a significant, albeit temporary, increase in student performance, but at a high cost. They cannot simply be used as a drop-in replacement for ineffective recitation materials, but require a restructuring of several other course components. Since (at least with premedical students), the introduction of tutorials leads to increases in student complaints and attrition, the department chair or dean should be involved in their introduction, especially if the instructor is pre-tenure. Is the gain worth the pain? We decided to continue using tutorial-style sessions, as we cannot, in good conscience, go back to an ineffective recitation format.

Reference

1. Lillian C. McDermott and Peter S. Schaffer, *Tutorials in Introductory Physics* (Prentice Hall, Upper Saddle River, NJ, 2001).
2. Lillian C. McDermott and Peter S. Schaffer, *Tutorials in Introductory Physics Homework* (Prentice Hall, Upper Saddle River, NJ, 2001).
3. Lillian C. McDermott and Peter S. Schaffer, *Tutorials in Introductory Physics Instructors' Guide* (Prentice Hall, Upper Saddle River, NJ, 2001), available from the publisher on request.
4. Lillian C. McDermott and Peter S. Schaffer, "Research as a guide for curriculum development: An example from introductory electricity. Part 1: Investigation of student understanding," *Am. J. Phys.* **60**, 994–1003 (Nov. 1992); "Part 2: Design of Instructional Strategies," *Am. J. Phys.* **60**, 1003–1013 (Nov. 1992), and "Erratum, 'Research as a guide for curriculum development: An example from introductory electricity. Part I: Investigation of student understanding," *Am. J. Phys.* **61**, 81 (Jan. 1993).
5. Lillian C. McDermott, Peter S. Shaffer, and M. D. Somers, "Research as a guide for teaching introductory mechanics: An illustration in the context of Atwood's machine," *Am. J. Phys.* **62,** 46–55 (Jan. 1994).
6. www.lon-capa.org/.
7. G. M. Novak, E. T. Patterson, and A. D. Gavrin, "Just-in-Time Teaching: Blending active learning with web technology," (Prentice Hall, Upper Saddle River, NJ, 1999).
8. C. H. Crouch and E. Mazur, "Peer Instruction: Ten years of experience and results," *Am. J. Phys.* **69**, 970–977 (Sept. 2001).
9. Gerd Kortemeyer, Edwin Kashy, Walter Benenson, and Wolfgang Bauer, "Experiences using the open-source learning content management and assessment system LON-CAPA in introductory physics courses," *Am. J. Phys.* **76**, 438–444 (April/May 2008).
10. D. Hestenes, M. Wells, and G. Swackhamer, "Force Concept Inventory," *Phys. Teach.* **30**, 141–158 (March 1992).
11. E. F. Redish, R. N. Steinberg, and J. M. Saul, "Student expectations in introductory physics," *Am. J. Phys.* **66**, 212–224 (March 1998).
12. Gerd Kortemeyer, "The challenge of teaching introductory physics to premedical students," *Phys. Teach.* **45**, 552–557 (Dec. 2007).
13. N. D. Finkelstein and S. J. Pollock, "Replicating and under-

standing successful innovations: Implementing tutorials in introductory physics," *Phys. Rev. ST-PER* **1**, 010101 (2005).
14. R. Steinberg and K. Donnelly, "PER-based reform at a multicultural institution," *Phys. Teach.* **40,** 4–10 (Jan. 2002).
15. Heidi V. Mauk and Dan Hingley, "Student understanding of induced current: Using tutorials in introductory physics to teach electricity and magnetism," *Am. J. Phys.* **73,** 1164-1171 (Dec. 2005).
16. Mel S. Sabella, "Implementing Tutorials in Introductory Physics at an Inner-City University in Chicago," *Physics Education Research Conference Proceedings* (PERC), 79-82 (2002).

Gerd Kortemeyer *is an associate professor of physics education and the director of the LON-CAPA Project at Michigan State University.*

Lyman-Briggs College, Michigan State University, East Lansing, MI 48825; korte@lite.msu.edu, www.lite.msu.edu/kortemeyer

Brian O'Shea *is an assistant professor of physics and astronomy at Michigan State University.*

Émerson Cruz *and* **Steven Wolf** *are graduate students of physics education;* **Werner Schaffenberger** *(when he is not helping out with tutorials) is a postdoctoral researcher in solar dynamics.*